Mathematical economics using Mathcad and Excel

Salmanov, Oleg

Moscow Region University of Technology (UNITECH)

7 July 2003

О. Н. Салманов

Математическая экономика с применением Mathcad и Excel

УДК 681.3.06
ББК 32.973.26-018.2
С16

Салманов О. Н.

Математическая экономика с применением Mathcad и Excel. — 2003. — 464 с.: ил.

ISBN 5-94157-262-X

В книге дано систематическое изложение основных базовых математических методов, применяемых в экономике. Приведены общая методология использования математического инструментария и математических моделей в экономике, а также конкретное изложение основных математических понятий и методов: применение дифференциального исчисления в экономике, матричная алгебра, интегральное исчисление, дифференциальные уравнения, основы корреляционно-регрессионного анализа и другие методы эконометрики, математическое моделирование и анализ экономических процессов, методы оптимизации и решение оптимизационных задач. Отличительной особенностью книги является соединение изучения математических методов и использования для их применения электронных таблиц Excel и математической системы Mathcad — систем, наиболее подходящих для экономических расчетов и взаимно дополняющих друг друга.

Для специалистов в области экономики, бизнеса и управления и всех, изучающих экономику

УДК 681.3.06
ББК 32.973.26-018.2

ISBN 5-94157-262-X

Содержание

Введение

Современная экономическая наука характеризуется широким использованием математики, статистики и эконометрики. Расширенное применение математических методов в последнее десятилетие обусловлено также распространением персональных компьютеров и их широким применением в экономической практике. Применение математических методов в единстве с экономическим анализом открывает новые возможности для экономической науки и практики.

При изучении различных экономических явлений используют их формальные описания, называемые экономическими моделями. Формализация основных особенностей функционирования экономических объектов позволяет оценить возможные последствия и использовать такие оценки в управлении, в частности предсказывать будущее поведение объекта при изменении каких-либо параметров. Для любого экономического объекта возможность прогнозирования ситуации означает получение лучших результатов или избежание потерь.

Важнейшей проблемой в управлении предприятиями является своевременное принятие правильных решений в связи с изменениями экономической ситуации. В то же время экономические системы и объекты могут быть представлены в достаточно строгой математической форме, т. е. формализованы. Поэтому возникающие в практической деятельности предприятий ситуации могут быть смоделированы, а варианты наиболее целесообразных решений по их управлению могут быть получены из анализа результатов моделирования.

Овладение методами моделирования экономических систем и принятие на их основе оптимальных решений по управлению предприятием является необходимым условием обеспечения эффективности их функционирования.

Любое экономическое исследование всегда предполагает объединение теории (экономической модели) и практики (статистических данных). Теоретические модели используются для описания и объяснения наблюдаемых процессов, а статистические данные – для эмпирического построения и обоснования моделей.

Известно, что существует большой разброс в математических знаниях как среди студентов, так и среди экономистов, работающих на предприятиях. Поэтому при изучении или желании применить математические методы приходится обращаться к учебникам математики, статистики и эконометрики. Однако больший объем материала в них посвящен доказательствам теорем, что естественно с точки зрения науки математики, но затрудняет использование этих знаний в моделировании. В этой книге приведены основные сведения по дифференциальному, матричному, интегральному исчислению в сжатом виде, без доказательств.

Экономисты используют количественные данные для наблюдения за ходом развития экономики, ее анализа и выдачи прогнозов. Набор статистических методов, используемых для этих целей, называется в совокупности эконометрикой. Для успешного применения указанных методов требуется правильное моделирование поведения экономических агентов; необходимо также понимание процессов, породивших имеющиеся данные, и того, насколько эти данные отражают явления, которые мы пытаемся исследовать. *Поскольку наши модели неполны,* а данные несовершенны, значительная часть эконометрики посвящена методам, которые могли бы работать с такими моделями и данными. Качество моделей и данных, а также то, как мы их используем, определяет результаты нашего анализа.

В книге приведены современные базовые идеи и методы эконометрики. Не зная хорошо этого предмета, не владея его инструментарием, невозможно получить эмпирические зависимости, а значит, и выдвинуть новые теории. Без эконометрических методов нельзя построить сколько-нибудь надежного прогноза, а следовательно под вопросом будет успех в бизнесе. Методы эконометрики являются инструментами для изучения окружающего экономического мира.

Кроме желания использовать компьютер для решения экономических задач, требуются знания по методам оптимизации, моделированию экономических систем, информационным системам, их возможностям. Необходимо иметь возможность изучить примеры решения задач и моделирования с применением информационных систем.

В настоящей книге рассматриваются две самые массовые информационные системы – Excel и Mathcad, а также их применение в экономико-математическом моделировании. Excel является самой простой в использовании и широко распространенной системой. Mathcad по праву может называться современной, универсальной и массовой математической системой. Она позволяет выполнять как численные, так и аналитические (символьные) вычисления, имеет удобный математико-ориентированный интерфейс. Исключительно велика роль таких систем в образовании. Они не только облегчают решение сложных математических задач и снимают психологический барьер в изучении математических методов, но и позволяют перейти к реальному моделированию экономических систем.

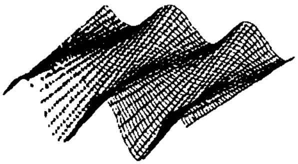
В книге дано систематическое изложение основных базовых математических методов, используемых в экономике. Приведены общая методология использования математического инструментария и математических моделей в экономике, а также конкретное изложение основных математических понятий и методов.

Отличительной особенностью книги является соединение изучения математических методов и использования для их применения электронных таблиц Excel и математической системы Mathcad – систем, наиболее подходящих для экономических расчетов и взаимно дополняющих друг друга.

В первую очередь книга будет полезна студентам, изучающим математические методы и моделирование экономических систем, математическую экономику и эконометрику, аспирантам любой специальности, а также специалистам по прикладной экономике.

Глава 1. Дифференциальное исчисление в экономическом анализе

1.1. Функции в экономическом моделировании



Многочисленные наблюдения и исследования показывают, что в окружающем нас мире величины (например, цена какого-либо товара и величина спроса на этот товар, прибыль фирмы и объем производства этой фирмы, инфляция и безработица и т. п.) не существуют изолированно друг от друга, а, напротив, они связаны между собой определенным образом. Понятие функции, или функциональной зависимости, – одно из основных математических понятий, при помощи которого моделируются взаимосвязи между различными величинами, количественные и качественные отношения между различными экономическими характеристиками и показателями.

Говорят, что задана функция f, если дан закон согласно которому каждому значению x из некоторого числового множества A ставится в соответствие одно определенное значение y из некоторого числового множества B. Функциональная зависимость между величинами x и y символически обозначается $y = f(x)$, и говорят, что x – аргумент (независимая переменная), а y – функция (зависимая переменная).

Совокупность всех значений аргумента, каждому из которых соответствует вполне определенное значение функции, называется областью определения функции. Множество значений, принимаемых y, называется областью изменения функции. Функцию можно задавать различными способами – формулой, таблицей, графиком.

Графиком функции f называется геометрическое место (множество) точек на координатной плоскости, имеющих координаты $(x; f(x))$, у которых абсциссами служат рассматриваемые значения независимой переменной x, а ординатами – соответствующие значения функции $y = f(x)$.

Перед построением графика функции необходимо провести исследования ее по следующим пунктам: область определения функции, область изменения функции, периодичность функции, четность или нечетность функции, монотонность функции, точки пересечения с осями координат, интервалы знакопостоянства, асимптоты.

Важнейшие функции, встречающиеся в экономических исследованиях

Линейная функция. Линейной функцией называется функция, определяемая формулой: $y = ax + b$. График этой функции есть прямая, пересекающая ось ординат в точке с ординатой b.

Пример. Затраты производства на промышленном предприятии, где изготавливается однородная продукция, можно разделить на две группы:

1) переменные затраты, пропорциональные объему продукции, например, материальные затраты;

2) постоянные затраты, т. е. не зависящие от объема продукции, например, затраты на содержание администрации.

Если постоянные затраты обозначить b, а пропорциональные затраты на единицу продукции – a, то при объеме x единиц полные затраты производства составят: $y = ax + b$.

Если линейная функция выражается формулой $y = ax$, то говорят, что она определяет прямую пропорциональность между y и x. В случае прямой пропорциональности график функции есть прямая, проходящая через начало координат. Величина a называется коэффициентом пропорциональности. Пусть x_1 и x_2 – два значения независимой переменной. Разность $x_2 - x_1 = \Delta x$ называется приращением независимой переменной. Соответствующая разность значений зависимой переменной $y_2 - y_1 = \Delta y$ называется приращением зависимой переменой.

$$\Delta y = a\Delta x.$$

Линейная функция имеет следующие свойства: приращение зависимой переменной пропорционально приращению независимой переменной, причем коэффициент пропорциональности есть a.

Степенная функция. Функция вида $f(x) = x^a$, где a – произвольное действительное число, отличное от нуля, называется степенной функцией. Если a – натуральное число, то степенная функция определена для всех действительных чисел, т. е. в интервале $-\infty < x < +\infty$. Если a – целое отрицательное число, то подставляя $a = -n$, где n – натуральное число, получим:

$$f(x) = x^a = x^{-n} = 1/x^n.$$

В этом случае степенная функция определена для всех значений, кроме 0. Если a будет числом, обратным натуральному, т. е. $a=1/n$, то получим:

$$f(x) = x^a = x^{1/n} = \sqrt[n]{x}\,.$$

Показательная функция. Функция $f(x) = a^x$, где $a > 0$, называется показательной функцией. Она определена в интервале $(-\infty,+\infty)$. Для $0 < a < 1$ показательная функция убывающая, а для $a > 1$ – возрастающая. Большое значение имеет функция при основании $a = e$, т.е. функция $y = e^x$, нередко обозначаемая также как $y = exp\ x$.

Логарифмическая функция. Функция, определяемая формулой

$$f(x) = \log_a x,$$

где $a > 1$, называется логарифмической функцией, она определена только для $x > 0$. Логарифмическая функция есть непрерывная и возрастающая для всех положительных значений x. Как и в случае показательной функции, большое значение имеет логарифмическая функция при основании $a = e$, т. е. $f(x) = \log_e x = \ln x$.

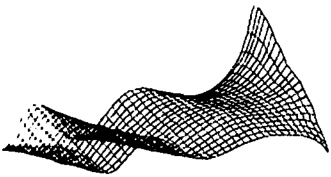
Гиперболы и параболы. Функция, определяемая формулой

$\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$, называется гиперболой.

Параболой называется функция, определяемая формулами:

$$y^2 = 2px \quad \text{при} \quad p > 0;\ x^2 = 2py,\ p > 0;\ y = ax^2 + bx + c.$$

1.2. Экономические задачи, решаемые методами дифференциального исчисления



При изучении экономических процессов – например: спроса, затрат, дохода – необходимо решать задачи определения скорости изменений соответствующих величин. В большинстве случаев такие задачи не могут быть решены с помощью элементарной математики. Чтобы решить их, необходимо использовать дифференциальное исчисление.

В экономике очень часто требуется найти наилучшее, или оптимальное, значение того или иного показателя: наивысшую производительность труда, максимальную прибыль, максимальный выпуск, минимальные издержки и т. д. Каждый показатель представляет собой функцию одного или нескольких аргументов. Например, выпуск можно рассматривать как функцию затрат труда и капитала. Таким образом, нахождение оптимального значения показателя сводится к нахождению экстремума (максимума или минимума) функции одной или нескольких переменных. Подобные задачи порождают класс экстремальных задач в экономике, решение которых требует использования методов дифференциального исчисления.

Экономические показатели обычно зависят от многих факторов. Поэтому в экономике часто приходится решать задачи нахождения экстремума функций нескольких переменных. Такие задачи хорошо изучены в теории функций нескольких переменных, использующей методы дифференциального исчисления. Многие задачи включают не только максимизируемую (минимизируемую) функцию, но и ограничения (скажем, бюджетное ограничение в задаче потребительского выбора). Это – задачи математического программирования, для решения которых разработаны специальные методы, также опнрающиеся на дифференциальное исчисление.

Важный раздел методов дифференциального исчисления, используемых в экономике, – это методы предельного анализа. Предельный анализ в экономике представляет собой совокупность приемов исследования изменяющихся величин затрат или результатов при изменениях объемов производства, потребления и т. п. на основе анализа их предельных значений. Предельный показатель функции $y = f(x)$ – это ее производная или частные производные в случае функции нескольких переменных.

В экономике широко используются средние величины: средняя производительность труда, средние издержки, средний доход, средняя прибыль и т. д. Но часто требуется узнать, на какую величину вырастет результат, если будут увеличены затраты, или, наоборот, на сколько уменьшится результат, если затраты сократятся. С помощью средних величин ответ на этот вопрос получить невозможно. В подобных задачах необходимо определить предел отношения приростов результата и затрат, т. е. найти предельный эффект. Следовательно, для их решения необходимо применение методов дифференциального исчисления – нахождение производной в случае одной переменной и частных производных, если функция зависит от нескольких аргументов.

Показатель предельного эффекта в оптимизационных моделях применяется для нахождения оптимального объема производства при заданных ресурсах, а также для определения оптимального распределения ограниченных ресурсов по различным направлениям их использования. Если максимизируемый показатель (например, прибыль) есть разность выручки и затрат, то в оптимальной точке предельная выручка должна равняться предельным затратам. Такое равенство должно выполняться по каждому из факторов, определяющих выручку и затраты, что вытекает из необходимости равенства нулю частных производных прибыли по всем этим факторам.

Методы дифференциального исчисления позволяют не только решить различные экономические задачи, но и записать необходимые и достаточные условия оптимума в этих задачах, которые позволяют дать ответ на те или иные конкретные вопросы.

1.3. Определение производной и ее смысл

Для функции $y = f(x)$ возьмем определенное значение независимой переменной x_0 и значение $x_0 + \Delta x$. Значение функции в точке x_0 составляет $f(x_0)$, а значение функции в точке $x_0 + \Delta x$ составляет $f(x_0 + \Delta x)$. Приращению независимой переменной Δx соответствует приращение зависимой переменной, т. е.

$$\Delta y = f(x_0 + \Delta x) - f(x_0).$$

Частное от деления приращения зависимой переменной на приращение независимой переменной называется дифференциальным отношением

$$\frac{\Delta y}{\Delta x} = \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}.$$

Если при $\Delta x \to 0$ существует предел дифференциального отношения, то этот предел называется производной функции $y = f(x)$ в точке x_0. Производная функции $y = f(x)$ обозначается y' или $f'(x)$; dy/dx; $df(x)/dx$. Отсюда следует:

$$y' = f'(x) = dy/dx = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}.$$

Если функция $y = f(x)$ имеет производную для какого-либо значения независимой переменой x, то для этого значения такая функция непрерывна.

Геометрическая интерпретация производной

На рис. 1.1 приведен график функции $y = f(x)$.

Дифференциальное отношение $\Delta y/\Delta x$ равно тангенсу угла β, который образует секущая, проходящая через точки A и B, которые соответствуют абсциссам x и $(x+\Delta x)$ с положительным направлением оси $0x$. Если приращение $\Delta x \to 0$, то точка B стремится к точке A, а угол β стремится к углу α, который образует касательная к кривой в точке x с положительным направлением $0x$, и, следовательно,

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \operatorname{tg}\alpha.$$

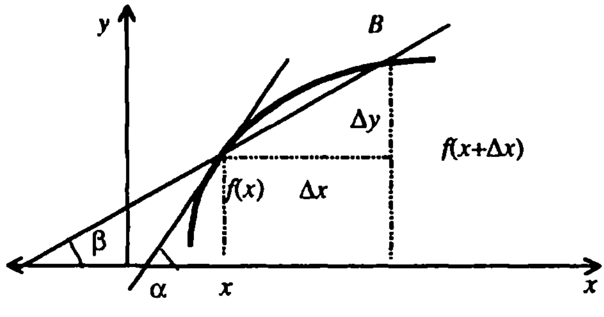


Рис. 1.1

Значение производной в данной точке равняется тангенсу угла, который образует касательная к данной точке кривой. Поэтому, зная производную, нетрудно вычертить касательную к кривой, являющейся геометрическим изображением функции.

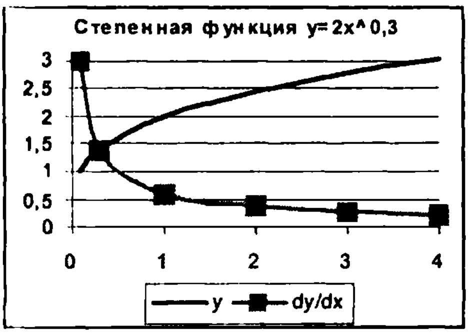
В экономике важно физическое понимание производной. Производная функции показывает ***скорость изменения*** функции.

1.4. Дифференцирование основных функций

Производная степенной функции

Производная функции $y = f(x) = x^n$ будет равна $y' = nx^{n-1}$.

Например, если $f(x) = x^5$, то $f'(x) = 5x^4$. Графическое представление производных функции $y = 2x^{0,3}$ и $y = 0,5x^3$ дано на рис. 1.2.



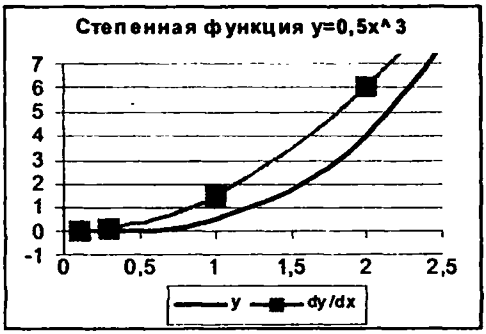


Рис. 1.2. Примеры производной степенной функции

Производная логарифмической функции

Определим производную логарифмической функции $f(x) = \log_a x$ при $a > 0$, $a \neq 1$. Она равна:

$$y' = \left(\frac{\ln x}{\ln a}\right)' = \frac{1}{x \ln a}.$$

Графическое представление производных логарифмической функции дано на рис. 1.3.

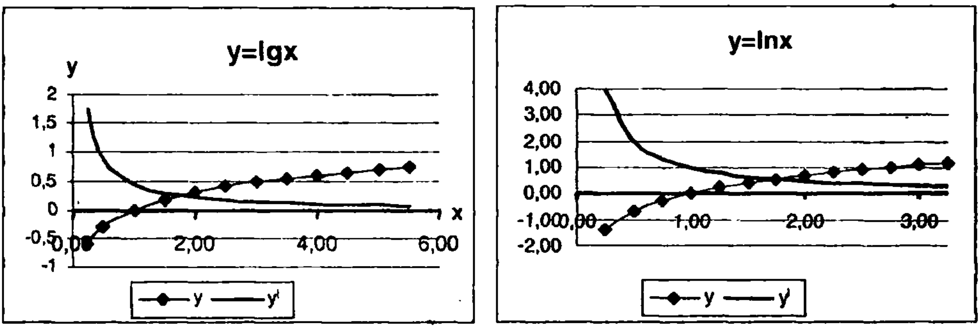


Рис. 1.3. Логарифмические функции и их производные

Логарифм с произвольным основанием можно заменить натуральным логарифмом и наоборот. Допустим, что имеется логарифм $\log_a x$. Надо вычислить натуральный логарифм числа x. Введем обозначение $\ln x = m$, откуда $e^m = x$. Логарифмируя обе части этого равенства при основании a, получим: $m\log_a e = \log_a x$, откуда $m = \log_a x / \log_a e$. Следовательно, $\ln x = \log_a x / \log_a e$. Таким образом, натуральный логарифм числа определяется делением логарифма числа при данном основании на логарифм числа e при данном основании. И наоборот, если задан $\ln x$, то получим $\log_a x = \ln x \log_a e$.

Производная показательной функции

Для функции $y = a^x$, где $a > 0$, производная равна:

$$y' = dy / dx = a^x / (1/\ln a) = a^x \ln a.$$

Производная функции $y = e^x$ просто равна $y' = e^x \ln e = e^x$ (рис. 1.4).

Рис. 1.4. Показательная функция и ее производная

Логарифмическая производная

Дана функция $y = \ln(f(x))$, где $f(x)$ – постоянно положительная функция. Найдем ее производную. Поскольку функция y состоит из функций $y = \ln u$ и $u=f(x)$ и учитывая, что для функции $y = \ln x$ производная равна $y' = 1/x$, имеем:

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx} = \frac{1}{x} \cdot \frac{du}{dx} = \frac{f'(x)}{f(x)}.$$

Выражение $f'(x)/f(x)$ называется логарифмической производной.

Логарифмическая производная, т. е. производная натурального логарифма функции переменной x, равняется производной функции, деленной на саму функцию, или $y' = (\ln(f(x)))' = f'(x)/f(x)$.

Темп роста функции

Предположим, что определена функция времени t: $y = f(t)$. Как известно, скорость изменения функции определяется ее производной

$$y' = f'(t).$$

Относительной скоростью называется отношение y'/y.

Относительная скорость называется также темпом роста функции.

Как раньше было показано,

$$(\ln y)' = y'/y,$$

следовательно, темп роста функции:

$$y'/y = (\ln y)'.$$

Темп роста функции равняется логарифмической производной функции.

1.5. Правила дифференцирования. Производные высших порядков

Правила дифференцирования

1. Производная суммы (разности) функций равна сумме (разности) производных этих функций:

$$(u(x) \pm v(x))' = u'(x) \pm v'(x).$$

2. Постоянный множитель выносится за знак производной:

$$(cu(x))' = cu'(x).$$

3. Правило дифференцирования произведения функций:

$$(u(x) \cdot v(x))' = u'(x)v(x) + u(x)v'(x).$$

4. Правило дифференцирования частного функций:

$$\left(\frac{u(x)}{v(x)}\right)' = \frac{u'(x) \cdot v(x) - u(x) \cdot v'(x)}{(v(x))^2}.$$

5. Правило дифференцирования сложной функции $f(u(x))$:

$$f'_x(u(x)) = f'_u(u)u'_x(x) \; (y'_x = y'_u \, u'_x).$$

6. Правило дифференцирования обратной функции $x(y)$:

$$x' = 1/y', \text{ где } y = f(x).$$

Производные высших порядков

Допустим, что функция $y = f(x)$ имеет производную в каждой точке интервала $(a, в)$.

Введем обозначение:

$$g(x) = f'(x).$$

Если функция $g(x)$ имеет производную в точке x, находящейся в интервале (a, b), то ее производная $g'(x)$ называется производной второго порядка (или коротко – второй производной) функции $f(x)$ в точке x.

Вторая производная функции $y = f(x)$ обозначается символами:

$$y'' \text{ или } f''(x), \text{ или } \frac{d^2 y}{dx^2}, \text{ или } \frac{d^2 f(x)}{dx^2}.$$

На рис. 1.5 приведены вторые производные некоторых функций.

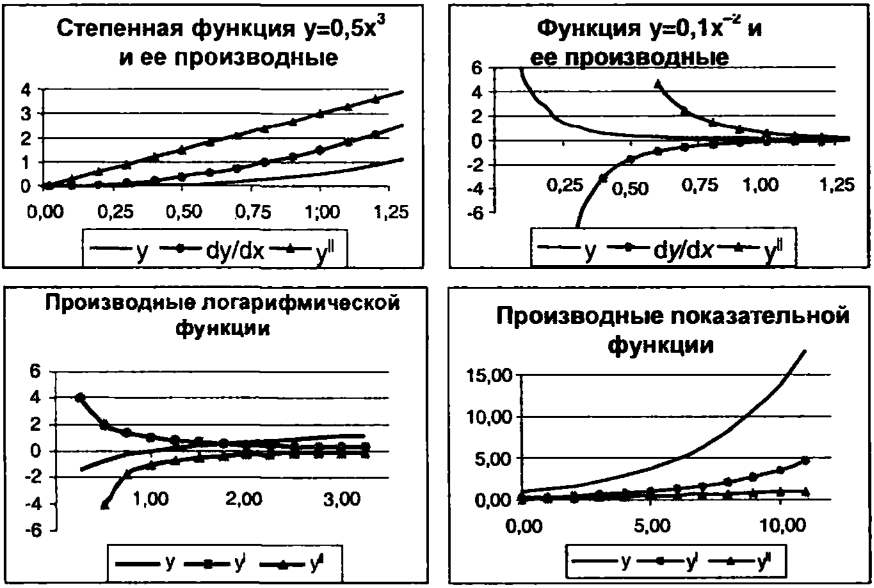


Рис. 1.5. Вторые производные некоторых функций

Может оказаться, что вторая производная также имеет производную в точке x. Тогда говорят о существовании производной третьего порядка, которая обозначается символами:

$$y''' \text{ или } f'''(x), \text{ или } \frac{d^3 y}{dx^3}, \text{ или } \frac{d^3 f(x)}{dx^3}.$$

Дифференциал функции

Дифференциалом функции называется главная, линейная относительно Δx, часть приращения функции, равная произведению производной на приращение независимой переменной:

$$dy = f'(x)\,\Delta x.$$

Дифференциал независимой переменной равен приращению этой переменной $dx = \Delta x$. Поэтому формулу для дифференцирования функции можно записать в виде:

$$dy = f'(x)dx.$$

Геометрически дифференциал функции есть приращение ординаты касательной, проведенной к графику функции $y = f(x)$ в данной точке, когда x получает приращение Δx.

При достаточно малых значениях Δx, $\Delta y \approx dy$, откуда

$$\Delta y = f'(x)\,\Delta x \text{ или } f(x + \Delta x) - f(x) \approx f'(x)\,\Delta x.$$

Данными формулами дифференциала можно пользоваться в приближенных вычислениях.

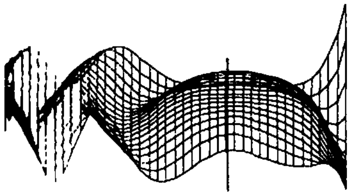
Дифференциалом второго порядка d^2y функции $y = f(x)$ называется дифференциал от дифференциала первого порядка этой функции

$$d^2y = d(dy), \quad \text{или} \quad d^2y = f''(x)dx^2,$$

где $dx^2 = (dx)^2$. Аналогично дифференциалом n-го порядка называется дифференциал от дифференциала $(n-1)$-го порядка этой функции, т. е.

$$d^ny = d(d^{n-1}y), \text{ или } d^ny = f^{(n)}(x)dx^n.$$

1.6. Применение дифференциального исчисления для исследования динамики функций



Многие явления и процессы экономики можно описать с помощью функций. Знание таких функций позволяет изучить, в каких интервалах эти функции возрастают или убывают, в каких точках они принимают наибольшие и наименьшие значения. Подобные исследования позволяют познать динамику явлений.

Критерии монотонности функции

Функция $y = f(x)$, определенная в интервале (a, b), называется:

- возрастающей, если из неравенства $x_1 < x_2$ следует, что $f(x_1) < f(x_2)$;
- убывающей, если из неравенства $x_1 < x_2$ следует, что $f(x_1) > f(x_2)$;
- неубывающей, если из неравенства $x_1 < x_2$ следует, что $f(x_1) \le f(x_2)$;
- невозрастающей, если из неравенства $x_1 < x_2$ следует, что $f(x_1) \ge f(x_2)$.

Функции возрастающие, убывающие, невозрастающие, неубывающие называются монотонными, а функции возрастающие и убывающие – строго монотонными функциями.

Если функция является возрастающей, то для каждого значения x производная $f'(x)$ будет положительна.

Возрастающая функция $\Rightarrow f'(x) > 0$.

Если функция является убывающей, то для каждого значения x производная $f'(x)$ будет отрицательна.

Убывающая функция $\Rightarrow f'(x) < 0$.

Темп возрастания и убывания функции

Во многих практических задачах необходимо определить, возрастает ли функция быстрее или медленнее. Пусть $y = f(x)$ есть функция, определенная и возрастающая в интервале (a, b).

Функция возрастает все быстрее, если равным приращениям переменной x соответствуют все бóльшие приращения зависимой переменной y. В этом случае говорят, что кривая выпуклая вниз (рис. 1.6).

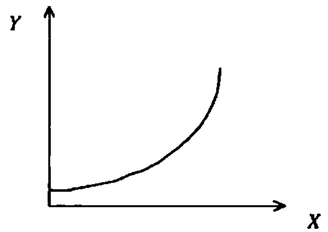


Рис. 1.6

Для функции возрастающей все быстрее первая и вторая производные положительны:

$$f'(x) > 0,\ f''(x) > 0.$$

Функция возрастает все медленнее, если равным приращениям x соответствуют все меньшие приращения функции, тогда говорят, что кривая выпуклая вверх (рис. 1.7).

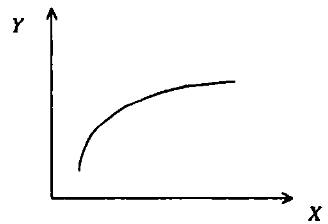


Рис. 1.7

Для функции, возрастающей все медленнее, первая производная положительна, а вторая производная отрицательна:

$$f'(x) > 0,\ f''(x) < 0.$$

Пусть $y = f(x)$ – функция, убывающая в интервале (a, b). Если равным приращениям независимой переменной x соответствуют все меньшие приращения зависимой переменной y, то функция убывает все быстрее, тогда говорят, что функция выпуклая вверх (рис. 1.8).

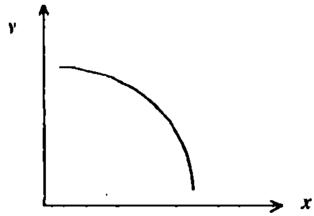


Рис. 1.8

Для убывающей все быстрее функции первая и вторая производные отрицательны:

$$f'(x) < 0,\ f''(x) < 0.$$

Если же равным приращениям переменной x соответствуют все большие приращения переменной y, то функция убывает все медленнее, тогда кривая выпуклая вниз (рис. 1.9).

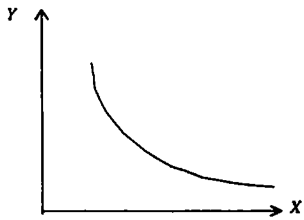


Рис. 1.9

Для убывающей все медленнее функции первая производная отрицательна, а вторая – положительна:

$$f'(x) < 0,\ f''(x) > 0.$$

1.7. Экстремумы функций одной переменной

Если функция $y = f(x)$ в точке $x = a$ принимает экстремальное значение, то производная в этой точке (если эта производная существует) равняется нулю: $f'(a) = 0$. Иными словами, если существует экстремум в точке $x = a$, то касательная к этой точке, если она существует, параллельна оси $0x$ (рис. 1.10).

Если производная функции в какой-либо точке равняется нулю, то в этой точке может и не существовать экстремума. Например, функция $y = x^3$. Производная $y' = 3x^2$ этой функции в точке $x = 0$ равняется нулю. Однако для $x = 0$ функция не имеет экстремума.

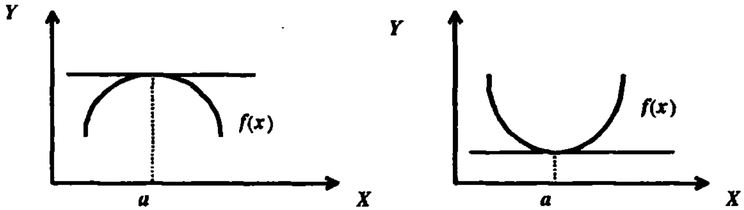


Рис. 1.10

Если функция $y = f(x)$ имеет в некотором интервале (a, b) производную, то внутри этого интервала экстремумы могут иметь место только в тех точках, где первая производная принимает нулевое значение. Следовательно, достаточно определить все значения x в интервале (a, b), для которых первая производная принимает нулевое значение, и определить, у каких из них имеет место экстремум. Однако экстремум функции может существовать и в таких точках, где функция не имеет производной. Достаточное условие существования экстремума заключается в следующем. Если функция $y = f(x)$ в окрестностях точки $x = a$ имеет непрерывную первую и вторую производные и если $f'(a) = 0$, то для $x = a$ имеет место:

1) максимум, если $f''(a) < 0$;

2) минимум, если $f''(a) > 0$.

Предположим, что $f''(a) < 0$. Поскольку по условию $f''(x)$ непрерывная функция, должна существовать такая окрестность $[(a - \alpha); (a + \alpha)]$, что для всех значений x в этой окрестности есть $f''(x) < 0$. Отсюда следует, что в этой окрестности производная $f'(x)$ – убывающая функция. Поскольку функция $f'(x)$ – непрерывна и равна нулю, проходя через точку a, эта производная меняет знак с положительного на отрицательный. Это означает, что для $x < a$ из окрестности точки a функция возрастающая, а для $x >$ a – убывающая. Следовательно, ее максимум находится в точке $x = a$.

На рис. 1.11 приведена функция $y = 11x - 2x^2 + 0{,}1x^3$. В точках максимума и минимума ее первая производная $y' = 11 - 4x + 0{,}3x^2$ равна нулю. Вторая производная $y'' = -4 + 0{,}6x$ отрицательна в точке максимума и положительна в точке минимума.

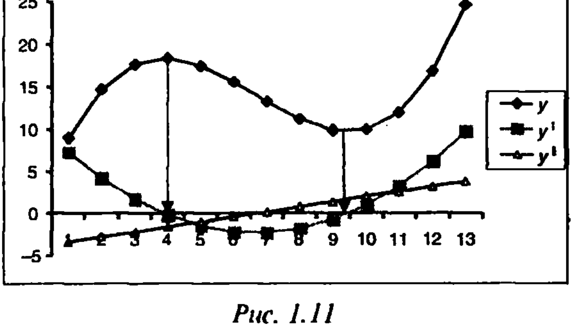


Рис. 1.11

Пример 1. Определить экстремумы функции $y = 3x^5 - 5x^3 + 10$.

Находим производную: $f'(x) = 15x^4 - 15x^2$.

Необходимым условием существования экстремума является $y' = 0$, т. е.

$$15x^4 - 15x^2 = 0.$$

Решив это уравнение, получим:

$$15x^2(x^2 - 1) = 0 \text{ или } 15x^2(x - 1)(x + 1) = 0.$$

Это уравнение имеет 3 корня, а именно: $f'(0) = f'(-1) = f'(1) = 0$.

В этих точках может существовать экстремум. Чтобы определить, в каких из этих точек имеется экстремум, вычислив вторую производную, получим:

$$f'(x) = 60x^3 - 30x.$$

Откуда $f''(0) = 0; f''(-1) = -30; f''(1) = 30$.

Следовательно, в точке $x = -1$ функция принимает максимальное значение, а в точке $x = 1$ – минимальное.

Пример 2. Определить экстремум функции $y = x^3 - 3x + 1$.

Имеем $y' = 3x^2 - 3x = 3(x^2 - 1) = 3(x - 1)(x + 1)$, откуда

$$3x^2 - 3 = 0, \text{ если } x = -1 \text{ или } x = 1.$$

Вычисляем далее $y''(x) = 6x$, откуда

$$y''(-1) = -6; y''(1) = 6.$$

Следовательно, при $x = -1$ функция принимает максимальное значение, а при 1 – минимальное.

1.8. Максимизация прибыли

Предположим, что предприятие производит x единиц некоторого продукта. Тогда цена, при которой спрос также составит x единиц, определяется уравнением $p = p(x)$. Обозначим суммарные затраты на производство x единиц продукции через $C(x)$, в этом случае прибыль $PR(x)$

$$PR(x) = R(x) - C(x)$$

есть также функция x. Таким образом,

$$PR(x) = x \cdot \rho(x) - C(x).$$

Прибыль предприятия максимальна, если выполняются два условия

1) необходимое условие

$$PR''(x)=0,$$

2) достаточное условие

$$PR''(x) < 0.$$

Из первого условия следует

$$R'(x) - C'(x) = 0, \text{ или } R'(x) = C'(x),$$

т. е. монополистическое предприятие получает максимальную прибыль при таком объеме производства x, когда **предельная выручка равна предельным затратам.**

Из второго условия следует, что

$$R''(x) - C''(x) < 0, \text{ или } R''(x) < C''(x).$$

Это означает, что предприятие получает максимальную прибыль, если **темп роста предельной выручки меньше темпа роста предельных затрат.** Другими словами, угловой коэффициент касательной к графику предельного дохода должен быть меньше углового коэффициента касательной к графику предельных затрат.

Пример 1. Завод производит x единиц продукции в месяц, а суммарные затраты производства составляют:

$$C = 0{,}2\,x^2 + 20x + 120.$$

Зависимость между удельной ценой p и количеством единиц продукции x, которое можно продать по этой цене, такова:

$$p = 80 - 0{,}1x.$$

Просчитать, при каких условиях прибыль будет максимальной.

Выручка составляет: $R = x\,p = 60x - 0{,}1x^2$.

Предельные суммарные затраты равны: $dC/dx = 0{,}4x + 20$.

Предельная выручка: $dR/dx = 80 - 0{,}2x$.

Если прибыль максимальна, то должно быть:

$$80 - 0{,}2x = 0{,}4x + 20.$$

Откуда $x = 100$ единиц.

Проверка по условию достаточности:

$$R''(x) = -0{,}2,\ C''(x) = 0{,}4,\ R''(x) < C''(x).$$

При выпуске 100 единиц продукции прибыль завода будет наибольшей. При таком объеме производства цена единицы продукции составит:

$$p = 80 - 0{,}1\ 100 = 70.$$

Пример 2. Функция суммарных затрат на изготовление продукции x: $C(x) = 4x^2 + 600$; функция спроса: $p = 200 - 6x$.

Рассчитать, при каком объеме производства прибыль будет наибольшей.

Суммарные предельные затраты есть $C'(x) = 8x$.

Суммарная выручка есть

$$R = xp = x(200 - 6x) = 200x - 6x^2.$$

Предельная суммарная выручка составляет:

$$R'(x) = 200 - 12x.$$

Находим вторые производные:

$$R''(x) = -12,\ C''(x) = 8.$$

Имеет место неравенство:

$$R''(x) < C''(x).$$

Следовательно, прибыль максимальна, если

$$200 - 12x - 8x = 0, \text{ откуда } x = 10.$$

Монопольная цена составляет:

$$p = 200 - 6 \cdot 10 = 140.$$

Тогда полные затраты $C(x) = 4x^2 + 600 = 4 \cdot 100 + 600 = 1000$.

Выручка равна:

$$R(x) = p \cdot x = 140 \cdot 10 = 1400,$$

а максимальная прибыль:

$$PR(x) = R(x) - C(x) = 1400 - 1000 = 400.$$

1.9. Функции нескольких переменных

Понятие функции нескольких переменных

Если каждой паре чисел x и y, называемых независимыми переменными, однозначно соответствует число z, называемое зависимой переменной, то говорят, что z есть функция двух переменных x и y: $z = f(x, y)$ (рис. 1.12).

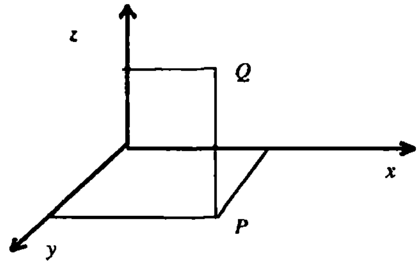


Рис. 1.12

Множество всех точек плоскости $0xy$, для которых эта функция определена, называется областью задания функции.

Для графического изображения функции двух переменных необходима система координат $0xyz$ в трехмерном пространстве.

Каждой паре чисел x и y соответствует точка $P(x, y)$ плоскости $0xy$. В точке $P(x, y)$ проводим прямую, перпендикулярную плоскости $0x$, и отмечаем на ней соответствующее значение функции z; получаем в пространстве точку Q с координатами x, y, z, которая обозначается символом $Q(x, y, z)$. Точки Q, соответствующие разным значениям независимых переменных, образуют некую поверхность в пространстве. Такая поверхность и есть геометрическое изображение функции $z = f(x, y)$ (рис. 1.13).

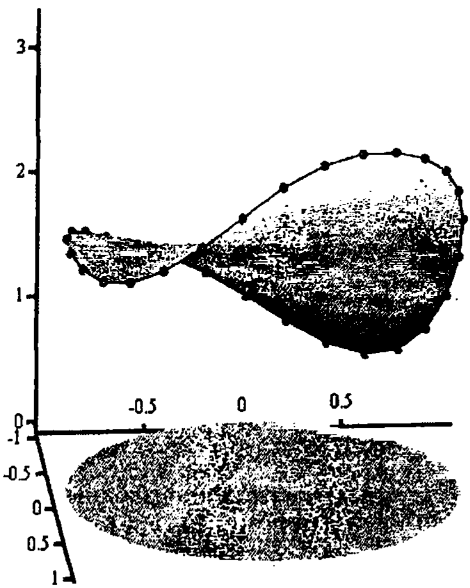


Рис. 1.13

Функция трех переменных и вообще функция n переменных $x_1, x_2, ..., x_n$

$$u = f(x_1, x_2, ..., x_n)$$

определяется аналогично функции двух переменных.

Частные производные первого порядка

Функция $z = f(x,y)$ зависит только от переменной x, если переменная y принимается постоянной. Обозначим: Δx – приращение переменной x; введем также обозначение:

$$\Delta z_x = f(x + \Delta x, y) - f(x,y).$$

Если существует предел отношения

$$\lim_{\Delta x \to 0} \frac{\Delta z_x}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x},$$

то этот предел называется частной производной первого порядка или первой частной производной по x; она обозначается символом:

$$\frac{\partial z}{\partial x} = f'_x(x, y).$$

Аналогично определяется частная производная по y:

$$\frac{\partial z}{\partial y} = f'_x(x, y)$$

как предел отношения

$$\lim_{\Delta y \to 0} \frac{\Delta z_y}{\Delta y} = \lim_{\Delta y \to 0} \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y}.$$

Поскольку определение частной производной вполне сходно с определением производной для функции одной переменной, теоремы о производных соответствуют и частным производным функции двух переменных.

Пример. Найти первые частные производные функции:

$$z = x^2 y + 4x - 2y + 5.$$

Чтобы найти частную производную по x, принимаем y за постоянную и находим производную по x:

$$\frac{\partial z}{\partial x} = 2x \cdot y + 4.$$

Чтобы найти частную производную по y, принимаем x за постоянную и находим производную по y:

$$\frac{\partial z}{\partial y} = x^2 - 2.$$

Градиент

Упорядоченная пара (первых) частных производных

$$\left[\frac{\partial f(x, y)}{\partial x}, \frac{\partial f(x, y)}{\partial y}\right]$$

функции $z = f(x,y)$ двух переменных x и y обозначается символом grad $f(x,y)$ и называется *градиентом* функции $z = f(x,y)$ двух переменных. Градиент функции двух переменных есть двухмерный вектор.

Градиент grad $f(x^0,y^0)$ функции $f(x,y)$ в точке (x^0,y^0) показывает направление самого быстрого роста функции $f(x,y)$ в точке (x^0,y^0).

Частной производной функции $f(x_1, x_2, ..., x_n)$ относительно переменной x называется производная функция одной переменной x, когда остальные переменные принимаются за постоянные.

Градиент функции $f(x_1, x_2,...,x_n)$ n переменных представляет собой n-мерный вектор и определяется:

$$\operatorname{grad} f(x_1,...,x_n) = \left[\frac{\partial f(x_1,...,x_n)}{\partial x_1},...,\frac{\partial f(x_1,...,x_n)}{\partial x_n}\right].$$

Применение частных производных в экономике

Частные производные широко применяются в экономическом анализе при использовании функции нескольких переменных. Так, все показатели, перечисленные в предельном анализе, в случае функции нескольких переменных будут устанавливаться с помощью частных производных. Остановимся на нескольких характерных показателях.

Перекрестная эластичность. Предположим, что функция $x = f(p_1,p_2,r)$ выражает спрос на товары 1 и 2, зависящий от цен на эти товары p_1 и p_2, а также доходов потребителей r. И если эластичность спроса относительно цены является отношением процентного изменения спроса к процентному изменению цены, то перекрестная эластичность определяется как отношение процентного изменения спроса к процентному изменению цены альтернативного товара и устанавливается с помощью частной производной.

Перекрестная эластичность относительно цены p_1 равна:

$$\varepsilon_1 = p_1/x \cdot \partial x/\partial p_1$$

и означает процент роста (или снижения) спроса на товар 1, если цена товара 1 возрастает на 1%, а цена товара 2 останется неизменной.

Перекрестная эластичность относительно цены p_2 равна:

$$\varepsilon_2 = p_2/x \cdot \partial x/\partial p_2$$

и означает процент роста (или снижения) спроса на товар 2, если цена товара 2 возрастает на 1%, а цена товара 1 останется неизменной.

Предельный продукт фактора производства. Предположим, что функция выпуска продукции зависит от двух ресурсов: капитала K и труда L, т. е. $Q = f(K, L)$.

Тогда предельным продуктом фактора производства будет являться предел соотношения приростов результата и затрат, которые его вызвали, т.е. частная производная функции выпуска по соответствующей переменной (ресурсу).

Предельный продукт капитала равен:

$$Q'_K = \frac{\partial f(K,L)}{\partial K}.$$

Предельный продукт труда равен:

$$Q'_L = \frac{\partial f(K,L)}{\partial L}.$$

Частные производные высших порядков

Частные производные производных $f'_x(x,y)$ и $f'_y(x,y)$ называются частными производными второго порядка или вторыми частными производными.

Пример. Найти вторые частные производные функции:

$$z = x^4 + 5x^2y^2 + 6xy + 5.$$

Имеем:

$$\frac{\partial z}{\partial x} = 4x^3 + 10xy^2 + 6y, \quad \frac{\partial z}{\partial y} = 10x^2y + 6x,$$

откуда $\dfrac{\partial^2 z}{\partial x^2} = 12x^2 + 10y^2$, $\dfrac{\partial^2 z}{\partial x \partial y} = 20xy + 6$,

$$\frac{\partial^2 z}{\partial y^2} = 10x^2, \quad \frac{\partial^2 z}{\partial y \partial x} = 20xy + 6.$$

Частные производные

$$f''_{xy} = \frac{\partial^2 z}{\partial x \partial y} \text{ и } f''_{yx} = \frac{\partial^2 z}{\partial y \partial x}$$

функции $z = f(x,y)$ называются *смешанными производными.*

Если существуют непрерывные смешанные производные f''_{xy} и f''_{yx}, то они равны одна другой.

Производная неявной функции

Если существует такая непрерывная функция $y = f(x)$, что соответствующие пары (x,y) удовлетворяют условию $F(x,y) = 0$, то это условие называется неявной формой функции $f(x)$; о самой же функции $f(x)$ говорят, что это есть *неявная функция*, удовлетворяющая условию $F(x,y) = 0$.

Пример. Неявная форма функции

$$y = \frac{1-x}{1+x} \quad \text{есть} \quad xy + x + y - 1 = 0.$$

Предположим, что непрерывная функция $y = f(x)$ дана в неявной форме $F(x,y) = 0$ и что $f'_x(x,y) = 0$. Производная функции будет равна:

$$\frac{dy}{dx} = -\frac{F'_x(x,y)}{F'_y(x,y)}.$$

Сложные функции двух переменных

Предположим, что имеется функция z двух переменных u и v: $z = f(u,v)$, причем эти переменные суть функции переменных x, y, т. е. $u = g(x,y)$, $v = h(x,y)$.

Формула $z = f(g(x,y), h(x,y))$ определяет сложную функцию переменных x и y.

Если функции z, u, v поддаются дифференцированию, то

$$\frac{\partial z}{\partial x} = \frac{dz}{\partial u}\cdot\frac{\partial u}{\partial x} + \frac{\partial z}{\partial v}\cdot\frac{\partial v}{\partial x}, \qquad \frac{\partial z}{\partial y} = \frac{dz}{\partial u}\cdot\frac{\partial u}{\partial y} + \frac{\partial z}{\partial v}\cdot\frac{\partial v}{\partial y}.$$

Дифференциал функции

Для функции двух независимых переменных дифференциалом функции называется сумма произведений частных производных этой функции на приращение соответствующих независимых переменных

$$dz = z'_x\Delta x + z'_y\Delta y.$$

Поскольку $dx = \Delta x$, $dy = \Delta y$, формулу дифференциала можно записать в виде

$$dz = z'_x dx + z'_y dy$$

или

$$dz = \frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy.$$

Экстремумы функции двух переменных

Если дифференцируемая функция $f(x,y)$ имеет в точке (x_0,y_0) экстремум (рис.1.14–1.15), то

$$f'_x(x_0,y_0) = 0, f'_y(x_0,y_0) = 0$$

и градиент функции равен нулю.

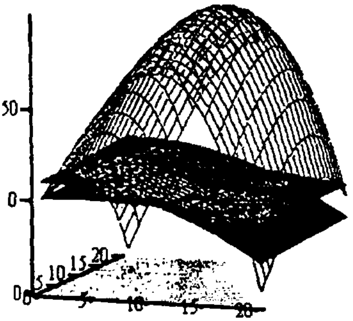


Рис. 1.14. График функции двух переменных с частными производными

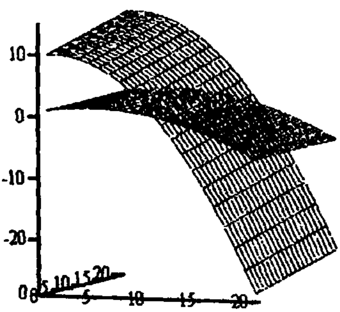


Рис. 1.15. Графики частных производных той же функции

Однако данных условий недостаточно для существования экстремума.

Если функция $z = f(x,y)$ имеет в окрестности точки (x_0,y_0) первые и вторые непрерывные частные производные, то в точке (x_0,y_0), в которой $f'_x = 0$ и $f'_y = 0$, имеет место экстремум в случае, когда в этой точке

$$\det[G] = f''_{xx}f''_{yy} - (f''_{xy})^2 > 0.$$

Функция имеет максимум в точке (x_0,y_0), если в этой точке $f''_{xx} < 0$, и минимум, если $f''_{xx} > 0$.

Выражение $\det[G]$ является определителем матрицы вторых частных производных функции $f(x,y)$.

Подробнее данный вопрос рассматривается в главе 14.

Однородные функции

Функция n переменных $f(x_1, x_2,...,x_n)$ называется однородной функцией k-й степени, если

$$f(tx_1, tx_2,...,tx_n) = t^k f(x_1, x_2,...,x_n)$$

для всех значений переменных $x_1, x_2,...,x_n$ и каждого значения t.

Пример 1. Рассмотрим функцию

$$f(x,y) = ax + by.$$

Если t – произвольное число, то

$$f(tx, ty) = atx + bty = t(ax + by) = tf(x, y).$$

Следовательно, данная функция – однородная функция первой степени.

Пример 2. Рассмотрим функцию

$$f(x, y, z) = x^2 + y^2 + z^2.$$

Если t – произвольное число, то

$$f(tx, ty, tz) = (tx)^2 + (ty)^2 + (tz) = t^2x^2t^2y + t^2z^2 = t^2(x^2 + y^2 + z^2) = t^2f(x,y,z).$$

Следовательно, эта функция есть однородная функция второй степени. Если функция переменных $f(x_1, x_2,....., x_n)$ есть однородная k-й степени, то

$$x_1f'_{x1} + x_2f'_{x2} +...+ x_nf'_{xn} = kf(x_1, x_2,..., x_n).$$

Пример 3. Дана функция предложения

$$x = f(p_1, p_2,..., p_n),$$

где $p_1, p_2,...., p_n$ – цены товаров 1,2,..., n соответственно. Если это есть однородная функция k-й степени, то

$$p_1 \frac{dx}{dp_1} + p_2 \frac{dx}{dp_2} + ... + p_n \frac{dx}{dp_n} = kx$$

или

$$\frac{p_1}{x}\frac{dx}{dp_1} + \frac{p_2}{x}\frac{dx}{dp_2} + ... + \frac{p_n}{x}\frac{dx}{dp_n} = k.$$

Величина $\frac{p_i}{x} \cdot \frac{dx}{dp_i}$ есть эластичность предложения товара относительно цены i-го товара.

Таким образом, когда предложение есть однородная функция k-й степени, то сумма частных эластичностей отдельных товаров равняется степени однородности функции предложения.

Глава 2. Работа с электронными таблицами

2.1. Предварительные сведения

Вся рабочая область окна Excel (рис. 2.1) занята чистым рабочим листом, разделенным на отдельные ячейки. Столбцы озаглавлены буквами, строки – цифрами. Ячейка выделяется щелчком мыши, при этом она помечается рамкой и в поле имени, находящемся на строке формул, расположенной под панелями инструментов, будет показан ее адрес, например A1.

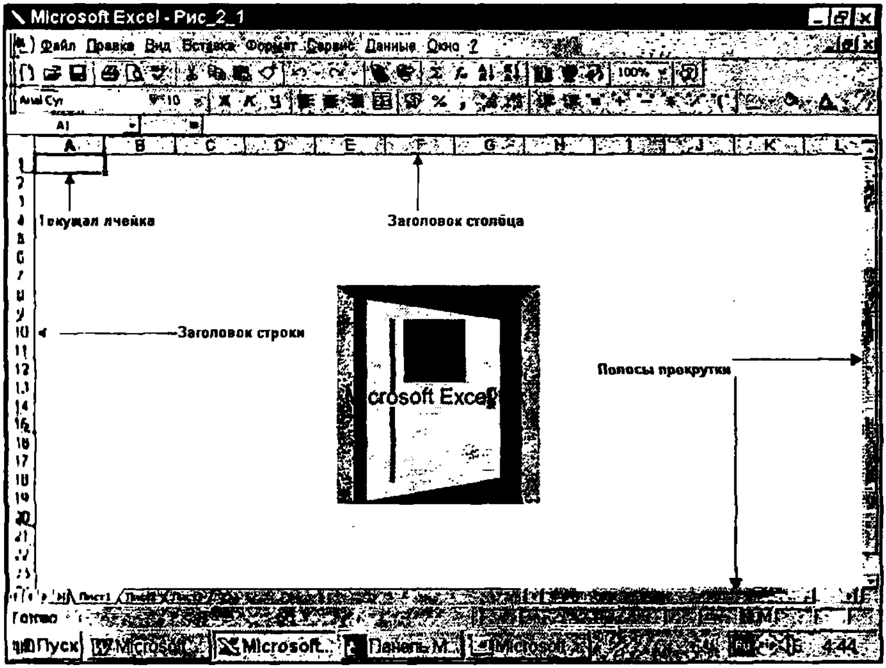


Рис. 2.1. Окно Excel

Одна рабочая страница содержит 256 столбцов и 16 384 строки. В окне Excel, как и в других программах под Windows, под заголовком окна находится строка меню. Ниже находятся панели инструментов – **Стандартная** и **Форматирование**. При необходимости панель инструментов настраивается нужным образом. Требуемая панель выводится по команде меню **Вид – Панели инструментов** с помощью щелчка мыши перед названием нужной вам панели. Инструменты (команды) настраиваются командой меню **Вид – Панели инструментов – Настройка** или командой меню **Сервис – Настройка**. В появившемся диалоге (рис. 2.2) выбирается пункт **Команды**, нужная панель, выбирается значок команды и мышью переносится на панель инструментов. Полезно перенести на панель инструментов знаки арифметических действий, чтобы работать мышью.

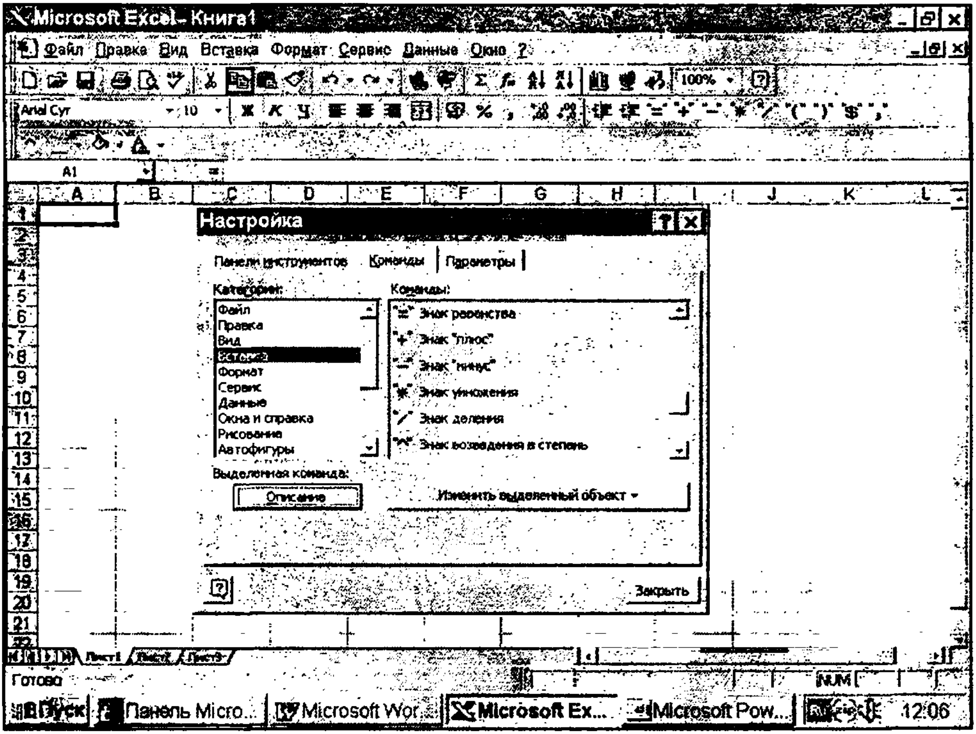


Рис. 2.2. Диалоговое окно настройки панели инструментов

Кнопки на панелях инструментов позволяют быстро вызывать функции Excel. Почти все функции можно вызвать и через меню.

Справа и снизу от рабочего листа находятся полосы прокрутки, с помощью которых можно быстро передвигаться по рабочему листу.

Внизу окна рабочей книги находятся ярлычки рабочих листов. Можно выбрать нужный рабочий лист, щелкнув мышью по соответствующему ярлычку. Вы можете присвоить каждому листу название. Название текущего листа всегда выделяется полужирным шрифтом. При необходимости ярлычки листов выводятся командой меню **Сервис – Параметры – Вид – Ярлычки листов.**

Передвижение по рабочему листу осуществляется с помощью мыши или клавиш управления курсором. Ячейка, в которой находится курсор, является выделенной.

Для исправления опечаток во время заполнения ячейки, до того как ввод текста подтвержден, в вашем распоряжении клавиша <Backspace> – . Можно установить текстовый курсор перед ошибкой и нажать клавишу <Delete>. Если вы хотите удалить весь введенный текст, то нажмите кнопку с крестом, находящуюся в строке формул перед полем ввода. При необходимости исправить уже подтвержденное содержание ячейки, выделите ее щелчком мыши. Затем щелкните мышью в строке формул и произведите исправления. При помощи команды меню **Правка – Очистить** имеется возможность выбрать, что удалить: введенный текст, формат, формулу или комментарий.

Ввод и форматирование данных. Ввод и форматирование данных требуют предварительного выделения ячеек. Выделение строк производится щелчком мыши по номеру строки. Выделение столбца производится щелчком мыши по имени столбца. Выделение блока ячеек производится щелчком мыши в углу блока, и далее, не отпуская кнопку мыши, следует протащить курсор в другой угол блока. Выделение нескольких блоков можно сделать следующим способом. Выделяется первый блок. Затем нажимается <Ctrl> и выделяется следующий блок. Выделение отменяется щелчком мыши в любом месте таблицы. Выделение всего рабочего листа производится с помощью меню или нажатием кнопки, которая находится в левом верхнем углу окна рабочей книги над заголовками строк и перед заголовками столбцов.

Ввод символов греческого алфавита. Ввод текста производится с помощью основных алфавитов (кириллицы и латиницы). Буквы греческого алфавита можно ввести несколькими способами. Во-первых, с помощью команды меню **Вставка – Символ**. Во-вторых – с помощью шрифта Symbol клавишами с буквами латинского алфавита (рис. 2.3).

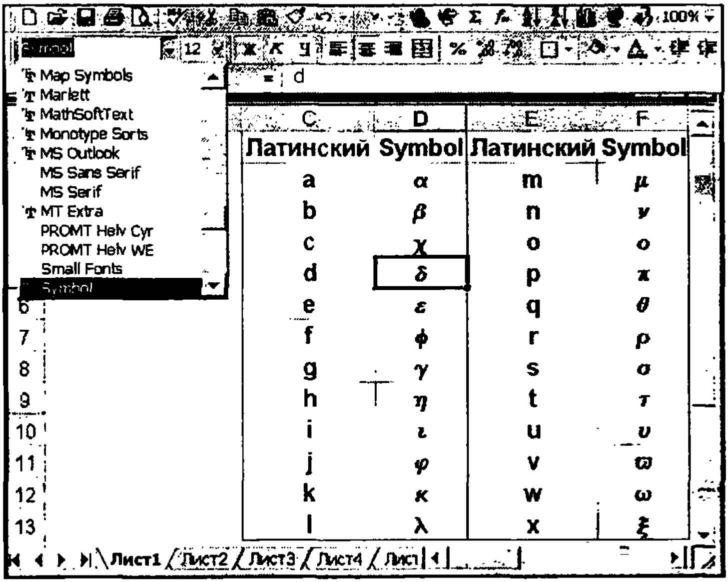


Рис. 2.3. Ввод символов греческого алфавита с помощью шрифта Symbol

Вставка верхних и нижних индексов. В математике часто используются переменные с индексами, ввод которых производится следующим образом. Выбирается команда меню **Формат – Ячейки – Шрифт**. В появившемся диалоговом окне назначается эффект – верхний или нижний индекс, назначается размер шрифта, нажимается **ОК**, вводится верхний или нижний индекс, нажимается <Enter>.

Размер ячейки определяется шириной столбца и высотой строки. Часто возникает необходимость изменить ширину столбцов, что делается следующим образом. Необходимо подвести курсор к правой границе имени столбца, где курсор приобретает вид вертикальной черты с двунаправленными стрелками. После щелчка мыши следует переместить курсор в направлении изменения ширины столбца. Изменение ширины нескольких столбцов производится аналогично, после предварительного выделения изменяемых столбцов. Изменение ширины столбцов по окончании заполнения производится после их выделения двойным щелчком мыши.

Форматирование чисел. Форматирование чисел в ячейках производится с помощью команды меню **Формат – Ячейки**. В диалоговом окне выбирается вкладка **Число** (рис. 2.4). Если выбирается числовой формат, то необходимо указать число десятичных знаков и разделитель групп разрядов.

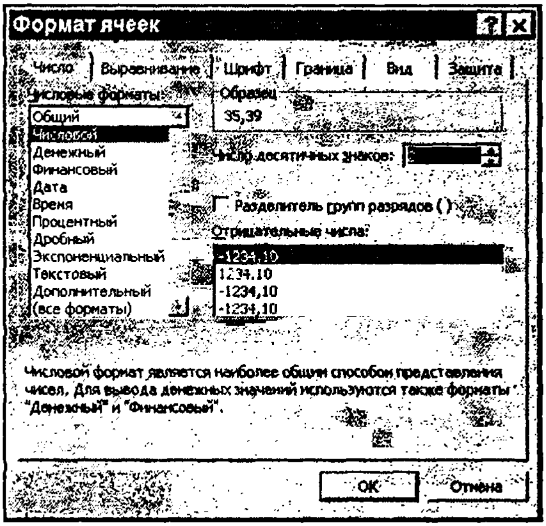


Рис. 2.4. Форматирование чисел

Перед началом работы лучше выделить весь лист и произвести форматирование чисел в ячейках.

2.2. Расчеты в Excel

Вставка и редактирование формул

Ввод формул начинается со знака =

Это означает, что за данной командой последует математическая операция. С помощью знака равенства отличается ввод текста от ввода формул. Буквы, вводимые в формулы, могут быть как строчными, так и прописными. Достаточно часто аргументом вводимых формул являются адреса ячеек. Использование арифметических операций производится достаточно просто. Введем в ячейку A1 значение 256, в ячейку A2 – 34, в A3 – 59, в A4 – 6980, в A5 – 78, в A6 – 5. В ячейку B1 введем результат вычислений =(((256 + 34) – 59) / 6980 * 78) ^ 5, используя адреса ячеек. Введем формулу в ячейку B1. Щелкнем мышью на ячейке B1, тем самым выделив ее, и введем знак . Введем скобки и щелкнем на ячейке A1. Введем знак «+» . Щелкнем мышью на ячейке A2 и закроем скобку. Введем знак «–». Щелкнем на ячейке A3. Закроем скобку. Введем знак "/" . Щелкнем на ячейке A4. Введем знак «*». Щелкнем мышью на ячейке A5 и закроем скобку. Введем знак возведения в степень « ^ ». Таким образом, в ячейке B1 будет введена формула:

=(((A1 + A2) – A3) / A4 * A5) ^ A6.

После нажатия клавиши <Enter> в ячейке появится результат – 114,6188.

Напомню, что разделителем знаков в Excel является запятая.

Формула является основным средством для анализа данных. С помощью формул можно складывать, умножать и сравнивать данные, а также объединять значения. Формулы могут ссылаться на ячейки текущего листа, листов той же книги или других книг. В следующем примере складывается значение ячейки A4 с числом 50,5. Полученный результат делится на сумму ячеек B1, B2, B3.

Ссылка на ячейку

Числовая константа

Функция листа

= (A4 + 50,5) / СУММ (B1:B3)

Ссылка на диапазон ячеек

Оператор деления

Оператор сложения

Операторы. Все математические функции описываются в программах с помощью специальных символов, называемых операторами. Список операторов Excel приведен в табл. 2.1.

Таблица 2.1. Операторы Excel

Оператор	**Функция**	**Оператор**	**Функция**
Арифметические операторы		**Операторы сравнения**	
+	сложение	=	равно
–	вычитание	<	меньше

Таблица 2.1 (окончание)

Оператор	Функция	Оператор	Функция
Арифметические операторы		**Операторы сравнения**	
*	умножение	>	больше
/	деление	<=	меньше или равно
%	процент	>=	больше или равно
^	возведение в степень	<>	не равно
Операторы связи		**Текстовый оператор объединения**	
:	диапазон	&	соединение текстов
;	объединение		

Если в формуле содержится несколько операторов, то при обработке данных Excel использует операторы в той же иерархической последовательности, как этого требуют правила математики. Иерархия операторов приведена ниже:

: ; операторы связи

– оператор отрицания

% оператор процентов

^ оператор возведения в степень

***, / операторы умножения, деления**

+, – операторы сложения, вычитания

& оператор текста

<, >, <=, =>, операторы сравнения

Еще один знак подчиненности операций – скобки. Скобки, как известно из математики, используются для указания приоритетов той или иной операции.

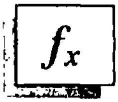
Функции в Excel. Применение функций значительно ускоряет работу с электронными таблицами. Рассмотрим применение функций на примере операции сложения. Допустим, нам необходимо сложить величины, содержащиеся в ячейках A1:A5. Если мы будем вводить в формулу адрес каждой ячейки в отдельности и знак сложения, то затратим время на формулу:

=A1+A2+A3+A4+A5

Имеющуюся в Excel функцию сложения можно ввести несколькими способами. Так, допустим, в ячейку B2 можно вручную набрать формулу следующим образом: вводится знак =, имя функции СУММ, открывается скобка, вводится адрес ячеек (в данном случае диапазон A1:A5) и закрывается скобка:

=СУММ(A1:A5).

Более быстрый способ – это нажать кнопку с изображением знака суммы с панели инструментов **Стандартная** – Σ. Если мы захотим ввести результат сложения в ячейку A6, т. е. непосредственно под столбцом слагаемых, то в диалоговом окне сразу появится диапазон A1:A5. Если требуется ввести результат сложения A1:A5 в ячейку B3, то после нажатия кнопки суммы необходимо ввести нужную область адресов ячеек, что проще всего сделать мышью, нажав левую кнопку на ячейке *A1* и, не отпуская ее, провести курсор до ячейки A5. Заканчивается ввод нажатием клавиши <Enter>. Вместо нажатия клавиши <Enter> можно использовать знак ✓, появляющийся в командной строке.



Мастер функций. Различные функции вводятся с помощью диалогового окна, называемого Мастером функций. Он вызывается командой меню **Вставка – Функция** или кнопкой на панели инструментов **Стандартная**, приведенной в начале абзаца.

В первом диалоговом окне Мастера функций приведены категории функций по тематическому принципу (рис. 2.5). В списке слева находятся имена тематических групп. Щелкнув на нужной категории функций, в правой части можно вывести список имен функций, содержащихся в данной категории.

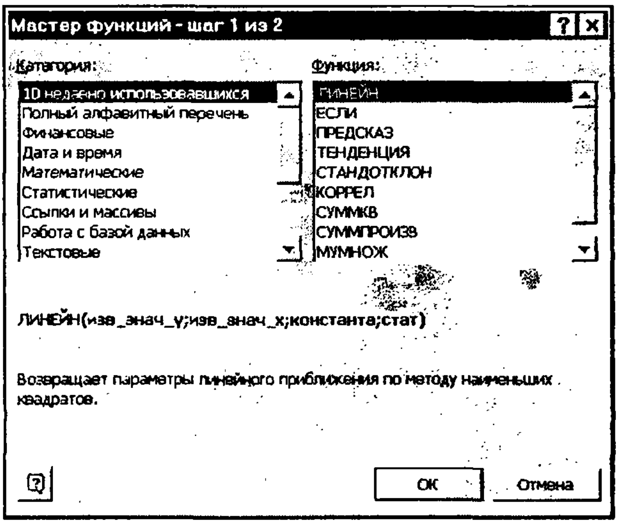


Рис. 2.5. Мастер функций

Вызов функции осуществляется двойным щелчком по ее имени. Также можно вызвать функцию, выделив ее и щелкнув на кнопке **ОК** или нажав клавишу <Enter>.

С помощью полос прокрутки можно просматривать списки функций.

Введем в ячейки A1:A10 значения 5; 7; 10; 3; 15; 20; 32; 16; 40; 49 (рис. 2.6). Рассчитаем стандартное отклонение. В ячейку B1 введем название функции. Результат выведем в ячейку B2. Вызовем Мастер функций, откроем категорию **Статистические**, с помощью полосы прокрутки найдем функцию СТАНДОТКЛОН, щелкнув по которой, вызовем диалоговое окно функции (см. рис. 2.6).

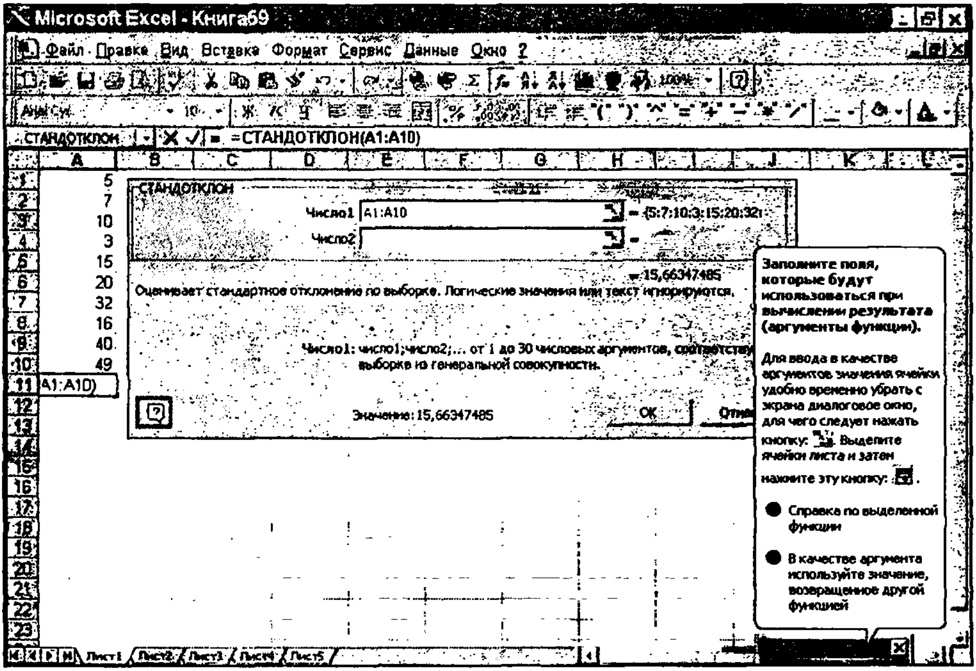


Рис. 2.6. Диалоговое окно функции СТАНДОТКЛОН

Для ввода аргументов можно передвинуть диалоговую панель мышью или нажать кнопку справа от окна ввода аргументов и убрать с экрана диалоговое окно, оставив только окно ввода аргументов. После этого вводятся аргументы – адреса ячеек, что проще сделать мышью, выделив блок ячеек A1:A10. После нажатия кнопки **OK** в ячейке B2 выводится результат. Иногда требуется редакция введенных аргументов. Для этого выделяется ячейка, где записана функция (B2), и вызывается Мастер функций.

В качестве аргументов могут быть использованы значения функций (вложенные функции). Для вставки функции в качестве аргумента в окно ввода аргументов вводится требуемая функция с вызовом Мастера функций. Для удобства он выводится слева от командной строки.

Числовые ряды. Во многих задачах встречаются числовые ряды. Задаются они в Excel просто: вводятся два первых элемента ряда, один под другим, и выделяются. После фиксации указателя мыши на маленьком черном квадратике в правом углу отмеченной области, он протягивается вниз до тех пор, пока не получится числовой ряд нужной длины. Точно так же можно создать и ряд в горизонтальном направлении. Существует и еще одна возможность создания ряда. Для этого нужно выполнить команду меню **Правка – Заполнить – Прогрессия** (рис. 2.7).

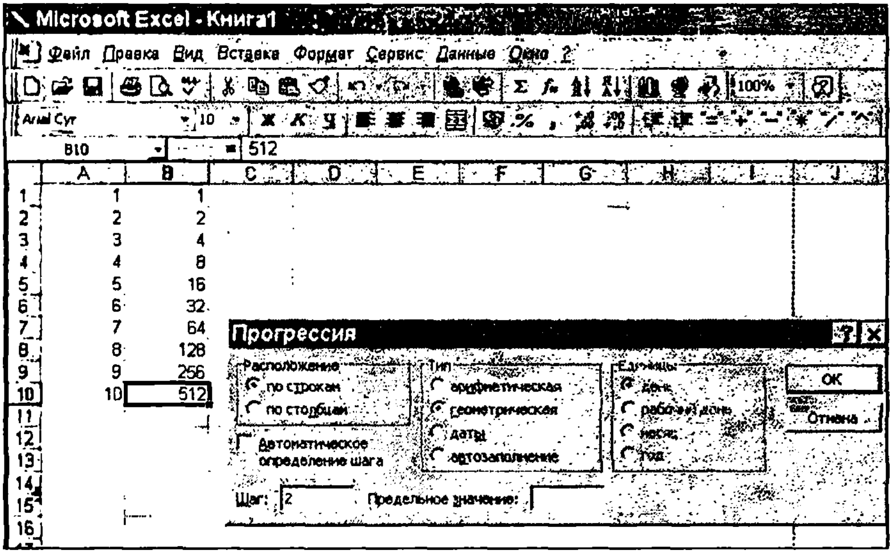


*Рис. 2.7. Диалоговое окно **Прогрессия***

При этом откроется диалоговое окно, в котором можно ввести параметры ряда. Здесь можно выбрать тип ряда – арифметическую или геометрическую прогрессию, его расположение – по строкам или по столбцам, приращение (по умолчанию 1) и конечное значение ряда. Если последняя величина не указана, то длина ряда определяется по числу выделенных ячеек. Геометрическая прогрессия строится путем умножения текущего элемента на значение ряда.

Вставка и удаление ячеек, столбцов, строк. Для вставки ячейки выделите ячейку, перед которой она должна быть вставлена, и выберите команду меню **Вставка – Ячейки.** Для вставки столбца выделите столбец, перед которым должен быть вставлен новый, и выберите команду меню **Вставка – Столбец.** Аналогично вставляются строки. Для удаления строки или столбца используйте команду меню **Правка – Удалить.**

Копирование, вырезание и вставка. Эти команды можно выполнить с помощью меню. Соответствующие команды: **Вырезать, Копировать, Вставить** – находятся в меню **Правка.**

Если копируемые ячейки уже выделены, то после щелчка правой кнопкой мыши над выделенным блоком откроется контекстно-зависимое меню. В этом меню есть все команды, необходимые для копирования, вырезания и вставки.

Можно использовать технику «перетащить-и-оставить» для копирования и перемещения содержимого выделенных ячеек. Для перемещения содержимого блока ячеек выполняются следующие действия:

1. Выделите блок ячеек.
2. Установите указатель мыши на рамку выделенного блока.

3. Нажмите левую кнопку мыши и, удерживая ее, перетащите рамку в нужное место.

При необходимости скопировать *содержание* блока ячеек, перед тем как нажать левую кнопку мыши, нажмите клавишу <Ctrl>. При этом рядом с указателем мыши появится символ «+». Необходимо удерживать клавишу <Ctrl> в течение всего процесса перетаскивания рамки.

Специальная вставка. Существует возможность осуществить выборочную, или специальную, вставку, т. е. скопировать или только формулы, или только значения, или только форматы. Соответствующее диалоговое окно вызывается с помощью команды меню **Правка – Специальная вставка** (рис. 2.8).

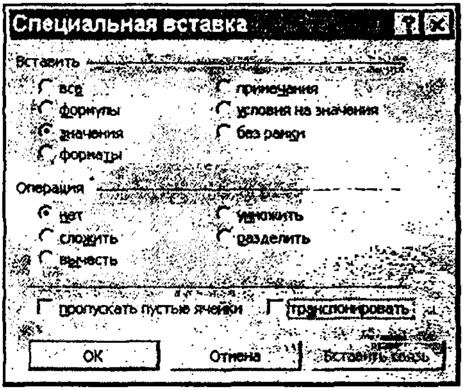


Рис. 2.8. Диалоговое окно ***Специальная вставка***

Одновременно в этом диалоговом окне можно выполнить арифметические операции и операцию транспонирования.

Прокрутка окна с фиксацией заголовков строк и столбцов. Часто документы бывают большими, и при этом необходимо видеть заголовки строк и столбцов. Для этого выделяется ячейка, выше и левее которой надо зафиксировать строки и столбцы при прокрутке. Затем выполняется команда меню **Окно – Закрепить области**. Выполняется прокрутка стрелками или с помощью полосы прокрутки. Отмена фиксации производится командой **Окно – Снять закрепление областей**.

Копирование формул. При копировании содержимого ячейки копируется и содержащаяся в ней формула. Это значительно ускоряет работу. Копирование формул идентично общей процедуре копирования. Копирование отдельно формулы или значения осуществляется командой **Специальная вставка**. Рассмотрим простой пример с использованием функции суммы. Допустим в ячейках A1:A3 введены значения: 30; 45; 64; в ячейках B1:B3 – значения 50; 60; 70; в ячейках C1:C3 – значения 100; 200; 300. В ячейку A4 введем функцию суммы и просуммируем значения в ячейках A1:A3. Закончим операцию нажатием клавиши <Enter>. Выделим ячейку A4. Теперь подведем курсор в правый нижний угол выделенной ячейки. Курсор приобретает вид черного крестика. Переместим курсор в ячейки

В4, С4. Отпустим кнопку мыши. В ячейках В4 и С4 будет результат суммирования соответствующих столбцов. Выделив ячейку В4, в командной строке увидим формулу =СУММ(В1:В3). В ячейке С4 будет записано =СУММ(С1:С3). Аналогичное копирование производится с любыми формулами. Например, в ячейку D1 введем формулу =C1–B1–A1. Скопируем ее в ячейки D2; D3; D4. Выделив затем ячейку D2, увидим формулу =C2–B2–A2. В ячейке D3 – формулу =C3–B3–A3.

Относительные и абсолютные адреса ячеек. При копировании и переносе формул автоматически изменяются адреса ячеек в формулах. Во многих случаях это очень удобно, но при использовании функций эта ситуация выглядит с точностью до наоборот. Рассмотрим такой пример. Построим степенную функцию $y = 2x^{0.5}$. Введем параметры функций. В ячейку A1 введем значение 2, в ячейку A2 – 0,5. В ячейки B1:B10 – ряд значений от 1 до 10. В ячейку C1 введем формулу =A1*B1^A2 и скопируем ее в ячейки C2:C10. Результат неудовлетворительный. Выделим, например, ячейку C10. Там введена формула =A10*B10^A11. Поэтому такой и результат. Для выполнения подобных операций необходимо зафиксировать адрес ячейки. Операция фиксирования адреса ячейки выполняется с помощью ввода символа $. Ввод символа $ перед именем столбца предохраняет от изменений по столбцам, ввод символа $ перед номером строки – от изменений по строкам. Таким образом, наша формула должна выглядеть так: =A$1*B1^A$2. Для ввода абсолютных (фиксированных) адресов ячеек удобней пользоваться клавишей <F4>. То есть после щелчка мыши по адресу ячейки нажимается клавиша <F4>, что приводит к появлению символа $ перед именем столбца и номером строки, и наша формула будет выглядеть так: =A1*B1^A2. Теперь скопируем ее в ячейки C2:C10 и получим правильный результат.

Другой способ абсолютной адресации заключается в назначении ячейкам имен и использовании их в формулах. Для назначения ячейке имени применяется команда меню **Вставка – Имя – Определить.**

Редактор формул. Редактор формул – программа, которая входит в комплект поставки Word для Windows и служит для записи сложных математических формул (но не для работы с ними). Эта программа часто используется для документирования формул, с которыми идет работа в Excel, т. к. программа не обладает развитыми средствами для этого.

Объект (математическая формула) создается в Редакторе формул и вставляется в Excel. Для того чтобы вставить объект в Excel, выполняется команда меню **Вставка – Объект**. Откроется диалоговое окно (рис. 2.9), где в списке **Тип объекта** нужно выбрать строку **Microsoft Equation 3.0**. Так обозначается Редактор формул. В окне приложения Excel появится новый объект, обрамленный штриховой рамкой, а также меню и панели инструментов Редактора формул (рис. 2.10).

После того как редактирование формулы закончено, щелкните мышью в любом месте рабочего листа. При этом выполнение Редактора формул будет завершено, и в Excel останется вставленный объект. Данный объект можно редактировать, вызывая Редактор формул.

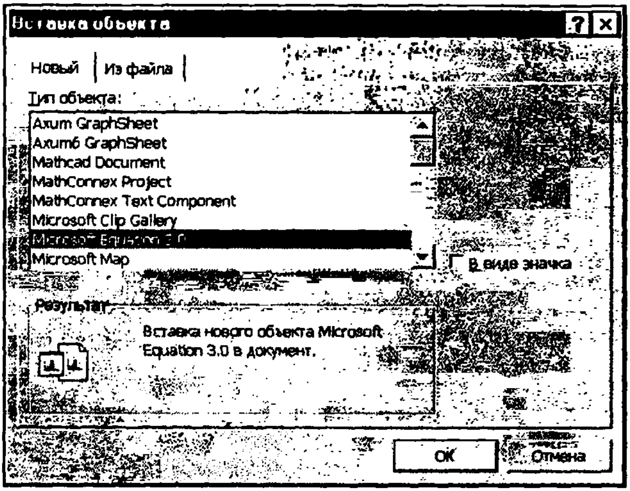


Рис. 2.9. *Диалоговое окно* ***Вставка объекта***

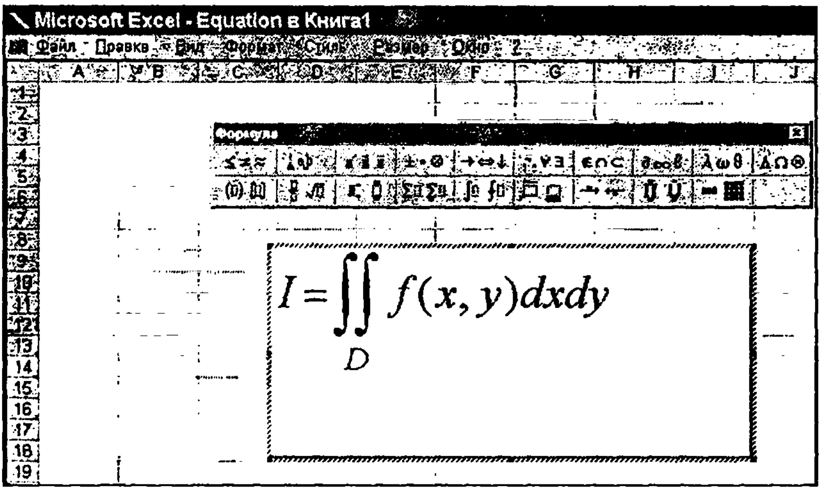


Рис. 2.10. Excel с активным объектом Редактора формул

2.3. Создание диаграмм

Создание диаграмм осуществляется с помощью Мастера диаграмм, кнопка которого находится на панели инструментов **Стандартная**. Введем данные для создания диаграммы. Построим логарифмические функции с основанием логарифма больше единицы, например натуральный логарифм, и с основанием меньше единицы, скажем, 0,5. В строке A1:A3 введем обозначения – x; $\ln x$; $\log_{0,5}x$. В столбец A2:A11 введем ряд чисел от 1 до 10. В ячейку B2 с помощью Мастера функций, категории **Математические**, введем функцию =lnA2 и скопируем ее в ячейки столбца B3:B11. В ячейку C2 с помощью Мастера функций введем функцию =log0,5A2 (логарифмическая функция с произвольным основанием – LOG – см. рис. 2.11) и скопируем функцию в ячейки C3:C11.

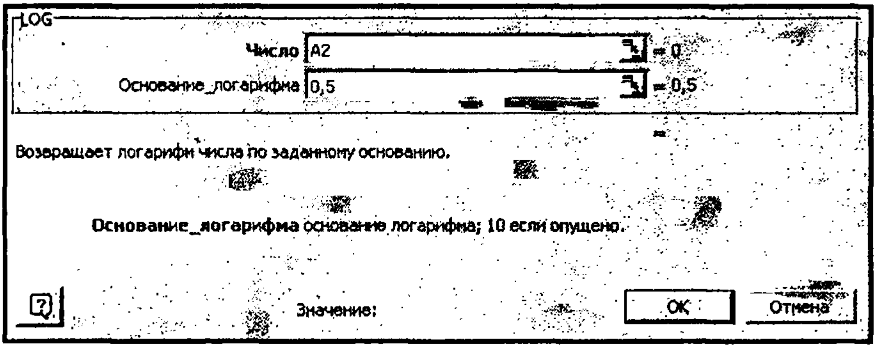


Рис. 2.11. Диалоговое окно функции LOG с заданным основанием

Теперь построим графики. Для этого выделим область данных вместе с названиями рядов и нажмем кнопку Мастера диаграмм. Откроется первое диалоговое окно Мастера диаграмм (рис. 2.12).

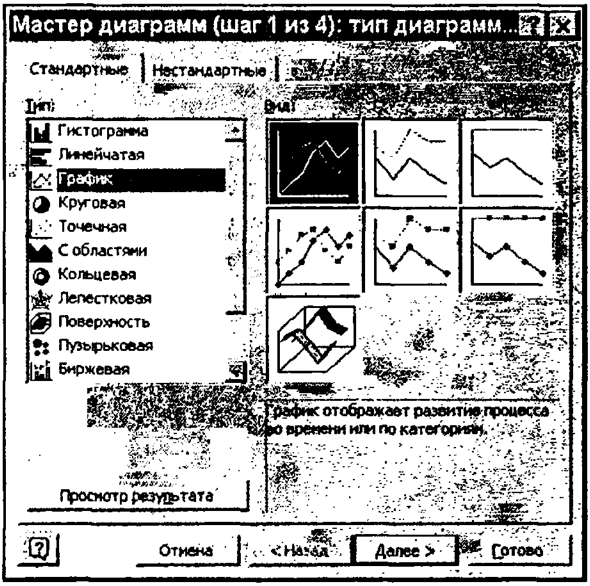


Рис. 2.12 Первое диалоговое окно Мастера диаграмм

Выберем тип диаграммы – график, выберем вид диаграммы и нажмем кнопку Далее. Откроется второе диалоговое окно Мастера диаграмм – рис. 2.13.

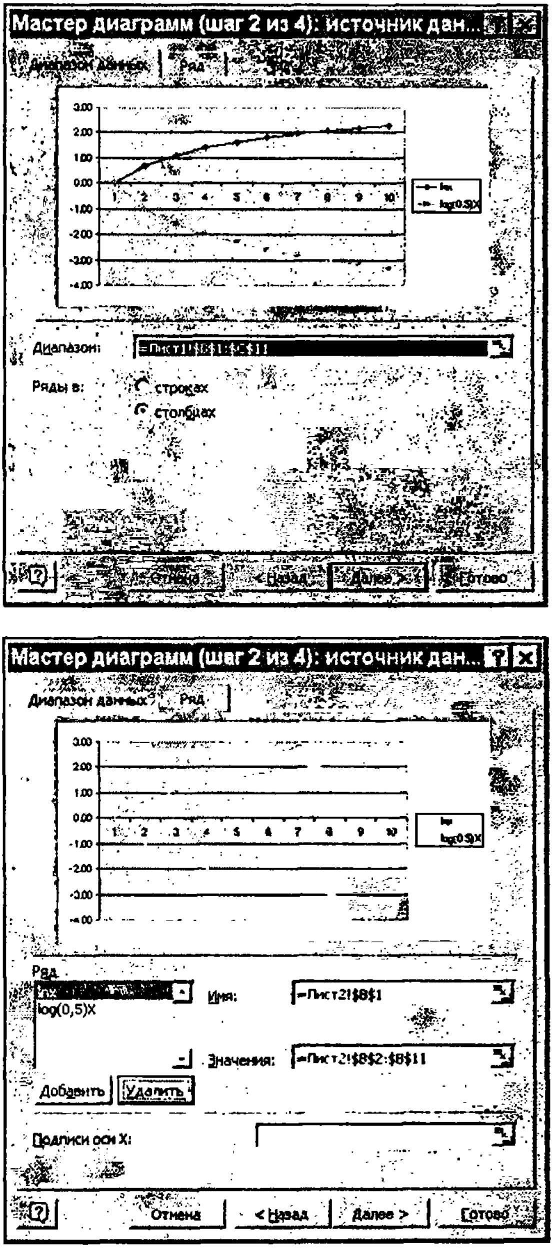


Рис. 2.13. Второе диалоговое окно Мастера диаграмм

На вкладке **Диапазон данных** можно отредактировать введенную область данных или ввести их, если они первоначально не были выделены. На вкладке **Ряд**, можно отредактировать данные каждого ряда, добавить или удалить ряд. Удалим ряд *X* и нажмем кнопку **Далее**. Откроется третье диалоговое окно – рис. 2. 14.

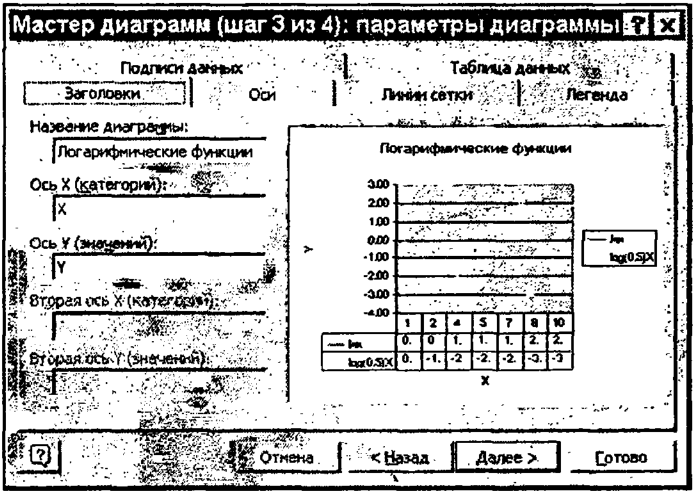


Рис. 2.14. Третье диалоговое окно Мастера диаграмм

Внесем заголовки – название диаграммы, оси *X* и *Y*, вставим таблицу данных. На вкладке **Легенда**, следует проверить, стоит ли флажок **Добавить легенду**, и выбрать размещение таблицы. Нажмем кнопку **Далее**. В четвертом диалоговом окне можно выбрать размещение диаграммы – на отдельном листе или имеющемся.

Осталось вывести готовую диаграмму и отформатировать ее (рис. 2.15). Для форматирования диаграмм существуют различные опции. Удобней всего выделить щелчком мыши необходимый элемент и правой кнопкой мыши вызвать контекстно-зависимое меню. Так, щелчком на общем поле диаграммы выводится меню, в котором, выбрав формат области построения, можно задать необходимый фон заливки. Щелчком на области построения диаграммы правой кнопкой вызывается меню, выводится формат области построения, в котором можно убрать фон заливки или задать необходимый. После щелчка на графике функции тот выделяется с появлением маркеров, и, после нажатия правой кнопки мыши, вызывается меню, в котором, выбирая формат рядов данных, можно задать толщину линии, тип маркеров или цвет. Аналогично форматируются оси.

Приведенные краткие сведения по работе с электронными таблицами Excel не отражают все возможности программы. В дальнейшем изложении мы будем раскрывать возможности Excel и использование программы в математических методах экономики.

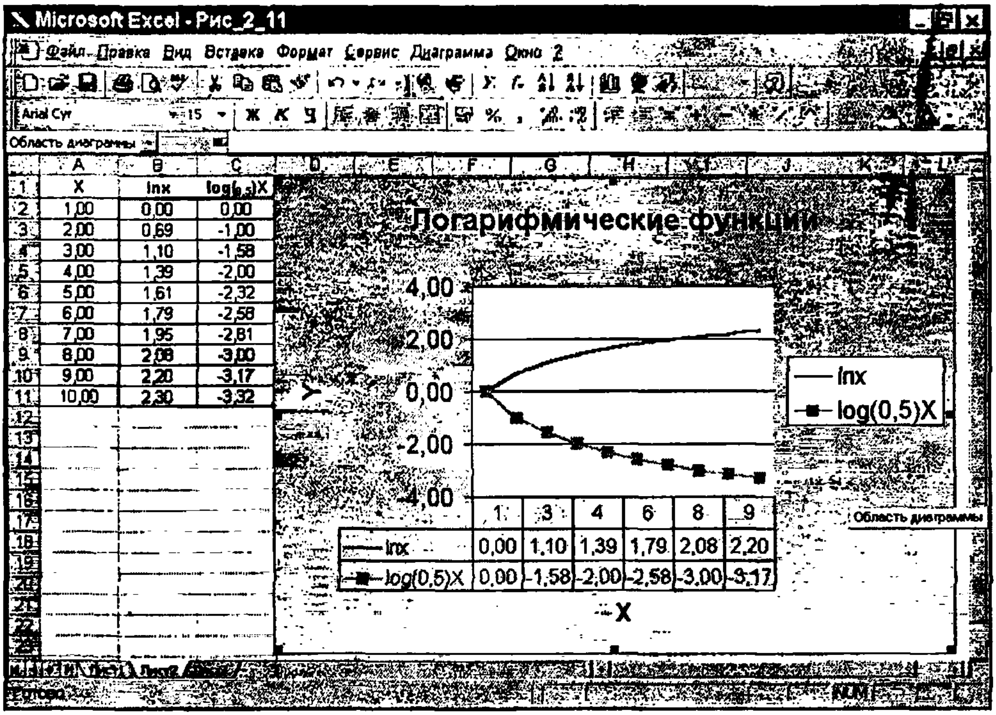
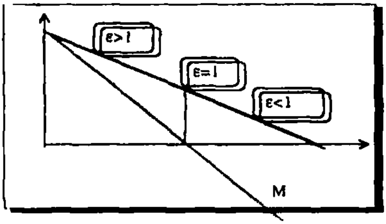


Рис. 2.15. Готовая диаграмма

Глава 3. Предельный анализ в экономике

3.1. Основной инструментарий предельного анализа



Применение дифференциального исчисления в экономике называется предельным анализом. Он предназначен для анализа прироста переменных величин.

Предельные затраты

Затраты производства C однородной продукции есть функция количества продукции x. Поэтому можно записать:

$$C = C(x).$$

Предположим, что количество продукции увеличивается на Δx. Продукции $x + \Delta x$ соответствуют затраты производства продукции $C(x + \Delta x)$. Следовательно, приращению количества продукции Δx соответствует приращение затрат производства продукции

$$\Delta C = C(x + \Delta x) - C(x).$$

Среднее приращение затрат производства есть $\Delta C/\Delta x$. Это есть приращение затрат производства на единицу приращения количества продукции. Предел

$$\lim_{\Delta x \to 0} \frac{\Delta C}{\Delta x} = C'(x)$$

называется предельными затратами производства. Величина предельных затрат показывает величину изменения в уровне полных затрат производства в результате увеличения (или уменьшения) выпуска продукции на единицу.

Предельный доход

Если мы обозначим через $R(x)$ выручку (доход) от продажи x единиц товара, то предел

$$\lim_{\Delta x \to 0} \frac{\Delta R(x)}{\Delta x} = R'(x)$$

называется предельной выручкой. Предельная выручка показывает изменение совокупной выручки фирмы в результате продажи одной дополнительной единицы продукции.

Пример 1. Затраты производства C зависят от объема продукции x по формуле:

$$C = 100x - x^3/30.$$

Допустим, необходимо определить предельные затраты, если объем производства составляет: а) 5 единиц; б) 10 единиц продукции.

Имеем

$$C' = 100 - 0{,}1x^2,$$

откуда а) $C'(5) = 100 - 0{,}1 \cdot 5^2 = 97{,}5$;

б) $C'(10) = 100 - 0{,}1 \cdot 10^2 = 90$.

Это означает, что при объеме производства в 5 единиц продукции, затраты на изготовление следующей (шестой) единицы продукции составят 97,5; при объеме производства в 10 единиц они составят 90.

Пример 2. Функция цен спроса на какой-либо товар определяется формулой

$$p = 10 - 2x,$$

где x – спрос, а p – цена. Выручка от продажи товара есть

$$R = x(10 - 2x) = 10x - 2x^2,$$

откуда $R' = 10 - 4x$. Если, например, $x = 2$, то $R'(2) = 2$.

Это означает, что если спрос возрастет с 2 до 3 единиц, то выручка возрастет приблизительно на 2 единицы.

Предельная прибыль

Если $R(x)$ – есть функция выручки от объемов производства; $C(x)$ – функция затрат от объемов производства, то прибыль $PR(x)$ представляет собой разницу между совокупной выручкой и совокупными затратами:

$$PR(x) = R(x) - C(x).$$

Предельная прибыль представляет собой производную функции прибыли по объемам производства или разницу между предельной выручкой и предельными затратами:

$$PR'(x) = R'(x) - C'(x).$$

Максимизация прибыли производится в соответствии с необходимыми и достаточными условиями существования экстремума функций.

Предельная полезность

Предельная полезность – величина изменения полезности данного блага с увеличением объема его потребления на единицу при условии, что количество других потребляемых благ остается без изменения. Предельная полезность измеряется производной функции полезности:

$$MU = U'(x) = dU/dx.$$

Предельный продукт

Предельным продуктом фактора производства называется изменение объема выпускаемой продукции в результате использования одной дополнительной единицы этого фактора при условии, что количество прочих применяемых факторов производства остается неизменным. Величина предельного продукта определяется как производная функции объема выпускаемой продукции Q по данному фактору производства x:

$$MP = dQ/dx.$$

Данная формула выражает также величину предельной производительности фактора производства.

Предельная производительность труда

Если мы упростим функцию выпуска продукции и представим ее только как функцию численности персонала (L) при условии, что остальные факторы производства остаются неизменными, т. е. $Q = f(L)$, то предельная производительность труда будет представлять собой производную функции выпуска продукции по величине численности персонала:

$$ML = Q'(L) = dQ / dL.$$

Предельная производительность труда показывает величину изменения объема выпускаемой продукции при изменении численности персонала на единицу.

Предельная производительность капитала

Если аналогично представить функцию выпуска продукции только от капитала I, принимая неизменными остальные факторы производства, т. е. $Q = f(I)$, то предельная производительность капитала будет определяться производной функции выпуска по величине капитала:

$$MI = Q'(I) = dQ / dI.$$

Предельная норма замещения факторов производства

Предельная норма замещения факторов производства представляет собой коэффициент, характеризуемый определенным соотношением прироста затрат одного фактора (капитала) и сокращения затрат другого (труда), при котором данный уровень производства обеспечивается при наименьших затратах. Конкретные формулы предельной нормы замещения будут приведены при рассмотрении производственных функций.

3.2. Эластичность функции

С помощью производной можно вычислить приращение зависимой переменной, соответствующее приращению независимой переменной. То есть производная характеризует скорость изменения функции с изменением аргумента x. Данный показатель бывает неудобен тем, что зависит от выбора единиц измерения. Во многих задачах удобнее вычислить процент прироста (относительное приращение) зависимой переменной, соответствующее проценту прироста независимой переменной.

Это приводит нас к понятию эластичности функции (иногда ее называют относительной производной).

Дана функция $y = f(x)$. Предположим, что приращение независимой переменной x есть Δx, а относительное приращение независимой переменной x есть $\Delta x/x$. Соответствующее приращение зависимой переменной будет равно:

$$\Delta y = f(x + \Delta x) - f(x),$$

а относительное приращение зависимой переменной – $\Delta y / y$.

Отношение относительного приращения функции (зависимой переменной) к относительному приращению независимой переменной равно:

$$\Delta y/y \,/\, \Delta x/x.$$

Это соотношение показывает, во сколько раз относительное приращение функции больше относительного приращения независимой переменной. Его можно записать также в следующем виде:

$$\Delta y/y\,/\,\Delta x/x = x/y \cdot \Delta y/\Delta x.$$

Если существует производная функции $y = f(x)$, то

$$\lim_{\Delta x\to 0}\left(\frac{\Delta y}{y}\Big/\frac{\Delta x}{x}\right) = \lim_{\Delta x\to 0}\left(\frac{x}{y}\cdot\frac{\Delta y}{\Delta x}\right) = \frac{x}{y}\lim_{\Delta x\to 0}\frac{\Delta x}{\Delta y} = \frac{x}{y}f'(x) = \frac{x}{y}\cdot\frac{dy}{dx}.$$

Данный предел отношения между относительным приращением независимой переменной и зависимой переменной, когда $\Delta x \to 0$, называется эластичностью функции $y = f(x)$ относительно переменной x. Эластичность функции $y = f(x)$ обозначается так:

$$\varepsilon_x(y) = x/y \cdot dy/dx.$$

Эластичность относительно x есть приближенный процентный прирост функции, соответствующий приращению независимой переменной на 1%.

Пример. Рассчитать эластичность функции $y = 3x - 6$.

По определению эластичности имеем:

$$\varepsilon_x(y) = \frac{x}{y}\cdot\frac{dy}{dx} = \frac{x}{3x-6}\cdot 3 = \frac{x}{x-2}.$$

Если, например, $x = 10$, то эластичность функции равна:

$10/(10 - 2) = 10/8 = 1{,}25$. Это означает, что если x возрастет на 1%, то y возрастет на 1,25%.

Свойства эластичности

1. Эластичность – безразмерная величина, значение которой не зависит от того, в каких единицах измеряются y и x.

2. Эластичности взаимно обратных функций – взаимно обратные величины:

$$\varepsilon_x(y) = \varepsilon_y(x).$$

Например, эластичность величины спроса по цене обратна эластичности цены по величине спроса.

3. Эластичность произведения двух функций (u и v) равняется сумме показателей эластичности сомножителей, т. е.

$$\varepsilon_x(uv) = \frac{x}{uv}(uv)' = \frac{x}{uv}(uv' + u'v) = \frac{x}{uv}uv' + \frac{x}{uv}u'v = \frac{x}{v}v' + \frac{x}{u}u' = \varepsilon_x(u) + \varepsilon_x(v).$$

Пример. Вычислить эластичность функции $y = xe^x$.

Можно написать $u = x$; $v = e^x$. Следовательно

$$\varepsilon_x(y) = \frac{x}{x}(x') + \frac{x}{e^x}(e^x)' = 1 + x.$$

Эластичность частного двух функций равняется разности показателей делимого и делителя:

$$\varepsilon_x\left(\frac{u}{v}\right) = \frac{x}{u}u' - \frac{x}{v}v' = \varepsilon_x(u) - \varepsilon_x(v).$$

Эластичности элементарных функций

1. Эластичность степенной функции $y = x^a$ постоянна и равна показателю степени a:

$$\varepsilon_x(x^a) = \frac{dx^a}{dx} \cdot \frac{x}{x^a} = \frac{ax^{a-1} \cdot x}{x^a} = a.$$

2. Эластичность показательной функции $y = a^x$ пропорциональна x:

$$\varepsilon_x(a^x) = x\ln(a).$$

3. Эластичность линейной функции $y = ax + b$:

$$\varepsilon_x(ax + b) = ax / (ax + b).$$

3.3. Виды эластичности в экономике

Эластичность спроса относительно цены

Во многих экономических исследованиях необходимо узнать изменение спроса, вызываемое определенным изменением цены. Иначе говоря, определяется эластичность спроса относительно цены (ε_c).

Предположим, что есть зависимость между спросом на товар q и его ценой P при условии, что цена других товаров, доходы потребителей и структура потребностей – это постоянные величины: $q = f(p)$. Тогда: $\varepsilon_c = P/q \cdot dq/dP$.

Эластичность спроса относительно цены приблизительно показывает, как изменится спрос на данный товар, если его цена возрастет на 1%.

В большинстве случаев функция спроса есть понижающаяся функция, т. к. с повышением цены на товар спрос на него понижается.

Следовательно, в таких случаях $dq/dP < 0$.

Чтобы избежать отрицательных чисел, при изучении эластичности спроса принимается, что

$$\varepsilon_c = - P/q \cdot dq/dP.$$

Если $\varepsilon_c > 1$, т. е. если повышение цены на 1% соответствует снижению спроса более чем на 1% , говорят, что спрос эластичен.

Если $\varepsilon_c = 1$, т. е. повышение цены на 1% соответствует понижению спроса на 1%, то спрос – нейтрален.

Если $0 <= \varepsilon_c <= 1$, т. е. если повышение цены на 1% соответствует понижению спроса менее чем на 1% , то спрос – неэластичен.

Эластичность спроса относительно дохода

Если факторы, от которых зависит спрос на данный товар, не изменяются, а изменяется только доход потребителей, то спрос q есть функция дохода r:

$$q = f(r).$$

Пусть Δr есть прирост дохода, а Δq – соответствующее повышение спроса на товар; частное

$$\frac{\Delta q}{q} : \frac{\Delta r}{r}$$

означает в этом случае относительный прирост спроса, соответствующий повышению спроса на единицу.

Эластичностью спроса относительно дохода называется выражение:

$$\varepsilon_r(q) = \lim_{\Delta r \to 0}\left(\frac{\Delta q}{q} \div \frac{\Delta r}{r}\right) = \frac{r}{q}\cdot\frac{dq}{dr}.$$

Эластичность спроса относительно дохода есть мера реакции спроса на изменение доходов потребителей.

Эластичность предложения

Под предложением понимается количество некоторого товара, предлагаемого на продажу в единицу времени. Как правило, предложение какого-либо товара в данный период есть возрастающая функция цены. При прочих равных условиях предложение при данной цене больше, чем при более низкой цене Эластичность предложения можно определить аналогично эластичности спроса.

Пусть $S = S(p)$ есть функция предложения. Пусть ΔP – приращение цены, а ΔS – соответствующее приращение предложения. Относительное приращение цены составляет $\Delta p/p$, а относительный прирост предложения $\Delta S/S$.

Предел

$$\varepsilon_p(s) = \lim_{\Delta p \to 0}\left(\frac{\Delta s/s}{\Delta p/p}\right) = \frac{p}{s}\cdot\frac{ds}{dp}$$

называется эластичностью предложения относительно цены.

Следовательно, эластичность предложения относительно цены приблизительно определяет процент прироста предложения на 1% прироста цены.

Эластичность полных и удельных затрат

Если предприятие производит x единиц какого-либо товара и определена функция полных затрат, то эластичность полных затрат составляет

$$\varepsilon_x(C_n) = x/C \cdot dC/dx = dC/dx \,/\, C/x,$$

и, следовательно, эластичность полных затрат есть отношение предельных затрат к средним удельным затратам.

Эластичность средних удельных затрат ($C_{уд} = C_{пол}/x$) составляет:

$$\varepsilon_x(C_y) = \frac{x}{C_y}\cdot\frac{dC_e}{dx} = \frac{x}{\frac{C_п}{x}}\cdot\left(\frac{x\cdot\frac{dC_п}{dx} - C_п}{x^2}\right) = \frac{x^2}{C_п}\cdot\left(\frac{x\cdot\frac{dC_п}{dx} - C_п}{x^2}\right) =$$

$$= \frac{x}{C_п}\cdot\frac{dC_п}{dx} - 1 = \varepsilon_{C_п} - 1.$$

и, следовательно, эластичность удельных затрат на единицу меньше эластичности полных затрат. Если $\varepsilon_x(C_п) = 1$, то $\varepsilon_x(C_{уд}) = 0$, а это означает, что удельные затраты постоянны. Отсюда следует, что $xdC \,/\, dx - C = 0$; $dC/dx = C/x$. Таким образом, если эластичность полных затрат равняется единице, то полные предельные затраты равны полным средним затратам.

3.4. Определение и графическое представление предельных показателей и эластичностей функций в Excel

Для определения предельных показателей в Excel необходимо ввести данные, рассчитать функции и построить график. Пример определения функции предельных затрат дан на рис. 3.1 для функции затрат $C = 2 + 3x - 0{,}25x^2$ и предельных затрат $C' = 3 - 0{,}5x$.

В столбец А введены данные по объему производства. В ячейку В3 введена формула предельных затрат и скопирована с помощью мыши до конца ряда.

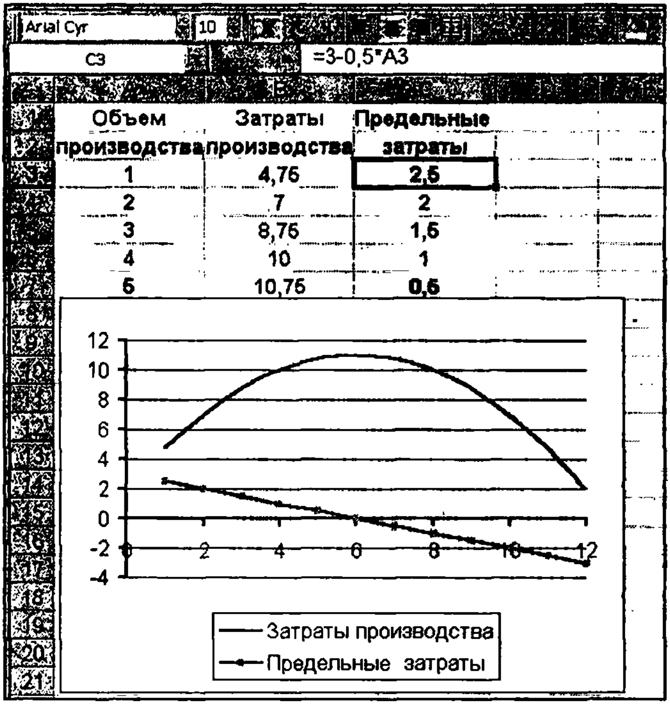


Рис. 3.1. Определение предельных затрат

В ячейку С3 введена формула предельных затрат и также скопирована до конца ряда. По рассчитанным данным построены графики затрат и предельных затрат (см. рис. 3.1).

Пример определения эластичности функции приведен на рис. 3.2. В этом примере взята функция спроса от дохода вида:

$$q = ar/(r + b).$$

Эластичность этой функции относительно дохода будет равна:

$$\varepsilon r(q) = b/(r + b).$$

Приняв численные значения $a = 2$; $b = 5$, рассчитываем функции спроса и эластичности спроса относительно дохода, по которым строится график (рис. 3.2).

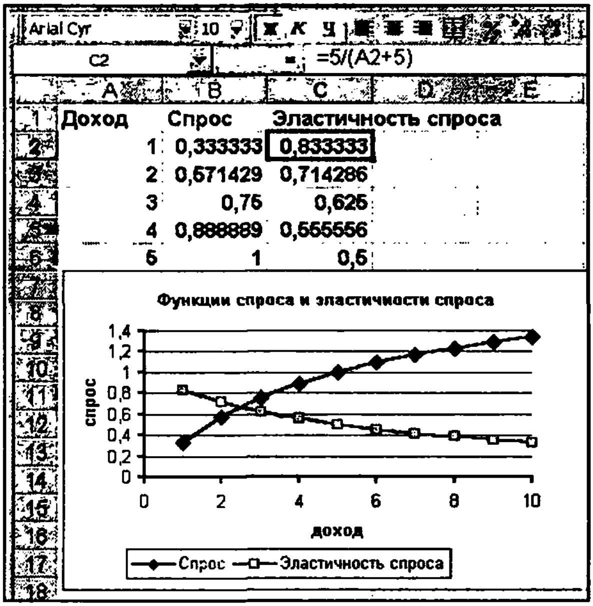


Рис. 3.2. Определение эластичности функции

Глава 4. Работа с системой Mathcad

4.1. Основные возможности системы Mathcad

Mathcad 2000 Professional – мощная универсальная массовая математическая система. Она позволяет выполнять как численные, так и аналитические (символьные) вычисления, имеет удобный математически-ориентированный интерфейс и прекрасные средства графики.

Уникальное свойство Mathcad – возможность описания математических алгоритмов в естественной математической форме с применением общепринятой символики для математических знаков. Такой подход значительно облегчает восприятие математической сущности решаемой задачи и избавляет пользователя от необходимости изучения некоторого промежуточного языка программирования. Можно сказать, что в Mathcad методы решения математических задач без использования программирования доведены до совершенства.

Это не означает, что в системе нет своего языка программирования. В действительности он есть, но это математически-ориентированный особый язык программирования сверхвысокого уровня, используемый в основном как входной язык для диалога с системой. В подавляющем большинстве расчетных задач этот язык позволяет задать их решение в виде вводимых математических формул и указать тип желаемых результатов (таблицы или графики). Специальные приемы применяются лишь для задания циклического изменения переменных и создания ранжированных переменных. По существу, входной язык системы – это промежуточное звено между скрытым от пользователя истинным языком программирования и текстом документа, видимым в окне дисплея. Важно подчеркнуть, что от пользователя не требуется знать язык программирования, достаточно освоить приближенный к естественному входной язык системы.

4.2. Главное меню системы

Основы пользовательского интерфейса. Пользовательский интерфейс системы разработан так, что пользователь, имеющий элементарные навыки работы с Windows-приложениями, может сразу начать работу с Mathcad. Это позволяет сократить описание общепринятых для Windows-приложений деталей работы с ними.

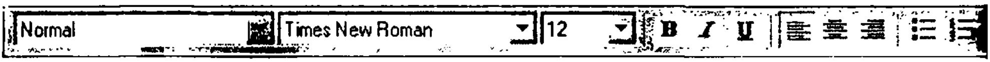
Сразу после запуска система готова к созданию документа с необходимыми пользователю вычислениями. Соответствующее окно редактирования получает название Untitled:N, где N – порядковый номер документа. Он начинается с цифры 1, максимальный номер окна равен 8.

Основную часть экрана занимает окно редактирования, изначально пустое. Вверху окна находятся строки с основными элементами интерфейса. Верхняя строка – строка заголовка. Она отображает название загруженного или вводимого с клавиатуры документа.

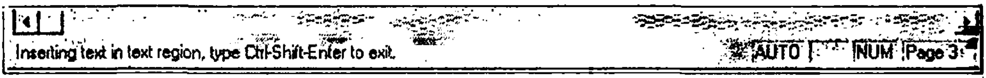
Вторая строка окна системы главное меню. Третью строку окна системы занимает панель инструментов. Она содержит несколько групп пиктограмм управления, каждая их которых дублирует одну из важнейших операций главного меню.

В этой же строке, справа, в более ранних версиях пакета, могут быть расположены пиктограммы вывода наборных панелей.

Четвертая строка верхней части экрана содержит типовые средства управления шрифтами: переключатели типа символов, набора фонтов и размеров шрифта, три пиктограммы типа шрифтов (полужирный, курсив и подчеркивание), выравнивание, список.

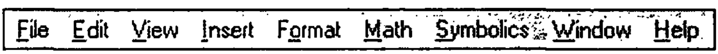


Внизу экрана, кроме полосы горизонтальной прокрутки, расположена еще одна строка – строка состояния. В ней выводится служебная информация, краткие комментарии, номер страницы документа и др.



Вид окна системы может меняться. Возможно удаление отдельных деталей интерфейса и их перемещение по экрану с помощью мыши.

Главное меню системы. Рассмотрим назначение опций главного меню:



- **File** (Файл) – работа с файлами;
- **Edit** (Правка) – редактирование документов;
- **View** (Вид) – управление видом экрана и панелями инструментов;
- **Insert** (Вставка) – управление вставкой объектов;
- **Format** (Формат) – управление форматированием;
- **Math** (Математика) – управление процессом вычислений;
- **Symbolics** (Символика) – выбор операций символьного процессора;
- **Window** (Окно) – управление окнами системы;
- **Help** (Помощь) – работа со справочной базой данных по системе.

Если сделать опцию **File** главного меню активной, появится раскрывающееся меню, которое содержит следующие команды (табл. 4.1):

Таблица 4.1

File		
New	Создать	*Открыть окно для нового* документа
Open	Открыть	Открыть существующий документ
Close	Закрыть	Закрыть текущий документ
Save	Сохранить	Сохранить на диске текущий документ
Save As	Сохранить как	Сохранить на диске текущий документ под новым именем
Send	Отправить	Соединение с почтовым сервером
Page Setup	Параметры страницы	Установить на странице
Print Previev	Просмотр	Предварительный просмотр документа
Print	Печать	Распечатать документ
Exit	Выход	Выйти из среды Mathcad

Опция **Edit** (Правка) содержит следующие команды (табл. 4.2):

Таблица 4.2

Edit		
Undo	Отменить	Отменить последнюю операцию редактирования
Redo	Восстановить	Восстановить отмененную операцию
Cut	Вырезать	Переместить выделенное в буфер обмена
Copy	Копировать	Скопировать выделенное в буфер обмена
Past	Вставить	Вставить выделенное из буфера обмена в документ
Past Spesial	Специальная вставка	Вставить из буфера обмена в различном формате
Delete	Удалить	Удалить выделенное
Select All	Выделить все	Выделить все области в документе
Find	Найти	Найти заданную текстовую или математическую строку

Таблица 4.2 (окончание)

Edit		
Replace	Заменить	Найти и заменить текстовую или математическую строку
Go to Page	Перейти к странице	Расположить начало указанной страницы в начале документа
Check Spelling	Проверка правописания	Проверка правописания выделенного текста
Links	Связи	Использование связанных и внедренных объектов для обмена данными
Object	Объект	Открытие и редактирование внедренных объектов

Опция **View** (Вид) содержит следующие команды (табл. 4.3):

Таблица 4.3

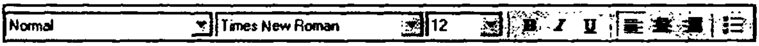
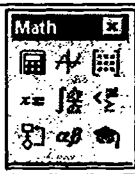
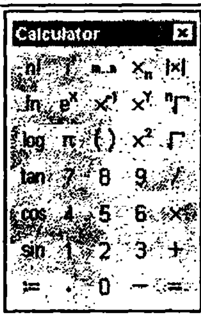
View			
Toolbars	Панели инструментов	Выводит (убирает) панели инструментов	
	Standart	Выводит (убирает) стандартную панель инструментов	
			
	Formatting	Выводит (убирает) панель форматирования	
			
	Math	Выводит (убирает) панель математических символов	
	Calculator	Выводит (убирает) панель операторов вычислений	

Таблица 4.3 (продолжение)

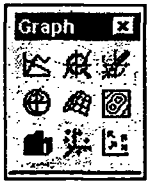
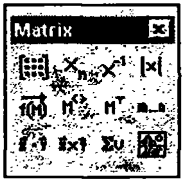
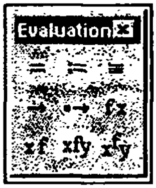
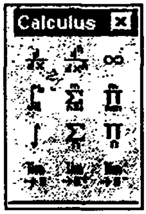
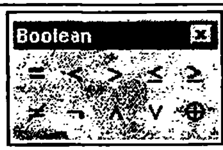
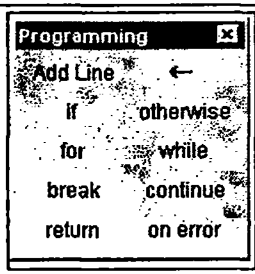
View			
	Graph	Выводит (убирает) панель операторов графики	
	Matrix	Выводит (убирает) паиель операторов матричных вычислений	
	Evaluation	Выводит (убирает) панель операторов расчетов	
	Calculus	Выводит (убирает) панель операторов объектов высшей математики	
	Boolean	Выводит (убирает) панель операторов отношений	
	Programming	Выводит (убирает) панель операторов программирования	

Таблица 4.3 (продолжение)

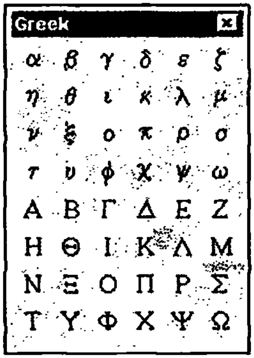
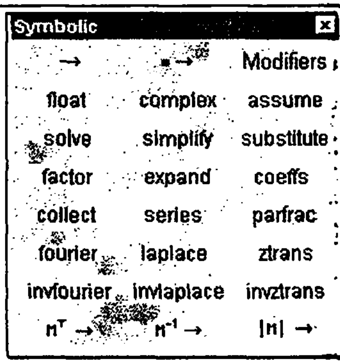
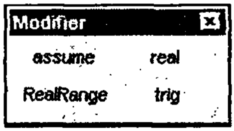
View			
	Greek	Выводит (убирает) панель греческого алфавита	
	Symbolic	Выводит (убирает) панель операторов символьных вычислений	
	Modifier	Выводит (убирает) панель операторов видоизменений	
Ruler	Линейка	Вывод линейки для рабочего листа	
Regions	Показать области	Показать области путем их подсветки	
Zoom	Масштаб	Изменить масштаб изображения рабочего документа	
Refresh	Реставрация окна	Восстановить изображение документа, пострадавшее после неудачных копирования и перемещения блоков	
Animate	Анимация	Создание и проигрывание анимационных файлов	

Таблица 4.3 (окончание)

View		
Playback	Воспроизвести	Воспроизвести существующий анимационный клип

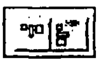
Опция **Insert** (Вставка) содержит следующие команды (табл. 4.4):

Таблица 4.4

Insert		
Graph	Графика	Создание графиков
Matrix	Матрицы	Создать матрицу или вектор, изменить размеры
Function	Вывод функций 	Показать прокручивающийся список имеющихся функций
Unit	Вставить единицы	Вставить единицы измерений
Picture	Изображение	Создать шаблон области для импорта рисунка
Area	Область	Создание закрытой области
Math Region	Математическая область	Создать математическую область внутри текстовой области
Text Region	Создать текстовую область	Создать текстовую область с началом в месте расположения курсора
Page Break	Разрыв страницы	Вставить разрыв страницы в месте нахождения курсора
Hyperlink	Гиперссылка	Вставить связь с другим документом или интернет-адресом
Reference	Закладка	Вставка ссылки в месте расположения курсора
Component	Компонент	Вставить связь документа с другими источниками или приложениями с использованием OLE (Excel, Matlab, Axum graphs, Filt Read/Write и др.)
Object	Объект	Вставка в документ нового OLE-объекта (документ Word, лист Excel, диаграмма Excel и др.)

Опция **Format** (Формат) содержит следующие команды (табл. 4.5):

Таблица 4.5

Format			
Equation		Выражение	Форматирование математических выражений
Result		Результат	Форматирование численных результатов
Text		Текст	Форматирование текста
Paragraph		Текстовая область	Изменение выравнивания и отступов выделенной текстовой области
Style		Стиль	Изменение стиля текстовой области
Properties		Свойства	Выводит диалоговое окно **Свойства**
Properties	**Highlight Region**	Подкраска области	Подкраска области, выражения, изменение цвета
	Disable Evaluation	Отключить выражение	Делает формульный блок неисполнимым
	Enable Optimization	Возможность прямой оптимизации	Дает возможность рассчитать выражение без первичного упрощения. Опция **Optumization** в меню **Math** сначала упрощает выражение через символьный процессор, а затем рассчитывает численный результат
Graph		Графика	Выводит диалоговое меню создання графиков
Color		Цвет	Выводит диалоговое окно изменения цвета
Separate Regions		Разделить области	Разделить перекрывающиеся области
Align Regions		Выровнять области 	Выровнять выделенные области
Area		Защита области	Закрыть/включить защиту области
Headers/Footers		Колонтитулы	Определить верхние/нижие колонтитулы в документе

Опция **Math** (Математика) содержит следующие команды (табл. 4.6):

Таблица 4.6

Math			
Calculate		Пересчитать	Провести расчеты по формулам, выдать результаты и обновить графики в пределах экрана
Calculate Worksheet		Пересчитать все	Провести расчеты по всем формулам и обновить его целиком
Automatic Calculation		Автоматические вычисления	Включить/выключить автоматический режим вычислений
Optimization		Оптимизация	Включить/выключить режим оптимизации численных расчетов
Options		Математические опции	
Options	**Built-In Variable**	Встроенные переменные	Установить значения встроенных (системных) переменных
	Unit System	Система единиц	Изменить систему единиц
	Dimensions	Единицы измерения	Изменить названия основных единиц измерения

Опция **Symbolics** (Символика) содержит следующие команды (табл. 4.7):

Таблица 4.7

Symbolics			
Evaluate		Вычислить	Преобразовать выражение с выбором вида преобразований из меню
Evaluette	**Symbolically**	Вычислить в символах	Вычислить символьное вычислеиие выражений
	Floating Point	С плавающей запятой	Символьное вычисление проводить по возможности с численным вычислением результата
	Complex	В комплексном виде	Выполнить преобразование выражений с комплексными числами
Simplify		Упростить	Упростить выделенное выражение с выполнением таких операций, как сокращение подобных слагаемых, приведение к общему знаменателю, использование основных тригонометрических тождеств и т. д.

Таблица 4.7 (продолжение)

<table>
<tr><th colspan="4">Symbolics</th></tr>
<tr><td colspan="2">Expand</td><td>Разложить по степеням</td><td>Раскрыть выражение</td></tr>
<tr><td colspan="2">Factor</td><td>Разложить на множители</td><td>Разложить выражение на множители</td></tr>
<tr><td colspan="2">Collect</td><td>Разложить по подвыражению</td><td>Собрать слагаемые, подобные выделенному выражению, которое может быть отдельной переменной или функцией со своим аргументом (результатом будет выражение, полиномиальное относительно выбранного выражения)</td></tr>
<tr><td colspan="2">Polynomial Coefficient</td><td>Полиномиальные коэффициенты</td><td>Найти коэффициенты полинома по заданиой переменной</td></tr>
<tr><td colspan="2">Variable</td><td>Переменная</td><td>Операции с выделенной переменной</td></tr>
<tr><td rowspan="6">Va-riadle</td><td>Solve</td><td>Решить относительно переменной</td><td>Найти значения выделенной переменной, при которых содержащее ее выражение становится равным нулю (решить уравнение или неравенство относительно выделенной переменной)</td></tr>
<tr><td>Substitute</td><td>Заменить переменную</td><td>Заменить указанную переменную содержимым буфера обмена</td></tr>
<tr><td>Differentiate</td><td>Дифференцировать по переменной</td><td>Дифференцировать все выражение, содержащее выделенную переменную, по отношению к этой переменной (остальные переменные рассматриваются как константы)</td></tr>
<tr><td>Integrate</td><td>Интегрировать по переменной</td><td>Интегрировать все выражение, содержащее выделенную переменную, по этой переменной</td></tr>
<tr><td>Expand to Series</td><td>Разложить в ряд</td><td>Найти несколько членов разложения выражения в ряд Тейлора относительно выделенной переменной</td></tr>
<tr><td>Convert to Partial Fraction</td><td>Разложить на элементарные дроби</td><td>Разложить на элементарные дроби выражение, которое рассматривается как рациональная дробь относительно выделенной переменной</td></tr>
<tr><td colspan="2">Matrix</td><td>Матричные операции</td><td>Операции с выделенными матрицами</td></tr>
</table>

Таблица 4.7 (окончание)

Symbolics			
Matrix	**Transpose**	Транспонировать	Получить транспонированную матрицу
	Invert	Обратить	Создать обратную матрицу
	Determinate	Определитель	Вычислить детерминант (определитель) матрицы
Transform		Преобразование	Операции преобразования
Evaluation Style		Расположение результата	Выбрать формат отображения результата символьных преобразований (наличие комментариев и вертикальное либо горизонтальное размещение по отношению к преобразуемому выражению)

Опция **Window** (Окно) содержит следующие команды (табл. 4.8):

Таблица 4.8

Window		
Cascade	Каскад	Расположить окна документов друг под другом так, чтобы были видны заголовки
Tile Horizontal	Окна по горизонтали	Расположить окна документов горизонтально, чтобы они не перекрывались
Tile Vertical	Окна по вертикали	Расположить окна документов вертикально, чтобы они не перекрывались

Опция **Help** (Помощь) содержит следующие команды (табл. 4.9):

Таблица 4.9

Help		
Mathcad Help	Справка	Открывает окно, содержащее все темы справки
Resourse Center	Электронная библиотека 	Открывает мультимедийную электронную книгу, содержащую информацию по использованию Mathcad, научным и техническим темам

3 Зак. 1030

Таблица 4.9 (окончание)

Help			
Resourse Center	**Overview Tutorials**	Введение	Открывает введение в вычислительные особенности Mathcad
		Самоучитель	Открывает обучающую систему использования Mathcad
	QuiksSheets Reference Tables	Быстрые примеры	Открывает собрание шаблонных образцов применения Mathcad
		Справочные таблицы	Открывает книгу с математической и научно-технической информацией
	Extending Mathcad	Расширение возможностей Mathcad	Примеры использования Mathcad с другими пакетами
Tip of the Day		Совет	Отображает сведения из справочной системы и советы по выполнению различных действий
Open Book		Открыть книгу	Открыть электронную книгу
About Mathcad		О программе	

4.3. Начальные сведения о работе с системой

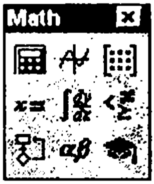
В простом случае работа с системой Mathcad сводится к подготовке в окне редактирования вычислительных блоков, текстовых комментариев и графиков.

Формульный редактор

Система Mathcad интегрирует в себе три редактора: формульный, текстовый и графический. Для запуска формульного редактора достаточно установить курсор мыши в любом свободном месте окна редактирования и щелкнуть левой кнопкой. Появится визир в виде маленького красного крестика. Его можно перемещать клавишами перемещения курсора. Визир указывает место, с которого можно начинать набор формул – вычислительных блоков.

Наборные панели

Подготовка вычислительных блоков облегчается благодаря выводу шаблона при задании того или иного оператора. Для этого служат наборные панели с набором шаблонов различных математических символов (знаков арифметических операций, матриц, знаков интегралов, производных и т. д.). Наборные панели появляются в окне редактирования документов при активизации соответствующих пиктограмм панели Math (**View – Toolbars**) – рис. 4.1 или прямым заданием по выводу набора шаблонов из **View – Toolbars**. Все они перечислены при рассмотрении меню **View – Toolbars**.



Calculator toolbar
Арифметические операторы

Graph toolbar
Операторы графики

Evaluation toolbar
Операторы вычислений и формул

Calculus toolbar
Операторы объектов высшей математики

Programming toolbar
Операторы программирования

Greek Symbol toolbar
Греческие буквы

Vector and Matrix toolbar
Матричные операторы

Boolean toolbar
Операторы отношения

Symbolic Keyword toolbar
Операторы символьных вычислений

Рис. 4.1. Пиктограммы панели Math

Применение пиктограмм для вывода математических знаков очень удобно, поскольку не надо запоминать разнообразные сочетания клавиш, используемые для ввода специальных математических символов.

Любую наборную панель можно переместить в удобное место экрана, подцепив ее верхнюю часть курсором мыши. Перемещая панель, левую кнопку мыши нужно держать нажатой. В верхнем правом углу каждой наборной панели помещена кнопка, служащая для устранения панели с экрана, как только она становится ненужной.

Шаблоны

Большинство кнопок на наборных панелях вводят общепринятые и специальные математические знаки и операторы, вставляя их шаблоны в месте расположения курсора. Для ввода данных можно указать курсором мыши на нужный шаблон данных и щелчком левой кнопки ввести их. В составе сложных шаблонов часто встречаются шаблоны для ввода данных. Они имеют вид небольших черных квадратиков. В шаблоне производной их два: для указания дифференцируемой функции и имени переменной. В шаблоне неопределенного интеграла имеются два шаблона для ввода данных: задания подынтегральной функции и имени переменной, по которой идет интегрирование. В шаблоне определенного интеграла добавляются еще два: для ввода верхнего и нижнего предела интегрирования. Данные шаблоны имеют следующий вид:

$$\frac{d}{d\blacksquare}\blacksquare \qquad \int \blacksquare\, d\blacksquare \qquad \int_{\blacksquare}^{\blacksquare} \blacksquare\, d\blacksquare$$

После заполнения они будут иметь вид:

$$\frac{d}{dx}y \qquad . \int x\,dx \qquad \int_{2}^{3} e^{x}\,dx$$

Ввод объектов

Ввод объектов (математических выражений, тестовых комментариев, графиков) в текущее окно редактирования производится по-разному. Для ввода текстовых комментариев в месте расположения курсора вызывается вставка текстовой области. Для ввода графиков в месте расположения курсора вызывается вставка графика и производится дальнейшее его форматирование. Рассмотрим ввод математических выражений на простом примере. Допустим, необходимо задать значение переменной $x = 2$ и вычислить значение функции $\ln(x)$. Для этого можно воспользоваться вводом с клавиатуры (табл. 4.10):

Таблица 4.10

Нажимаемая клавиша	Изображение в окне	Комментарий
<x>	x	Ввод имени переменной
:	x:=	Ввод символа присвоения
<2>	x:=2	Ввод числовой константы 2
<Enter>		Фиксация ввода, скачок курсора
<ln>	ln	Ввод имени функции ln
<(x)>	ln(x)	Ввод открывающейся скобки, имени переменной, закрывающейся скобки
<=>	ln(x)=	Ввод знака вывода =
<Enter>	ln(x)=0.693	Вычисление ln(x)=ln(2)

В системе Mathcad знак используется как знак вывода результатов вычислений, а для присваивания переменным значений используется знак :=.

Все спецсимволы можно вводить из наборных панелей. Достаточно установить графический курсор в нужном месте окна и выбрать мышью пиктограмму нужного спецсимвола из выбранной палитры символов. Нажатие левой кнопки мыши вызовет вывод на место курсора шаблона спецзнака. Далее надо заполнить выделенные темными прямоугольниками места шаблона, и соответствующий объект будет полностью введен.

Для заполнения шаблонов операндов сначала надо отметить нужный шаблон с помощью клавиш перемещения курсора. Удобнее сделать это с помощью мыши, указав перемещаемым маркером мыши на нужный шаблон и нажав ее левую кнопку. Активный шаблон помечается синим уголком. Теперь можно вводить операнд. Аналогичные действия производятся и для других шаблонов ввода операндов.

Выделение и редактирование объектов

При редактировании математических выражений существует возможность выделения их целиком или в виде отдельных фрагментов. Выделения в выражениях отмечаются синим уголком. Расширение выделения можно произвести мышью или клавишей <Space>.

Выделение фрагментов математических выражений необходимо для их редактирования, копирования, проведения операций, а также изменения шрифтов, которыми набирается выражение. Например, для изменения шрифта в математических формулах достаточно выделить одну букву, установив выделение в виде синего уголка сразу после буквы. Затем можно воспользоваться средствами модификации шрифтов.

Изменение параметров и типов шрифтов для определенных объектов (например, переменных или констант) действует для них глобально. Так, если увеличить размеры обозначения какой-либо переменной, то все обозначения других переменных тоже будут увеличены. Для локального изменения размеров надо выделить соответствующий блок документа и только затем изменить шрифт.

Редактирование документов возможно как с помощью мыши, так и с клавиатуры. В любом случае полезно знать функции графического курсора-маркера. При перемещении по документу он может принимать одну из трех приведенных ниже форм.

- Крестообразный курсор (визир) служит для указания места для новых блоков – текстовых, формульных или графических. Курсор имеет такой вид только вне пространства блоков, т. е. на пустом месте экрана, и может перемещаться клавишами управления курсором или устанавливаться мышью.
- Курсор в виде синего уголка (маркер ввода) служит для указания на отдельные элементы блоков и обычно используется для ввода данных и заполнения шаблонов. В текстовых блоках он используется для указания места вставки или удаления отдельных символов.
- Курсор в виде длинного синего уголка предназначен для выделения отдельных частей выражения или выражения целиком.

Математические выражения не столько набираются, сколько конструируются. При этом учитывается определенная структура выражений и иерархия выполняемых операций.

Конструирование выражений облегчается средствами выделения отдельных фрагментов выражений. Полезно помнить, что все, что попадает в ходе выделения в рамку, оказывается операндом для следующего вводимого оператора. Операторы возведения в степень, извлечения корня и деления являются "цепкими" операторами. После их ввода все, что набирается следом, становится показателем степени, подкоренным выражением или знаменателем. Для отмены этого обычно приходится использовать клавишу <Space>.

Удобно сочетать отметку начала выделения с использованием мыши, поскольку она позволяет сразу и точно указать место выделения, с последующим расширением или сужением места выделения с помощью управляющих клавиш. При определенном навыке все выделения можно выполнить, используя мышь.

Для замены числа или буквы в выражении можно отметить их курсором мыши и щелкнуть левой кнопкой. Появится маркер ввода. Если он находится справа от стираемого символа, для стирания нажмите клавишу ←, а если слева – клавишу <Del>.

Нажатием клавиши <F3> убирается выделяемый фрагмент и помещается в буфер обмена. Нажатие клавиши <F4> помещает фрагмент из буфера в место, на которое указывает текущее положение курсора.

Нажатие клавиш <Shift> + <F1> вызывает контекстную справку. При выполнении этой операции у курсора-стрелки появляется знак вопроса. Если теперь указать на какой-либо объект или его часть (например, на имя функции или оператор), будет выведено окно с сообщением о назначении и правилах использования указанного объекта. Для отмены этого режима помощи достаточно нажать клавишу <Esc>.

Работа с блоками документа

Документ состоит из отдельных блоков. Они могут быть различного типа: тексты (комментарии), формулы, графики, таблицы и т. д. Каждый блок занимает в текущем окне определенную область прямоугольной формы. Для конструирования блоков служат три встроенных в систему редактора: текстовый, формульный и графический. Блоки просматриваются системой, интерпретируются и исполняются. Выделить области блоков можно мышью, перемещая мышь по диагонали через нужные блоки, начиная с нерабочей части экрана. При выполнении этого действия появится прямоугольник из пунктирных черных линий. В результате все области блоков окажутся обведенными прямоугольниками из пунктирных линий. Эти прямоугольники являются обычно невидимыми границами областей объектов. Блоки можно выровнять по горизонтали и вертикали. Блоки можно скопировать и поставить на нужное место. Блоки можно переместить в требуемую позицию. Зацепив курсором мыши любой из выделенных блоков и нажав левую кнопку и удерживая ее, можно перетаскивать блоки с одного места на другое.

Расположение блоков имеет принципиально важное значение. Их выполнение происходит справа налево и сверху вниз. Поэтому блоки не должны взаимно перекрываться (небольшое перекрытие несущественно). Указанный порядок исполнения блоков означает, что, например, при построении графика функции или блицы ее значений сначала должны исполняться блоки, задающие саму функцию и пределы изменения аргумента, а затем блок, задающий вывод таблицы или построение графика функции.

4.4. Входной язык системы Mathcad

Алфавит, константы и переменные системы

Mathcad требует от пользователя корректного описания алгоритма решения математической задачи на входном языке, очень напоминающем общепринятый язык описания математических расчетов. Входной язык системы состоит из таких понятий и объектов, как идентификаторы, константы, переменные, массивы и другие типы данных, операторы и функции, управляющие структуры и т. д. Четкое представление об их возможностях и правилах применения (синтаксисе) необходимо при решении различных задач.

Алфавит входного языка системы определяет совокупность специальных знаков и слов, которые используются при задании команд, необходимых для решения задач. Алфавит системы Mathcad содержит: малые и большие латинские буквы; малые и большие греческие буквы; арабские цифры от 0 до 9; системные переменные; операторы; имена встроенных функций; спецзнаки; малые и большие буквы кириллицы (при работе в среде русифицированной системы Windows).

К укрупненным элементам языка относятся типы данных, операторы, функции и управляющие структуры.

К типам данных относятся числовые константы, обычные и системные переменные, массивы (векторы и матрицы) и данные файлового типа.

Операторы – это элементы языка, с помощью которых можно создавать математические выражения. К ним относятся символы арифметических операций, знаки вычислений сумм, произведений, производной и интеграла и т. д. После указания операндов (параметров операндов) операторы становятся исполняемыми по программе блоками.

При указании имени функции с указанием аргумента она возвращает некоторое значение – символьное, числовое, вектор или матрицу. Наряду со встроенными функциями могут задаваться и функции пользователя, описывающие произвольные функции, отсутствующие в наборе встроенных в Mathcad функций.

Функции могут (наряду с операторами) входить в математическое выражение.

Числовые константы задаются с помощью арабских цифр и *десятичной точки* (а не запятой). Например, десять целых пять десятых будет выглядеть так: 10.5. Порядок числа вводится умножением мантиссы на десять в степени, определяющей порядок. Например, $10.5 \cdot 10^{-3}$ – десятичная константа с мантиссой 10.5 и порядком –3.

Константами называют поименованные объекты, хранящие некоторые значения, которые не могут быть изменены. В качестве имени числовых констант используются их числовые значения.

Переменные являются поименованными объектами, имеющими некоторое значение, которое может изменяться по ходу выполнения программы. Имена констант, переменных и иных объектов называют идентификаторами.

Идентификаторы в системе Mathcad могут иметь практически любую длину, и в них могут входить любые латинские и греческие буквы, а также цифры. Однако начинаться идентификатор может только с буквы, например: у, Y2, α, β1. Кроме того, идентификатор должен быть слитным. Некоторые спецсимволы (например, знак объединения _) могут входить в состав идентификаторов, другие (например, знаки операторов арифметических действий) недопустимы. Нельзя использовать для идентификаторов буквы русского языка. Малые и большие буквы в идентификаторах различаются.

В Mathcad есть группа объектов, которые считаются системными переменными, имеющими предопределенные системой начальные значения. Это следующие объекты: π – число «пи» – (3.14); е – основание натурального логарифма (2.71); ∞ – системная бесконечность (10^307); % – процент (0.01); **TOL** – погрешность численных методов (0.001); **ORIGIN** – нижняя граница индексации массивов (0); **PRNCOLWIDTH** – число столбцов оператора WRITEPRN (8); **PRNPRECISSION** – число десятичных знаков, используемых оператором WRITEPRN (4); **FRAME** – переменная счетчика кадров при работе с анимационными рисунками (0).

Пять последних переменных вводятся набором их имен. Значения системных переменных могут быть в дальнейшем изменены, как и значения обычных переменных, присваиванием им новых значений.

При выполнении символьных операций переменные π и е используются только в символьном виде.

:= оператор присваивания

Обычные переменные отличаются от системных тем, что они должны быть предварительно определены пользователем, т. е. им необходимо хотя бы однажды присвоить значение. В качестве оператора присваивания используется знак :=, тогда как знак = отведен для вывода значения переменной. Попытка использовать неопределенную переменную ведет к выводу сообщения об ошибке. Существует также знак равенства, выделенный жирным шрифтом, который используется либо как признак неравенства в операциях сравнения, либо как оператор приближенного равенства.

= оператор вывода значений

Если переменной присваивается начальное значение с помощью оператора :=, такое присваивание называется локальным. До этого присваивания переменная не определена, и ее нельзя использовать. Однако с помощью знака ≡ можно обеспечить глобальное присваивание, т. е. оно может производиться в любом месте документа. К примеру, если переменной присвоено таким образом значение в самом конце документа, то она будет иметь это же значение и в начале документа.

Переменные могут использоваться в математических выражениях, быть аргументами функций или операндом операторов. Переменные могут быть индексированными (элементы векторов и матриц), а также иметь заданные пределы изменения (ранжированные переменные).

Ранжированные переменные имеют ряд фиксированных значений, либо целочисленных, либо в виде чисел, с определенным шагом от начального значения до конечного.

Ранжированные переменные характеризуются именем и индексом каждого своего элемента. Для создания ранжированной переменной целочисленного типа используется выражение:

Name:=*Nbegin*..*Nend*, пример: i:=1..10,

где Name – имя переменной, *Nbegin* - ее начальное значение, *Nend* - конечное значение, .. – символ, указывающий на изменение переменной в заданных пределах (он вводится клавишей с изображением точки с запятой (;) или с панели арифметических операторов, после чего появляется шаблон ▪..▪. Если *Nbegin* < *Nend*, то шаг изменения переменной будет равен +1, в противном случае –1.

Для создания ранжированной переменной общего вида используется выражение:

Name:=*Nbegin*,(*Nbegin* + *Step*)..*Nend*.

Здесь *Step* – заданный шаг изменения переменной.

Ранжированные переменные широко применяются для представления численных значений функций в виде таблиц, а также для построения их графиков. Любое выражение с ранжированными переменными после знака равенства инициирует таблицу вывода. Несколько таких таблиц показано на рис. 4.2.

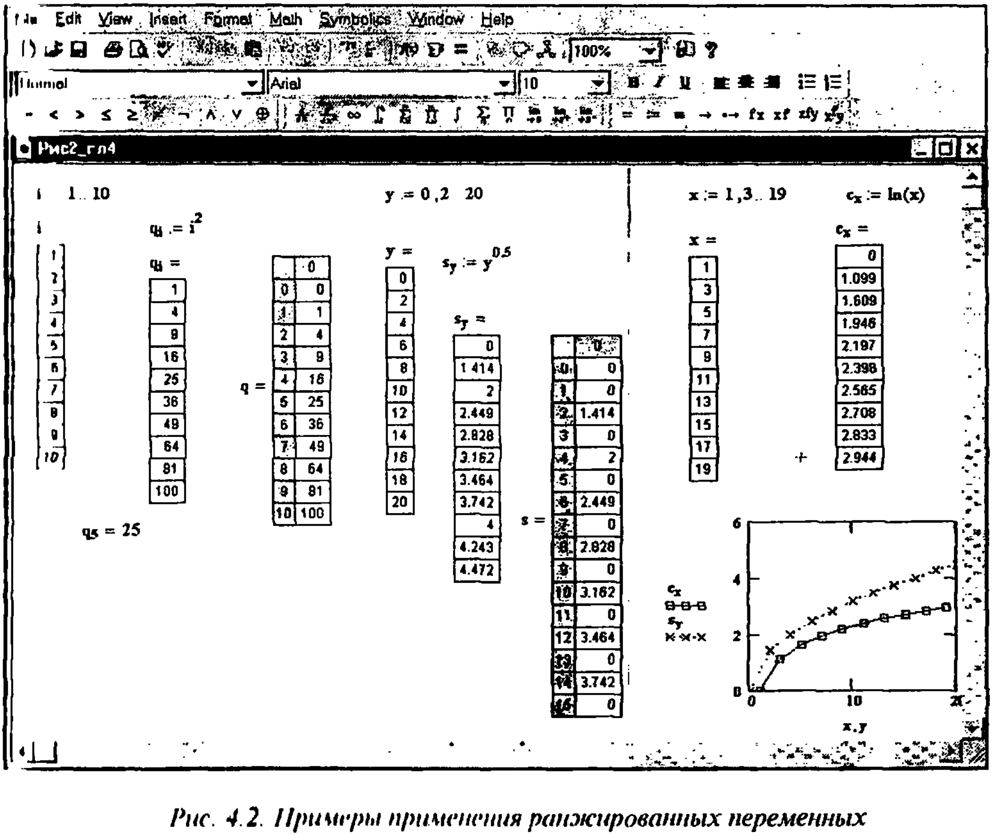


Рис. 4.2. Примеры применения ранжированных переменных

Таблицы вывода имеют следующие свойства:

- число строк в них не может быть больше 50;
- числа в таблицах можно задавать в требуемом формате с помощью операций задания формата чисел;
- при использовании в таблице единиц размерности все данные таблицы будут содержать единицы размерности;
- есть три способа показать значения векторов:
  - $X_i=$ выводится обычная таблица вывода;
  - X= выводится вектор, если число элементов вектора меньше 10;
  - X = выводится таблица вывода со слайдером, если число элементов вектора больше 10.

Задание ранжированных переменных эквивалентно заданию конечных циклов. Сами ранжированные переменные являются векторами. Выдача их значений производится в форме столбца со всеми значениями переменных. Индексированные переменные, образующиеся в результате задания ранжированных переменных, могут применяться в последующих формульных блоках. Однако в этих блоках необходимо соблюдать соответствие результатов (конечных и промежуточных) векторному типу этих переменных. С ранжированными переменными нельзя выполнять действия, корректные лишь для обычных (скалярных) переменных. Допустимы только действия с векторами.

Ранжированные переменные широко применяются при построении графиков. Например, для построения графика некоторой функции $f(x)$ прежде всего надо создать ряд значений переменной x – для этого она должна быть ранжированной переменной.

Скалярные и логические операторы

Скалярные операторы предназначены для выполнения арифметических действий над скалярными величинами.

Оператор локального присваивания выводится клавишей <:>, оператор глобального присваивания – клавишей <~>, ввод пары круглых скобок с шаблоном – клавишей с одиночной кавычкой <'>, ввод нижнего индекса клавишей < [>, ввод верхнего индекса клавишами <Ctrl>+<6>.

Система Mathcad содержит ряд расширенных скалярных операторов:

- $\sum_{\blacksquare} \blacksquare$ вычисление суммы ряда, вызываемой клавишей <$>,
- $\prod_{\blacksquare} \blacksquare$ произведения ряда, вызываемой клавишей <#>,
- $\frac{d}{d\blacksquare}\blacksquare$ вычисления производной, вызываемой клавишей <?>,
- $\int_{\blacksquare}^{\blacksquare} \blacksquare \, d\blacksquare$ определенного интеграла, выводимого клавишей <&>.

Система Mathcad содержит следующие скалярные операторы:

$X:=$ **Локальное присваивание X значения y**

$X \equiv y$ **Глобальное присваивание X значения y**

$X=$ **Вывод значения X**

$-X$ **Смена знака X**

$+, -, *, /,$ **Арифметические операторы**

X^n **Возведение в степень**

$X!$ **Вычисление факториала**

$|Z|$ **Вычисление модуля комплексного Z**

$\overline{Z}$ **Вычисление комплексно сопряженного числа**

(■) **Ввод пары круглых скобок с шаблоном**

( ) **Ввод скобок**

X_n, $X^{<n>}$ **Ввод нижнего и верхнего индексов**

Разумеется, для вызова шаблона можно использовать и соответствующую панель набора математических спецсимволов.

Ряд операторов предназначен для сравнения двух величин. Они называются операторами отношения или логическими операторами: >, <, ≥, ≤, ≠, =. Они выводятся с панели отношений – .

Не следует путать оператор сравнения (знак равенства) с похожим знаком вывода значений переменных. Знак равенства, как оператор отношения, имеет больший размер и более жирное начертание (**=**), чем обычный знак равенства (=) – оператор вывода.

Выражения с логическими операторами возвращают логическое значение, соответствующее выполнению или невыполнению условия, заданного оператором. Это значение является логической единицей <1>, если условие выполнено, и логическим нулем <0>, если оно не выполнено. Математически значения логической единицы и нуля совпадают со значениями числовых констант 1 и 0. Например:

$4 > 2 = 1$ – условие выполнено, результат 1;

$2 > 4 = 0$ – условие не выполнено, результат 0.

Указанное свойство логических операторов позволяет строить не совсем обычные выражения, содержащие в себе логические операторы, например:

$$3*(4 > 2) = 3.$$

Выражение (4 > 2) возвращает единицу, поэтому результат вычисления такого выражения даст число 3. Логические операторы часто используются совместно с условными функциями.

Функции

Система Mathcad содержит расширенный набор встроенных элементарных функций. Встроенные функции вызываются с помощью меню **Insert Function** (рис. 4.3).

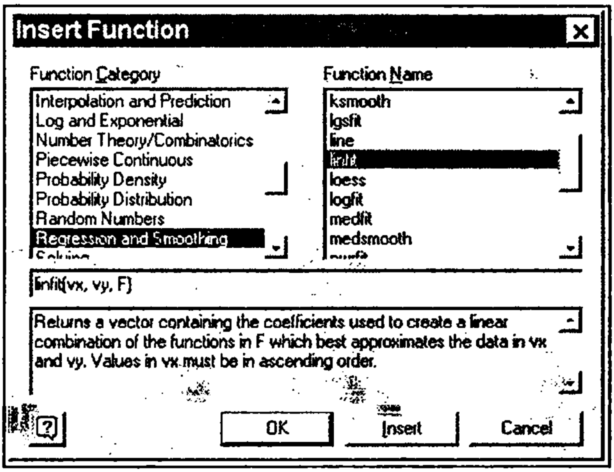


Рис. 4.3. Диалоговое окно ***Insert Function***

В частности, в окне Insert Function (Вставить функцию) встроенные функции разбиты на категории (табл. 4.11):

Таблица 4.11

Bessel	функции Бесселя
Complex Numbers	функции комплексных чисел
Curve Fitting	**подбор эмпирического уравнения кривой**
Differential Equation Solving	**функции решения дифференциальных уравнений и систем, задача Коши, краевая задача, уравнения в частных производных**
Expression Type	функции типа выражения
File Access	функции работы с файлами
Finans	**финансовые функции**
Fourier Transform	функции преобразований Фурье
Hyperbolic	гиперболические функции
Image Processing	функции обработки образов
Interpolation and Prediction	функции интерполяции и экстраполяции
Log and Exponential	**логарифмические и экспоненциальные функции**
Numbers Theory/ Combinatorics	функции теории чисел и комбинаторики
Numerical Recipes	функции, реализующие численные методы
Piecewise Continuous	**функции создания условных выражений**
Probably Density	**функции плотности вероятности**
Probably Distribution	**функции распределения вероятности**
Random Numbers	функции случайных чисел
Regression and Smoothing	**функции регрессии и сглаживания**
Solving	**функции решения алгебраических уравнений и систем, а также оптимизации**
Sorting	функции сортировки
Special	специальные функции
Statistics	**статистические функции**
String	текстовые функции
Trigonometric	тригонометрические функции
Truncation and Round-Off	функции округления и работы с частью числа

Таблица 4.11 (окончание)

User defined	определяемые пользователем (произведение Кронекера)
Vector and Matrix	**функции работы с векторами и матрицами**
Wavelet Transform	функции волнового преобразования

Это существенно облегчает поиск нужной функции из почти трехсот встроенных.

Функции задаются своим именем и значением аргумента в круглых скобках. В ответ на обращение к ним функции возвращают вычисленные значения. Рис. 4.4 иллюстрирует работу с элементарными функциями.

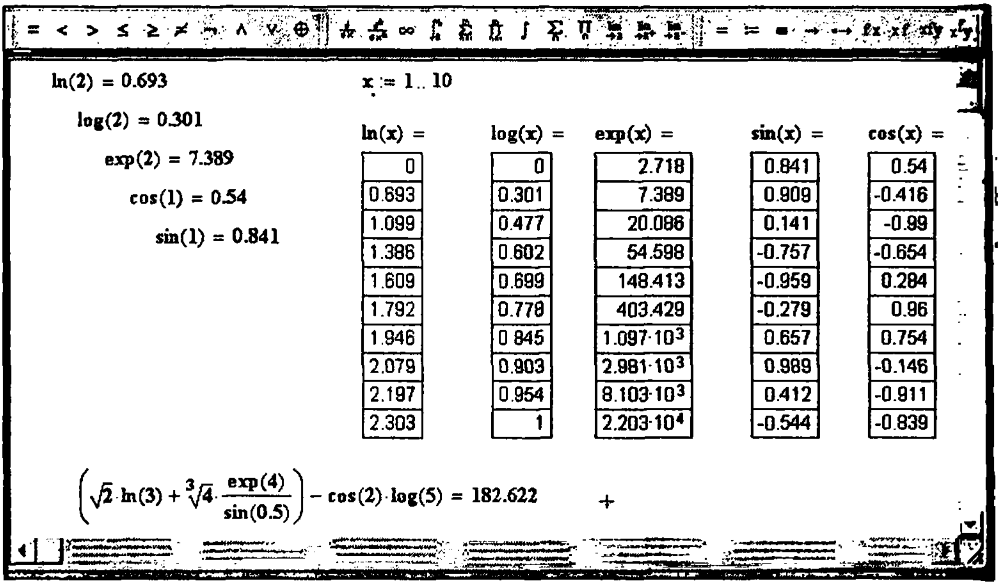


Рис. 4.4. Примеры вычислений с элементарными функциями

Большинство элементарных функций можно набрать с клавиатуры малыми латинскими символами. Расширенный набор функций создает заметные удобства при проведении многих вычислений.

Наряду с элементарными функциями, в системе Mathcad содержится ряд встроенных специальных математических функций. К важнейшим из них принадлежит функции Бесселя, являющиеся решениями дифференциальных уравнений второго порядка.

Существует ряд встроенных функций, у которых возвращаемый результат зависит от знака или значения аргумента. Так, при их вычислении производится сравнение аргумента с некоторыми числовыми константами, например с нулем или целыми числами. Широкие возможности дает функция *if* для создания условных выражений:

if(условие, Выражение 1, Выражение 2)

Если в этой функции условие выполняется, то будет вычисляться Выражение 1, в противном случае – Выражение 2.

Несмотря на довольно широкий набор встроенных функций, всегда возникает необходимость расширить систему новыми функциями, представляющими интерес для пользователя. Функции пользователя вводятся с применением следующего выражения:

Имя функции (Список параметров) := Выражение

Имя функции задается как любой идентификатор, например, имя переменной. В скобках указывается список параметров функции – перечень используемых в выражении переменных, разделенных запятыми. Выражение – это любое выражение, содержащее доступные системе операторы и функции с операндами и аргументами, указанными в списке параметров.

Примеры задания функций двух переменных:

- $funk(x,y):=x^2+y^2$
- $a_0:=1.2$ $a_1:=0.25$ $a_2:=0.75$

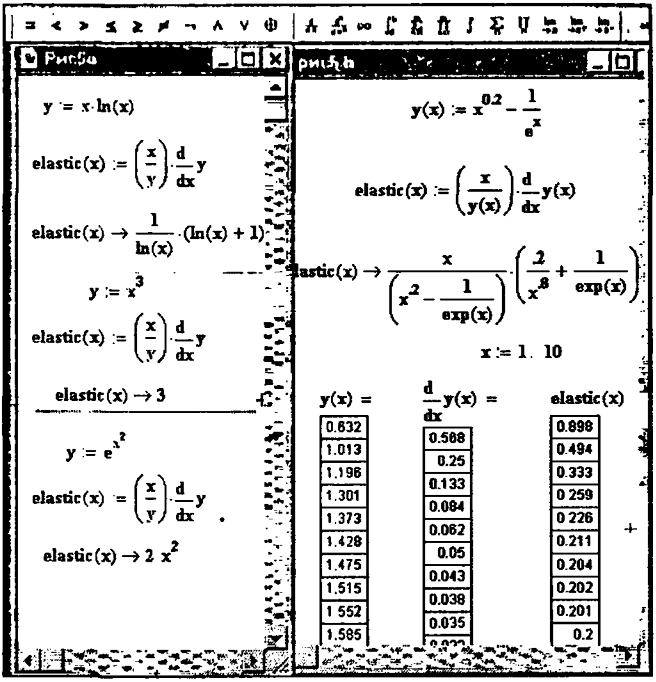
  $CobbDouglus(K,L):=a_0 \cdot K^{a1} \cdot L^{a2}$

Следует отметить особый статус переменных, указанных в списке параметров функций пользователя. Эти переменные являются локальными, поэтому они могут не определяться до задания функций; фактически их указание в списке параметров и является заданием этих переменных. Естественно, что локальные переменные могут использоваться только в выражении, описывающем функцию. Их имена могут совпадать с именами глобальных переменных, введенных ранее. Но при этом после выхода из блока задания функции значения этих переменных будут сохранять ранее заданные (для глобальных переменных) значения.

Функции пользователя можно вводить для задания новых специальных математических функций.

Заданные с применением знака := функции являются заданными локально. Поэтому они должны быть заданы в документе до того, как будут использованы. С помощью знака ≡ можно задать функции пользователя как глобальные. Тогда они могут задаваться в любом месте документа, в том числе в конце. Если же функция нигде и никак не задана, то применять ее, естественно, нельзя. Попытки такого применения будут сопровождаться сообщениями об ошибке – имя функции окажется выделенным черным фоном. Ряд примеров задания и применения функций пользователя приведен на рис. 4.5.

В этих примерах в качестве функции пользователя задана эластичность функции, обозначенная **elastic(x)**. Для наглядности примеры приведены в различных окнах. В левом окне один раз заданная функция эластичности и ее символьное преобразование применены к трем функциям путем копирования функции эластичности и символьного преобразования. В центральном окне приведен пример, как последовательный ввод, определение функции *y* и размещение ее блока ниже первоначально заданной приводит к соответствующему результату. Численное значение функции эластичности получаем, определяя *x*. В правом окне приведен пример задания функции пользователя, ее символьного преобразования, символьного упрощения и построения графика.



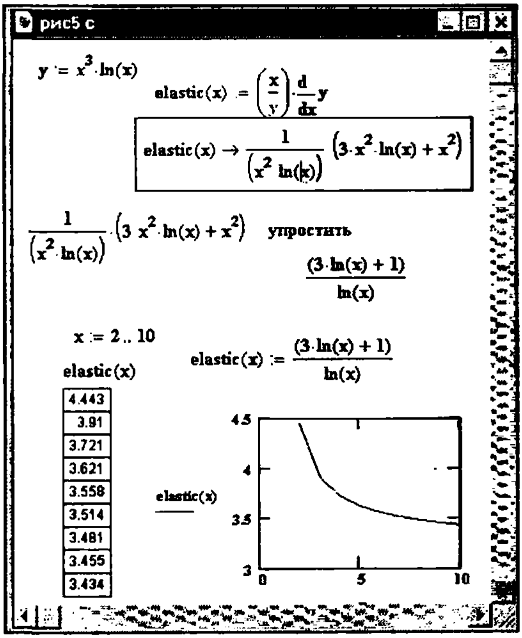


Рис. 4.5. Примеры задания и применения функций пользователя

Ряд функций – операторы и функции для работы с векторами и матрицами, функции статистики и линейной регрессии, функции для решения уравнений, символьное дифференцирование и интегрирование будут рассмотрены дальше, после изложения соответствующих теоретических вопросов.

Сообщения об ошибках

При вводе, обработке или вычислениях возможны ошибки, которые Mathcad обнаруживает и выводит диагностические сообщения. Система выводит сообщения об ошибках в красных прямоугольниках с линией, выходящей от места ошибки к прямоугольнику с сообщением о ней.

Если Mathcad находит ошибку при попытке вычисления функции, определенной пользователем, он помечает сообщением об ошибке имя функции, а не ее определение. В этом случае проверьте определение функции, чтобы понять, что вызвало ошибку.

Глава 5. Методы численного решения уравнений

Пусть $f(x) = 0$ – некоторое уравнение. Число $x_* = \eta$ называется корнем или решением данного уравнения, если подстановка его в уравнение обращает его в равенство, т. е. $f(\eta) = 0$.

Число η называют нулем функции $y = f(x)$. Если функция $y = f(x)$ непрерывна и принимает на концах отрезка $[a, b]$ значения разных знаков, т. е. $f(a)f(b) < 0$, то внутри этого промежутка найдется нуль функции.

5.1. Метод деления отрезка пополам

Пусть дано уравнение $f(x) = 0$, причем функция $f(x)$ непрерывна на отрезке $[a, b]$ и $f(a)f(b) < 0$.

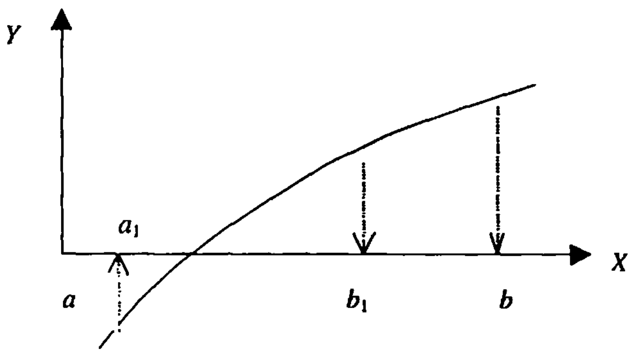


Рис. 5.1

Для вычисления корня уравнения, принадлежащего отрезку $[a, b]$ найдем середину этого отрезка:

$$x_1 = (a + b)/2.$$

Если $f(x_1) \neq 0$, то для продолжения вычислений выберем ту из частей данного отрезка $[a, x_1]$ или $[x_1, b]$, на концах которой функция $f(x)$ имеет противоположные знаки. Концы нового отрезка обозначим через a_1 и b_1 (рис. 5.1).

Новый суженый промежуток $[a_1, b_1]$ снова делим пополам и повторяем вычисления. Описанный процесс деления отрезка пополам можно продолжать до тех пор, пока не получится отрезок $[a_n, b_n]$ длины $b_n - a_n < \varepsilon$, где ε – заданная точность.

Пример. Найти решение уравнения $f(x) = 10 - x^2 = 0$ с точностью $\varepsilon = 0{,}1$.

Принимаем первоначальные значения: $a = 3{,}00$ и $b = 3{,}50$. Значения функции в этих точках: $f(a) = 1$, $f(b) = -2{,}25$.

$$x_1 = (a + b)/2 = (3 + 3{,}5)/2 = 3{,}25; f(x_1) = -0{,}5625.$$

Значения нового отрезка выбираем с помощью условия $f(a)f(b) < 0$. Устанавливаем:

$$f(a)f(x_1) < 0; f(x_1)f(b) > 0.$$

Следовательно, новый отрезок $[a_1, b_1] = [a, x_1] = [3{,}0, 3{,}25]$.

$$x_2 = (3 + 3{,}25)/2 = 3{,}125; f(x_2) = 0{,}234375$$

$$f(a)f(x_2) > 0; f(x_2)f(b) < 0.$$

Следовательно, $[a_2, b_2] = [x_2, b] = [3{,}125; 3{,}5]$.

$$x_3 = (3{,}125 + 3{,}5)/2 = 3{,}3125\ f(x_3) = -0{,}9726.$$

$$f(x_2)f(x_3) < 0, f(x_3)f(b) > 0.$$

Следовательно, $[a_3, b_3] = [x_2, x_3] = [3{,}125; 3{,}3125]$.

$$x_4 = (3{,}125 + 3{,}3125)/2 = 3{,}21875; f(x_4) = -0{,}36.$$

$$f(x_2)f(x_4) < 0; f(x_4)f(x_3) > 0.$$

Следовательно, $[a_4, b_4] = [x_2, x_4] = [3{,}125; 3{,}21875]$.

Заданная точность достигнута: $3{,}21875 - 3{,}125 = 0{,}094 < \varepsilon = 0{,}1$

Решение уравнения находится в середине данного отрезка:

$$x_5 = (3{,}125 + 3{,}21875)/2 = 3{,}1719; f(x_5) = -0{,}06.$$

5.2. Метод секущих

Метод секущих (метод хорд). Рассмотрим уравнение $f(x) = 0$, где $y = f(x)$ – непрерывная функция. Сущность метода заключается в том, что предполагается интервал [a, b], в котором находится корень уравнения, принимаются два близких значения и методом итерации подбираются более точные значения.

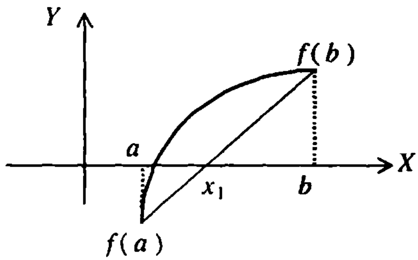


Рис. 5.2

В некотором интервале $[a, b]$, в котором находится корень уравнения, устанавливаем значения функции, т. е. $f(a)$ и $f(b)$, причем они имеют противоположные знаки (рис. 5.2). Соединяем точки прямой. Уравнение прямой имеет вид:

$$y - f(a) = \frac{f(b) - f(a)}{(b - a)}(x - a).$$

В качестве приближенного значения уравнения $f(x) = 0$ принимается точка x_1, в которой прямая пересекается с $0x$. Подставляя в уравнение $x = x_1$ и $y = 0$, получим:

$$x_1 = a - f(a)\frac{b - a}{f(b) - f(a)}.$$

Продолжая такой процесс, можно получить решение любой точности.

Пример. Решить приближенно уравнение $y = f(x) = 15\ln x + 0,5x^2 - 5x$;

$$15\ln x + 0,5x^2 - 5x = 0.$$

Корни уравнения есть абсциссы точек пересечения кривой с осью $0x$. Подберем возможно меньший отрезок, у которого значения функции имеют разные знаки:

$$f(1) = -4,5 < 0;$$
$$f(4) = 8,7944 > 0.$$

Попробуем уменьшить интервал:

$$f(2) = 2,3973 > 0.$$

Следовательно, искомый корень находится в интервале $[a, b] = [1, 2]$. Применяя метод секущих, рассчитываем по формуле:

$$x_1 = 1 - (-4,5)(2 - 1)/(2,3973 - 8,7944) = 1,6524.$$

Подставив $x_1 = 1,6524$ в исходную формулу, получим: $f(1,6524) = 0,6386$. Проверяем $f(a)f(x_1) < 0$. Следовательно, искомый корень находится в интервале $[a, x_1] = [1, 1,6524]$. Применяя далее метод секущих к этому интервалу, находим:

$$x_2 = 1 - (-4,5)(1,6524 - 1)/(0,6386 - 9,7944) = 1,5715.$$

Для $x_2 = 1,5715$ получаем $f(1,5715) = 0,1581$. Результаты дальнейших расчетов приведены в табл. 5.1.

Таблица 5.1

n	a	$f(a)$	b	$f(b)$	x_n
1	1	–4,50	2	2,397207708	1,652438
2	1	–4,50	1,652438	0,636861934	1,571549
3	1	–4,50	1,571549	0,158067387	1,552154
4	1	–4,50	1,552154	0,038478468	1,547473
5	1	–4,50	1,547473	0,009321337	1,546341
6	1	–4,50	1,546341	0,002255394	1,546068
7	1	–4,50	1,546068	0,000545559	1,546002
8	1	–4,50	1,546002	0,000131956	1,545986
9	1	–4,50	1,545986	0,000031916	1,545982

Необходимую точность расчетов устанавливаем по отличию от нуля значения функции в точке решения уравнения.

5.3. Метод Ньютона

Метод Ньютона (метод касательных). Данный метод применяется в том случае, когда функция $y = f(x)$ имеет в интервале (a, b) непрерывные первую и вторую производные, причем в этом интервале производные сохраняют свои знаки.

Знак второй производной определяет направление выпуклости кривой, поэтому в зависимости от сочетания знаков первой и второй производных, $f'(x)$ и $f''(x)$, возможны четыре случая (рис. 5.3).

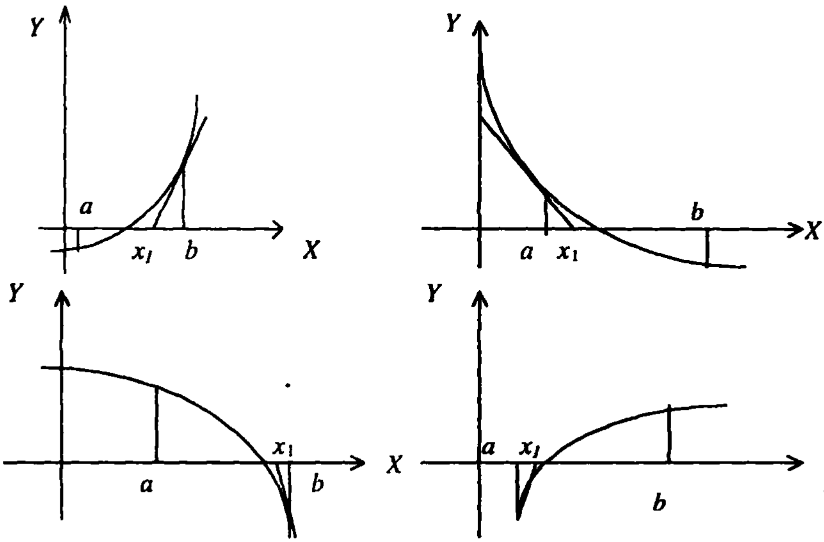


Рис. 5.3

Определим конец x_0 интервала (a, b), в котором функция $f(x)$ и производная $f'(x)$ имеют одинаковый знак. В качестве приближенного решения уравнения принимается точка пересечения оси абсцисс с касательной, проведенной к кривой $y = f(x)$ в точке с координатами $(x_0; f(x_0))$.

Уравнение этой касательной имеет вид:

$$y - f(x_0) = f'(x_0)(x - x_0).$$

Подставив значения $x = x_1$ и $y = 0$, получаем приближенное решение:

$$x_1 = x_0 - [f(x_0) / f'(x_0)].$$

Взяв далее точку x_1, получим аналогично:

$$x_2 = x_1 - [f(x_1) / f'(x_1)]$$

и более точное решение.

В общем случае расчетное выражение равно:

$$x_n = x_{n-1} - [f(x_{n-1}) / f'(x_{n-1})].$$

Пример. Вернемся к условию предыдущего примера. Там было установлено, что $f(2) = 2{,}3973$, поскольку $f''(2) = 4{,}75$, и, следовательно, $x = 2$ можно принять в качестве исходной точки x_0.

Получаем:

$$15\ln x + 0,5x^2 - 5x = 0; f'(x) = 15/x + x - 5;$$
$$x_1 = 2 - [f(2) / f'(2)];$$
$$f(2) = 2,3973; f'(2) = 4,5;$$
$$x_1 = 2 - (2,3973)/4,5 = 1,4673.$$

Проверяем точность решения, для чего устанавливаем значение функции в точке 1,4673. Оно равно –0,5087.

Повторяем расчет для точки (1,4673). Значение производной в этой точке равно 6,6902:

$$x_2 = 1,4673 - [(-0,5987) / 6,6902] = 1,5433.$$

Функция для этого значения равна $f(1,5433) = -0,0166$. Дальнейшие вычисления представлены ниже:

x	$f(x_n)$	$f'(x)$	x_n
2	2,397207708	4,500	1,467287
1,467287	–0,508741492	6,690235	1,54333
1,543330	–0,016582637	6,262576	1,545977
1,545977	–1,85464E-05	6,248577	1,545980
1,545980	–0,000000000023	6,248561	1,545980
1,545980	0,0000000000000000	6,248561	1,545980

Таким образом, приближенное значение корня, полученное методом Ньютона, точнее, чем полученное ранее методом секущих. Для класса дифференцируемых функций метод Ньютона имеет более высокую скорость сходимости. Графическое представление решения уравнения методом Ньютона представлено на рис. 5.4.

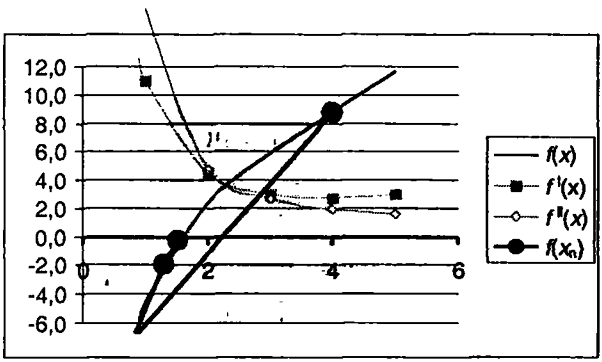


Рис. 5.4. Путь итерационного решения уравнения методом Ньютона

Нередко для более быстрого получения по возможности более точного решения уравнения применяется комбинированный метод, состоящий в переменном применении метода хорд и метода Ньютона. Корень уравнения всегда будет находиться между последовательными приближениями, рассчитанными обоими методами.

5.4. Решение уравнения в Excel

Подбор параметра – средство Excel для решения уравнений. Его еще называют средством «Что – если» анализа. При этом значения ячеек-параметров изменяются так, чтобы число в целевой ячейке стало равным заданному. Алгоритм действий следующий. Одну ячейку обозначают искомым параметром и присваивают ей какое-либо первоначальное значение, например 1. В другую, целевую, ячейку вводят уравнение, где искомый параметр записывается через его адрес ячейки. Вызывается меню *подбора параметра* (**Сервис – Подбор Параметра**) – рис. 5.5.

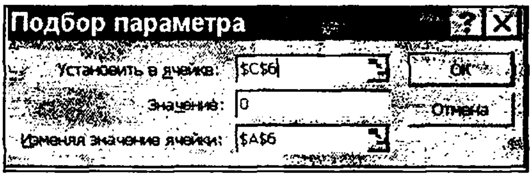


Рис. 5.5. Диалоговое окно ***Подбор параметра***

В диалоговом окне нужно указать целевую ячейку в списке **Установить в ячейке**, значение, которое должно быть достигнуто в поле **Значение**, и изменяемую ячейку в списке **Изменяя значение ячейки**.

Рассмотрим решение уравнения:

$$x^3 + x + 1 = 0.$$

Обозначим ячейку А6 – x и введем первоначальное значение =1. В ячейку С6 введем формулу:

=A6^3 +A6+1

Вызовем диалоговое окно **Подбор параметра** и выберем в списке **Установить в ячейке** адрес ячейки С6, в поле **Значение** – 0, в поле **Изменяя значение ячейки** – адрес ячейки А6. На рис. 5.6 показан результат выполнения операции подбора.

Если задача обладает плохой сходимостью, т. е. требуется много итераций, чтобы найти решение с заданной точностью, то можно воспользоваться кнопками **Пауза** и **Шаг** и с их помощью контролировать процесс и прервать его при необходимости.

Excel дает возможность выполнять подбор параметров для данных, которые представлены в виде диаграмм. Введем данные для расчета прибыли (рис. 5.7). Здесь в столбце В представлены формулы, по которым подсчитаны показатели и которые вводить, разумеется, надо со знаком =.

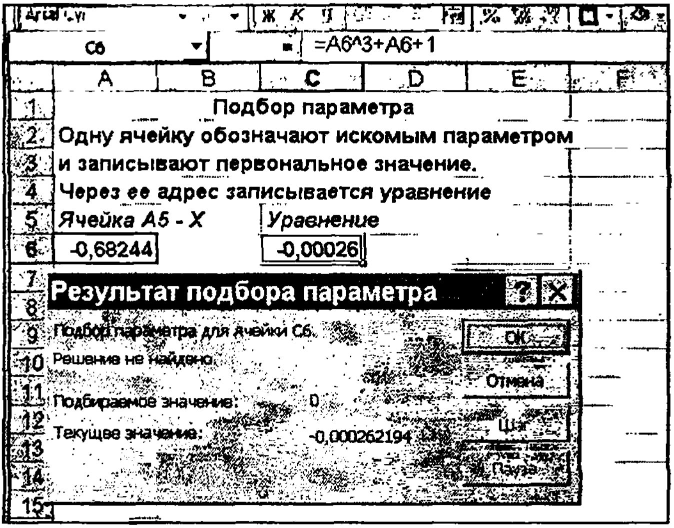


Рис. 5.6. Решение уравнения с помощью опции ***Подбор параметра***

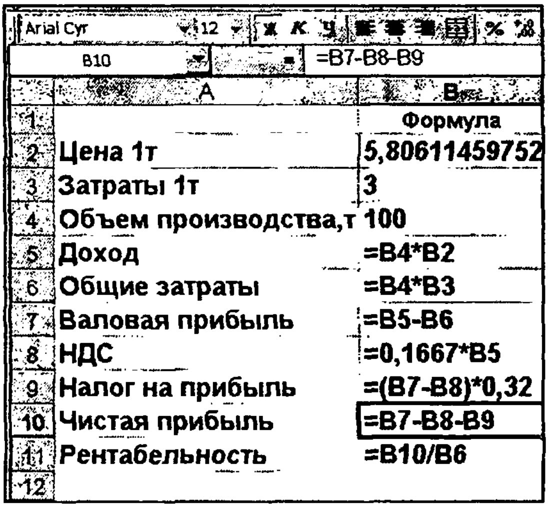


Рис. 5.7. Данные для расчета прибыли

По этим данным построим диаграмму (рис. 5.8). Выделим область диаграммы и столбиковую диаграмму щелчком мыши по любому столбику. Затем выберем столбец **Чистая прибыль** и щелкнем по нему левой кнопкой мыши при нажатой клавише <Ctrl>. Excel выделяет данный столбец восемью маленькими квадратиками, при этом средний квадрат имеет наверху цвет, отличный от остальных. При подведении к нему курсора мыши, стрелка становится разнонаправленной. Если тащить этот квадрат, удерживая левую кнопку мыши, то можно изменить высоту столбца диаграммы. Как только клавиша будет отпущена, появится диалоговое окно **Подбор параметра**, где нужно указать за счет какой ячейки произвести изменение.

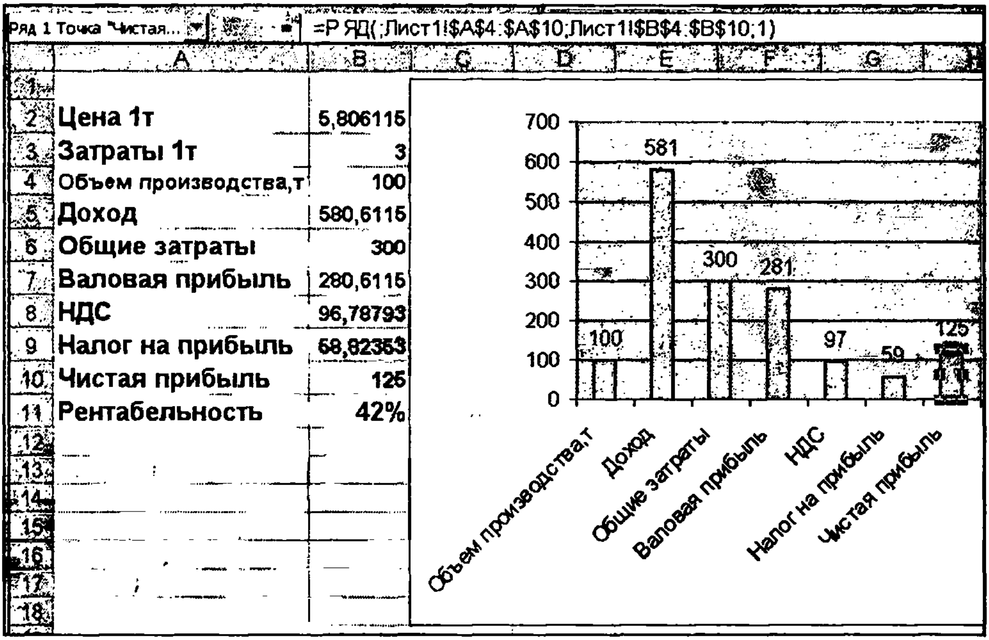


Рис. 5.8. Результат подбора параметра с использованием диаграммы

Excel позволяет сохранять промежуточные результаты. Это позволяет делать Диспетчер сценариев, диалоговое окно которого открывается командой меню **Сервис – Сценарии**.

5.5. Двумерный метод Ньютона–Рафсона

Для численного решения систем нелинейных уравнений с двумя переменными служит двумерный метод Ньютона–Рафсона.

Предположим, что необходимо решить систему из двух нелинейных уравнений:

$$f_1(x,y) = 0 \text{ и } f_2(x,y) = 0,$$

неизвестное решение обозначим через (x_0,y_0). Начинаем процесс решения с исходной точки (x_0,y_0), которую можно отыскать методом проб и ошибок. Далее требуется найти поправки p_0 и k_0, такие, что

$$x_0 + p_0 = x_1 \to x_q$$
$$y_0 + k_0 = y_1 \to y_q;$$

и мы получим решение системы нелинейных уравнений. В общем случае на n-й итерации находится точка (x_n, y_n), и мы ищем поправки p_n и k_n, такие, что

$$x_n + p_n = x_q$$
$$y_n + k_n = y_q.$$

Используя двумерный ряд Тейлора и приравнивая к нулю разложения функций f_1 и f_2 в точке $(x_n + p_n, y_n + k_n)$, получаем:

$$\left.\begin{aligned}(\frac{\partial}{\partial x} f_1(x_n, y_n)p_n + (\frac{\partial}{\partial y} f_1(x_n, y_n)k_n = -f_1(x_n, y_n);\\ (\frac{\partial}{\partial x} f_2(x_n, y_n)p_n + (\frac{\partial}{\partial y} f_2(x_n, y_n)k_n = -f_2(x_n, y_n).\end{aligned}\right\}$$

Далее ищем решение полученной системы из двух линейных уравнений с неизвестными p_n и k_n и получаем улучшенный вариант решения системы нелинейных уравнений:

$$x_{n+1} = x_n + p_n$$
$$y_{n+1} = y_n + k_n.$$

Этот процесс продолжается до тех пор, пока не будет достигнута желаемая степень точности. Аналогично производится обобщение метода на случай трех и более переменных. Алгоритм поиска решения системы из двух нелинейных уравнений следующий. Принимаются первоначальные значения x_0 и y_0. Рассчитываются значения функций f_1 и f_2 в этих точках. Устанавливаются производные f'_{1x}, $f'_{1y}, f'_{2x}, f'_{2y}$. Решаемая система уравнений:

$$f'_{1x}\, p_0 + f'_{1y}\, k_0 = -f_1;$$
$$f'_{2x}\, p_0 + f'_{2y}\, k_0 = -f_2,$$

откуда в общем случае:

$$k_0 = \frac{f_1\, f'_{2x} - f_2\, f'_{1x}}{f'_{1x}\, f'_{2y} - f'_{2x}\, f'_{1y}};$$

$$p_0 = -\frac{f_1}{f'_{1x}} - \frac{f'_{1y}}{f'_{1x}} \cdot k_0.$$

Устанавливаются значения:

$$x_1 = x_0 + p_0,$$
$$y_1 = y_0 + k_0$$

и расчет повторяется до желаемой точности f_1 и f_2.

Пример. Решим систему из двух нелинейных уравнений:

$$x^2 + (y-5)^2 - 6 = 0,$$
$$(x-3)^2 + y^2 - 12 = 0.$$

Примем начальные значения $x_0 = 1$ и $y_0 = 2$. Заметим, что

$$f'_{1x} = 2x, f'_{2x} = 2(x-3),$$
$$f'_{1y} = 2(y-5), f'_{2y} = 2y.$$

Таким образом, на первой итерации необходимо решить систему линейных уравнений:

$$2p_0 - 6k_0 = -4,$$
$$-4p_0 + 4k_0 = 4.$$

Находим, что $k_0 = 0{,}5$ и $p_0 = -0{,}5$ и улучшенное приближение задается значениями $x_1 = 0{,}5$ и $y_1 = 2{,}5$.

При повторении этой операции получены результаты, приведенные в табл. 5.2.

Таблица 5.2

n	f_1	f_2	k	p	x_n	y_n
2	0,5	0,5	0,15	0,25	0,75	2,65
3	0,085	0,085	0,038636	0,064394	0,814394	2,688636
4	0,005641	0,005637	0,002956	0,004926	0,819320	2,691592
5	0,000033	0,000033	0,000017	0,000029	0,819349	2,691609

5.6. Решение уравнений в Mathcad

Многие уравнения и системы из них не имеют аналитических решений. Однако они могут решаться численными методами с заданной погрешностью (не более значения, заданного системной переменной TOL).

Для простых уравнений вида $F(x) = 0$ решение находится с помощью функции:

root(Выражение, Имя переменной)

Эта функция возвращает значение переменной с указанным уровнем, при котором выражение дает 0. Функция реализует вычисления итерационным методом, причем можно задать начальное значение переменной. Это особенно полезно, если возможны несколько решений. Тогда выбор решения определяется выбором начального значения переменной. На рис. 5.9 приведен пример применения функции root для вычисления корня уравнения.

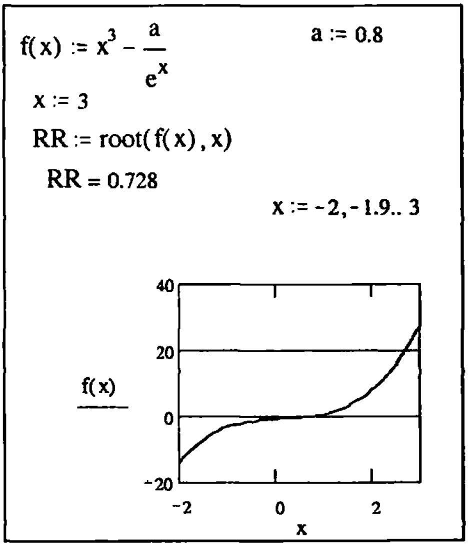


Рис. 5.9. Решение уравнения с использованием функции root

Чтобы правильно применять функцию **root**, необходимо знать, что в программе данной функции заложено, что корень функции – это не то значение аргумента, при котором выражение равно нулю, а то значение аргумента, при котором значение выражения не превышает значения системной переменной TOL. Чтобы функция root сработала правильно, необходимо переменной TOL присвоить новое значение (10^{-7}, например), заменив им предопределенное значение 10^{-3}, и построить график, визуально проверив решение.

Для поиска корней полинома степени *n* Mathcad содержит функцию:

polyroots(V)

Она возвращает вектор корней многочлена (полинома) степени *n*, коэффициенты которого находятся в векторе *V*, имеющим длину, равную *n* + 1. Вектор коэффициентов заполняется в обратном порядке. Включаются все коэффициенты многочлена, даже если они равны нулю. На рис. 5.10 приведен пример применения функции **polyroots**.

$$f(x) := \left(x^5 - 20\cdot x^3 - 10\cdot x^2\right) - x + 5$$

$$V := \begin{pmatrix} 5 \\ -1 \\ -10 \\ -20 \\ 0 \\ 1 \end{pmatrix}$$

$$r := \mathrm{polyroots}(V)$$

$$r^T = \begin{pmatrix} -4.213 & -0.486 - 0.537i & -0.486 + 0.537i & 0.481 & 4.704 \end{pmatrix}$$

$$x := -5, -4.99 .. 5 \qquad j := 0, 1 .. 4$$

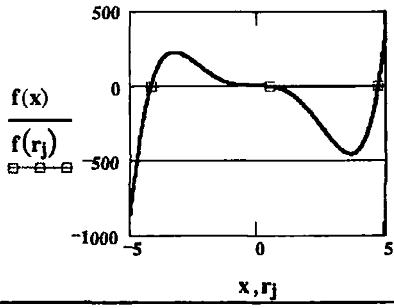


Рис. 5.10. Вычисление корней полинома

Заметим, что корни полинома могут быть как вещественными, так и комплексными числами. Не рекомендуется пользоваться этой функцией, если степень полинома выше пятой-шестой, т. к. тогда трудно получить малую погрешность вычисления корней.

При решении систем нелинейных уравнений используется специальный вычислительный блок, открываемый служебным словом **Given** и имеющий следующую структуру:

Given
Уравнения
Ограничительные условия
Выражения с функциями **find** и **minerr**

Рекомендуется дополнить блок проверкой решения системы.

В блоке используется одна из следующих двух функций:

- **find**($v_1, v_2, \ldots, v_n$) – возвращает значение одной или ряда переменных для точного решения;
- **minerr**($v_1, v_2, \ldots, v_n$) – возвращает значение одной или ряда переменных для приближенного решения.

Между этими функциями существует принципиальное различие. Первая функция используется, когда решение реально существует (хотя и не является аналитическим). Вторая функция пытается найти максимальное приближение даже к несуществующему решению путем минимизации среднеквадратической погрешности решения.

Фактически функция **minerr** применима для решения задач оптимизации.

Ограничительные условия задаются, как обычно, с помощью операторов – знаков отношения величин. Уравнения задаются через жирный знак равенства '**=**', вызываемый из панели или клавишами <Ctrl>+<=>. Рис. 5.11–5.15 иллюстрируют применение функций **find** и **minerr** для решения системы нелинейных уравнений.

Рассмотрим решение системы нелинейных уравнений

$$\begin{cases} \dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1; \\ y - x^2 = -1 \end{cases}$$

с помощью функции **find**. График этих функций представлен на рис. 5.11.

Учитывая, что график показывает, что есть два решения, первое решение найдем, приняв ограничение $x < 0$ (рис. 5.12).

Второе решение найдем, приняв ограничение $x > 0$ (рис. 5.13).

Теперь рассмотрим решение системы нелинейных уравнений:

$$\begin{cases} \dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1; \\ y - (x-2)^2 = 2{,}5 \end{cases}$$

с помощью функции **minerr**. График этих функций представлен на рис. 5.14.

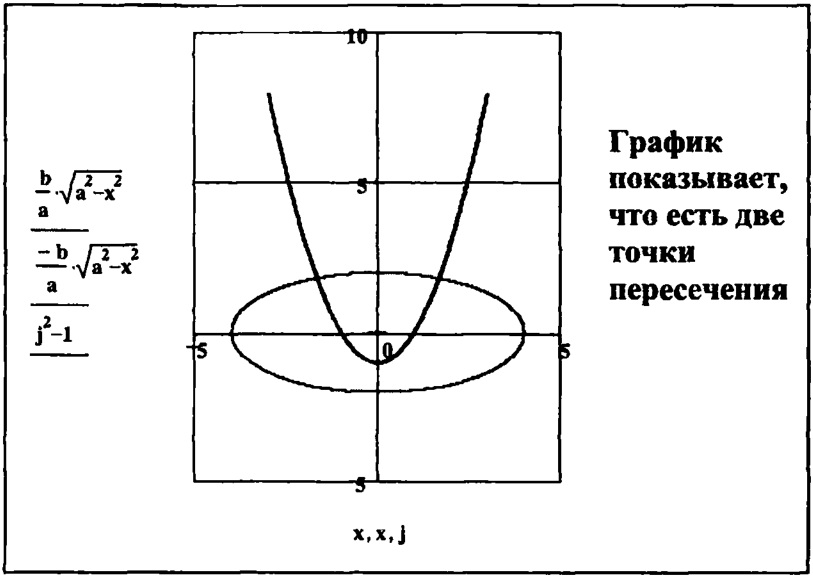


Рис. 5.11. Графики функций

Поиск первого решения

$x := -1$ $\quad y := 1$ $\quad a := 4$ $\quad b := 2$

Given

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad x < 0$$

$$y - x^2 = -1$$

Решаемая система уравнений при ограничениях x<0 и начальных условиях x=-1 и y=1

$V := \mathrm{Find}(x, y)$

$$V = \begin{pmatrix} -1.678 \\ 1.816 \end{pmatrix}$$

Найденное первое решение

*Рис. 5.12. Решение системы нелинейных уравнений с использованием функции **find**. Первое решение*

Поиск второго решения

$x := 1 \qquad y := 1 \qquad a := 4 \qquad b := 2$

Given

$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad x > 0$

$y - x^2 = -1$

Решаемая система уравнений при ограничениях x>0 и начальных условиях x=+1 и y=1

$V := \mathrm{Find}(x, y)$

$V = \begin{pmatrix} 1.678 \\ 1.816 \end{pmatrix}$. **Найденное второе решение**

*Рис. 5.13. Решение системы нелинейных уравнений с использованием функции **find**. Второе решение*

$x := -4.0, -3.99 .. 4.0 \qquad j := -2, -1.9 .. 5$

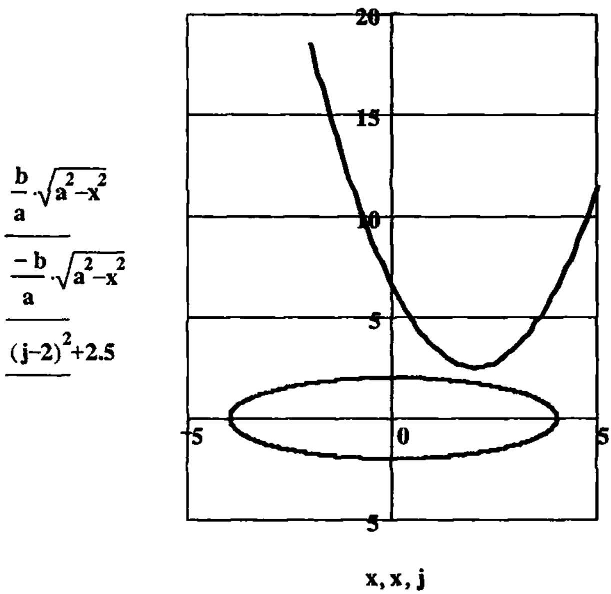


Рис. 5.14. Графики функций

Из рис. 5.14 видно, что кривые не пересекаются. Поиск решения в данном случае – поиск максимального приближения кривых путем минимизации среднеквадратической погрешности решения. Такое решение с использованием функции **minerr** приведено на рис. 5.15.

$x := 1.7$ $\quad y := 1.9$ $\quad a := 4$ $\quad b := 2$

Given

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad x > 0$$

$$y - (x - 2)^2 = 2.5$$

$V3 := \mathrm{Minerr}(x, y)$

$$V3 = \begin{pmatrix} 1.889 \\ 2.131 \end{pmatrix}$$

Решаемая система уравнений при ограничениях x>0 и начальных условиях x=1.7 и y=1.9

Найденное решение

*Рис. 5.15. Решение системы нелинейных уравнений с использованием функции **minerr***

Глава 6. Численные методы минимизации функции одной переменной

6.1. Основные понятия экстремальных задач

Понятия максимума и минимума объединяет термин экстремум (от латинского *extremum*, означающий «крайнее»). Термин оптимальный от латинского *optimum* означает «наилучший». Экстремум функции необязательно равен оптимуму задачи с учетом ограничений.

Теорию задач на отыскание наибольших и наименьших величин называют теорией экстремальных задач. Теорию отыскания оптимумов – теорией оптимизации [3, 6].

Экстремальные задачи, возникающие в экономике, обычно ставятся словесно, в терминах экономической науки. Чтобы можно было воспользоваться теорией, необходим перевод задач на математический язык.

Этот перевод называется формализацией.

Одна и та же задача может быть формализована разными способами, и простота решения зачастую сильно зависит от того, насколько удачно она формализована.

Любая формализованная задача включает в себя следующие элементы:

функционал $f: x \to \overline{R}, \qquad x \in C;$

x – область определения функционала f и ограничение $x \in C$; $\overline{R}$ – это расширенная действительная (вещественная) прямая, т. е. совокупность всех действительных чисел, дополненная значениями $+\infty$ и $-\infty$.

Запись $F: X \to Y$ означает, что отображение F имеет область определения X, и $F(x)$ для каждого элемента x из X лежит в множестве Y.

Таки образом, формализовать экстремальную задачу – значит точно описать ее элементы f, x и C.

Для формализованной задачи употребляется запись:

$$f(x) \to \inf; \quad x \in C \qquad\qquad f(x) \to \sup; \quad x \in C$$

или проще:

$$f(x) \to \min; \quad x \in C \qquad\qquad f(x) \to \max; \quad x \in C. \tag{6.1}$$

Точки $x \in C$ называют допустимыми. Если $C = x$, то задача называется задачей без ограничений.

Задачу на максимум всегда можно свести к задаче на минимум, заменив задачу $f(x) \to \max$; $x \in C$ задачей $\tilde{f}(x) \to \min$, $x \in C$, где $\tilde{f}(x) = -f(x)$.

И наоборот, задачу на минимум можно аналогичным образом свести к задаче на максимум. Если необходимо исследовать обе задачи, то пишут:

$$f(x) \to \text{extr}; \quad x \in C.$$

Допустимая точка x^* называется абсолютным (или глобальным) минимумом, если $f(x) \geq f(x^*)$ для любого $x \in X$ (соответственно, максимумом, если $f(x) \leq f(x^*)$ для любого $x \in C$). Абсолютный минимум (максимум) задачи называется решением задачи. Величина $f(x^*)$, где x^* решение задачи, называется численным значением задачи.

Точка x^* является локальным минимумом (максимумом), если $x^* \in C$, и существует $\delta > 0$, такое, что для любой допустимой точки x, для которой $\|x - x^*\| < \delta$, выполняется неравенство $f(x) \geq f(x^*)$; (для максимума $- f(x) \leq f(x^*)$).

Теория экстремальных задач дает правила нахождения решений экстремальных задач. В большинстве своем эти правила выделяют некоторое подмножество точек, среди которых должно содержаться решение задачи. Это подмножество точек, которое называют критическим, возможно, несколько шире, чем множество абсолютных и даже локальных экстремумов. После нахождения всех критических точек надо выделить из них решения.

Унимодальиые функции

Введем понятие унимодальной функции. Непрерывная функция $y = f(x)$ называется унимодальной на отрезке $[a, b]$, если точка x^* локального минимума функции принадлежит отрезку $[a, b]$; и для любых двух точек отрезка x_1 и x_2, взятых по одну сторону от точки минимума, точке x_1, более близкой к точке минимума, соответствует меньшее значение функции, т. е. как при $x_2 < x_1 < x^*$ (рис. 6.1), так и при $x^* < x_1 < x_2$ (рис. 6.2) справедливо неравенство $f(x_1) < f(x_2)$.

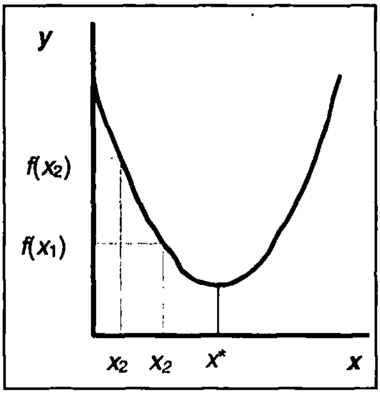


Рис. 6.1

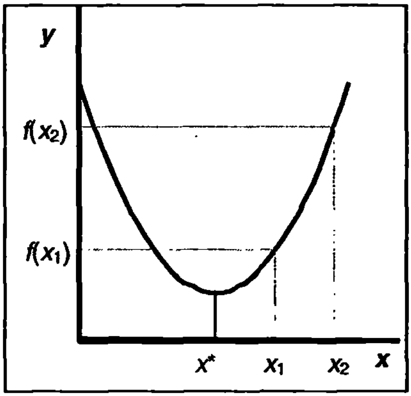


Рис. 6.2

Достаточное условие унимодальности функции $f(x)$ на отрезке $[a, b]$ формулируется следующим образом. Если функция дважды дифференцируема на отрезке $[a, b]$ и $f''(x) < 0$ в любой точке этого отрезка, то $f(x)$ – унимодальная функция на $[a, b]$.

Отметим, что условие $f''(x) > 0$ определяет множество точек, на котором функция является выпуклой вниз. Сооответственно, если функция является выпуклой вверх, условие формулируется как $f''(x) < 0$. Поведение второй производной представлено на рис. 6.3.

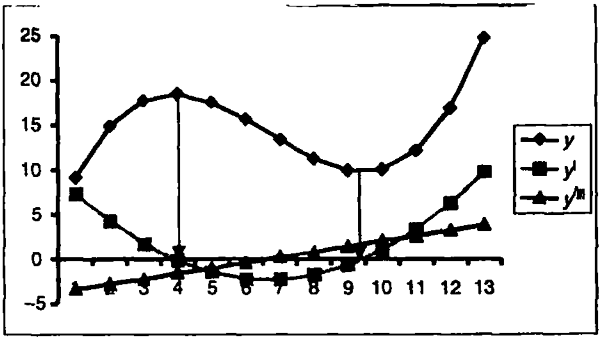


Рис. 6.3

Применение численных методов для отыскания точек локального минимума функции $f(x)$ предполагает:

1) определение промежутков унимодальности функции, т. е. нахождение отрезков, которым принадлежит одна точка локального минимума;

2) вычисление значения x^*, принадлежащего выбранному отрезку, с заданной точностью ε.

Желательно построить график функции и выбрать отрезок $[a, b]$ по возможности малым.

При вычислении точки минимума точность достигается последовательным уменьшением отрезка $[a, b]$, содержащего точку x^*, до размеров, не превышающих заданную точность $(b - a) \le \varepsilon$.

Если функция имеет производную в области определения, то для отыскания стационарных точек функции $f(x)$ нужно найти корни уравнения $f'(x) = 0$. Для решения этого уравнения можно использовать численные методы решения уравнений, уже рассмотренные нами: метод деления отрезка пополам, метод секущих, метод Ньютона. Однако для решения задачи минимизации функции проще применять прямые численные методы поиска минимума функции $f(x)$.

Рассмотрим прием, позволяющий сузить отрезок унимодальности функции.

Пусть функция $f(x)$ унимодальна на отрезке $[a, b]$. Возьмем две произвольные точки x_1 и x_2, принадлежащие отрезку $[a, b]$ и такие, что $x_1 < x_2$. В каждом из двух следующих случаев можно указать отрезок меньших размеров $[a_1, b_1]$, содержащий точку минимума x^* и принадлежащий первоначальному отрезку (рис. 6.4).

1. Если $f(x_1) < f(x_2)$, то примем $a_1 = a$ и $b_1 = x_2$ и получим меньший отрезок унимодальности $[a_1, b_1]$.

2. Если $f(x_1) > f(x_2)$, то принимаем $a_1 = x_1$ и $b_1 = b$.

Методы вычисления значения точки минимума функции одной переменной отличаются алгоритмами выбора точек x_1 и x_2 для локализации точки x^* с заданной точностью. Рассмотрим методы половинного деления, золотого сечения и Фибоначчи.

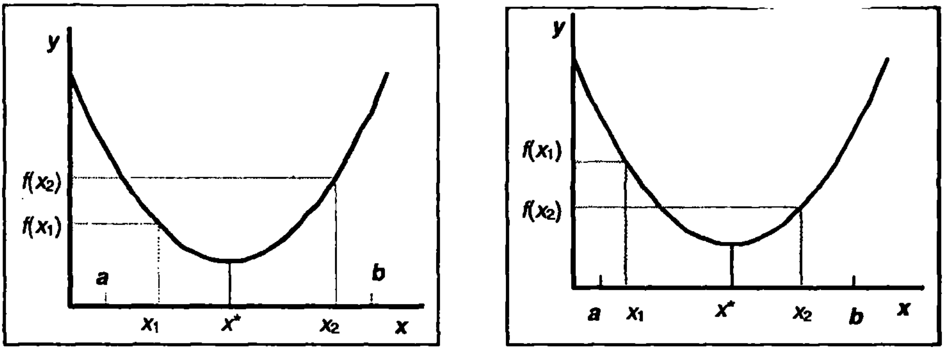


Рис. 6.4

6.2. Метод половинного деления

Пусть при решении задачи

$$f(x) \to \min, \quad x \in R$$

определен отрезок $[a, b]$, которому принадлежит точка локального минимума x^*, и функция $f(x)$ является унимодальной на этом отрезке.

Для сужения отрезка унимодальности используем точки x_1 и x_2, расположенные симметрично относительно середины отрезка $[a, b]$:

$$x_{2,1} = (a + b)/2 \pm k(b - a)/2, \quad k << 1.$$

Поскольку k гораздо меньше 1, точки x_1 и x_2 принадлежат отрезку $[a, b]$, получим новый суженный отрезок $[a_1, b_1]$, выбрав его длину в одном из трех возможных случаев:

1. $f(x_1) < f(x_2)$, $\quad a_1 = a$, $\quad b_1 = x_2 = (a + b)/2 + k(b - a)/2$.
2. $f(x_1) > f(x_2)$, $\quad a_1 = x_1 = (a + b)/2 - k(b - a)/2$, $\quad b_1 = b$.
3. $f(x_1) = f(x_2)$, $\quad a_1 = x_1$, $\quad b_1 = x_2$.

В новом суженном промежутке $[a_1, b_1]$ выберем точки $x_{1,1}$ и $x_{2,1}$, симметричные относительно его середины:

$$x_{1,1} = (a_1 + b_1)/2 - k(b_1 - a_1)/2.$$
$$x_{2,1} = (a_1 + b_1)/2 + k(b_1 - a_1)/2.$$

Проведя вычисления функции, устанавливаем отрезок $[a_2, b_2]$. Далее продолжаем итерационные вычисления. В общем случае

$$x_{2_n;\,1_n} = (a_{n-1} + b_{n-1})/2 \pm k(b_{n-1} - a_{n-1})/2.$$

Точность решения на n-м шаге вычислений оценивается неравенством:

$$0 \le x^* - a_n \le b_n - a_n \le (1 + k)^n (b - a)/2^n < \varepsilon.$$

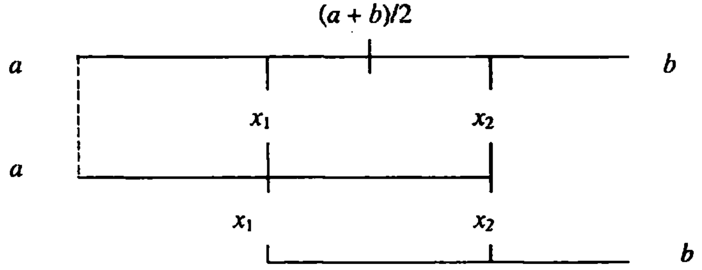
Минимальное количество итераций n для достижения точности ε можно оценить по выражению:

$$n > \ln[\varepsilon/(b - a)] / \ln[(1 + k)/2].$$

6.3. Метод золотого сечения и реализация его в Excel

Точка x_1 является золотым сечением отрезка $[a, b]$, если отношение длины всего отрезка $(b - a)$ к длине большей части $(b - x_1)$ равно отношению длины большей части отрезка к длине ($x_1 - a$) меньшей части. То есть точка x_1 является золотым сечением, если справедливо отношение:

$$(b-a)/(b-x_1)=(b-x_1)/(x_1-a)$$



Аналогично, точка x_2, симметричная точке x_1 относительно середины отрезка $[a, b]$, является вторым золотым сечением этого отрезка. Так как точки x_1 и x_2 расположены симметрично относительно середины отрезка $[a, b]$, то можно записать:

$$x_{1,2}=\frac{a+b}{2}\mp k\frac{b-a}{2}.$$

Учитывая, что

$$b-x_1=b-\frac{a+b}{2}+k\frac{b-a}{2}=(1+k)\frac{b-a}{2};$$

$$x_1-a=(1-k)\frac{b-a}{2},$$

и используя определение золотого сечения, вычислим число $k > 0$:

$$\frac{b-a}{b-x_1}=\frac{b-x_1}{x_1-a},\qquad \frac{b-a}{b-x_1}=\frac{2}{1+k},\qquad \frac{b-x_1}{x_1-a}=\frac{1+k}{1-k},\qquad \frac{2}{1+k}=\frac{1+k}{1-k},$$

$$k^2+4k-1=0,\qquad k=\sqrt{5}-2.$$

Отметим свойство золотого сечения: пусть x_1 и x_2 – два золотых сечения отрезка $[a, b]$; тогда точка x_1 одновременно является и золотым сечением отрезка $[a, x_2]$, а другая точка x_2 – золотым сечением отрезка $[x_1, b]$.

Алгоритм метода следующий: сначала на исходном отрезке $[a, b]$ по формуле

$$x_{1,2}=\frac{a+b}{2}\mp k\frac{b-a}{2}$$

при $k=\sqrt{5}-2$ находятся точки x_1 и x_2, а затем разность $\Delta_1 x = x_2 - x_1$. Далее вычисляем значения функций $y_1 = f(x_1)$ и $y_2 = f(x_2)$ и образуем суженный отрезок $[a, b]$.

Используя свойство золотого сечения на отрезке $[a_1, b_1]$, находим два сечения: x^1_1, x^1_2. При этом возможны три случая:

1. $y_1 < y_2$, $a_1 = a$, $b_1 = x_2$, $x^1_2 = x_1$, $x^1_1 = a_1 + \Delta_1 x$, $y^1_2 = y_1$.
2. $y_1 > y_2$, $a_1 = x_1$, $b_1 = b$, $x^1_1 = x_2$, $x^1_2 = b_1 - \Delta_1 x$, $y^1_2 = y_1$.
3. $y_1 = y_2$, $a_1 = x_1$, $b_1 = x_2$, $x^1_{1,2} = (a + b)/2 \mp k(b - a)/2$.

Далее выполняются итерационные расчеты по нахождению отрезков $[a_2, b_2]$, $[a_3, b_3]$ и т. д., учитывая при этом, что в случаях 1 или 2 значение $y^1{}_2$ или $y^1{}_1$ целевой функции уже получено на предыдущем шаге ($i = 1,2,\ldots,n$).

Точность приближенного равенства $x^* \approx a_n \approx b_n$ на n-м шаге вычислений можно оценить неравенством:

$$0 \le x^* - a_n \le (b-a)/\tau^n < \varepsilon,$$

$$0 \le x^* - a_n \le ((\sqrt{5} - 1)/2)^n (b-a) < \varepsilon,$$

$$\tau = 2/(1+k) = 2/(\sqrt{5} - 1) \approx 1{,}618\ldots$$

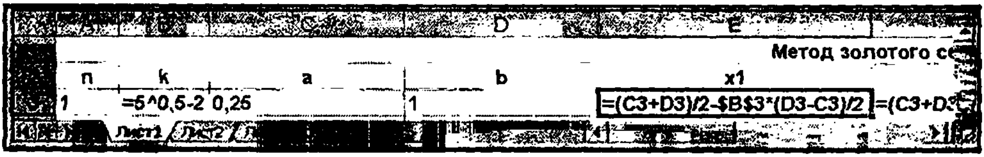
Минимальное количество итераций n для достижения точности ε устанавливается из формулы:

$$(b-a)/\tau^n < \varepsilon,$$

$$n < \ln[(b-a)/\varepsilon] / \ln\tau.$$

Метод золотого сечения в Excel. Метод реализуется в Excel с помощью функции ЕСЛИ. Рассмотрим реализацию метода на примере нахождения локального минимума функции $f(x) = 2x^2 - \ln x$ на отрезке [0,25; 1] с точностью 0,1.

Введем в ячейку B3 величину k, в ячейку C3 – a, в ячейку D3 – b. Далее в ячейки E3 и F3 формулы соответственно для x_1 и x_2, а в ячейки G3 и H3 выражения функции y_1 и y_2 через адреса ячеек x_1 и x_2 (рис. 6.5).



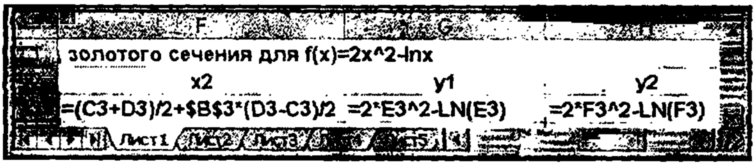


Рис. 6.5

Дальнейшее сужение интервала $[a, b]$ производим с помощью функции ЕСЛИ (рис. 6.6).

В ячейки C4 и D4 вводим функцию ЕСЛИ с соответствующими условиями:

=ЕСЛИ(G3<H3;C3;E3)	=ЕСЛИ(G3<H3;F3;D3)

Ячейки F3, G3, H3 выделяем и копируем формулы в ячейки F4, G4, H4 (рис. 6.7).

Теперь выделяем строку 4 (вторая строка с расчетами) и мышью, взяв за нижний правый квадратик, копируем формулы на другие строки, до требуемой точности расчетов (рис. 6.8).

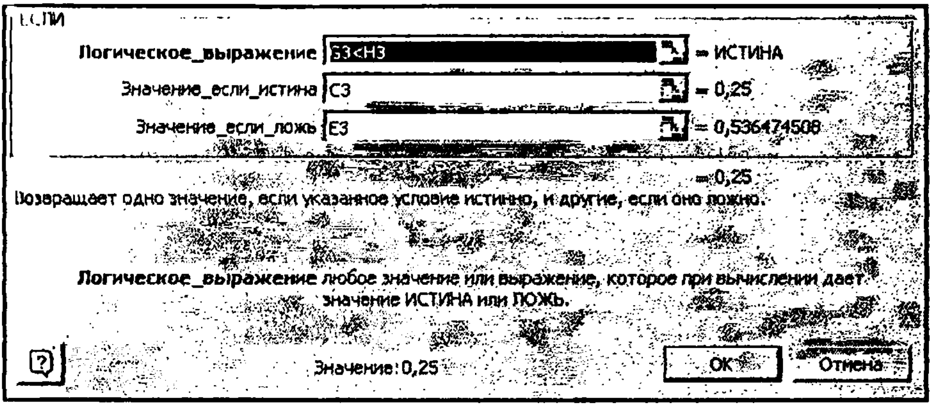


Рис. 6.6

x2	y1	y2
=(C3+D3)/2+$B$3*(D3-C3)/2	=2*E3^2-LN(E3)	=2*F3^2-LN(F3)
=(C4+D4)/2+$B$3*(D4-C4)/2	=2*E4^2-LN(E4)	=2*F4^2-LN(F4)

Рис. 6.7

	A	B	C	D	E	F	G	H	I
1					Метод золотого сечения для f(x)=2x^2-lnx				
2	n	k	a	b	x1	x2	y1	y2	b-a
3	1	0,2	0,25	1	0,536474508	0,713525492	1,1983	1,355774	0,75
4	2		0,25	0,7135255	0,427050983	0,536474508	1,2156	1,198346	0,463525
5	3		0,427051	0,7135255	0,536474508	0,604101966	1,1983	1,233891	0,286475
6	4		0,427051	0,604102	0,494678441	0,536474508	1,1933	1,198346	0,177051
7	5		0,427051	0,5364745	0,468847051	0,494678441	1,1971	1,193261	0,109424
8	6		0,4688471	0,5364745	0,494678441	0,510643118	1,1933	1,193597	0,067627
9									

Рис. 6.8

Установленная нами точность $\varepsilon = 0{,}1$ достигается после 5-й итерации. Используя Excel, можно, конечно, провести и очень точные расчеты. Например, в данном случае на 25-ой итерации точность достигает 6 знака, на 40-й итерации – 9 знака (рис. 6.9), а на 73-й итерации – 16 знака.

Выделив ряды a, b, x_1, x_2 и построив график (рис. 6.10), можно увидеть графическое решение задачи, причем отрезок $[a, b]$ всегда делится точками x_1, x_2 в пропорции золотого сечения.

	A	B	C	D	E	F	G	H	I
19	17		0,4998867	0,5002266	0,500016527	0,500096749	1,1931	1,193147	0,00034
20	18		0,4998867	0,5000967	0,499966947	0,500016527	1,1931	1,193147	0,00021
21	19		0,4999669	0,5000967	0,500016527	0,500047169	1,1931	1,193147	0,00013
22	20		0,4999669	0,5000472	0,499997589	0,500016527	1,1931	1,193147	8,02E-05
23	21		0,4999669	0,5000165	0,499985884	0,499997589	1,1931	1,193147	4,96E-05
24	22		0,4999859	0,5000165	0,499997589	0,500004822	1,1931	1,193147	3,06E-05
25	23		0,4999859	0,5000048	0,499993118	0,499997589	1,1931	1,193147	1,89E-05
26	24		0,4999931	0,5000048	0,499997589	0,500000352	1,1931	1,193147	1,17E-05
27	25		0,4999976	0,5000048	0,500000352	0,500002059	1,1931	1,193147	7,23E-06
28	26		0,4999976	0,5000021	0,499999296	0,500000352	1,1931	1,193147	4,47E-06
29	27		0,4999993	0,5000021	0,500000352	0,500001004	1,1931	1,193147	2,76E-06
30	28		0,4999993	0,500001	0,499999949	0,500000352	1,1931	1,193147	1,71E-06
31	29		0,4999993	0,5000004	0,4999997	0,499999949	1,1931	1,193147	1,06E-06
32	30		0,4999997	0,5000004	0,499999949	0,500000103	1,1931	1,193147	6,52E-07
33	31		0,4999997	0,5000001	0,499999854	0,499999949	1,1931	1,193147	4,03E-07
34	32		0,4999999	0,5000001	0,499999949	0,500000007	1,1931	1,193147	2,49E-07
35	33		0,4999999	0,5000001	0,500000007	0,500000044	1,1931	1,193147	1,54E-07
36	34		0,5	0,5000001	0,500000044	0,500000066	1,1931	1,193147	9,52E-08
37	35		0,5	0,5000001	0,50000003	0,500000044	1,1931	1,193147	5,88E-08
38	36		0,5	0,5000001	0,500000044	0,500000052	1,1931	1,193147	3,63E-08
39	37		0,5	0,5000001	0,500000039	0,500000044	1,1931	1,193147	2,25E-08
40	38		0,5	0,5000001	0,500000044	0,500000047	1,1931	1,193147	1,39E-08
41	39		0,5	0,5000001	0,500000047	0,500000049	1,1931	1,193147	8,58E-09
42	40		0,5	0,5	0,500000046	0,500000047	1,1931	1,193147	5,3E-09
43	41		0,5	0,5	0,500000047	0,500000048	1,1931	1,193147	3,28E-09

Лист1 Лист2 Лист3 Лист4 Лист5

Рис. 6.9

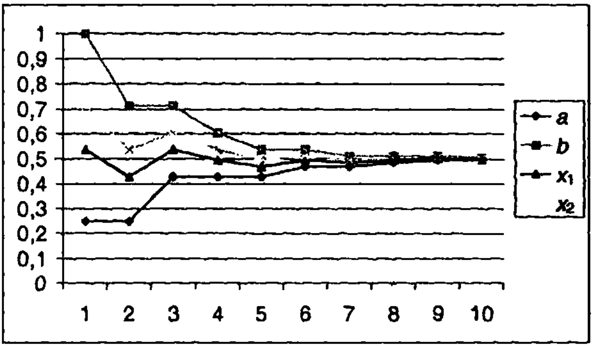


Рис. 6.10

6.4. Метод Фибоначчи и реализация его в Excel

Одним из лучших методов безусловной оптимизации без использования производных является метод Фибоначчи. Он является процедурой линейного поиска для минимизации выпуклой функции на замкнутом ограниченном интер-

...ние. Подобно методу золотого сечения процедура поиска Фибоначчи требует ...ух вычисления функции на первой итерации, а на каждой последующей – только по одному. Однако эта процедура отличается от метода золотого сечения тем, что сокращение интервала неопределенности меняется от итерации к итерации.

Числа Фибоначчи

Последовательность чисел F_i задается условиями:

$$F_0 = F_1 = 1, \quad F_{i+1} = F_i + F_{i-1}, \quad i = 1, 2, \ldots$$

Элементы введенной последовательности называются числами Фибоначчи. Выведем формулу, выражающую числа Фибоначчи в явном виде. Для этого будем искать решение уравнения $F_i = F_i + F_{i-1}$ среди геометрических прогрессий с i-м членом t^i. Имеем $t^{i+1} = t^i + t^{i-1}$. Ненулевые корни этого уравнения являются корнями квадратного уравнения $t^2 - t - 1$ и равны $(1 \pm \sqrt{5})/2$. Введем обозначение $\tau = (1 + \sqrt{5})/2 = 1{,}618\ldots$ Тогда $(1 - \sqrt{5})/2 = -1/\tau$. Последовательности τ^i и $(-1/\tau)^i$ удовлетворяют нашему уравнению. Ему же, очевидно, удовлетворяет и любая линейная комбинация этих последовательностей $c_1\tau^i + c_2(-1/\tau)^i$. Выберем коэффициенты c_1, c_2 так, чтобы выполнялись условия $F_0 = F_1 = 1$. Имеем

$$c_1 + c_2 = 1, \; c_1\tau + c_2(-1/\tau) = 1.$$

Решая эти уравнения, получаем следующую формулу для чисел Фибоначчи:

$$F_i = \frac{\tau^{i+1} - (-\tau)^{-(i+1)}}{\sqrt{5}} = \frac{\left(\frac{1+\sqrt{5}}{2}\right)^{i+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{i+1}}{\sqrt{5}}.$$

Последовательность чисел Фибоначчи представлена на рис. 6.11.

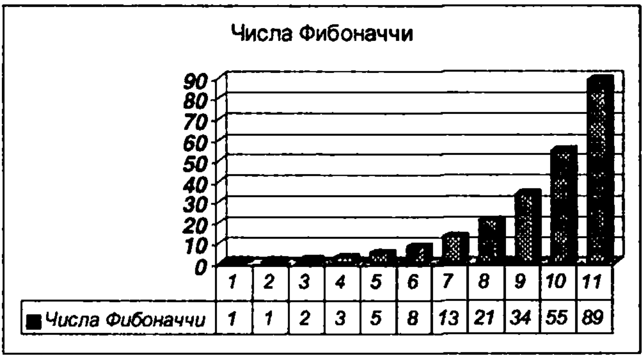


Рис. 6.11. Графическое представление начального ряда чисел Фибоначчи

Отношение предыдущих чисел Фибоначчи к последующим – $F_n - k/F_n$ образует ряд чисел, позволяющих симметрично локализовать середину отрезка (рис. 6.12).

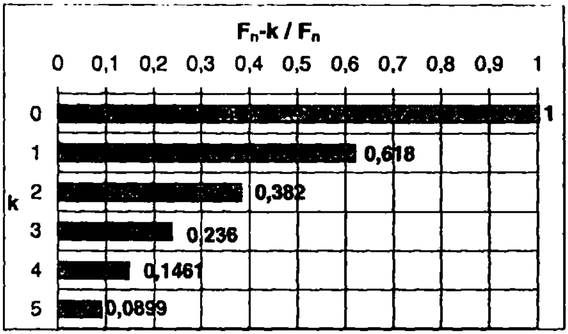


Рис. 6.12. Отношение чисел Фибоначчи

Метод Фибоначчи

Задается начальный интервал неопределенности $[a_1, b_1]$. Выбирается допустимый конечный интервал неопределенности $h > 0$ и константа различимости $\varepsilon > 0$. Выбирается число n вычислений функции так, чтобы

$$F_n > (b_1 - a_1)/h.$$

Рассмотрим частный случай $[a_1, b_1] = [0, 1]$. Первые два вычисления проводятся в точках $\lambda_1 = F_{n-2}/F_n$, $\mu_1 = F_{n-1}/F_n$, расположенных симметрично относительно середины отрезка [0, 1]. Если $f(\lambda_1) \le f(\mu_1)$, то отрезком локализации минимума является отрезок $[0, \mu_1]$, в случае $f(\lambda_1) > f(\mu_1)$ – отрезок $[\lambda_1, 1]$. Алгоритм определяется следующим условием: на каждом шаге точка очередного вычисления выбирается симметрично относительно середины отрезка локализации к лежащей внутри этого отрезка точке уже проведенного вычисления. Заметим, что на каждом шаге метода Фибоначчи точка, лежащая внутри отрезка локализации, делит его в отношении двух последовательных чисел Фибоначчи (рис. 6.13).

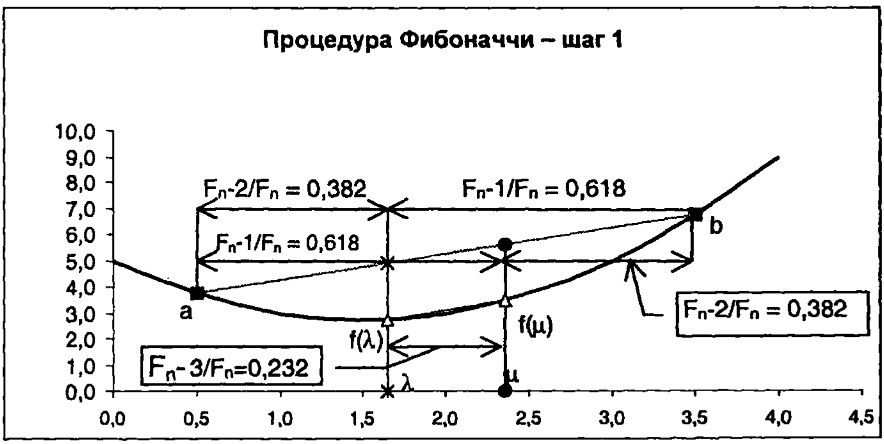


Рис. 6.13. Первый шаг алгоритма Фибоначчи

Перейдем к общему случаю. Первые две точки устанавливаются так:

$$\lambda_1 = a_1 + (F_{n-2}/F_n)(b_1 - a_1), \qquad \mu_1 = a_1 + (F_{n-1}/F_n)(b_1 - a_1).$$

Вычисляются $f(\lambda_1)$, $f(\mu_1)$ и устанавливаются точки следующего отрезка локализации. Если $f(\lambda_1) \le f(\mu_1)$, то отрезком локализации минимума является отрезок $[a_1, \mu_1]$, в случае $f(\lambda_1) > f(\mu_1)$ – отрезок $[\lambda_1, b_1]$ (рис. 6.13, 6.14).

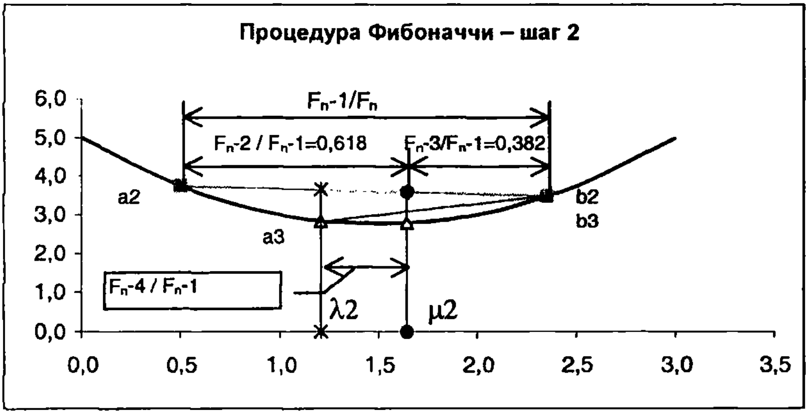


Рис. 6.14. Второй шаг алгоритма Фибоначчи

Для дальнейших вычислений примем, что k – номер итерации. Тогда:

$$\lambda_k = a_k + \frac{F_{n-k-1}}{F_{n-k+1}}(b_k - a_k), \qquad k = 1, \ldots, n-1;$$

$$\mu_k = a_k + \frac{F_{n-k}}{F_{n-k+1}}(b_k - a_k), \qquad k = 1, \ldots, n-1.$$

Отрезком локализации минимума является отрезок $[a_k, \mu_k]$, если $f(\lambda_k) \le f(\mu_k)$, и в случае $f(\lambda_k) > f(\mu_k)$ – отрезок $[\lambda_k, b_k]$.

Длина интервала неопределенности (отрезка локализации минимума) сжимается с коэффициентом F_{n-k}/F_{n-k+1}. Наглядно это для рассматриваемого ниже примера представлено на рис. 6.15 и 6.16.

На первой итерации требуются два вычисления функции, а на каждой последующей – только одно. Следовательно, в конце $(n-2)$-й итерации будет выполнено $(n-1)$ вычислений функции. Далее, для $k = n-1$, $\lambda_{n-1} = \mu_{n-1} = S(a_{n-1} + b_{n-1})$. Следовательно, либо $\lambda_{n-1} = \mu_{n-2}$, либо $\mu_{n-1} = \lambda_{n-2}$, т. е. теоретически не должно производиться новых вычислений функции на этой стадии. Однако чтобы обеспечить дальнейшее сокращение интервала неопределенности, точка последнего вычисления слегка перемещается вправо или влево от средней точки $\lambda_{n-1} = \mu_{n-1}$ на величину ε, так что $S(a_{n-1} + b_{n-1})$ есть длина конечного интервала неопределенности.

Таким образом, если $k = n-2$, то положить $\lambda_n = \lambda_{n-1}$ и $\mu_n = \lambda_n + \varepsilon$. Если $f(\lambda_n) > f(\mu_n)$, то положить $a_n = \lambda_n$ и $b_n = b_{n-1}$. Если $f(\lambda_n) < f(\mu_n)$, то положить $a_n = a_{n-1}$ и $b_n = \lambda_n$. Оптимальное решение содержится в интервале $[a_n, b_n]$.

Рис. 6.15. Сжатие интервала неопределенности

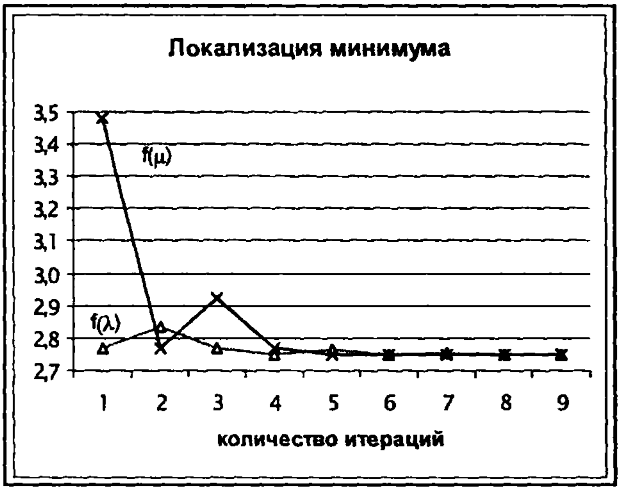


Рис. 6.16. Локализация минимума функции

Пример. Рассмотрим минимизацию функции $x^2 - 3x + 5$ методом Фибоначчи в Excel. Примем, что начальный интервал равен [0,5; 3,5], а длина конечного интервала неопределенности не превосходила 0,05, константа различимости равна 0,01. Тогда (3,5 – 0,5)/0,05 = 60. Больше этого числа ближайшее число Фибоначчи 89, следовательно: $n = 10$.

Для этого n на рис. 6.17 приведен расчет отношеиий F_{n-k-1}/F_{n-k+1} и F_{n-k}/F_{n-k+1}.

На рис.6.18 приведены результаты итерационных вычислений по методу Фибоначчи. Для расчета интервала неопределенности использованы логические функции. Так, начиная с ячейки I3 введено =ЕСЛИ(M2<=N2;J2;K2), начиная с ячейки J3 – =ЕСЛИ(M2<=N2;L2;J2).

	A	B	C	D	E	F	G	H
1	n	Fn	0,5	n-k-1	n-k+1	F(n-k-1)/F(n-k+1)	n-k	F(n-k)/F(n-k+1)
2	1	1	3,5	8	10	0,382022472	9	0,617977528
3	1	1	3,00	8	10	0,382022472	9	0,617977528
4	2	2	0,05	7	9	0,381818182	8	0,618181818
5	3	3	60,00	6	8	0,382352941	7	0,617647059
6	4	5	10,00	5	7	0,380952381	6	0,619047619
7	5	8		4	6	0,384615385	5	0,615384615
8	6	13		3	5	0,375	4	0,625
9	7	21		2	4	0,4	3	0,6
10	8	34		1	3	0,333333333	2	0,666666667
11	9	55		0	2	0,5	1	0,5
12	10	89						
13	11	144						
14	12	233						
15	13	377						
16	14	610						

Лист1

Рис. 6.17. Начальные вычисления процедуры Фибоначчи

a	b	λ	μ	f(λn)	f(μn)
0,500000	3,500000	1,646067	2,353933	2,771336	3,479201
0,500000	2,353933	1,208244	1,645689	2,835122	2,771225
1,208244	2,353933	1,645689	1,916488	2,771225	2,923462
1,208244	1,916488	1,479043	1,645689	2,750439	2,771225
1,208244	1,645689	1,374890	1,479043	2,765653	2,750439
1,374890	1,645689	1,479043	1,541535	2,750439	2,751725
1,374890	1,541535	1,437382	1,479043	2,753921	2,750439
1,437382	1,541535	1,479043	1,499874	2,750439	2,750000
1,479043	1,541535	1,499874	1,520704	2,750000	2,750429
1,479043	1,520704	1,499874	1,509874	2,750000	2,750097
1,479043	1,509874				
	1,4944584				2,750031
0,030831			0,010000		

Лист1

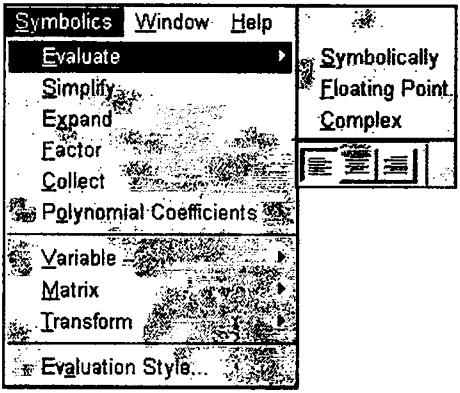
Рис. 6.18. Результаты минимизации функции по методу Фибоначчи

Для $k = 10$ $\lambda_{10} = 1{,}499874$ и $\mu_{10} = \lambda_{10} + \varepsilon$. Тогда конечный интервал неопределенности $[a_{10}, b_{10}]$ равен [1,479043; 1,509874], длина которого 0,030831. В качестве приближенного значения точки минимума выберем середину этого отрезка: 1,494484.

Глава 7. Работа с символьным процессором Mathcad

7.1. Выделение объектов символьных операций

Операции, относящиеся к работе символьного процессора, содержатся в меню опции **Symbolics** (Символика).



Чтобы символьные операции выполнялись, процессору необходимо указать, над каким выражением это должно проводиться, т. е. надо выделить выражение. Для ряда операций надо не только указать выражение, но и наметить переменную, относительно которой выполняется символьная операция. Само выражение в таком случае не выделяется.

Операции с выделенными выражениями:

Evaluate (Вычислить); **Simplify** (Упростить); **Expand** (Разложить по степеням); **Factor** (Разложить на множители); **Collect** (Разложить по подвыражению).

Операции с выделенными переменными:

Polynomial Coefficient (Полиномиальные коэффициенты); **Solve** (Решить относительно переменной); **Substitute** (Заменить переменную); **Differentiate** (Дифференцировать по переменной); **Integrate** (Интегрировать по переменной); **Expand to Series** (Разложить в ряд); **Convert to Partial Fraction** (Разложить на элементарные дроби).

Если объект отсутствует (не выделен), доступа к соответствующим операциям в опции **Symbolics** (Символика) главного меню нет и эти операции показываются «затененным» шрифтом.

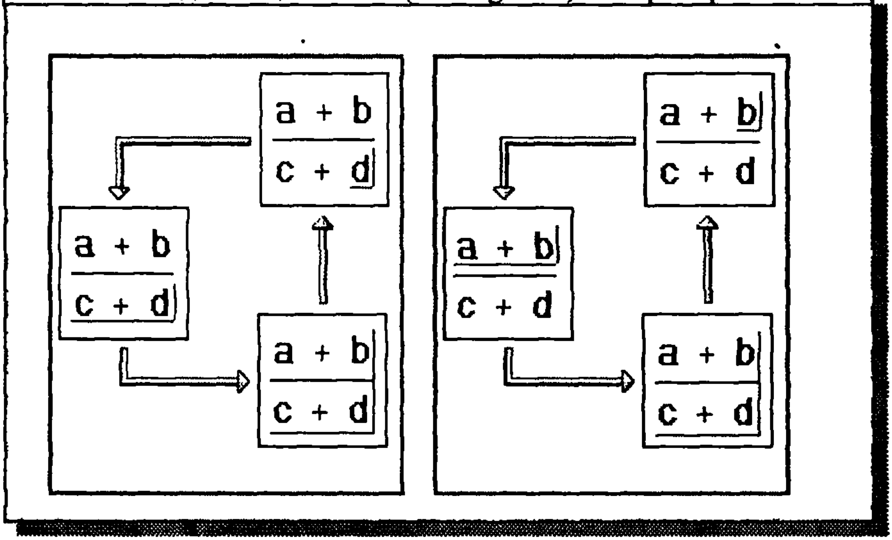
Напомню, что возможны два вида выделения – пунктирными линиями и сплошными. Чтобы отметить выражение пунктирной линией, достаточно установить курсор-крестик на нужном объекте, нажать клавишу <Shift> или <Ctrl> и воспользоваться кнопкой мыши. Выделение пунктирной линией используется для перемещения объектов по окну.

Для операций с символьным процессором нужно отметить объект синим уголком.

Для выделения некоторой переменной в объекте нужно подвести к его концу курсор мыши и нажать левую кнопку мыши. Переменная будет отмечена синим уголком, расположенным сразу после переменной. Перемещая курсор с нажатой кнопкой <Пробел>, можно расширить черту, т. е. выделить отдельную часть выражения или выражение целиком.

Использование клавиши <Пробел>

Используйте клавишу <Пробел> для выделения части или всего математического выражения.
Выделение зависит от начального положения выделяющей линии (editing lines). Например:



Если заданная операция невыполнима, система выводит в дополнительном окне сообщение об ошибке или просто повторяет выражение. Последнее означает, что операция задана корректно, но результат не может быть получен, например, если делается попытка разложить на множители объект, уже разложенный или не содержащий такого разложения в принципе.

7.2. Выполнение символьных вычислений

Символьная операция **Evaluate Symbolically** (Вычислить) обеспечивает работу с математическими выражениями, содержащими встроенные в систему функции и представленные в различном виде: полиномиальном, дробно-рациональном, в виде сумм и произведений, производных и интегралов и т. д.:

$$\sum_{i=1}^{n} i^2 \quad \text{Evaluate Symbolically} \quad \frac{1}{3}\cdot(n+1)^3 - \frac{1}{2}\cdot(n+1)^2 + \frac{1}{6}\cdot n + \frac{1}{6}$$

$$\prod_{n=2}^{20}\left(1-\frac{1}{n^4}\right) \quad \text{Evaluate Symbolically} \quad \frac{212151878941598283455885}{230836979524030949228544}$$

$$\text{Evaluate Floating Point} \quad .91905499447722811477$$

Операция стремится произвести все возможные числовые вычисления и представить результаты в наиболее простом виде. Она возможна над матрицами с символьными элементами. Производные и определенные интегралы, символьные значения которых вычисляются, должны быть представлены в своей естественной форме.

Существует возможность выполнения числовых вычислений с повышенной точностью – 20 знаков после запятой. Для перехода в такой режим вычислений нужно числовые константы в вычисляемых объектах задавать с обязательным указанием десятичной точки, например: 5.0 или 8.0, а не 5 или 8.

7.3. Упрощение выражений

Символьная операция Simplify (Упростить) позволяет упрощать математические выражения, содержащие алгебраические и тригонометрические функции, а также выражения со степенными многочленами (полиномами). Выражение упрощается с выполнением таких операций, как сокращение подобных слагаемых, приведение к общему знаменателю, использование основных тригонометрических тождеств. Например:

$$\frac{\dfrac{y+x}{x}+\dfrac{x}{y-x}}{\dfrac{x}{y-x}} \quad \text{simplifies to} \quad \frac{y^2}{x^2}$$

$$\frac{x^2-5\cdot x-6}{x-6}+3\cdot x-6 \quad \text{simplifies to} \quad 4\cdot x-5$$

$$e^{3\cdot\ln(x)} \quad \text{simplifies to} \quad x^3$$

Эта команда открывает широкие возможности для упрощения сложных и плохо упорядоченных алгебраических выражений.

7.4. Расширение выражений

Действие команды **Expand** (Разложить по степеням) в известном смысле противоположно действию команды **Simplify**. Подвергаемое преобразованию выражение расширяется с использованием известных соотношений:

$$(a+b)^3 \text{ expand } \to a^3+3\cdot a^2\cdot b+3\cdot a\cdot b^2+b^3$$

Разумеется, расширение происходит только в том случае, когда его результат однозначно возможен. При преобразовании выражений команда **Expand** старается более простые функции представить через более сложные, свести алгебраические выражения, представленные в сжатом виде, к выражениям в развернутом виде.

7.5. Разложение выражений

Команда **Factor** (Разложить на множители) используется для факторизации — разложению выражений на простые множители:

$$a^3+3\cdot a^2\cdot b+3\cdot a\cdot b^2+b^3 \text{ factor } \to (a+b)^3$$

Она способствует раскрытию математической сущности выражений, например, наглядно выявляет представление полинома через его действительные корни. В большинстве случаев операция факторизации ведет к упрощению выражений.

7.6. Комплектование по выражениям

Команда **Collect** (Разложить по подвыражению) обеспечивает замену указанного выражения выражением, скомплектованным по базису указанной переменной, если такое представление возможно. Например:

$$x^2-a\cdot y\cdot x^2+2\cdot y^2\cdot x-x$$

by collecting terms, yields

$$(1-a\cdot y)\cdot x^2+\left(2\cdot y^2-1\right)\cdot x$$

Эта команда особенно удобна, когда заданное выражение есть функция ряда переменных и нужно представить его в виде функции заданной переменной, имеющей вид степенного многочлена. При этом другие переменные входят в сомножители указанной переменной, представленной в порядке уменьшения ее степени.

7.7. Вычисление коэффициентов полиномов

Команда Polynomial Coefficients (Полиномиальные коэффициенты) служит для вычисления коэффициентов полинома. Результатом операции является вектор с коэффициентами полинома. Например:

Click on x and choose
Polynomial Coefficients
from the **Symbolics** menu:

$$3\cdot b\cdot x^4-\pi\cdot x^2+\frac{2}{3}\cdot x-.3\cdot a\cdot b \quad \text{has coefficients} \quad \begin{pmatrix}-.3\cdot a\cdot b\\ \frac{2}{3}\\ -\pi\\ 0\\ 3\cdot b\end{pmatrix}$$

Операция полезна при решении задач полиномиальной аппроксимации и регрессии.

7.8. Решение уравнения относительно заданной переменной

Если задано некоторое выражение $F(x)$ и отмечена переменная x, то команда Solve (Решить относительно переменной) возвращает символьные значения указанной переменной x, при которых $F(x) = 0$. Это очень удобно при решении алгебраических уравнений, например квадратных и кубических, или нахождения корней полинома четвертой степени и др.

$$a\cdot x^2+b\cdot x+c \quad \text{solves} \quad \begin{bmatrix}\frac{1}{(2\cdot a)}\cdot\left[-b+\left(b^2-4\cdot a\cdot c\right)^{\left(\frac{1}{2}\right)}\right]\\ \frac{1}{(2\cdot a)}\cdot\left[-b-\left(b^2-4\cdot a\cdot c\right)^{\left(\frac{1}{2}\right)}\right]\end{bmatrix}$$

$$2\cdot x^2+4\cdot x+1 \quad \text{solves} \quad \begin{pmatrix}-1+\frac{1}{2}\cdot\sqrt{2}\\ -1-\frac{1}{2}\cdot\sqrt{2}\end{pmatrix}$$

floating point

$$\begin{pmatrix}-.29289321881345247560\\ -1.7071067811865475244\end{pmatrix}$$

Рис. 7.1. Примеры решения уравнений

Рис. 7.1 и рис. 7.2 содержат примеры решения уравнений в символьном, а также в численном виде с использованием команды **Floating Point**.

$$x^3 + x + 1 \quad \text{solves}$$

$$\begin{bmatrix} \frac{-1}{6}\cdot\left(108+12\cdot\sqrt{93}\right)^{\left(\frac{1}{3}\right)} + \frac{2}{\left(108+12\cdot\sqrt{93}\right)^{\left(\frac{1}{3}\right)}} \\ \frac{1}{12}\cdot\left(108+12\cdot\sqrt{93}\right)^{\left(\frac{1}{3}\right)} - \frac{1}{\left(108+12\cdot\sqrt{93}\right)^{\left(\frac{1}{3}\right)}} + \frac{1}{2}\cdot i\cdot\sqrt{3}\cdot\left[\frac{-1}{6}\cdot\left(108+12\cdot\sqrt{93}\right)^{\left(\frac{1}{3}\right)} - \frac{2}{\left(108+12\cdot\sqrt{93}\right)^{\left(\frac{1}{3}\right)}}\right] \\ \frac{1}{12}\cdot\left(108+12\cdot\sqrt{93}\right)^{\left(\frac{1}{3}\right)} - \frac{1}{\left(108+12\cdot\sqrt{93}\right)^{\left(\frac{1}{3}\right)}} - \frac{1}{2}\cdot i\cdot\sqrt{3}\cdot\left[\frac{-1}{6}\cdot\left(108+12\cdot\sqrt{93}\right)^{\left(\frac{1}{3}\right)} - \frac{2}{\left(108+12\cdot\sqrt{93}\right)^{\left(\frac{1}{3}\right)}}\right] \end{bmatrix}$$

$$\text{floating point}$$

$$\begin{pmatrix} -.68232780382801932744 \\ .34116390191400966368 - 1.1615413999972519362\cdot i \\ .34116390191400966368 + 1.1615413999972519362\cdot i \end{pmatrix}$$

Рис. 7.2. Пример решения уравнения

Усложнение уравнения, например переход от квадратного уравнения к кубическому, может вызвать существенное усложнение результата. Тогда система представляет решение в более компактном виде и предлагает занести его в буфер обмена. С помощью команды **Paste** (Вставить) можно перенести решение в основное окно системы, но оно будет иметь уже тип текстового комментария, а не математического выражения, пригодного для дальнейших преобразований.

7.9. Подстановка для заданной переменной

Команда **Substitute** (Заменить переменную) возвращает новое выражение, полученное путем подстановки на место указанной переменной некоторого другого выражения. Последнее должно быть подготовлено и скопировано в буфер обмена. Наряду с получением результата в символьном виде, эта команда позволяет найти и числовые значения функции некоторой переменной путем замены ее на числовое значение. На рис. 7.3 представлены примеры операций с подстановкой.

Выражение для подстаиовки	$z + b$	Copy
Исходное выраженне	$x^4 + 3 \cdot x^3 + 5$ by substitution	
Выраженне после замены		

$$\left[(z+b)^4 + 3\cdot(z+b)^3 + 5\right]^4 + 3\cdot\left[(z+b)^4 + 3\cdot(z+b)^3 + 5\right]^3 + 5$$

Число для подстановки	3	Copy
Исходное выражение	$x^4 + 3 \cdot x^3 + 5$	

by substitution

167 Выражение после замены

Чнсло для подстановки	–.68232780382801932734
Исходное выраженне	$x^3 + x + 1$

by substitution

Выраженне после замены $7. \cdot 10^{-20}$

Рис. 7.3. Примеры операций с подстановкой

7.10. Дифференцирование

Дифференцирование может выполняться с помощью разных команд: **Evaluete** (Вычислить), **Simplify** (Упростить), **Differentiate** из подменю **Variable** (Дифференцировать по переменной).

Примеры символьного дифференцирования с применением команд **Evaluete** и **Simplify** приведены на рис. 7.4. Основные правила использования этих команд для дифференцирования следующие: производные должны быть представлены в своем естественном виде, выражение должно быть выделено целиком.

Дифференцирование – Evaluate Symbolically

$$\frac{d}{dx}\ln(x) \quad \text{Evaluate Symbolically} \quad \frac{1}{x}$$

$$\frac{d^2}{dx^2}\ln(x) \quad \text{Evaluate Symbolically} \quad \frac{-1}{x^2}$$

$$\frac{d^3}{dx^3}\ln(x) \quad \text{Evaluate Symbolically} \quad \frac{2}{x^3}$$

Дифференцирование – Simplify

$$\frac{d}{dx}\frac{1}{x\cdot\ln(x)} \quad \text{Simplifies to} \quad \frac{-(\ln(x)+1)}{\left(x^2\cdot\ln(x)^2\right)}$$

$$\frac{d^2}{dx^2}\frac{1}{x\cdot\ln(x)} \quad \text{Simplifies to} \quad \frac{\left(2\cdot\ln(x)^2+3\cdot\ln(x)+2\right)}{\left(x^3\cdot\ln(x)^3\right)}$$

$$\frac{1}{x\cdot\ln(x)} \quad \text{Differentiate} \quad \frac{-1}{\left(x^2\cdot\ln(x)\right)}-\frac{1}{\left(x^2\cdot\ln(x)^2\right)}$$

$$\frac{-1}{\left(x^2\cdot\ln(x)\right)}-\frac{1}{\left(x^2\cdot\ln(x)^2\right)} \quad \text{Simplifies to} \quad \frac{-(\ln(x)+1)}{\left(x^2\cdot\ln(x)^2\right)}$$

Рис. 7.4. Примеры символьного дифференцирования с применением команд ***Evaluate*** *и* ***Simplify***

Дифференцирование с применением **Differentiate** из подменю **Variable** выполняется с выражениями, требующими указания переменной, по отношению к которой выполняется операция. Для указания переменной достаточно установить справа от нее маркер ввода. Само выражение при этом не выделяется отдельно, поскольку указание в нем на переменную является одновременно и указанием на само выражение. Если выражение содержит другие переменные, то они рассматриваются как константы. Для вычисления производных высшего порядка нужно повторить вычисление необходимое количество раз. На рис. 7.5 показано применение команды дифференцирования.

Дифференцирование – Differentiate on Variable

Alternatively, click on the variable of differentiationx and choose **Differentiate** under **Variable** in the **Symbolics** menu.

$a^x \cdot \ln(x)$ by differentiation, yields $a^x \cdot \ln(a) \cdot \ln(x) + \dfrac{a^x}{x}$

by differentiation, yields

$$a^x \cdot \ln(a)^2 \cdot \ln(x) + 2 \cdot a^x \cdot \frac{\ln(a)}{x} - \frac{a^x}{x^2}$$

$u(x) \cdot v(x)$ by differentiation, yields $\left(\dfrac{d}{dx}u(x)\right) \cdot v(x) + u(x) \cdot \dfrac{d}{dx}v(x)$

$\left(\dfrac{d}{dx}u(x)\right) \cdot v(x) + u(x) \cdot \dfrac{d}{dx}v(x)$ by differentiation, yields

$$\left(\frac{d^2}{dx^2}u(x)\right) \cdot v(x) + 2 \cdot \left(\frac{d}{dx}u(x)\right) \cdot \frac{d}{dx}v(x) + u(x) \cdot \frac{d^2}{dx^2}v(x)$$

$\dfrac{u(x)}{v(x)}$ by differentiation, yields $\dfrac{\frac{d}{dx}u(x)}{v(x)} - \dfrac{u(x)}{v(x)^2} \cdot \dfrac{d}{dx}v(x)$

simplifies to

$$\frac{\left[\left(\frac{d}{dx}u(x)\right) \cdot v(x) - u(x) \cdot \frac{d}{dx}v(x)\right]}{v(x)^2}$$

$\dfrac{\left[\left(\frac{d}{dx}u(x)\right) \cdot v(x) - u(x) \cdot \frac{d}{dx}v(x)\right]}{v(x)^2}$ by differentiation, yields

$$\frac{\left(\frac{d^2}{dx^2}u(x)\right) \cdot v(x) - u(x) \cdot \frac{d^2}{dx^2}v(x)}{v(x)^2} - 2 \cdot \frac{\left[\left(\frac{d}{dx}u(x)\right) \cdot v(x) - u(x) \cdot \frac{d}{dx}v(x)\right]}{v(x)^3} \cdot \frac{d}{dx}v(x)$$

*Рис. 7.5. Примеры символьного дифференцирования с применением команды **Differentiate***

7.11. Расширение возможностей символьных преобразований

Инструменты символьной математики системы Mathcad разделяются на три группы:

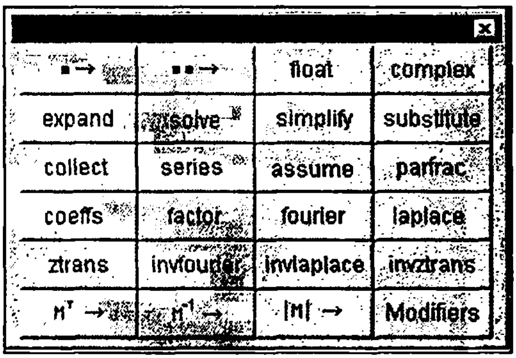
1. Команды символьной математики из меню **Symbolics**.
2. Режим непрерывных символьных преобразований (Live Symbolics).
3. Оптимизация численных выкладок через символьные преобразования (Optimize).

Первые мы уже в большинстве рассмотрели (интегрирование рассмотрим позже). Две других возможности символьных преобразований выполняются с помощью системы SmartMath, встроенной в систему Mathcad и отвечающей за режим автоматических символьных преобразований и за оптимизацию.

7.12. Режим автоматических символьных преобразований

Режим SmartMath позволяет работать со средствами символьной математики не только в ручном, но и в автоматическом режиме. В среде Mathcad знак = означает числовой, а знак → – символьный вывод значений переменной, функции, выражения.

Режим автоматических символьных преобразований включается с помощью панели **Symbolics keyword palette**.



Система SmartMath более полно использует ядро символьных операций, чем символьные вычисления на их выполнение. Например, возможно использование в преобразуемых выражениях функций пользователя. Еще важнее то, что результаты символьных преобразований в системе SmartMath автоматически меняются при изменении исходных символьных данных. Этого не происходит при символьных вычислениях с помощью команд **Symbolics** главного меню.

Преобразования выполняются с использованием соответствующих команд меню **Symbolics keyword palette**. Порядок применения команд приведен ниже. Лишние шаблоны удаляются.

$$(x+y)^3$$

$$(x+y)^3 \; \blacksquare \rightarrow$$

$$(x+y)^3 \text{ expand} \rightarrow$$

Press Enter

$$(x+y)^3 \text{ expand} \rightarrow x^3 + 3 \cdot x^2 \cdot y + 3 \cdot x \cdot y^2 + y^3$$

Дифференцирование:

$$\frac{d}{dx} x^3 \cdot \sin(x)$$

$$\frac{d}{dx} x^3 \cdot \sin(x) \rightarrow$$

Enter

$$\frac{d}{dx} x^3 \cdot \sin(x) \rightarrow 3 \cdot x^2 \cdot \sin(x) + x^3 \cdot \cos(x)$$

Примеры использования различных команд приведены на рис. 7.6.

$$\frac{a^2-b^2}{a+b}\ \text{simplify}\ \rightarrow a-b$$

$$\ln(x)\ \text{series}\ ,x=2\ \rightarrow \ln(2)+\frac{1}{2}\cdot x-1-\frac{1}{8}\cdot(x-2)^2+\frac{1}{24}\cdot(x-2)^3-\frac{1}{64}\cdot(x-2)^4+\frac{1}{160}\cdot(x-2)^5$$

Взятие производной

$$CD := a_0\cdot\ln\left(x^{a_1}\right)$$

$$\frac{d}{dx}CD \rightarrow a_0\cdot\frac{a_1}{x}$$

$$a^4-b^4\ \text{factor}\ \rightarrow (a-b)\cdot(a+b)\cdot\left(a^2+b^2\right)$$

$$(a-b)\cdot(a+b)\ \text{expand}\ \rightarrow a^2-b^2$$

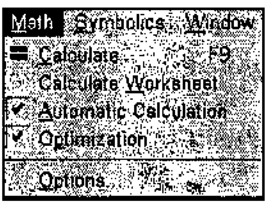
$$b^4\ \text{expand}\ \rightarrow b^4\ \text{series}\ ,b=a^2\ \rightarrow a^8+4\cdot a^6\cdot\left(b-a^2\right)+6\cdot a^4\cdot\left(b-a^2\right)^2+4\cdot a^2\cdot\left(b-a^2\right)^3+\left(b-a^2\right)^4$$

$$\frac{a\cdot(y-c)^2}{(y+b)^2}+\frac{4\cdot a\cdot(y-c)}{(y+b)^2}+\frac{4\cdot a}{(y+b)^2}\ \left|\begin{array}{l}\text{substitute},(y+b)=x,(y-c)=z\\ \text{factor}\end{array}\right.\ \rightarrow a\cdot\frac{(z+2)^2}{x^2}$$

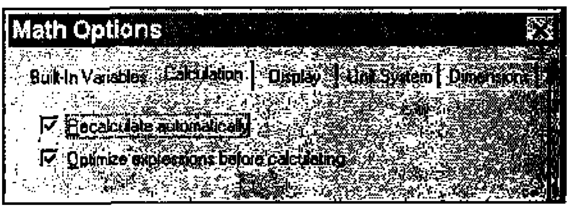
Рис. 7.6. Примеры символьных преобразований в автоматическом режиме

7.13. Совместная работа символьного и численного процессоров

Второй режим системы SmartMath позволяет проводить оптимизацию численных выкладок через символьные преобразования. При этом символьный процессор пытается упростить выражения, которые вычисляются численным процессором, разумеется, если такое упрощение возможно. Для реализации этой возможности необходимо использовать команду **Optimization** из меню **Math**.



или



При запуске система SmartMath пометит выражение красной звездочкой.

Глава 8. Интегральное исчисление

8.1. Неопределенный интеграл

Интегрирование есть действие, обратное вычислению производных. Интегральное исчисление решает задачу нахождения функции по ее производной, т. е. нахождения функции, производная которой равняется данной функции [5, 18, 32].

Функция $F(x)$, производная которой равняется данной функции $f(x)$, $F'(x) = f(x)$, называется первообразной функцией функции $f(x)$.

Совокупность всех первообразных для функции $f(x)$ во всех точках области определения функции $f(x)$ называется неопределенным интегралом от функции $f(x)$ и обозначается $\int f(x)dx$, где $f(x)$ – подынтегральная функция, $f(x)dx$ – подынтегральное выражение. Таким образом,

$$\int f(x)dx = F(x)+C,$$

где $F(x)$ – некоторая первообразная для $f(x)$, C – произвольная постоянная.

Основные свойства неопределенного интеграла следующие:

1. Производная от неопределенного интеграла равна подынтегральной функции:

$$\left(\int f(x)dx\right)' = f(x).$$

2. Дифференциал неопределенного интеграла равен подынтегральному выражению:

$$d\left(\int f(x)dx\right) = f(x)dx.$$

3. Неопределенный интеграл от дифференциала некоторой функции равен этой функции с точностью до постоянного слагаемого:

$$\int dF(x) = F(x)+C.$$

4. Постоянный множитель можно выносить за знак интеграла:

$$\int af(x)dx = a\int f(x)dx.$$

5. Интеграл от алгебраической суммы двух функций равен сумме интегралов от этих функций:

$$\int (f(x) \pm g(x)) = \int f(x)dx \pm \int g(x)dx.$$

Рассмотрим интегралы от элементарных функций.

$$\int dx = x + C,$$

$$\int xdx = \frac{x^2}{2} + C,$$

$$\int \frac{1}{x}dx = \ln|x| + C,$$

$$\int x^n dx = \frac{x^{n+1}}{n+1} + C,\ n \neq -1,$$

$$\int a^x dx = \frac{a^x}{\ln a} + C,\ a > 0,\ a \neq 1,$$

$$\int e^x dx = e^x + C,$$

$$\int e^{mx} dx = \frac{e^{mx}}{m} + C.$$

Неопределенный интеграл с логарифмической производной.
Известно, что

$$(\ln(f(x)))' = \frac{f'(x)}{f(x)}.$$

Из свойств неопределенного интеграла следует, что

$$\int \frac{f'(x)}{f(x)} dx = \ln|f(x)| + C.$$

Эта формула облегчает расчет неопределенного интеграла во всех тех случаях, когда подынтегральная функция есть дробь, числитель которой – производная знаменателя. Рассмотрим пример. Рассчитать интеграл

$$\int \frac{dx}{ax+b}.$$

Преобразуем подынтегральную функцию таким образом, чтобы в ее числителе оказалась производная знаменателя. Для этого достаточно числитель и знаменатель умножить на постоянную. Получаем:

$$\int \frac{a}{a(ax+b)} dx = \frac{1}{a}\int \frac{a}{ax+b} dx.$$

Поскольку числитель есть производная знаменателя, то можно записать:

$$\frac{1}{a}\int \frac{a}{ax+b} dx = \frac{1}{a}\ln|ax+b| + C.$$

Интегрирование по частям. Если $u(x)$ и $v(x)$ – функции, имеющие соответственно производные $u'(x)$ и $v'(x)$, то

$$(uv)' = u'v + uv',$$

и, следовательно,

$$uv' = (uv)' - vu'.$$

Интегрируя обе части, получаем:

$$\int uv'dx = \int (uv)'dx - \int vu'dx = uv - \int vu'dx + C.$$

Формула

$$\int uv'dx = uv - \int vu'dx + C$$

называется формулой интегрирования по частям. Она позволяет вычислить более трудный интеграл $\int uv'dx$, если мы можем рассчитать $\int vu'dx$.

Интегрирование подстановкой. Нередко интеграл $\int f(x)dx$ можно упростить, введя новую переменную t и приняв $x = g(t)$.

Если $F(x)$ – первообразная функция, то по теореме о производной сложной функции имеем:

$$F'(x) = F'(g(t)) \cdot g'(t)dt + C$$

и, следовательно,

$$F(x) = \int F'(g(t))g'(t)dt + C = \int f(g(t))g'(t)dt + C.$$

Данная формула называется формулой интегрирования подстановкой. Она показывает, как преобразуется интеграл, если вводится новая переменная.

Пример. Найти интеграл

$$\int (ax+b)^m dx \quad (m \neq -1).$$

Вводим новую переменную $ax + b = t$, откуда

$$ax = t - b,\ x = (t - b)/a,\ x' = 1/a.$$

Тогда

$$\int (ax+b)^m dx = \int t^m \frac{1}{a} dt = \frac{1}{a}\int t^m dt = \frac{1}{a}\frac{t^{m+1}}{m+1} + C = \frac{1}{a}\frac{(ax+b)^{m+1}}{m+1} + C.$$

8.2. Использование неопределенного интеграла в экономике

В главе 3 рассматривались различные предельные функции – предельные затраты, предельный доход и др. Предельные функции могут быть получены из соответствующих экономических функций посредством дифференцирования. Интегрирование позволяет решить обратную задачу: найти исходную экономическую функцию по известной предельной функции.

В одном из примеров в предельном анализе функция предельных затрат предприятия имеет вид (x – объем производства):

$$C'(x) = 0{,}04x + 15.$$

Требуется определить суммарные затраты при условии, что постоянные затраты равны 800.

Для решения задачи вспомним определение предельных затрат:

$$C'(x) = \frac{dC}{dx} = \frac{d}{dx}F(x),$$

где $C = F(x)$ и есть искомая функция суммарных затрат. Установим ее:

$$F(x) = \int f(x)dx = \int (0{,}04x + 15)dx = 0{,}02x^2 + 15x + C,$$

где C – постоянная интегрирования.

Разумеется, затраты не могут быть неопределенными. Вспомним, что постоянные затраты, т. е. не зависящие от объема производства, по условию задачи равны 800. Другими словами, если ничего не производится, то $x = 0$, а постоянные затраты все равно равны 800. Поэтому можно записать:

$$F(0) = 0{,}02 \cdot 0^2 + 15 \cdot 0 + C = 800,$$

откуда сразу находим постоянную интегрирования $C = 800$ и окончательный результат:

$$F(x) = 0{,}02x^2 + 15x + 800.$$

Следовательно, постоянная интегрирования равна постоянным издержкам. Этот пример типичен: любая прикладная задача всегда содержит условие, позволяющее однозначно определить постоянную интегрирования.

В следующем примере нужно найти суммарный доход R, если известен предельный доход:

$$R'(x) = 50 - 0{,}1x^2.$$

Поскольку предельный доход определяется как

$$R'(x) = \frac{dR}{dx} = \frac{d}{dx}F(x),$$

где $F(x)$ – искомая функция дохода, то

$$F(x) = \int f(x)dx = \int (50 - 0{,}1x^2)dx = 50x - 0{,}1x^2 + C,$$

где C – постоянная интегрирования.

При нулевом производстве и доход будет нулевым. Это дает

$$F(0) = 50 \cdot 0 - 0{,}1 \cdot 0^2 + C = 0,$$

т. е. постоянная интегрирования равна нулю. Окончательно суммарный доход равен:

$$R = 50x - 0{,}1x^2.$$

Для данного предприятия с указанными предельными затратами и предельным доходом предельная прибыль будет равна:

$$PR'(x) = 35 - 0{,}24x.$$

Установим суммарную прибыль:

$$F(x) = \int (35 - 0{,}24x)dx = 35x - 0{,}12x^2 + C.$$

При нулевом объеме производства прибыль будет в размере убытков от постоянных затрат:

$$F(0) = 35 \cdot 0 - 0{,}12 \cdot 0^2 + C = -800.$$

Окончательно суммарная прибыль будет равна:

$$PR = 35x - 0{,}12x^2 - 800.$$

8.3. Определенный интеграл

Понятие интегральной суммы. Пусть на $[a, b]$ задана функция $y = f(x)$. Разобьем отрезок $[a, b]$ на n элементарных отрезков точками $x_0, x_1, \ldots, x_n$: $a = x_0 < x_1 < x_2 < \ldots < x_n = b$. На каждом отрезке $[x_{i-1}, x_i]$ разбиения выберем некоторую точку ξ_i и положим $\Delta x_i = x_i - x_{i-1}$, где $i = 1, 2, \ldots, n$. Сумму вида

$$\sum_{i=1}^{n} f(\xi_i)\Delta x_i$$

будем называть интегральной суммой для функции $y = f(x)$ на $[a, b]$.

Геометрический смысл интегральной суммы. Отдельное слагаемое $f(\xi_i)\Delta x_i$ интегральной суммы равно площади S_i прямоугольника со сторонами $f(\xi_i)$ и Δx_i, где $i = 1, 2, \ldots, n$; $x_1 - x_0 = \Delta x_1$, $x_2 - x_1 = \Delta x_2$ и т. д. Другими словами, S_i – это площадь под прямой $y = f(x)$ на отрезке $[x_{i-1}, x_i]$. Поэтому вся интегральная сумма равна площади $S_n = S_1 + S_2 + ... + S_n$ под ломаной, образованной на каждом из отрезков $[x_{i-1}, x_i]$ прямой $y = f(\xi)$, параллельной оси абсцисс.

Понятие определенного интеграла. Предел интегральной суммы называется определенным интегралом от функции $y = f(x)$ на $[a, b]$:

$$\int_a^b f(x)dx = \lim_{\max\Delta x_i \to 0} \sum_{i=1}^{n} f(\xi_i)\Delta x_i.$$

При этом число a называется нижним пределом, число b – верхним пределом.

Несмотря на сходство в обозначениях и терминологии, определенный и неопределенный интегралы существенно различаются: в то время как $\int f(x)dx$ представляет семейство функций, $\int_a^b f(x)dx$ есть определенное число.

Геометрический смысл определенного интеграла. Понятие определенного интеграла введено таким образом, что в случае, когда функция $y = f(x)$ неотрицательна на отрезке $[a, b]$, где $a < b$, $\int_a^b f(x)dx$ численно равен площади S под кривой $y = f(x)$.

Рассмотрим свойства определенного интеграла.

1. Постоянный множитель можно выносить за знак интеграла.

2. Во введенном определении определениого интеграла $\int_a^b f(x)dx$ предполагается, что $a < b$. Тогда по определению

$$\int_a^b f(x)dx = -\int_b^a f(x)dx.$$

3. Интеграл от алгебраической суммы двух функций равен такой же сумме интегралов от этих функций, т. е.

$$\int_a^b (f(x) \pm g(x))dx = \int_a^b f(x)dx + \int_a^b g(x)dx.$$

4. Если отрезок интегрирования разбит на части, то интеграл на всем отрезке равен сумме интегралов для каждой из возникших частей, т. е. при любых a, b, c

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx.$$

5. Если на отрезке $[a, b]$, где $a < b, f(x) \le g(x)$, то и

$$\int_a^b f(x)dx \le \int_a^b g(x)dx,$$

т. е. обе части неравенства можно почленно интегрировать.

Определенный интеграл как функция верхнего предела. Если функция $y = f(x)$ интегрируема на отрезке $[a, b]$, то она интегрируема также на произвольном отрезке $[a, x]$, вложенном в $[a, b]$:

$$\Phi(x) = \int_a^x f(x)dx = \int_a^x f(t)dt,$$

где $x \in [a, b]$, а функция $\Phi(x)$ называется интегралом с переменным верхним пределом.

Формула Ньютона–Лейбница. Пусть функция $y = f(x)$ непрерывна на отрезке $[a, b]$ и $F(x)$ – любая первообразная для $f(x)$ на $[a, b]$. Тогда определенный интеграл от функции $f(x)$ на $[a, b]$ равен приращению первообразной $F(x)$ на этом отрезке, т. е.

$$\int_a^b f(x)dx = F(b) - F(a).$$

Нахождение определенных интегралов с использованием формулы Ньютона–Лейбница осуществляется в два шага: на первом шаге, используя технику нахождения неопределенного интеграла, находят некоторую первообразную $F(x)$ для подынтегральной функции $f(x)$; на втором применяется собственно формула Ньютона–Лейбница – находится приращение первообразной, равное искомому интегралу. В связи с этим вводится обозначение для приращения первообразной, которое удобно использовать при записи решений:

$$F(x)\Big|_b^a = F(b) - F(a).$$

При применении формулы Ньютона–Лейбница можно использовать первообразную $F(x)$ для подынтегральной функции $f(x)$, имеющую наиболее простой вид, например, при $C = 0$.

Вычислим $\int_2^4 x^3 dx$. Первообразная для функции $f(x) = x^3$ имеет вид $F(x) = \frac{x^4}{4} + C$. Для нахождения интеграла по формуле Ньютона–Лейбница возьмем первообразную, у которой $C = 0$. Тогда

$$\int_2^4 x^3 dx = \frac{x^4}{4}\Big|_2^4 = \frac{4^4}{4} - \frac{2^4}{4} = 60.$$

Несобственные интегралы. Ранее мы рассматривали интегралы от функций, ограниченных на конечных отрезках интегрирования. На практике возникает необходимость обобщения этих понятий на случаи, когда либо один из концов (или оба) отрезка интегрирования удален в бесконечность, либо функция не ограничена на отрезке интегрирования. Несобственным интегралом $\int_a^{\infty} f(x)dx$ от функции $f(x)$ на полуинтервале $[a, \infty]$ называется предел функции $\int_a^t f(t)dx$ при t, стремящемся к ∞, т. е.

$$\int_a^{+\infty} f(x)dx = \lim_{t \to +\infty} \int_a^t f(x)dx.$$

Если предел, стоящий в правой части равенства существует и конечен, то несобственный интеграл называется сходящимся (к данному пределу), в противном случае – расходящимся.

8.4. Использование определенного интеграла в экономике

В большинстве практических экономических задач операция вычисления определенного интеграла подменяется вычислением интегральной суммы, что связывается с дискретностью экономических показателей, а также большей доступностью вычислений. Однако при моделировании ситуация меняется на обратную. Рассмотрим несколько примеров применения определенного интеграла.

Необходимо определить суммарный выпуск продукции (Q) за ряд лет (T) по функции выпуска от времени t^n. Тогда суммарный выпуск будет равен:

$$Q = \int_0^T t^n dt = \left.\frac{t^{n+1}}{n+1}\right|_0^T .$$

Например, выпуск за 10 лет при функции $t^{0,9}$ составит:

$$Q = \int_0^{10} t^{0,9} dt = \frac{10^{1,9}}{1,9} = 41,806 .$$

Затраты имеют функцию $a + bQ$. Суммарные затраты (C) за ряд лет будут равны:

$$C = \int_0^T (a + bQ)dt = \int_0^T (a + bt^n)dt = at + b\left.\frac{t^{n+1}}{(n+1)}\right|_0^T .$$

Если параметры a и b равны соответственно 3 и 2, то суммарные затраты за 10 лет будут равны 113,613.

Пусть функция спроса равна P. Тогда суммарный доход (D) за ряд лет равен:

$$D = \int_0^T PQdt = \int_0^T Pt^n dt = P\left.\frac{t^{n+1}}{(n+1)}\right|_0^T .$$

Если $P = 5$, то суммарный доход за 10 лет равен 209,034.

Суммарная прибыль (PR) за ряд лет при исходных функциях выпуска, спроса и затрат будет равна:

$$PR = \int_0^T Pt^n - (a + bt^n) = P\frac{t^{n+1}}{(n+1)} - at - b\left.\frac{t^{n+1}}{n+1}\right|_0^T .$$

При исходных данных значениях параметров суммарная прибыль за 10 лет равна 95,418.

Рассмотрим применение определенного интеграла в области финансовых расчетов. Определим величину дисконтированного дохода, полученного за t лет, если годовой доход $D = f(t)$ есть функция времени, норма процента i, проценты начисляются непрерывно.

Для вычисления данной величины, разделим отрезок времени в t лет на n равных отрезков времени Δt. В весьма малом отрезке времени Δt доход можно считать неизменным и следовательно равным $f(t)\,\Delta t$.

Конечная величина дисконтированного дохода D_t начальной суммы D через t лет при ставке процента i и начислении процентов m раз в году составляет:

$$D_t = D(1 + i/m)^{-mt}.$$

Сумма дисконтированного дохода за t лет составит:

$$D_{sum} = \sum_{t=0}^{T} D_t (1+\frac{i}{m})^{-mt}.$$

При непрерывном начислении процентов при $m \to \infty$ выражение:

$$\lim_{m\to\infty}(1+\frac{1}{m})^{mt} = \lim_{m\to\infty}(1+\frac{i}{m/i})^{\frac{m}{i}it} = e^{it}.$$

Объем дисконтированного дохода при непрерывном начислении процентов составит:

$$\lim_{\Delta t\to 0}\sum_{t=0}^{T} f(t)e^{-it}\Delta t = \int_0^t f(t)e^{-it}dt.$$

Если годовой доход постоянен, $f(t) = R$, то объем дисконтированного дохода будет равен:

$$D_{sum} = \int_0^t \mathrm{Re}^{-it}dt = R\int_0^t e^{-it}dt = R\left[-\frac{1}{i}e^{-it}\right]_0^t = R\left(-\frac{1}{i}e^{-it} + \frac{1}{i}e^{0}\right) = \frac{R}{i}(1-e^{-it}).$$

Теперь рассмотрим вычисление одного из основных показателей эффективности инвестиций – чистого приведенного дохода (NPV). Данный показатель устанавливается как разница суммы дисконтированного (приведенного) дохода (D_t) за время эксплуатации предприятия ($T - T_c$) и суммы дисконтированных инвестиций (K_t) за время строительства предприятия (T_c) при ежегодном начислении процентов:

$$NPV = \sum_{t=Tc}^{T} D_t(1+i)^{-t} - \sum_{t=1}^{T_c} K_t(1+i)^{-t},$$

где $D_t = (B - C) - H + S$; B – объем выручки от продажи продукции; C – текущие расходы без амортизационных отчислений; H – налоговые отчисления; S – различного рода компенсации; T_c – продолжительность периода инвестиций или строительства объекта; T – продолжительность периода отдачи от инвестиций; i – процентная ставка; K_t – инвестиционные расходы в периоде $t_{стр}$.

Расчетное выражение NPV в случае равномерного потока дохода и инвестиций будет равно:

$$NPV = D_t(\frac{1-(1+i)^{-T}}{i}) - D_t(\frac{1-(1+i)^{-Tc}}{i}) - K_t(\frac{1-(1+i)^{-Tc}}{i})$$

или

$$NPV = D_t\frac{(1+i)^{-Tc}-(1+i)^{-T}}{i} - K_t\frac{1-(1+i)^{-T}}{i}.$$

Если доход будет функцией времени $D = f(t)$ и инвестиции $K = h(t)$, то при непрерывном начислении процентов NPV будет равен:

$$NPV = \int_{Tc}^{T} f(t)e^{-it}dt - \int_0^{Tc} h(t)e^{-it}dt.$$

Если поток дохода и инвестиций будет равномерным, т. е. $D_t = D$, $K_t = K$, то NPV будет равен:

$$NPV=\int_{Tc}^{T} De^{-it}dt-\int_{0}^{Tc} Ke^{-it}dt=\frac{D}{i}(1-e^{-it})\Bigg|_{Tc}^{T}-\frac{K}{i}(1-e^{-it})\Bigg|_{0}^{Tc}.$$

Откуда можно записать:

$$NPV=\frac{D}{i}(e^{-iTc}-e^{-iT})-\frac{K}{i}(1-e^{-iTc}).$$

Например, если D = 10 млн. руб., i = 0,1, T = 20 лет, T_c = 5 лет, K = 4 млн. руб., то:

$$NPV = 10/0{,}1(e^{-0{,}1\cdot5}-e^{-0{,}1\cdot20})-4/0{,}1(1-e^{-0{,}1\cdot5}) = 31{,}381 \text{ млн. руб.}$$

8.5. Интегрирование в Mathcad

Интегрирование в Mathcad может производиться с помощью разных команд. Во-первых, с использованием команд из меню **Symbolics – Evaluate** и **Simplify**. Примеры вычисления интегралов с помощью этих команд приведены на рис. 8.1 и 8.2.

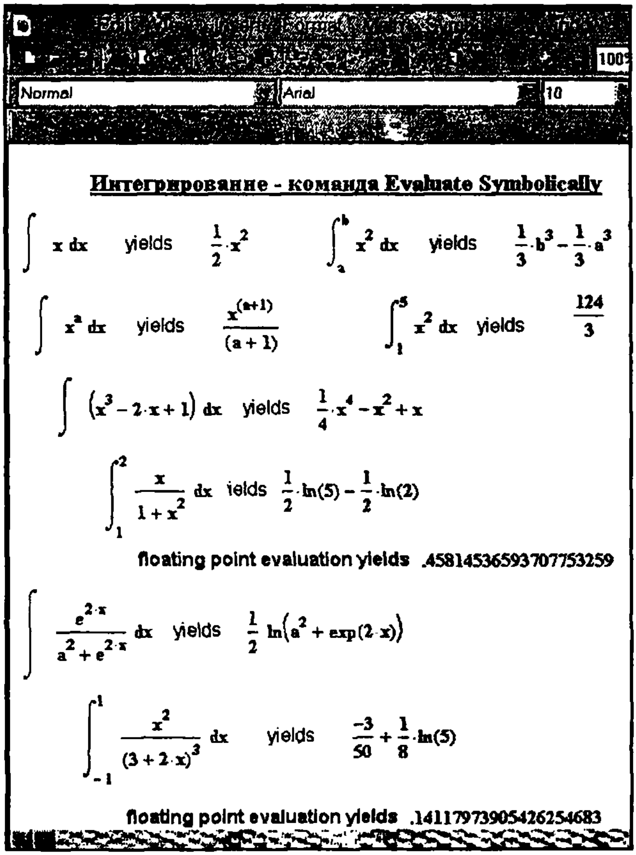


Рис. 8.1. Вычисление интегралов с помощью команды Evaluate

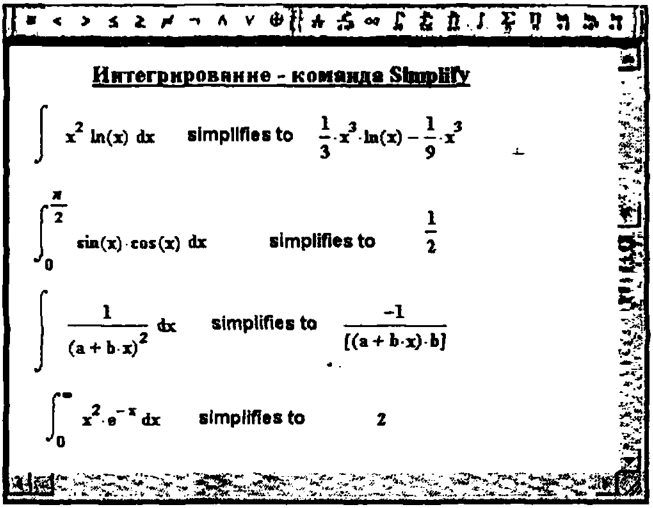


Рис. 8.2. Вычисление интегралов с помощью команды ***Simplify***

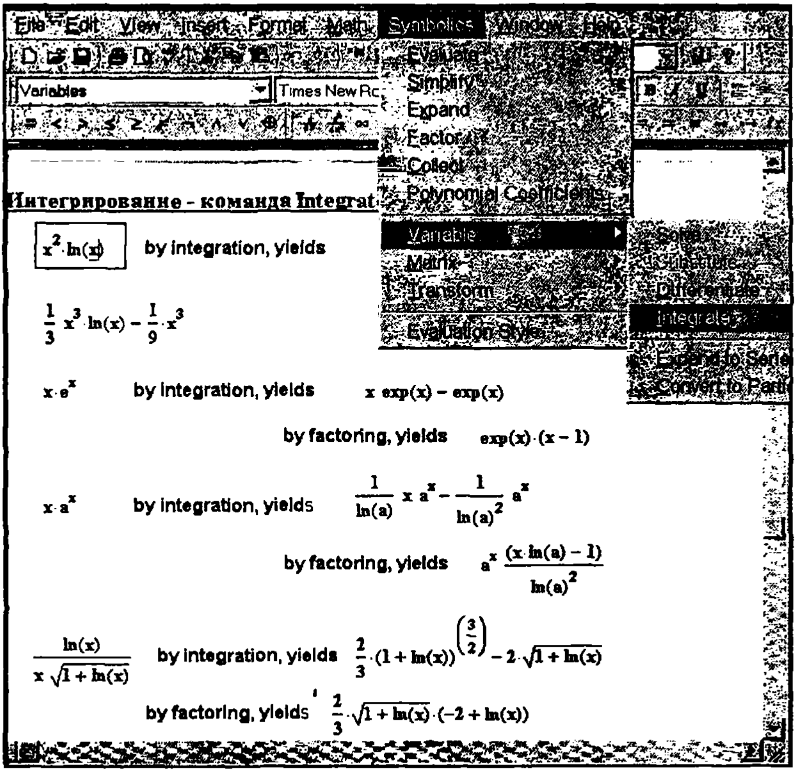


Рис. 8.3. Примеры вычисления интегралов с помощью команды ***Integrate***

Другая возможность вычисления интегралов – это использование команды **Integrate on Variable**. Она возвращает символьное значение неопределенного интеграла по указанной маркером ввода переменной. Выражение, в состав которого входит переменная, является подынтегральной функцией. На рис. 8.3 показаны примеры символьного интегрирования по переменной.

Кроме того, как мы уже видели, система SmartMath имеет возможность одновременной работы символьного и численного процессоров. Символьный процессор упрощает выражения, которые вычисляются численным процессором. Эта возможность резко уменьшает время вычислений. Реализуется данная возможность командой **Optimization** в меню **Math.** Примеры интегрирования с помощью команды **Optimization** приведены на рис. 8.4.

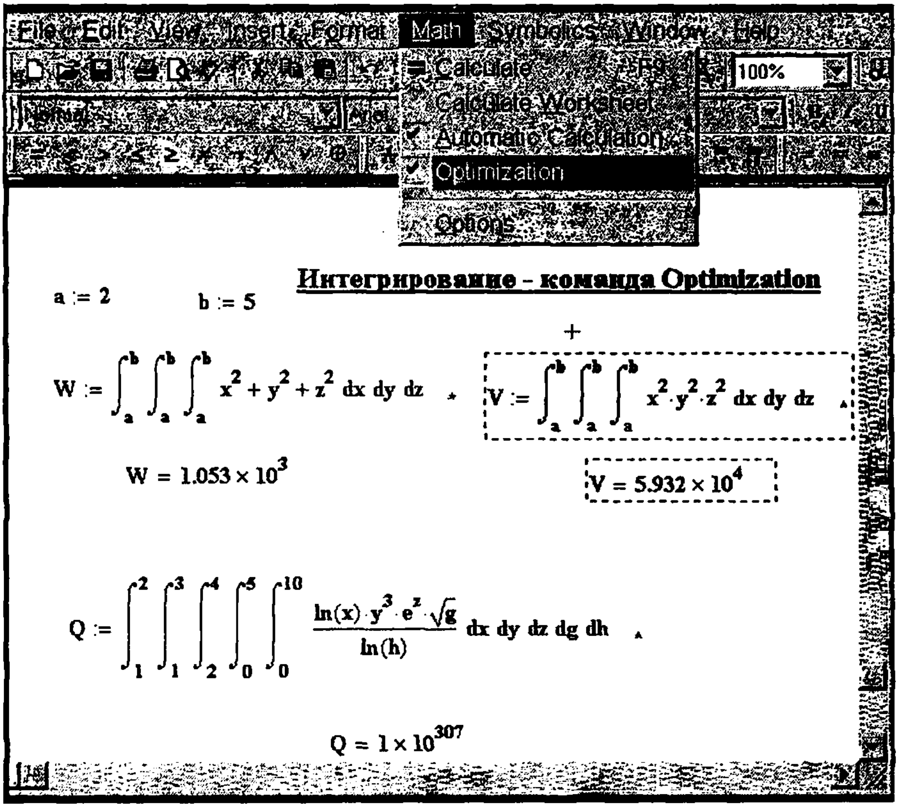


Рис. 8.4. Примеры вычисления интегралов с помощью команды ***Optimization***

Глава 9. Дифференциальные уравнения

9.1. Основные понятия

Уравнение, связывающее искомую функцию одной или нескольких переменных, эти переменные и производные различных порядков данной функции, называется дифференциальным уравнением [5, 18, 32]. Например, дифференциальными уравнениями будут:

$$y' = 6x;$$
$$y'' = 10 - 3x.$$

Примером дифференциального уравнения является задача о нахождении первообразной $F(x)$ для заданной функции $f(x)$, поскольку ее можно рассматривать как задачу о нахождении функции $F(x)$, удовлетворяющей уравнению $F'(x) = f(x)$.

Если искомая функция зависит от одной переменной, то дифференциальное уравнение называется обыкновенным, если от нескольких – то уравнением в частных производных.

В общем случае дифференциальное уравнение можно записать в виде

$$G(x, y, y', \ldots, y^{n}) = 0,$$

где G – некоторая функция от $n + 2$ переменных, при этом порядок n старшей производной, входящей в запись уравнения, называется порядком дифференциального уравнения.

Решением дифференциального уравнения называется такая функция $y = y(x)$, которая при подстановке ее в это уравнение обращает его в тождество.

Задача о нахождении решения некоторого дифференциального уравнения называется задачей интегрирования данного дифференциального уравнения. График решения дифференциального уравнения называется интегральной кривой.

Пример. Решить дифференциальное уравнение $y'' = 6x$.

Поскольку $y'' = dy'/dx$, то исходное уравнение равносильно следующему равенству дифференциалов: $dy' = xdx$. Выполняя интегрирование, получаем $y' = 3x^2 + C_1$, где C_1 – произвольная постоянная. Вновь записывая производную как отношение двух дифференциалов, приходим к равенству $dy = (3x^2 + C_1)dx$. Интегрируя, окончательно получаем $y = x^3 + C_1x + C_2$, где C_2 – произвольная постоянная.

Без дополнительных предположений решение данного уравнения принципиально неоднозначно. Дифференциальное уравнение задает семейство интегральных кривых на плоскости. Для выделения однозначно определенной кривой (решения) в нашем случае достаточно указать точку плоскости, через которую проходит интегральная кривая, и направление, в котором она проходит через эту точку. Дополнительные условия такого рода обычно называют *начальными*, поскольку часто дифференциальные уравнения используются для описания динамических процессов – процессов, проходящих во времени. В этих случаях независимая переменная x обозначает время. Например, если известно, что $y(0) = 1$ и $y'(0) = 2$, то приходим к решению: $y = x^3 + 2x + 1$.

Общим решением дифференциального уравнения n – го порядка называется такое его решение:

$$y = \varphi(x, C_1, \ldots, C_n),$$

которое является функцией переменной x и n произвольных независимых постоянных $C_1, C_2, \ldots, C_n$.

Частным решением дифференциального уравнения называется уравнение, получаемое из общего решения при некоторых конкретных числовых значениях постоянных $C_1, C_2, \ldots, C_n$.

В приведенном примере $y = x^3 + C_1x + C_2$ – общее решение, $y = x^3 + 2x + 1$ частное решение дифференциального уравнения.

9.2. Дифференциальные уравнения первого порядка

Дифференциальные уравнения первого порядка могут быть представлены в виде:

$$y' = f(x,y),$$

где f – некоторая функция двух переменных, определенная на плоскости $0xy$.

Рассмотрим геометрический смысл данного уравнения. Производная функции y' представляет угловой коэффициент (тангенс угла наклона) касательной к кривой $y = y(x)$ в точке с абсциссой x. Следовательно, данное уравнение каждой точке (x,y) плоскости $0xy$ сопоставляет направление $\mathrm{tg}\alpha = f(x,y)$ касательной к интегральной кривой $y = y(x)$, проходящей через эту точку.

Дифференциальное уравнение называется автономным, если функция f зависит только от переменной y, т. е. если уравнение имеет вид:

$$y' = f(y).$$

Уравнения такого типа часто встречаются на практике. Например, если дифференциальное уравнение описывает динамическое действие некоторого закона природы, то естественно предположить, что сам закон не будет изменяться с течением времени, и потому в запись правой части время x не входит.

Дифференциальное уравнение первого порядка называется *неполным*, если функция f явно зависит только от x или только от y. Рассмотрим методы решения таких уравнений.

1. Уравнение $y' = f(x)$. Его можно представить в виде $dy = f(x)dx$. Интегрируя, получим решение:

$$y = \int f(x)dx.$$

2. Уравнение $y' = f(y)$. Его решение удобно искать в виде $x = x(y)$, т. е. считать, что переменная y обозначает независимую переменную, а переменная x – функцию. Поскольку $y' = dy/dx$, то уравнение можно записать в виде $dy/f(x) = dx$. Отсюда:

$$x = \int \frac{dy}{f(y)}.$$

3. Если дифференциальное уравнение имеет вид:

$$\frac{1}{y} \cdot \frac{dy}{dx} = f(x),$$

то $(\ln|y|)' = f(x)$, и, интегрируя, получаем:

$$\ln y = \int f(x)\,dx + C,$$

откуда

$$y = e^{\int f(x)dx + C} = \exp(\int f(x)dx + C.$$

Пример. Решить дифференциальное уравнение: $y(3 - 2x)dx - xdy = 0$.

Можно записать $xdy = y(3 - 2x)dx$ или

$$\frac{1}{y}\cdot\frac{dy}{dx} = \frac{3-2x}{x}, \quad \text{или} \quad \frac{1}{y}\cdot\frac{dy}{dx} = \frac{3}{x} - 2,$$

откуда

$$\ln|y| = \int (\frac{3}{x} - 2)\,dx + C, \quad \text{или} \quad \ln|y| = 3\ln|x| - 2x + C.$$

Проводя преобразования:

$$\ln|y| - \ln|x^3| = -2x + C, \qquad \ln\left|\frac{y}{x^3}\right| = -2x + C.$$

$$\frac{y}{x^3} = e^{-2x+C},$$

получаем в итоге: $y = x^3e^{-2x+C}$.

Дифференциальное уравнение первого порядка называется уравнением с разделяющимися переменными, если оно может быть представлено в виде:

$$\frac{dy}{dx} = f(x)h(y),$$

где $f(x)$, $h(y)$ – функции переменных x и y.

Для решения такого уравнения его следует преобразовать к виду, в котором дифференциал и функции переменной x окажутся в одной части равенства, а переменной y – в другой. Затем проинтегрировать обе части полученного равенства. Так, для приведенного уравнения решение будет:

$$\frac{dy}{h(y)} = f(x)dx \quad \text{и} \quad \int\frac{dy}{h(y)} = \int f(x)dx.$$

9.3. Однородные дифференциальные уравнения первого порядка

Дифференциальное уравнение первого порядка называется однородным, если оно может быть представлено в виде, когда в левой части уравнения находится производная, а правая часть есть функция частного y/x:

$$y' = f(y/x).$$

Основной метод решения таких уравнений – это введение в рассмотрение вспомогательной функции $u = y/x$, что позволяет свести это уравнение к уравнению с разделяющимися переменными. Так как $y = ux$, то $y' = u'x + u$, поэтому уравнение приобретает следующий вид:

$$u'x + u = f(u),$$

откуда получим, что

$$\frac{du}{f(u)-u} = \frac{dx}{x}.$$

Пример. Рассмотрим дифференциальное уравнение:

$$(y^2 - x^2)dx + 2xydy = 0$$

и представим его в следующем виде:

$$\frac{dy}{dx} = \frac{y^2 - x^2}{2xy}.$$

Деля числитель и знаменатель на x^2, получаем:

$$\frac{dy}{dx} = \frac{\left(\frac{y}{x}\right)^2 - 1}{2\frac{y}{x}}.$$

Теперь правая часть уравнения представлена в виде функции одной переменной y/x; это оказалось возможным потому, что числитель и знаменатель дроби однородные функции одинаковой степени. Чтобы найти общий интеграл уравнения, введем вспомогательную переменную $u = y/x$. Отсюда следует, что

$$y = xu \text{ и } u' = u + xu'.$$

Подставляя, получим:

$$u + x\frac{du}{dx} = \frac{u^2 - 1}{2u},$$

а после переноса u в другую часть уравнения:

$$x\frac{du}{dx} = \frac{u^2 - 1 - 2u^2}{2u}, \quad \text{или} \quad x\frac{du}{dx} = \frac{-u^2 - 1}{2u}.$$

Применим метод разделения переменных. Получим:

$$\frac{2u}{-u^2 - 1}du = \frac{1}{x}dx.$$

Интегрируя, имеем:

$$\int \frac{2u}{-u^2 - 1}du = \int \frac{1}{x}dx, \quad \text{или} \quad -\int \frac{2u}{u^2 + 1}du = \int \frac{1}{x}dx.$$

Отсюда

$$-\ln(u2 + 1) + \ln C = \ln|x|, \text{ или } \ln|x| + \ln(u^2 + 1) = \ln C$$

и, следовательно,

$$\ln|x(u^2+1)| = \ln C, \text{ или } x(u^2 + 1) = C.$$

Возвращаясь к переменной $y = ux$, получим:

$$x\left(\frac{y^2}{x^2} + 1\right) = C, \quad \text{или} \quad \frac{y^2}{x} + x = C,$$

откуда, в конечном счете:

$$y^2 + x^2 = Cx.$$

Таким образом, полученное уравнение является общим интегралом исходного уравнения.

9.4. Дифференциальное линейное уравнение первого порядка

Дифференциальное уравнение первого порядка называется линейным, если оно имеет вид:

$$y' + f(x)y = g(x),$$

где $f(x)$ и $g(x)$ – некоторые функции переменной x. В случае, когда функция $g(x)$ тождественно равна нулю, уравнение называется однородным, в противном случае – неоднородным. Уравнение решается методом замены постоянной. Он основывается на том, что сначала определяется общий интеграл соответствующего однородного уравнения. Это уравнение – частный случай уравнения с разделенными переменными.

Из однородного уравнения $y' + f(x)y = 0$ получаем:

$$\frac{y'}{y} = -g(x), \quad \text{откуда} \quad \int \frac{1}{y} dy = -\int g(x)dx.$$

Какую-либо первообразную функцию функции $g(x)$ обозначим $F(x)$; тогда имеем:

$$\ln|y| = -F(x) + C, \quad \text{откуда} \quad y = e^C e^{-F(x)}.$$

Поскольку сомножитель e^C – постоянная, то, введя обозначение $e^C = k$, получим:

$$y = ke^{-F(x)}.$$

Неоднородное линейное дифференциальное уравнение решается путем замены постоянной k пока еще неизвестной функцией $k(x)$ переменной x, подобранной так, чтобы функция

$$y = k(x)e^{-F(x)}$$

была интегралом исходного уравнения.

Подставляя в исходное уравнение y и установленную производную y', определяем функцию $k(x)$. Подставив определенную функцию $k(x)$ в уравнение $y = k(x)e^{-F(x)}$, находим общий интеграл неоднородного уравнения:

$$y = k(x)e^{-\int f(x)dx}.$$

Пример. Решить дифференциальное уравнение:

$$y' - \frac{2}{x} y = 2x^3.$$

Сначала решаем соответствующее однородное уравнение:

$$y' - \frac{2}{x} y = 0.$$

Из этого уравнения находим:

$$\frac{y'}{y}=\frac{2}{x}.$$

Интегрируя это уравнение, получаем:

$$\int\frac{1}{y}dy=\int\frac{2}{x}dx,$$

откуда: $\ln|y|=2\ln x+C$, или $y=\exp C\cdot x^2$, где C – произвольная постоянная.

Теперь возьмем функцию:

$$y=C(x)x^2,$$

где $C(x)$ – пока неизвестная функция. Определяем производную:

$$y'=C'(x)x^2+C(x)2x.$$

Подставляя значения y и y' в исходное выражение, получаем:

$$C'(x)x^2+C(x)2x-(2/x)\,C(x)x^2=2x^3 \quad\text{или}\quad C'(x)x^2=2x^3,$$

откуда:

$$C'(x)=2x.$$

Интегрируя, имеем:

$$C(x)=\int 2xdx=x^2+A,$$

где A – произвольная постоянная.

Тогда общий интеграл дифференциального уравнения равен:

$$y=(x^2+A)x^2.$$

9.5. Дифференциальные уравнения второго порядка

Дифференциальным линейным уравнением второго порядка называется уравнение вида:

$$y''+h(x)\,y'+g(x)y=f(x),$$

где $h(x)$, $g(x)$,$f(x)$ – функции переменной x. Если $f(x)\neq 0$, уравнение неоднородное. Если $f(x)=0$, уравнение однородное.

Если функции $h(x)$ и $g(x)$ равны постоянным коэффициентам a и b, то полученные уравнения

$$y''+a\,y'+by=f(x);$$
$$y''+a\,y'+by=0$$

называются уравнениями с постоянными коэффициентами.

9.6. Дифференциальные уравнения второго порядка, допускающие понижение порядка

В некоторых случаях решение дифференциального уравнения второго порядка может быть сведено к последовательному решению двух дифференциальных уравнений первого порядка. В этом случае говорят, что данное дифференциальное уравнение допускает понижение порядка. Если дифференциальное уравнение имеет вид:

$$y'''=f(x),$$

то оно решается последовательным интегрированием.

Если в запись уравнения не входит искомая функция $y(x)$, т. е. оно имеет вид:

$$P(x, y', y'') = 0,$$

то такое уравнение можно решить, найдя сначала вспомогательную функцию y'.

Пример. Решить уравнение $xy'' + y' = 0$.

Примем $z = y'$. Тогда $y'' = z'$, и исходное уравнение принимает вид $xz' + z = 0$.

Проведя преобразование $z' = -z/x$, получаем $z'/z = -1/x$. Интегрируя, приходим к решению: $z = C_1/x$. Возвращаясь к исходной функции, получаем уравнение $y' = C_1/x$, или $dy = C_1dx/x$, решая которое, окончательно имеем: $y = C_1\ln|x| + C_2$.

Если в уравнение не входит переменная x, т. е. оно имеет вид:

$$P(y, y', y'') = 0,$$

то порядок уравнения можно понизить, если за независимую переменную взять y, а за неизвестную функцию взять $z = y'$.

9.7. Линейные дифференциальные уравнения второго порядка с постоянными коэффициентами

Линейной комбинацией функций $y_1(x)$ и $y_2(x)$ с коэффициентами C_1 и C_2 называется выражение вида $C_1y_1(x) + C_2y_2(x)$. Функции $y_1(x)$ и $y_2(x)$ являются линейно независимыми, если их отношение $y_1(x)/y_2(x)$ непостоянно. В противном случае функции $y_1(x)$ и $y_2(x)$ являются линейно зависимыми.

Для решения линейных дифференциальных уравнений второго порядка с постоянными коэффициентами используется следующая теорема. Если $y_1(x)$ и $y_2(x)$ – линейно независимые частные решения линейного дифференциального уравнения второго порядка с постоянными коэффициентами вида:

$$y'' + ay' + by = 0,$$

то общее решение этого уравнения является линейной комбинацией этих частных решений, т. е. имеет вид:

$$y = C_1y_1 + C_2y_2,$$

где C_1 и C_2 – некоторые действительные числа.

То есть, чтобы найти общее решение уравнения $y'' + ay' + by = 0$, надо знать два его частных решения.

Рассмотрим решение уравнения $y'' + ay' + by = 0$ в форме $y = e^{nx}$, где n – некоторое действительное число. Так как

$$(e^{nx})'' + a(e^{nx})' + be^{nx} = (n^2 + an + b)e^{nx},$$

то функция $y = e^{nx}$ является решением уравнения, если число n есть корень уравнения

$$n^2 + an + b = 0,$$

которое называется характеристическим уравнением исходного уравнения.

Решение уравнения зависит от того, имеет ли соответствующее характеристическое уравнение два различных корня, один корень или не имеет действительных корней.

Если характеристическое уравнение $n^2 + an + b = 0$ имеет действительные корни n_1 и n_2 ($\Delta = a^2 - 4b > 0$), причем $n_1 \neq n_2$, тогда общее решение уравнения имеет вид:

$$y = C_1e^{n_1x} + C_2e^{n_2x}.$$

Если характеристическое уравнение имеет один корень n (кратности 2), то общее решение уравнения имеет вид:

$$y = (C_1 + C_2x)e^{nx}.$$

Если характеристическое уравнение не имеет действительных корней, то общее решение уравнения имеет вид:

$$y = C_1e^{ux}\sin\beta x + C_2e^{ux}\cos\beta x,$$

где $u = -a/2$, $\beta=\sqrt{b-a^2/4}$, C_1 и C_2 – некоторые числа.

Пример. Решить дифференциальное уравнение:

$$y'' + 3y' + 2y = 0$$

Характеристическое уравнение имеет вид:

$$n^2 + 6n + 3 = 0.$$

Поскольку $\Delta = 36 - 9 = 25 > 0$, имеем:

$$n_1 = (-6-5)/2 = -5{,}5;\ n_2 = (-6+5)/2 = -0{,}5.$$

Следовательно, общий интеграл есть:

$$y = C_1e^{-5,5x} + C_2e^{-0,5x}.$$

Рассмотрим теперь решение линейного неоднородного уравнения с постоянными коэффициентами.

Данное уравнение может быть, в частности, решено методом вариации произвольных постоянных, который состоит в следующем. Сначала находится общее решение $y = C_1y_1 + C_2y_2$ однородного уравнения, имеющего ту же левую часть, что и исходное неоднородное уравнение. Затем решение уравнения находится в виде $y - C_1(x)y_1 + C_2(x)y_2$, т. е. предполагается, что постоянные C_1 и C_2 являются функциями независимой переменной x. При этом $C_1(x)$ и $C_2(x)$ могут быть найдены как решение системы:

$$\begin{cases} C_1'y_1 + C_2'y_2 = 0; \\ C_1'y_1 + C_2'y_2 = f(x). \end{cases}$$

Пример. Решить уравнение:

$$y'' - 3y' + 2y = e^x.$$

Решая соответствующее однородное уравнение:

$$y'' - 3y' + 2y = 0,$$

находим:

$$y = C_1e^x + C_2e^{2x}.$$

Полагая теперь, что C_1 и C_2 – функции переменной x, найдем первые производные этих функций, решая систему:

$$\begin{cases} C_1'e^x + C_2'e^{2x} = 0; \\ C_1'e^x + C_2'e^{2x} = e^x. \end{cases}$$

Найдем $C_1' = -1$, $C_2' = e^x$. Полученные дифференциальные уравнения – с разделяющимися переменными. Решая эти уравнения, получаем $C_1 = -x + C_3$, $C_2 = -e^x + C_4$, где C_3, C_4 – некоторые постоянные. Таким образом, окончательно решение уравнения имеет вид:

$$y = (-x + C_3)e^x + (-e^x + C_4)e^{2x} = C_3e^x + C_4e^{2x} + (-x - 1)e^x.$$

В структуре полученного решения первые два слагаемых – это общее решение однородного уравнения, соответствующее исходному дифференциальному уравнению. Последнее слагаемое – частное решение исходного уравнения. Аналогичное утверждение справедливо и в общем случае, т. е. справедлива следующая теорема: *общее решение линейного неоднородного дифференциального уравнения равно сумме общего решения соответствующего однородного уравнения и частного решения исходного неоднородного уравнения.*

Метод решения, основанный на данной теореме, сводится к следующему. Находится общее решение однородного дифференциального уравнения, затем отыскивается частное решение неоднородного уравнения. При этом вид частного решения устанавливается по виду правой части уравнения, и задача сводится к отысканию коэффициентов этого частного решения. Этим методом можно пользоваться в нескольких случаях; рассмотрим случай, когда

$$f(x) = e^{mx}P_k(x),$$

где $P(x)$ – многочлен k-й степени.

Если m не является корнем характеристического уравнения

$$n^2 + an + b = 0,$$

то частный интеграл имеет вид:

$$I = e^{mx}P_k(x),$$

если m – корень характеристического уравнения

$$I = x^r e^{mx}P^k(x),$$

где r =1 или 2, смотря по тому, имеет ли характеристическое уравнение двойной корень $n_1 = n_2 = m$ или же единственный корень $n = m$.

Пример. Решить дифференциальное уравнение:

$$2y'' - y' - y = 4xe^{2x}.$$

Соответствующее однородное уравнение имеет вид:

$$2y'' - y' - y = 0$$

Характеристическое уравнение этого уравнения имеет вид:

$$2n^2 - n - 1 = 0;$$

оно имеет два корня:

$$n_1 = -1/2,\ n_2 = 1;$$

следовательно, общее решение однородного уравнения есть:

$$y = C_1e^{-x/2} + C_2e^x.$$

Найдем теперь частное решение данного неоднородного уравнения. Правая часть этого уравнения имеет вид:

$$4xe^{2x} = e^{mx}P_k(x).$$

В нашем случае $m = 2$ не является корнем характеристического уравнения и $n = 1$; поэтому частное решение данного уравнения имеет вид:

$$I = e^{2x}(Ax + B).$$

$$I' = Ae^{2x} + 2e^{2x}(Ax + B) = e^{2x}(A + 2Ax + 2B),$$

$$I'' = e^{2x}2A + 2e^{2x}(A + 2Ax + B) = 2e^{2x}(2A + 2Ax + B).$$

Подставляя соответствующие величины в исходное уравнение, получим:

$$4e^{2x}(2A + 2Ax + B) - e^{2x}(A + 2Ax + B) - e^{2x}(Ax + B) = 4xe^{2x}.$$

Сокращая на e^{2x} и получая уравнение, которое будет удовлетворяться тождественно, т. е. при всех x, имеем:

$$5A = 4;\ 7A + 2B = 0,$$

откуда:

$$A = 0{,}8,\ B = -2{,}8.$$

Следовательно, частное решение уравнения есть:

$$I = e^{2x}(0{,}8x - 2{,}8),$$

а общее решение:

$$y = C_1e^x + C_2e^{-x/2} + e^{2x}(0{,}8x - 2{,}8).$$

9.8. Применение дифференциальных уравнений в экономике

Рассмотрим следующий пример. Затраты q являются функцией объема производства x. Необходимо найти функцию затрат, если известно, что предельные затраты для всех значений x равняются средним затратам.

Из условия задачи следует, что

$$\frac{dq}{dx} = \frac{q}{x} \text{ или } \frac{q'}{q} = \frac{1}{x}.$$

Интегрируя, получаем:

$$\int \frac{q'}{q}dx = \int \frac{1}{x}dx,$$

откуда

$$\ln|q| = \ln|x| + \ln|C|,$$

где C – произвольная постоянная. Отсюда:

$$\ln|q| = \ln|Cx|, \quad \text{или} \quad q = Cx.$$

Дифференциальные уравнения находят достаточно широкое применение в моделях динамической экономики, в которых отражается не только зависимость переменных от времени, но и их взаимосвязь во времени.

Рассмотрим некоторые задачи динамической экономики.

Пусть $y(t)$ – объем продукции некоторой отрасли, реализованной к моменту времени t. Вся производимая отраслью продукция реализуется по некоторой фиксированной цене p. Тогда доход к моменту времени t составит $Y(t) = py(t)$.

Величину инвестиций, направляемых на расширение производства, обозначим через $I(t)$. Будем полагать, что скорость выпуска продукции (акселерация) пропорциональна величине инвестиций, а инвестиционный лаг (время между окончанием производства продукции и ее реализацией) равен нулю:

$$y'(t) = lI(t),$$

где l – норма акселерации. Полагая, что величина инвестиций $I(t)$ составляет фиксированную часть дохода, получим:

$$I(t) = mY(t) = mpy(t),$$

где коэффициент пропорциональности m (норма инвестиций) – постоянная величина, $0 < m < 1$.

Подставляя данное выражение, приходим к выражению:

$$y' = ky, \text{ где } k = mpl.$$

Полученное дифференциальное уравнение – с разделяющими переменными. Решая его, приходим к функции:

$$y(t) = y_0 e^{k(t-t_0)}, \text{ где } y_0 = y(t_0).$$

В общем случае кривая спроса, т. е. зависимость цены p реализованной продукции от ее объема y является убывающей функцией $p = p(y)$ (с увеличением объема произведенной продукции ее цена падает). Поэтому модель роста в условиях конкурентного рынка примет вид:

$$y' = mlp(y)\,y,$$

оставаясь по-прежнему уравнением с разделяющимися переменными.

Так как все сомножители в правой части уравнения положительны, то $y' > 0$, и это уравнение описывает возрастающую функцию $y(t)$. При исследовании функции $y(t)$ на выпуклость используется понятие эластичности функции. Из предыдущего выражения следует, что

$$y'' = mly'\left(\frac{dp}{dy} + p\right).$$

Вспомним, что эластичность спроса относительно цены определяется формулой:

$$\varepsilon_p(y) = \frac{p}{y}\frac{dy}{dp}.$$

Тогда выражение для y'' можно записать в виде:

$$y'' = mly'p\left(\frac{1}{\varepsilon_p(y)} + 1\right),$$

и условие $y'' = 0$ равносильно равенству $\varepsilon_p(y) = -1$.

Таким образом, если спрос эластичен, т. е. $|\varepsilon_p(y)| > 1$ или $\varepsilon_p(y) < -1$, то $y'' > 0$ и функция $y(t)$ выпукла вниз; в случае, если спрос не эластичен, то $y'' < 0$ и функция $y(t)$ выпукла вверх.

Рассмотрим следующий пример. Нужно найти выражение для объема реализованной продукции $y = y(t)$, если известно, что кривая спроса $p(y)$ задается уравнением $p(y) = 2 - y$, норма акселерации $1/l = 2$, норма инвестиций $m = 0{,}5$, $y(0) = 0{,}5$.

В этом случае

$$y' = (2 - y)y,$$

или

$$\frac{dy}{(2-y)y} = dt.$$

Выполняя почленное интегрирование, получаем:

$$\ln\left|\frac{y-2}{y} = -2t + C_1\right|,$$

или

$$\frac{y-2}{y} = Ce^{-2t},$$

где $C = \pm e^{C_1}$.

Учитывая, что $y(0) = 0{,}5$, получаем, что $C = -3$. Выражая теперь y, окончательно имеем:

$$y = \frac{2}{1 + 3e^{-2t}}.$$

9.9. Решение дифференциальных уравнений в Mathcad

Анализ динамических систем и их математическое моделирование базируются на численных методах решения систем дифференциальных уравнений. В Mathcad есть возможность решения дифференциальных уравнений и систем с этими уравнениями в численном виде.

Нелинейные дифференциальные уравнения и системы с такими уравнениями, как правило, не имеют аналитических методов решения, поэтому важна возможность их решения численными методами. В большинстве случаев желательно представление решений в графическом виде, что и позволяет сделать Mathcad.

Для решения дифференциальных уравнений существует ряд функций, в первую очередь это функции:

Rkadapt, rkfixed, Bulstoer.

Рассмотрим сначала функции, дающие решения для систем обыкновенных дифференциальных уравнений:

- **Rkadapt(v,x1,x2,*n*,F)** – возвращает матрицу решений методом Рунге-Кутта с переменным шагом для системы обыкновенных дифференциальных уравнений с начальными условиями в векторе v, правые части первых производных уравнений записаны в символьном векторе F, на интервале от x1 до x2; *n* – число шагов.
- **rkfixed(v,x1,x2,*n*,F)** – возвращает матрицу решений методом Рунге-Кутта системы обыкновенных дифференциальных уравнений с начальными условиями в векторе v. Первые производные записаны в символьном векторе F, на интервале от x1 до x2 при фиксированном числе шагов *n*.
- **rkadapt(v,x1,x2,acc,*n*,F,k,s)** – возвращает матрицу, содержащую таблицу значений решения задачи Коши на интервале от x1 до x2 для системы обыкновенных дифференциальных уравнений, вычисленную методом Рунге-Кутта с переменным шагом и начальными условиями в векторе v, правыми частями производных системы в векторе F. Здесь *n* – число шагов, k – максимальное число промежуточных точек решения, s – минимально допустимый интервал между точками, acc – погрешность решения.

Функция **Rkadapt** благодаря автоматическому изменению шага решения дает более точный результат. По скорости вычислений она может проигрывать функции **rkfixed**, хотя и не всегда – если решение меняется медленно, это может привести к заметному уменьшению числа шагов. Таким образом, функция **Rkadapt** наиболее привлекательна для решения систем дифференциальных уравнений, дающих медленно изменяющиеся решения.

Требует пояснения заполнение вектора F, поскольку это связано с методом численного решения, запрограммированного в вычислительном блоке Mathcad решения дифференциальных уравнений. Коротко можно сказать, что вектор F должен состоять из правых частей первых производных, выраженных через новые функции.

Например, дифференциальное уравнение

$$\frac{d^2}{dx^2}y(x) + 4\cdot\frac{d}{dx}y(x) - 5\cdot y(x) = 6\cdot x$$

содержит вторую производную, которая может быть выражена через первую производную

$$\frac{d^2}{dx^2}y(x) = \frac{d}{dx}\frac{d}{dx}y(x)$$

Определим две функции, как

$$y_0(x) = y(x)$$

$$y_1(x) = \frac{d}{dx}y_0(x)$$

Теперь дифференциальное уравнение содержит две функции и представляет собой систему из двух дифференциальных уравнений

$$\frac{d}{dx}y_0(x) = y_1(x)$$

$$\frac{d}{dx}y_1(x) = 6\cdot x + 5\cdot y_0(x) - 4\cdot y_1(x)$$

Используя полученную систему, можно составить вектор F, который Mathcad использует в вычислительном блоке решения дифференциальных уравнений

$$F(x,y) := \begin{bmatrix} y_1(x) \\ 6\cdot x + 5\cdot y_0(x) - 4\cdot y_1(x) \end{bmatrix}$$

Рассмотрим первый пример решения дифференциального уравнения

$$y' = x + y,$$

используя функцию **rkfixed** (рис. 9.1). Для решения дифференциального уравнения надо задать начальные условия и составить вектор F. Решение можно вывести численно, в виде матрицы, и графически.

$$\frac{d}{dx}y = x + y$$ Решаемое уравнение

$x(0) = 0$ $y(0) = 1$ Задание начальных условий

$$v := \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$ Вектор начальных условий

$$F(x, y) := \begin{pmatrix} y_1 \\ x + y_0 \end{pmatrix}$$

Задание вектора F из системы

$$\frac{d}{dx}y(x) = y_1$$

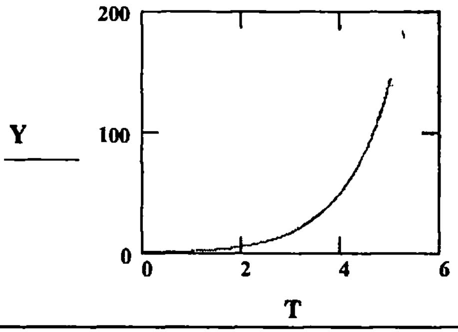
$$y_1 = x + y_0$$

Решение

$$Q := \text{rkfixed}(v, 0, 5, 400, F)$$

$T := Q^{\langle 0 \rangle}$ Значение независимой переменной

$Y := Q^{\langle 1 \rangle}$ Значение вычисляемой функции



$Q =$

	0	1	2
0	0	0	1
1	0.013	0.013	1
2	0.025	0.025	1.001
3	0.038	0.038	1.001
4	0.05	0.05	1.003
5	0.063	0.063	1.004
6	0.075	0.075	1.006
7	0.088	0.088	1.008
8	0.1	0.1	1.01
9	0.113	0.113	1.013
10	0.125	0.126	1.016
11	0.137	0.138	1.019
12	0.15	0.151	1.023
13	0.163	0.164	1.026
14	0.175	0.177	1.031
15	0.188	0.19	1.035

Рис. 9.1. Пример решения дифференциального уравнения с использованием функции rkfixed

Решение дифференциального уравнения с использованием функции **Rkadapt** приведено на рис. 9.2.

$$\frac{d^2}{dx^2}y = \frac{d}{dx}y + 2\cdot y$$ **Решаемое уравнение**

$$y(0) = 1 \quad \frac{d}{dx}y(0) = 3$$ **Задание начальных условий**

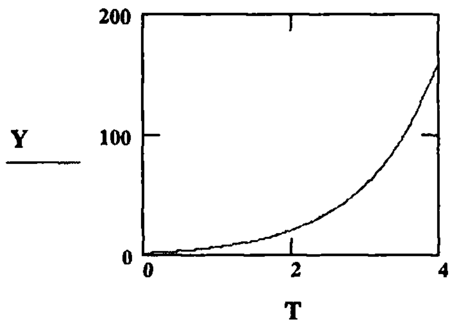
$$v := \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$ **Вектор начальных условий**

$$D(t, y) := \begin{pmatrix} y_1 \\ -y_1 + 2\cdot y_1 \end{pmatrix}$$ **Задание вектора**

$$Z := \mathrm{Rkadapt}(v, 0, 4, 400, D)$$

$$T := Z^{\langle 0 \rangle}$$

$$Y := Z^{\langle 1 \rangle}$$



Z =

	0	1	2
0	0	1	3
1	0.01	1.03	3.03
2	0.02	1.061	3.061
3	0.03	1.091	3.091
4	0.04	1.122	3.122
5	0.05	1.154	3.154
6	0.06	1.186	3.186
7	0.07	1.218	3.218
8	0.08	1.25	3.25
9	0.09	1.283	3.283
10	0.1	1.316	3.316
11	0.11	1.349	3.349
12	0.12	1.382	3.382
13	0.13	1.416	3.416
14	0.14	1.451	3.451
15	0.15	1.486	3.486

*Рис. 9.2. Пример решения дифференциального уравнения с использованием функции **Rkadapt***

Решение системы дифференциальных уравнений с применением функции **Rkadapt** приведено на рис. 9.3. и 9.4.

$$\frac{d}{dx}y_0(x) = -7\cdot y_0(x) + 5\cdot y_1(x) \qquad y_0(0) = -1$$

$$\frac{d}{dx}y_1(x) = 30\cdot y_0(x) + 2\,y_1(x) - y_0(x)\,y_2(x) \qquad y_1(0) = 0$$

$$\frac{d}{dx}y_2(x) = y_0(x)\cdot y_1(x) - 2\cdot y_2(x) \qquad y_2(0) = 1$$

$$F(x, y) := \begin{pmatrix} -7\cdot y_0 + 5\cdot y_1 \\ 30\cdot y_0 + 2\cdot y_1 - y_0\cdot y_2 \\ y_0\cdot y_1 - 2\cdot y_2 \end{pmatrix}$$

<u>Определение аргуметов для вычислительного блока</u>

$x_0 := 0$ Первоначальное значение независимой

$x_1 := 10$ Конечное значение независимой

$$V_0 := \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$ Вектор начальных значений

$N := 1000$ Количество решений (точек) на [x[0),x[1)]

<u>Решение</u>

$W := \mathrm{Rkadapt}(V_0, x_0, x_1, N, F)$

$x := W^{\langle 0 \rangle}$ Значения независимой

$y_0 := W^{\langle 1 \rangle}$ Значение функции (первое

$y_1 := W^{\langle 2 \rangle}$Значение функции (второе решение)

$y_2 := W^{\langle 3 \rangle}$ Значение функции (третье

Рис. 9.3. Решение системы дифференциальных уравнений

W =

	0	1	2	3
985	9.85	0.489	-0.488	21.977
986	9.86	0.433	-0.46	21.539
987	9.87	0.382	-0.434	21.111
988	9.88	0.336	-0.409	20.691
989	9.89	0.294	-0.387	20.28
990	9.9	0.256	-0.368	19.878
991	9.91	0.221	-0.35	19.483
992	9.92	0.19	-0.335	19.097
993	9.93	0.161	-0.322	18.718
994	9.94	0.135	-0.312	18.347
995	9.95	0.111	-0.304	17.983
996	9.96	0.089	-0.297	17.627
997	9.97	0.069	-0.293	17.278
998	9.98	0.05	-0.292	16.935
999	9.99	0.032	-0.292	16.6
000	10	0.016	-0.295	16.271

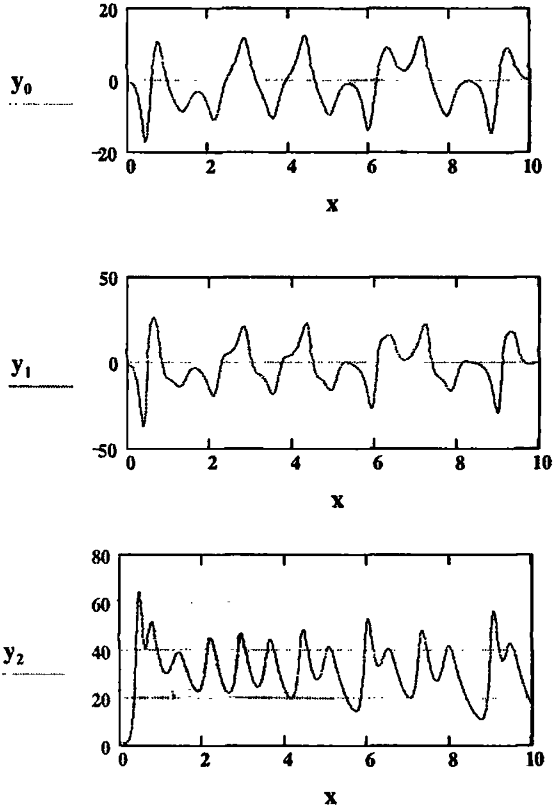


Рис. 9.4. Результат и графическое представление решения системы дифференциальных уравнений

Если решение системы дифференциальных уравнений имеет вид гладких функций, то целесообразно применять функцию:

Bulstoer(v,x1,x2,*n*,F).

Система дифференциальных уравнений, записанная в матричной форме y' = A·x, где A – почти вырожденная матрица, называется жесткой. Для решения жестких дифференциальных уравнений в Mathcad есть ряд функций:

- **bulstoer**(v,x1,x2,acc,n,F,k,s);
- **Stiffb**(v,x1,x2,*n*,F,J), где J – функция Якобиана;
- **stiffb**(v,x1,x2,acc,*n*,F,k,s);
- **Stiffr**(v,x1,x2,*n*,F,J);
- **stiffr**(v,x1,x2,acc,*n*,F,J,k,s).

Функции, начинающиеся с малой буквы, дают решения только для конечной точки. Различаются функции также и методом решения.

Для решения дифференциальных уравнений Пуассона и уравнений Лапласа в систему введены следующие функции:

- **relax(M1,M2,M3,M4,M5,A,U,r)** – возвращает квадратную матрицу решения уравнения Пуассона для спектрального радиуса r;
- **multigrid(M,n)** – возвращает матрицу решения уравнения Пуассона, у которого решение равно нулю на границах;
- **bvalfit(v1,v2,x1,x2,xi,F,L1,L2,score)** – задает начальные условия для краевой задачи, заданной в векторах F,v1,v2 на интервале [x1,x2], где решение известно в некоторой промежуточной точке xi;
- **sbval(v,x1,x2,F,L,score)** – дает установку начальных условий для краевой задачи, определенной в символьном векторе F.

Для решения двухточечных краевых задач предназначены функции:

- **sbval(v,x1,x2,D,load,score)**;
- **bvalfit(v1,v2,x1,x2,xf,S,load1,load2,score)**.

В них векторы v,v1,v2 задают начальные условия, x,x1,x2 – граничные точки интервала решений, D(x,y) – функция, возвращающая n – компонентный вектор с первыми производными неизвестных функций, load(x1,v), load(x1,v1), load(x2,v2) – векторзначные функции, возвращающие значения начальных условий в точках x1(x2), score(xf,y) – векторзначная функция, возвращающая n – элементный вектор соответствия. Он указывает, насколько значения решений, начинающихся из точек x1 и x2, должны соответствовать xf. Например, если нужно совпадение решений, то score(xf,y):=y.

Достаточное число примеров применения этих функций дано в справочной системе Mathcad.

Глава 10. Матричное исчисление

10.1. Векторы и матрицы

Матрицы находят широкое применение в экономических исследованиях, особенно в планировании производства. Они значительно облегчают работу, снижают трудоемкость, позволяют быстро разработать разные варианты плана и, кроме того, облегчают изучение зависимостей между разными экономическими показателями.

Таблица $m\times n$ чисел, записанных в виде прямоугольника и взятых в квадратные скобки:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ . & . & . & . \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

или кратко: $[a_{ij}]$ ($i = 1, 2, ..., m, j = 1, 2, ..., n$), называется прямоугольной матрицей с m строками и n столбцами или матрицей $m\times n$. Числа a_{ij} называются элементами матрицы. Каждый из элементов матрицы обозначается двумя индексами: первый индекс i обозначает номер строки, второй индекс j – номер столбца, в которых находится данный элемент.

Если число строк матрицы равняется числу ее столбцов, т. е. $m = n$, то матрица называется квадратной. В этом случае число $m = n$ называется порядком матрицы. Совокупность элементов квадратной матрицы, для которых $i = j$, называется главной диагональю матрицы.

Матрицы принято также обозначать буквами полужирного начертания:

$$A=\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ . & . & . & . \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}.$$

Размерность или порядок матрицы обозначается через соответствующий индекс. Например, прямоугольная матрица A размерностью $m\times n$ будет обозначаться $A_{m\times n}$, а квадратная матрица $\boldsymbol{B}$ порядка n будет $\boldsymbol{B}_n$. Матрица, состоящая из одной строки, т. е. матрица $A = [a_1\ a_2 \dots a_n] = A_{1\times n}$, называется вектором строкой.

Матрица, состоящая из одного столбца, т. е. матрица

$$B=\begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{bmatrix}=B_{m\times 1}$$

называется вектором-столбцом.

Матрица, нее элементы которой есть пули, называется nyjiciioh матрицей обозначается символом O или Om. Диагональной называется матрица. в котором все элементы, не находящиеся на главной диагонали, равняются нулю. Динго нальная матрица, в которой все элементы главной диагонали единицы, называется единичной матрицей; единичная матрица обозначается буквами /или /„.

Матрица

$$/=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}=/,$$

есть единичная матрица третьего порядка.

10.2. Операции над матрицами

Над матрицами можно производить некоторые математические операции сложение, вычитание, умножение матрицы на число, умножение матрицы на матрицу.

Две матрицы одинакового "размера mxn называются равными, если все элементы, стоящие на одних и тех же местах, равны между собой. Тогда пишут:

$$A=B.$$

Суммой двух матриц A и B, имеющих соответственно равные количества строк и столбцов, называется матрица, элементы которой равняются суммам соответствующих элементов матриц A и B.

Если

$$A=\begin{bmatrix}5 & 3 & 7\\ 4 & 9 & 2\end{bmatrix},\quad B=\begin{bmatrix}2 & -4 & 5\\ 3 & 6 & 7\end{bmatrix},$$

то сумма матриц есть матрица C:

$$C=A+B=\begin{bmatrix}5+2 & 3-4 & 7+5\\ 4+3 & 9+6 & 2+7\end{bmatrix}=\begin{bmatrix}7 & -1 & 12\\ 7 & 15 & 9\end{bmatrix}.$$

Сложение матриц обладает следующими свойствами:

$$A+B=B+A;$$
$$A+(B+Q=(A+B)+C;$$
$$A+O=O+A^A.$$

Произведением матрицы A на вещественное число $к$ называется матрица, в которой все элементы матрицы A умножены на число $к$:

$$B=кA=к\cdot\begin{bmatrix}a & b & c\\ f & g & h\end{bmatrix}=\begin{bmatrix}кa & kb & кc\\ кf & kg & kh\end{bmatrix}.$$

Разность двух матриц определяется формулой:

$$AB=A+(-\backslash)B.$$

Умножение матриц. Произведение матриц A и B определяется только в том случае, если число столбцов матрицы A равняется числу строк матрицы B. Умножение матриц производится по правилу перемножения вектор-строки на вектор-столбец, т. е. элементы /-ой строки матрицы A умножаются на соответствующие элементыу-го столбца матрицы B, и произведения складываются.

Пример перемножения вектор-строки на вектор-столбец. Предприятие закупает три вида сырья по соответствующим ценам. Количество сырья можно представить вектор-строкой, а цены – вектор-столбцом. Сырье вида Q_1, Q_2, и Q_3 закупается в количестве 10, 15, 20, что можно изобразить вектор-строкой:

$$Q = [10\ 15\ 20].$$

Цены D_1, D_2 и D_3 на сырье вида Q_1, Q_2, и Q_3 составляют соответственно 2, 3, 4, что можно представить вектор-столбцом:

$$D=\begin{bmatrix}2\\3\\4\end{bmatrix}.$$

Затраты на приобретение сырья будут представлять собой сумму произведений каждого вида сырья на его цену:

$$10\cdot 2 + 15\cdot 3 + 20\cdot 4 = 145.$$

Учитывая, что вектор-строка и вектор-столбец являются частными случаями матриц, перейдем к общему определению матриц.

Пусть даны две матрицы $A = A_{mp}$ и $B = B_{pn}$, т. е. матрица A имеет m строк и p столбцов, а матрица B имеет p строк и n столбцов:

$$A=\begin{bmatrix}a_{11} & a_{12} & \dots & a_{1p}\\ a_{21} & a_{22} & \dots & a_{2p}\\ . & . & . & .\\ a_{m1} & a_{m2} & \dots & a_{mp}\end{bmatrix},\quad B=\begin{bmatrix}b_{11} & b_{12} & \dots & b_{1n}\\ b_{21} & b_{22} & \dots & b_{2n}\\ . & . & . & .\\ b_{p1} & b_{p2} & \dots & b_{pn}\end{bmatrix}.$$

Тогда произведением матриц AB называется матрица:

$$C=AB=\begin{bmatrix}c_{11} & c_{12} & \dots & c_{1n}\\ c_{21} & c_{22} & \dots & c_{2n}\\ . & . & . & .\\ c_{m1} & c_{m2} & \dots & c_{mn}\end{bmatrix}.$$

размерностью $m\times n$, элементы которой c_{ij} вычисляются по формуле:

$$c_{ij} = a_{i1}\, b_{1j} + a_{i2}\, b_{2j} + \dots + a_{ip}\, b_{pj},$$

т. е. элементы i-ой строки матрицы A умножаются на соответствующие элементы j-го столбца матрицы B, и произведения складываются.

Умножим, для примера, матрицу A на матрицу B, где

$$A=\begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix}=\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix},\quad B=\begin{bmatrix}b_{11} & b_{12}\\ b_{21} & b_{22}\end{bmatrix}=\begin{bmatrix}5 & 6\\ 7 & 8\end{bmatrix}.$$

Если обозначить их произведение как

$$C=\begin{bmatrix}c_{11} & c_{12}\\ c_{21} & c_{22}\end{bmatrix},$$

то элементы матрицы C:

$$c_{11} = a_{11}\cdot b_{11} + a_{12}\cdot b_{21} = 1\cdot 5 + 2\cdot 7 = 19;$$
$$c_{12} = a_{11}\cdot b_{12} + a_{12}\cdot b_{22} = 1\cdot 6 + 2\cdot 8 = 22;$$
$$c_{21} = a_{21}\cdot b_{11} + a_{22}\cdot b_{21} = 3\cdot 5 + 4\cdot 7 = 43;$$
$$c_{22} = a_{21}\cdot b_{12} + a_{22}\cdot b_{22} = 3\cdot 6 + 4\cdot 8 = 50.$$

Окончательный результат можно записать в виде:

$$C=\begin{bmatrix}19 & 22\\ 43 & 50\end{bmatrix}.$$

В противоположность умножению чисел и умножению матрицы на число, произведение двух матриц A и B, как правило, некоммутативно, произведения AB и BA не всегда равны:

$$AB \neq BA.$$

Умножение матриц ассоциативно, т. е. если даны матрицы A, B и C, то

$$A(BC) = (AB)C.$$

Как и для чисел, для матриц справедливо правило *дистрибутивности* умножения относительно сложения:

$$(A + B)C = AC + BC,\ C(A + B) = CA + CB.$$

10.3. Операции над матрицами в Excel

При выполнении операций над матрицами в Excel необходимо соблюдать следующий порядок команд.

1. Выделение области ячеек, где будет записан ответ.
2. Операции начинаются со знака <равно>, даже при вводе формул.
3. Вводимые данные, т. е. матрицы, с которыми производятся операции, выделяются как блок (диапазон) ячеек.
4. Операции сложения, вычитания, умножения матрицы на число производятся с помощью аналогичных команд с клавиатуры или мыши, а остальные – умножение матрицы на матрицу, транспонирование, обращение и т. д. – с помощью матричных функций.
5. Заканчивать ввод нужно не нажатием клавиши <Enter>, а комбинацией клавиш <Shift>+<Ctrl>+<Enter>. Для правильного ввода данной команды необходимо при нажатых клавишах <Shift>+<Ctrl> нажать клавишу <Enter>.

Сложение матриц. При сложении матриц вводятся две матрицы, выделяется блок ячеек под ответ и вводится команда, например: <= A2:C4 +D2:F4; <Shift>+ +<Ctrl>+<Enter> (рис. 10.1).

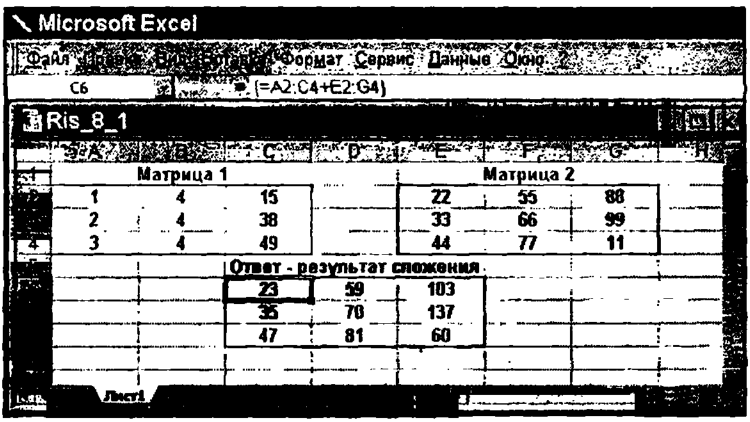


Рис. 10.1. Сложение матриц

Вычитание матриц. При вычитании действия аналогичны сложению. Команда =A2:C4 – D2:F4; <Shift>+<Ctrl>+<Enter> (рис. 10.2).

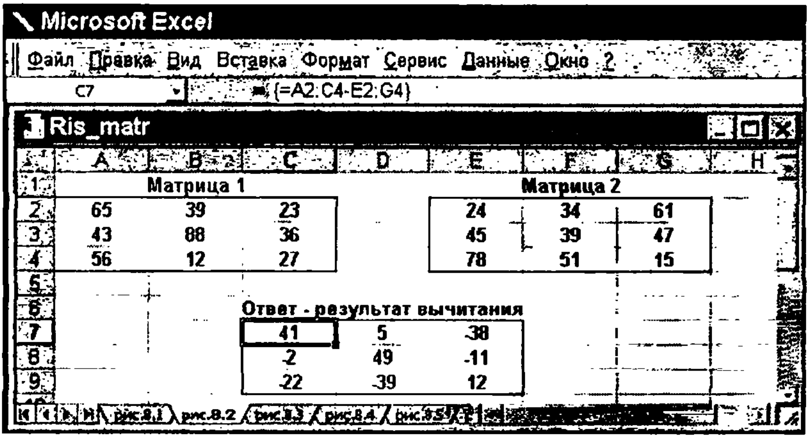


Рис. 10.2. Вычитание матриц

Умножение матрицы на число. При умножении матрицы на число также выделяется блок ячеек под ответ и вводится команда умножения на число, которая заканчивается нажатием <Shift>+<Ctrl>+<Enter>. Например: =A2:C5 * 255; <Shift>+<Ctrl>+<Enter> (рис. 10.3).

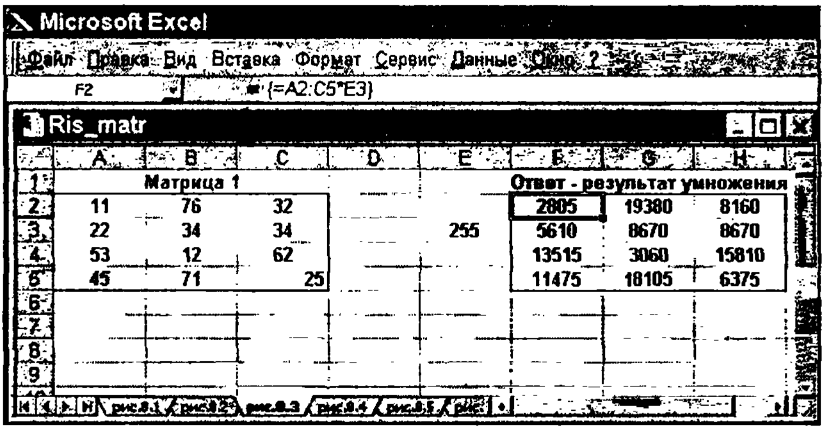


Рис. 10.3. Умножение матрицы на число

Умножение матрицы на матрицу. В данном случае используется матричная функция МУМНОЖ (*ММ*ULT) – рис. 10.4. Порядок действий следующий. Вводятся данные в виде матриц, выделяется область ячеек под ответ с числом строк, как у матрицы 1, и числом столбцов, как у матрицы 2. Вызывается функция МУМНОЖ (Мастер функций, категория **Все**).

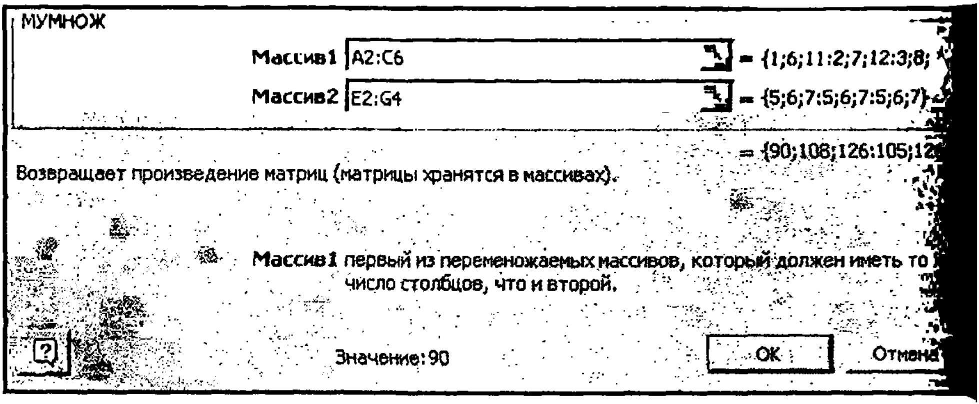


Рис. 10.4. Диалоговое окно функции МУМНОЖ

В поле **Массив1** вводятся данные первой матрицы, в поле **Массив2** – данные второй матрицы. Заканчивать ввод также необходимо командой <Shift>+<Ctrl>+<Enter> (рис. 10.4 и 10.5).

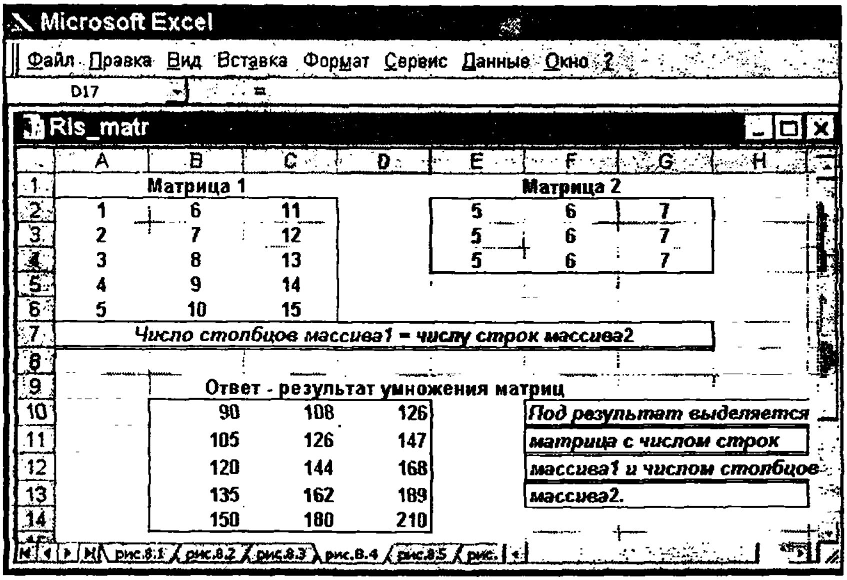


Рис. 10.5. Умножение матрицы на матрицу

Количество столбцов аргумента **Массив1** должно быть таким же, как количество строк аргумента **Массив2**. В противном случае функция МУМНОЖ возвращает значения ошибки – #ЗНАЧ!

10.4. Обращение матриц

Обратная матрица. Операцию деления для чисел можно свести к умножению, или ввести понятие обратного числа. Обратным относительно a называется такое число b, для которого выполняется условие $ab = 1$, в этом случае b является обратным a:

$$b = 1/a = a^{-1}.$$

Такого же подхода придерживаются и в матричной алгебре путем введения обратной матрицы.

Матрицей, обратной матрице A, будет являться матрица X, умножение на которую матрицы A даст единичную матрицу того же порядка I:

$$XA = AX = I.$$

Матрицу, обратную матрице A, принято обозначать символом A^{-1}. Тогда предыдущее равенство будет выглядеть:

$$A^{-1} A = A A^{-1} = I.$$

Пример. Дана матрица,

$$A=\begin{bmatrix} 7 & 8 \\ 5 & 6 \end{bmatrix},$$

обратная ей матрица есть

$$A^{-1}=\begin{bmatrix} 3 & -4 \\ -2{,}5 & 3{,}5 \end{bmatrix}.$$

Следовательно, имеем:

$$A\cdot A^{-1}=\begin{bmatrix} 7 & 8 \\ 5 & 6 \end{bmatrix}\cdot\begin{bmatrix} 3 & -4 \\ -2{,}5 & 3{,}5 \end{bmatrix}=\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}=I.$$

Транспонированная матрица. Операция замены строк на столбцы, а столбцов на строки называется транспонированием. Если провести операцию транспонирования с вектором-строкой

$$A = [a_1 a_2 \dots a_n],$$

то получим вектор-столбец

$$A^T=\begin{bmatrix} a_1 \\ a_2 \\ \dots \\ a_n \end{bmatrix}.$$

Матрица, в которой произведена замена строк на столбцы, а столбцов на строки называется транспонированной и обозначается символом A^T. Возьмем матрицу

$$A=\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ . & . & . & . \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix},$$

тогда транспонированная матрица будет иметь вид:

$$A^T = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ . & . & . & . \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{bmatrix}.$$

Определители (детерминанты). Каждой квадратной матрице однозначно соответствует некоторое число, которое называется определителем этой матрицы.

Для квадратной матрицы A n-го порядка

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ . & . & . & . \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

определитель будет обозначаться такой же квадратной матрицей, ограниченной двумя вертикальными чертами и состоящей из тех же элементов, и символьно - как $\det A$ или $\det[a_{ij}]$:

$$\det A = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ . & . & . & . \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}.$$

Число строк или столбцов определителя называется порядком этого определителя.

Определитель первого порядка есть число, равное его элементу:

$$\det A = \det[a] = |a| = a.$$

Определитель второго порядка устанавливается по формуле:

$$\det A = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

Введем два дополнительных понятия – минора определителя и алгебраического дополнения.

Минором называется определитель, у которого исключена произвольная строка (i-я) и произвольный столбец (j-й).

Алгебраическим дополнением D_{ij} элемента a_{ij} определителя называется число, полученное умножением минора на $(-1)^{i+j}$.

Теперь можно определить значение детерминанта произвольного порядка. Определитель

$$\det A = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ . & . & . & . \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

можно выразить через разложение по элементам i-й строки:

$$\det A = a_{i1}\,D_{i1} + a_{i2}\,D_{i2} + \ldots + a_{in}\,D_{in}$$

или по элементам j-го столбца:

$$\det A = a_{1j}\,D_{1j} + a_{2j}\,D_{2j} + \ldots + a_{nj}\,D_{nj}.$$

Установим определитель третьего порядка:

$$\det A = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.$$

Например, по элементам первой строки:

$$\det A = a_{11}\,D_{11} + a_{12}\,D_{12} + a_{13}\,D_{13},$$

где алгебраические дополнения равны (исключается первая строка и поочередно столбцы):

$$D_{11} = (-1)^{1+1}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix};\ D_{12} = (-1)^{1+2}\ D_{12} = (-1)^{1+2}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix};$$

$$D_{13} = (-1)^{1+3}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

или по элементам второй строки:

$$\det A = a_{21}\,D_{21} + a_{22}\,D_{22} + a_{23}\,D_{23},$$

где алгебраические дополнения равны (исключается вторая строка и поочередно столбцы):

$$D_{21} = (-1)^{2+1}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix};\ D_{22} = (-1)^{2+2}\begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix};\ D_{23} = (-1)^{2+3}\begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix},$$

аналогично по элементам третьего столбца:

$$\det A = a_{13}\,D_{13} + a_{23}\,D_{23} + a_{33}\,D_{33},$$

где алгебраические дополнения равны (исключается третий столбец и поочередно строки):

$$D_{21} = (-1)^{1+3}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix};\ D_{22} = (-1)^{2+3}\begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix};\ D_{23} = (-1)^{3+3}\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}.$$

Рассмотрим численный пример. Вычислим значение определителя

$$\begin{vmatrix} 1 & 4 & 3 \\ 2 & 5 & 1 \\ 3 & 6 & -4 \end{vmatrix},$$

например, по элементам второй строки:

$$\det A = 2\cdot D_{21} + 5\cdot D_{22} + 1\cdot D_{23}$$

$$D_{21} = (-1)^{2+1}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} = (-1)(a_{12}\cdot a_{33} - a_{13}\cdot a_{32}) = (-1)\cdot(4\cdot(-4) - 3\cdot 6) = 34;$$

$$D_{22} = (-1)^{2+2}\begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} = 1\cdot(a_{11}\cdot a_{33} - a_{13}\cdot a_{31}) = 1\cdot(1\cdot(-4) - 3\cdot 3) = -13;$$

$$D_{23} = (-1)^{2+3}\begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} = (-1)\cdot(a_{11}\cdot a_{32} - a_{12}\cdot a_{31}) = (-1)\cdot(1\cdot 6 - 4\cdot 3) = 6.$$

$$\det A = 2\cdot 34 + 5\cdot(-13) + 1\cdot 6 = 9.$$

Квадратная матрица, определитель которой равняется нулю, называется особой матрицей; в противном случае матрица называется неособой матрицей.

Вычисление обратных матриц с помощью определителей. Для квадратной матрицы n-го порядка можно записать матрицу D алгебраических дополнений, т. е. матрицу, где каждый элемент D_{ij} есть алгебраическое дополнение элемента a_{ij}:

$$D = \begin{bmatrix} D_{11} & D_{12} & \dots & D_{1n} \\ D_{21} & D_{22} & \dots & D_{2n} \\ . & . & . & . \\ D_{n1} & D_{n2} & \dots & D_{nn} \end{bmatrix}.$$

Транспонированная матрица матрицы D обозначается символом A^d:

$$A^d = \begin{bmatrix} D_{11} & D_{21} & \dots & D_{n1} \\ D_{12} & D_{22} & \dots & D_{n2} \\ . & . & . & . \\ D_{1n} & D_{2n} & \dots & D_{nn} \end{bmatrix}.$$

Произведение квадратной неособой матрицы и транспонированной матрицы алгебраических дополнений равняется произведению единичной матрицы на определитель матрицы, т. е.:

$$A \quad A^d = I \quad \det A.$$

Из свойств матриц следует, что

$$A \quad A^d = A^d \quad A = I \quad \det A.$$

Умножая обе части равенства на A^{-1}, получим:

$$A \quad A^{-1} \quad A^d = A^d \quad A \quad A^{-1} = I \quad \det A \quad A^{-1}.$$

Поскольку $A \quad A^{-1} = I$, имеем: $A^d = \det A \cdot A^{-1}$, откуда:

$$A^{-1} = \frac{1}{\det A} A^d.$$

То есть обратная матрица равна произведению транспонированной матрицы алгебраических дополнений на величину, обратную определителю матрицы.

Пример. Рассмотрим предыдущий пример – матрицу, определитель которой равен 9. Необходимо рассчитать еще 6 алгебраических дополнений:

$$D_{11} = \begin{vmatrix} 5 & 1 \\ 6 & -4 \end{vmatrix} = -26; \ D_{12} = -\begin{vmatrix} 2 & 1 \\ 3 & -4 \end{vmatrix} = 11; \ D_{13} = \begin{vmatrix} 2 & 5 \\ 3 & 6 \end{vmatrix} = -3;$$

$$D_{31} = \begin{vmatrix} 4 & 3 \\ 5 & 1 \end{vmatrix} = -11; \ D_{32} = -\begin{vmatrix} 1 & 3 \\ 2 & 1 \end{vmatrix} = 5; \ D_{33} = \begin{vmatrix} 1 & 4 \\ 2 & 5 \end{vmatrix} = -3.$$

Запишем матрицу алгебраических дополнений:

$$D=\begin{bmatrix}-26 & 11 & -3\\ 34 & -13 & 6\\ -11 & 5 & -3\end{bmatrix}.$$

Транспонированная матрица алгебраических дополнений будет иметь вид:

$$A^{d}=\begin{bmatrix}-26 & 34 & -11\\ 11 & -13 & 5\\ -3 & 6 & -3\end{bmatrix},$$

Следовательно, обратная матрица будет равна:

$$A^{-1}=\frac{1}{9}\begin{bmatrix}-26 & 34 & -11\\ 11 & -13 & 5\\ -3 & 6 & -3\end{bmatrix}=\begin{bmatrix}-2{,}89 & 3{,}78 & -1{,}22\\ 1{,}22 & -1{,}44 & 0{,}55\\ -0{,}33 & 0{,}67 & -0{,}33\end{bmatrix}.$$

10.5. Транспонирование, вычисление определителя и обращение матриц в Excel

Транспонирование матрицы. Операция замены строк на столбцы, а столбцов на строки называется транспонированием. Для выполнения этой операции имеется функция ТРАНСП (MTRANS). Ввод нужно также заканчивать комбинации <Shift>+<Ctrl>+<Enter>. На рис. 10.6 показана матрица 5×3, которая при транспонировании превращается в матрицу 3×5.

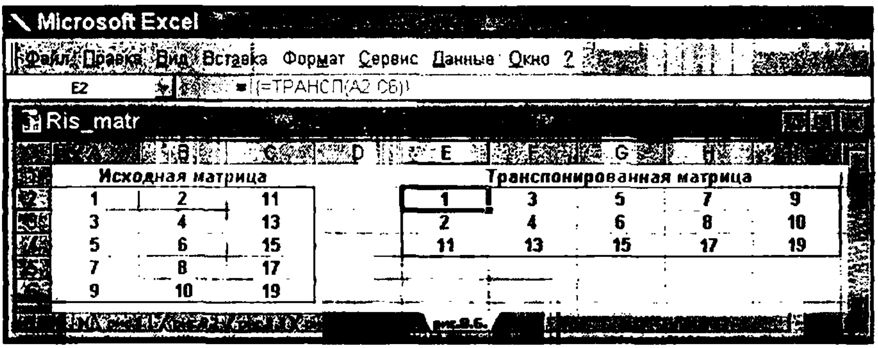


Рис. 10.6. Транспонирование матрицы

Кроме того, операцию транспонирования можно выполнить командой **Специальная вставка**. Для этого необходимо скопировать исходную матрицу, из меню **Правка** вызвать окно **Специальная вставка**, выбрать переключатель **Значения** и установить флажок **Транспонировать** (рис. 10.7).

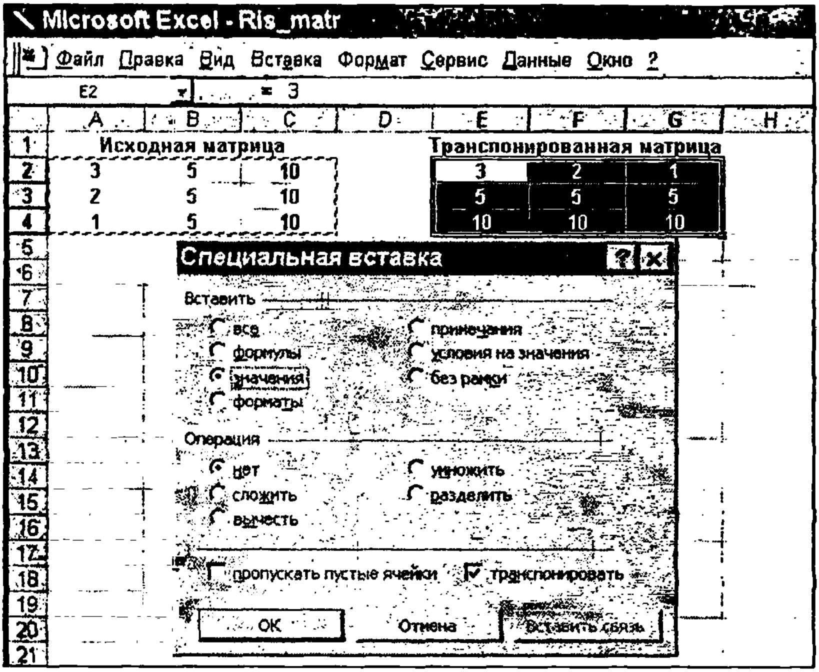


Рис. 10.7. Транспонирование матрицы через специальную вставку

Вычисление определителя матрицы. Для выполнения этой операции в Excel существует функция МОПР (MDETERM). На рис.10.8 приведен пример вычисления определителя.

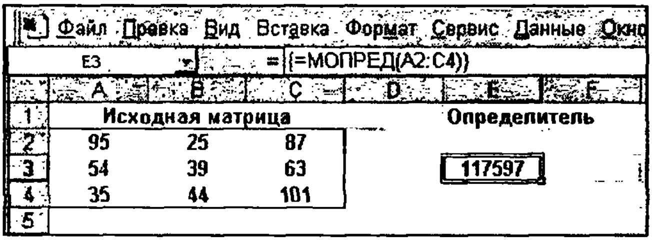


Рис. 10.8. Вычисление определителя матрицы

Функция предназначена для квадратных матриц, поэтому в противном случае функция выдает значение ошибки.

Обращение матриц. Эта операция выполняется с помощью функции МОБР (MINVERSE). Ввод нужно также заканчивать нажатием комбинации клавиш <Shift>+<Ctrl>+<Enter> (рис. 10.9).

Если перемножить исходную матрицу и обратную ей, то получим единичную матрицу. Для матриц, которые не могут быть обращены и определитель которых равен нулю, будет выводиться значение ошибки – #Число!

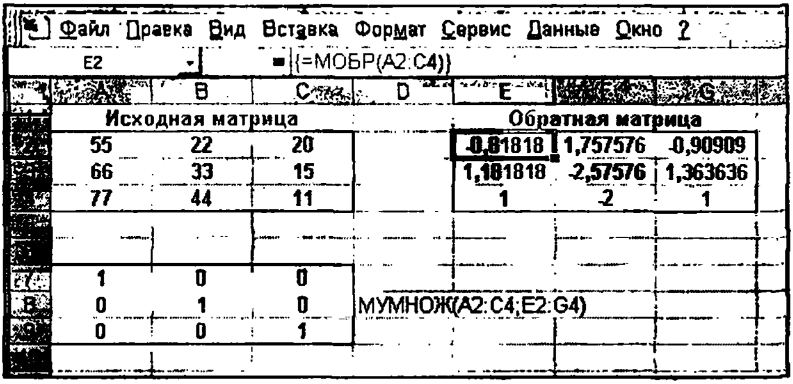


Рис. 10.9. Вычисление обратной матрицы

Редактирование матричных формул. Поскольку матричные формулы действуют на все ячейки матрицы, то изменять часть матрицы нельзя. При таких попытках выводится сообщение: "Нельзя изменять часть массива". Чтобы выполнить операцию по изменению части массива, необходимо активизировать любую ячейку в матрице и щелкнуть мышью в строке формул. При этом пропадут фигурные скобки. После этого выполняется редактирование, которое нужно закончить комбинацией клавиш <Shift>+<Ctrl>+<Enter>.

10.6. Системы линейных уравнений

Решение системы линейных уравнений. Рассмотрим систему n линейных уравнений с n неизвестными:

$$a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1;$$
$$a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2;$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n.$$

Решением системы называется совокупность n значений неизвестных $x_1, x_2, \ldots, x_n$, при подстановке которых все уравнения системы обращаются в тождества. Используя матричное представление системы линейных уравнений и понятие обратной матрицы, можно сразу получить решение системы линейных уравнений.

Введем матрицы:

$$A=\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ . & . & . & . \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix},\quad X=\begin{bmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{bmatrix},\quad B=\begin{bmatrix} b_1 \\ b_2 \\ \ldots \\ b_n \end{bmatrix},$$

где A – матрица коэффициентов при неизвестных; X – матрица неизвестных; B – матрица свободных членов.

Теперь можно записать систему линейных уравнений в матричном представлении:

$$AX = B.$$

Умножая обе части последнего равенства на обратную матрицу A^{-1}, получим:

$$A^{-1}AX = A^{-1}B.$$

Учитывая, что

$$A^{-1}A = I, IX = X,$$

окончательно имеем:

$$X = A^{-1}B.$$

Пример 1. Решить систему уравнений:

$$2x_1 + 4x_2 = 12,$$
$$3x_1 + 5x_2 = 20.$$

Матрица A коэффициентов при неизвестных есть:

$$A=\begin{bmatrix} 2 & 4 \\ 3 & 5 \end{bmatrix}.$$

Определитель будет равен:

$$\det A=\begin{vmatrix} 2 & 4 \\ 3 & 5 \end{vmatrix} = 2\cdot5-3\cdot4=-2.$$

Алгебраические дополнения матрицы A:

$$D_{11} = 5; D_{12} = -3; D_{21} = -4; D_{22} = 2.$$

Матрица алгебраических дополнений:

$$D=\begin{bmatrix} 5 & -3 \\ -4 & 2 \end{bmatrix}.$$

Транспонированная матрица равна:

$$A^d=\begin{bmatrix} 5 & -4 \\ -3 & 2 \end{bmatrix}.$$

Обратная матрица равна:

$$A^{-1}=\frac{1}{\det A}A^d=-\frac{1}{2}\begin{bmatrix} 5 & -4 \\ -3 & 2 \end{bmatrix}=\begin{bmatrix} -2{,}5 & 2 \\ 1{,}5 & -1 \end{bmatrix}.$$

Следовательно, решение системы уравнений есть:

$$X=\begin{bmatrix} -2{,}5 & 2 \\ 1{,}5 & -1 \end{bmatrix}\cdot\begin{bmatrix} 12 \\ 20 \end{bmatrix}=\begin{bmatrix} 10 \\ -2 \end{bmatrix},$$

т. е. $x_1 = 10$; $x_2 = -2$.

Пример 2. Решить систему уравнений:

$$x_1 + 4x_2 + 3x_3 = 10;$$
$$2x_1 + 5x_2 + x_3 = 20;$$
$$3x_1 + 6x_2 - 4x_3 = 5.$$

Матрица коэффициентов при неизвестных системы есть:

$$A=\begin{bmatrix} 1 & 4 & 3 \\ 2 & 5 & 1 \\ 3 & 6 & -4 \end{bmatrix}.$$

Данную матрицу мы уже рассматривали. Определитель для нее равен:

$$\det A = 9.$$

Матрица алгебраических дополнений:

$$D=\begin{bmatrix} -26 & 11 & -3 \\ 34 & -13 & 6 \\ -11 & 5 & -3 \end{bmatrix}.$$

Транспонированная матрица алгебраических дополнений имеет вид:

$$A^d=\begin{bmatrix} -26 & 34 & -11 \\ 11 & -13 & 5 \\ -3 & 6 & -3 \end{bmatrix},$$

обратная матрица равна:

$$A^{-1}=\frac{1}{9}\begin{bmatrix} -26 & 34 & -11 \\ 11 & -13 & 5 \\ -3 & 6 & -3 \end{bmatrix}=\begin{bmatrix} -2{,}89 & 3{,}78 & -1{,}22 \\ 1{,}22 & -1{,}44 & 0{,}55 \\ -0{,}33 & 0{,}67 & -0{,}33 \end{bmatrix}.$$

Решение матричного уравнения $X = A^{-1} B$ есть матрица:

$$X=\begin{bmatrix} -2{,}89 & 3{,}79 & -1{,}22 \\ 1{,}22 & 1{,}44 & 0{,}55 \\ -0{,}33 & 0{,}66 & -0{,}33 \end{bmatrix}\cdot\begin{bmatrix} 10 \\ 20 \\ 5 \end{bmatrix}=\begin{bmatrix} 40{,}56 \\ 13{,}88 \\ 8{,}34 \end{bmatrix},$$

следовательно, решение системы есть:

$$x_1 = 40{,}56;\ x_2 = 13{,}88;\ x_3 = 8{,}34.$$

10.7. Решение системы линейных уравнений в Excel

Решим систему уравнений:

$$5x_1 + 6x_2 - x_3 + x_4 = 3;$$
$$4x_1 - x_2 - 5x_3 + 2x_4 = 2;$$
$$9x_1 + 3x_2 - 10x_3 + x_4 = 6;$$
$$10x_1 + 2x_2 + 3x_3 + 2x_4 = 1.$$

Решение будет заключаться в умножении обратной матрицы коэффициентов при неизвестных на матрицу свободных членов. Эти операции можно выполнить последовательно, т. е. сначала определить обратную матрицу коэффициентов при неизвестных при помощи функции МОБР, а затем полученную обратную матрицу умножить на матрицу свободных членов при помощи функции МУМНОЖ. Но можно выполнить и быстрее: сначала вызывается функция МУМНОЖ, в диалоговом окне которой вызывается встроенная функция у первого массива, где в свою очередь вызывается функция обращения и вводится матрица коэффициентов. Для второго массива диалогового окна функции МУМНОЖ вводится диапазон матрицы свободных членов. Ввод заканчивается комбинацией клавиш <Shift>+<Ctrl>+<Enter>. Например, если матрица коэффициентов записана в диапазоне A2:D5, а матрица свободных членов – в диапазоне G2:G5, то формула выглядит так:

{=МУМНОЖ(МОБР(A2:D5); G2:G5)}

Решение системы уравнений приведено на рис. 10.10.

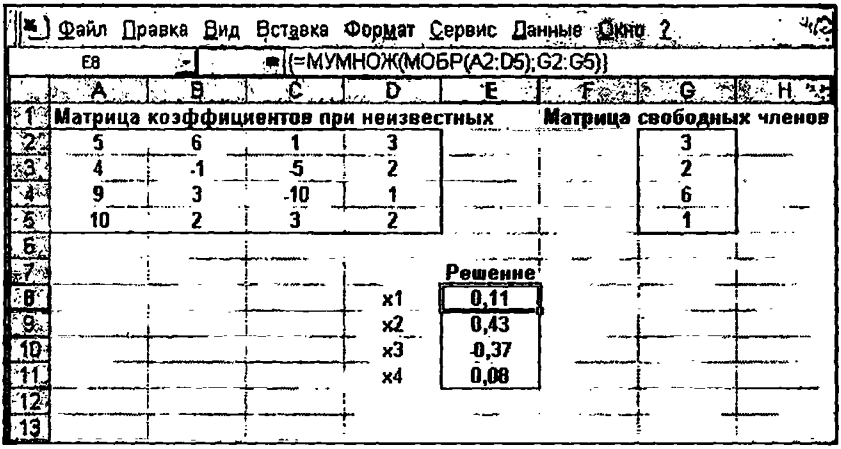
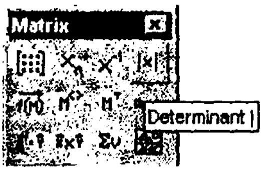


Рис. 10.10. Решение системы линейных уравнений

10.8. Работа с матрицами в Mathcad



Для задания векторов и матриц в Mathcad можно воспользоваться командой **Matrices...** (Матрицы...) подменю **Math** (Математика) основного меню, нажатием клавиш <Ctrl>+<V> или выбором пиктограммы с изображением шаблона матрицы.

Это вызывает появление диалогового окна (рис. 10.11), в котором надо указать размерность матрицы, т. е. количество строк (Rows) и столбцов (Columns). Нажатием клавиши <Enter> или щелчком по кнопке **Insert** в диалоговом окне выводится шаблон матрицы или вектора (вектор будет иметь один из параметров размерности, равный 1).

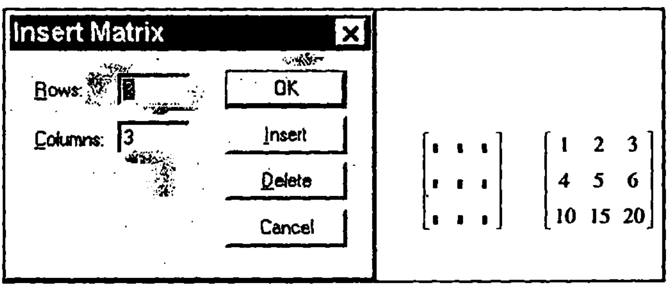


Рис. 10.11. Вывод шаблонов матриц и их заполнение

Шаблон имеет обрамляющие рамки и места ввода значений (числовых или символьных) для элементов матрицы. Одно из мест ввода значений активизируется курсором мыши. При этом оно выделяется синим уголком. Теперь вводятся все элементы матрицы путем перемещения по шаблону с помощью мыши или клавиш перемещения курсора. Если ввод закончен, но не заполнены все элементы шаблона, система выведет сообщение об ошибке – «**missing operand**» (пропущенный операнд). Это сообщение дается красным цветом в рамке с линией, указывающей на незаполненный шаблон или на несколько незаполненных шаблонов.

Каждый элемент матрицы характеризуется индексированной переменной, и его положение в матрице определяется двумя индексами: один указывает номер строки, другой – номер столбца. Для указания подстрочных индексов после имени переменной (V – вектор, M – матрица) вводится знак открывающей квадратной скобки «[». Для элементов матрицы подстрочные индексы вводятся в круглых скобках с разделением их запятыми:

Ввод (изображение в окне):

$$M[(3,4) : M_{3,4} :=$$

Индексы могут начинаться с нуля или с единицы. Нижняя граница индексов задается значением системной переменной ORIGIN. Для того чтобы нумерация индексов начиналась с единицы, надо, выполнив команду **Math – Options – Build – In Variable...** (Встроенные переменные...), открыть опцию **ORIGIN** и установить необходимую нижнюю границу значения индексации – 1.

Если введены две матрицы, то путем ввода соответствующей операции (умножения, сложения, вычитания, деления) матрицы могут перемножаться, складываться, вычитаться, умножаться на число.

Для работы с векторами и матрицами система Mathcad содержит ряд операторов и функций. Вначале рассмотрим операторы, введя следующие обозначения: для векторов – V, для матриц – M и для скалярных величин – Z. Для того чтобы воспользоваться операторами, необходимо изначально матрице дать соответствующее обозначение, например:

$$M := \begin{bmatrix} 3 & 7 \\ 4 & 9 \end{bmatrix}.$$

В табл. 10.1 представлены некоторые операторы для работы с матрицами.

Таблица 10.1

Оператор	**Ввод**	**Назначение оператора**
M	–M	Смена знака у элементов матрицы
Z*M, M*Z	Z*M, M*Z	Умножение матрицы на скаляр Z
M1*M2	M1*M2	Умножение двух матриц M1 и M2
$\frac{M}{Z}$	M/Z	Деление матрицы на скаляр Z
M^{-1}	M^–1	**Обращение матрицы M**
M^{n}	M^*n*	Возведение матрицы в степень *n*

Таблица 10.1 (окончание)

Оператор	Ввод	Назначение оператора
$\lvert M\rvert$	$\lvert M\rvert$	Вычисление определителя матрицы
M^T	M <Ctrl>+<!>	Транспонирование матрицы M
$\vec{M}$	M <Ctrl> +<–>	Векторизация матрицы M
$M^{<n>}$	M <Ctrl>^*n*	Выделение *n*–го столбца матрицы M
$M_{m,n}$	M [(*m*,*n*)	Выделение элемента (*m*, *n*) матрицы M

Применение операторов для работы с матрицами показано на рис. 10.12.

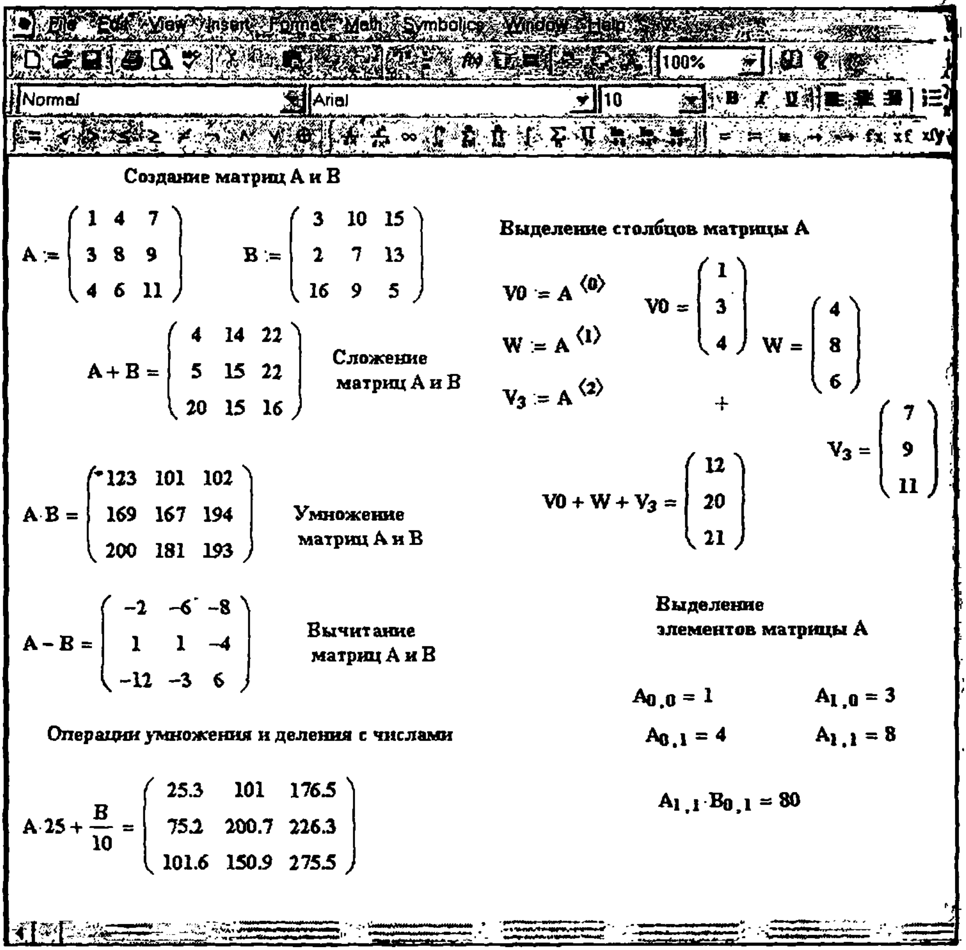


Рис. 10.12. Матричные операции

Для работы с матрицами также существует ряд встроенных функций:

identity(*n*) – создает единичную квадратную матрицу размером $n \times n$;

augment(M1, M2) – объединяет в одну матрицы M1 и M2, имеющие одинаковое число строк (объединение идет сторона к стороне);

stack(M1, M2) – объединяет две матрицы M1 и M2, имеющие одинаковое число столбцов, размещая M1 над M2;

submatrix(A, ir,jr, ic, jc) – возвращает субматрицу, состоящую из всех элементов, содержащихся в строках с ir по jr и столбцах с ic по jc (ir ≤ jr, ic ≥ jc);

diag(v) – создает диагональную матрицу, элемент главной диагонали которой – вектор V.

Функции, возвращающие специальные характеристики матриц:

cols(M) – возвращает число столбцов матрицы M;

rows(M) – возвращает число строк матрицы M;

rank(M) – возвращает ранг матрицы M;

tr(M) – возвращает след (сумму диагональных элементов) квадратной матрицы M;

mean(M) – возвращает среднее значение элементов массива M;

median(M) – возвращает медиану элементов массива M;

lu(M) – выполняет треугольное разложение матрицы M: P*M=L*U, где L и U соответственно нижняя и верхняя треугольные матрицы. Все четыре матрицы квадратные, одного порядка;

qr(A) – дает разложение матрицы A, A=Q*R, где Q – ортогональная матрица и R – верхняя треугольная матрица;

geninv(A) – левая обратная матрице A. L*A=I, где I – единичная матрица размером n×n, L – прямоугольная матрица размером n×m, A – прямоугольная матрица размером m×n;

sort(V) – сортировка элементов векторов в порядке возрастания их значений;

eigenvals(M) – возвращает вектор, содержащий собственные значения матрицы M.

На рис. 10.13. приведен ряд примеров применения матричных функций.

Для решения *систем линейных уравнений в Mathcad имеется* функция

lsolve(A,B),

которая возвращает вектор *X* для системы линейных уравнений $A \cdot X = B$ при заданной матрице коэффициентов *A* и векторе свободных членов *B*.

Для примера возьмем систему уравнений:

$$x_1 + 2x_2 + 3x_3 - 2x_4 = 6;$$
$$2x_1 - x_2 - 2x_3 - 3x_4 = 8;$$
$$3x_1 + 2x_2 - x_3 + 2x_4 = 4;$$
$$2x_1 - 3x_2 + 2x_3 + x_4 = -8.$$

Пример решения данной системы уравнений дан на рис. 10.14.

$C := A^T \quad C = \begin{pmatrix} 1 & 3 & 4 \\ 4 & 8 & 6 \\ 7 & 9 & 11 \end{pmatrix}$ **Транспонирование матрицы А**

$D := A^{-1} \quad D = \begin{pmatrix} -0.654 & 0.038 & 0.385 \\ -0.058 & 0.327 & -0.231 \\ 0.269 & -0.192 & 0.077 \end{pmatrix}$ **Обращение матрицы А**

$Q1 := |A| \quad Q1 = -52$

$Q2 := |B| \quad Q2 = 324$

Вычисление определителей матрицы А и В

$\text{cols}(A) = 3$ **Определение числа столбцов матрицы А**

$\text{rows}(A) = 3$ **Определение числа строк матрицы А**

$\text{rank}(A) = 3$

$A2 := \begin{pmatrix} 0 & -1 & 3 & 0 & 2 \\ 2 & -4 & 1 & 5 & 3 \\ -4 & 5 & 7 & -10 & 0 \\ -2 & 1 & 8 & -5 & 3 \end{pmatrix}$ **Определение ранга матриц А, А2**

$\text{rank}(A2) = 2$

$M1 := \text{identity}(2) \quad M1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

$M2 := \text{identity}(3) \quad M2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

Создание единичных матриц с помощью функции identity

$\text{tr}(A) = 20 \quad \text{tr}(M2) = 3$ **Вычисление следа матриц**

$V4 := \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$

$\text{augment}(V4, M2) = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 \\ 3 & 0 & 0 & 1 \end{pmatrix}$ **Объединение вектора V4 и матрицы М2**

$\text{augment}(M2, V4) = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{pmatrix}$

Рис. 10.13. Примеры матричных операций

Матрица коэффициентов системы

$$A := \begin{pmatrix} 1 & 2 & 3 & -2 \\ 2 & -1 & -2 & -3 \\ 3 & 2 & -1 & 2 \\ 2 & -3 & 2 & 1 \end{pmatrix}$$

Вектор свободных членов

$$B := \begin{pmatrix} 6 \\ 8 \\ 4 \\ -8 \end{pmatrix}$$

$$X := A^{-1} \cdot B$$

Решение системы

$$X = \begin{pmatrix} 1 \\ 2 \\ -1 \\ -2 \end{pmatrix}$$

Результат решения

или **Решение системы с помощью функции lsolve**

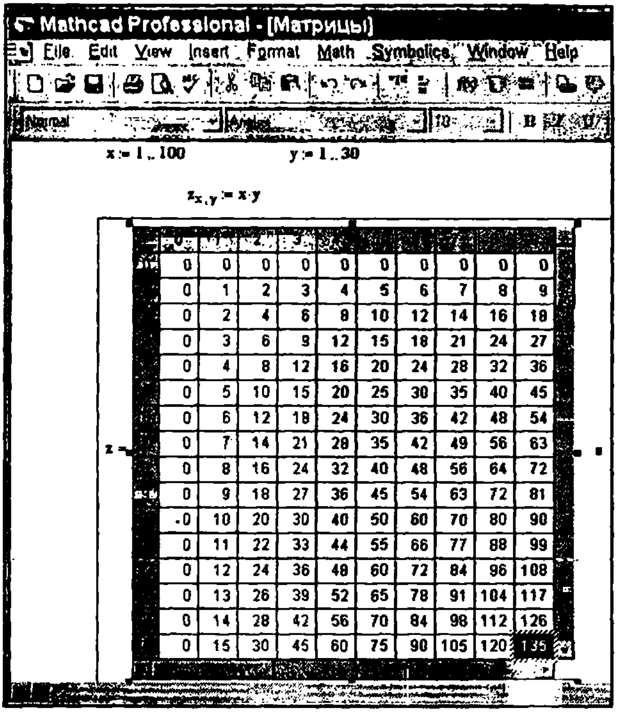
$$X1 := \mathrm{lsolve}(A, B)$$

$$X1 = \begin{pmatrix} 1 \\ 2 \\ -1 \\ -2 \end{pmatrix}$$

Рис. 10.14. Решение системы линейных уравнений

Система Mathcad способна работать с векторами и матрицами больших размеров. Однако вывод таблиц со значениями их элементов может вызвать определенные трудности. Поэтому обычная форма вывода таблицы действует, только если число элементов вектора не выше 50. При большем числе элементов (у матрицы в столбцах и в строках) форма вывода таблицы меняется. Это показано на рис. 10.15.

Таблица вывода в этом случае содержит затемненные графы со значениями индексов элементов отображаемой матрицы. Если выделить таблицу рамкой, появляются полосы прокрутки, с помощью которых можно перемещать видимую часть таблицы по общему массиву элементов большой матрицы.



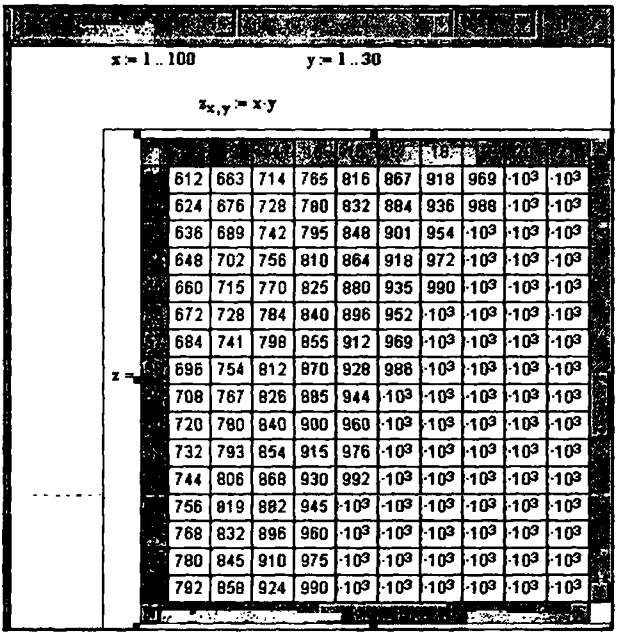


Рис. 10.15. Вывод матрицы большого размера

10.9. Дифференцирование по векторному аргументу

Часто приходится дифференцировать выражения, содержащие векторы и матрицы. Рассмотрим сначала:

$$a^T x = [a_1 a_2 \ldots a_n]\{x_1 \ldots x_n\} = a_1 x_1 + a_2 x_2 + \ldots + a_n x_n.$$

Если мы вычислим частные производные выражения $a^T x$ по скалярной переменной x_i, то получим:

$$\frac{\partial(a^T x)}{\partial x_1} = a_1, \quad \frac{\partial(a^T x)}{\partial x_2} = a_2, \quad \ldots, \quad \frac{\partial(a^T x)}{\partial x_n} = a_n;$$

$$\frac{\partial}{\partial x_1}(x^T A x) = 2(a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n)$$

и увидим, что частные производные есть просто элементы вектора a. Поэтому если мы образуем поочередно n частных производных и затем расположим их в виде вектора a, то сможем рассматривать этот процесс как дифференцирование по вектору, определяемое следующим образом:

$$\frac{\partial(a^T x)}{\partial x} = a,$$

где левая часть уравнения обозначает операцию дифференцирования по элементам вектора x.

Рассмотрим теперь квадратичную форму. Квадратичная форма называется симметрической, если A – симметрическая матрица ($a_{mn} = a_{nm}$):

$$x^T A x = [x_1 x_2 \ldots x_n] \cdot \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{12} & a_{22} & \ldots & a_{2n} \\ . & . & \ldots & . \\ a_{1n} & a_{2n} & \ldots & a_{nn} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ . \\ x_n \end{bmatrix} =$$

$$= a_{11}x_1^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + \ldots + 2a_{1n}x_1x_n +$$
$$+ a_{22}x_2^2 + 2a_{23}x_2x_3 + 2a_{23}x_2x_3 + \ldots + 2a_{2n}x_2x_n +$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots + a_{nn}x_n^2.$$

Возьмем частные производные по элементам вектора x и получим:

$$\frac{\partial}{\partial x_1}(x^T A x) = 2(a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n);$$

$$\frac{\partial}{\partial x_2}(x^T A x) = 2(a_{12}x_1 + a_{22}x_2 + a_{23}x_3 + \ldots + a_{2n}x_n);$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\frac{\partial}{\partial x_n}(x^T A x) = 2(a_{1n}x_1 + a_{2n}x_2 + a_{3n}x_3 + \ldots + a_{nn}x_n).$$

Правые части этих уравнений содержат элементы матричного произведения Ax, которые образуют n-мерный вектор-столбец. С другой стороны, мы можем рассматривать правые части этих уравнений как элементы матричного произведения $x^T A$, образующие вектор-строку из n элементов. Итак,

$$\frac{\partial}{\partial x}(x^T Ax) = 2Ax,$$

$$\frac{\partial}{\partial x}(x^T Ax) = 2x^T A.$$

Выбор между данными уравнениями обычно определяется контекстом, в рамках которого осуществляется дифференцирование, поскольку в матричных уравнениях мы можем приравнивать только матрицы одного порядка и не можем вектор-строку приравнять к вектору-столбцу.

Рассмотрим теперь случай, когда y обозначает n-мерный вектор-столбец, каждый из элементов которого есть функция m элементов вектора x, т. е.

$$y_i = f_i(x_1, x_2, \ldots, x_m), \quad i = 1, \ldots, n.$$

Каждую из функций y_i можно продифференцировать частным образом по каждому из x_i, что приведет всего к mn частным производным. Если эти частные производные расположить в виде матрицы порядка $m \times n$, то результат будет выглядеть так:

$$\frac{\partial y}{\partial x} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_2}{\partial x_1} & \cdots & \frac{\partial y_n}{\partial x_1} \\ \frac{\partial y_1}{\partial x_2} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_n}{\partial x_2} \\ \cdot & \cdot & \cdots & \cdot \\ \frac{\partial y_1}{\partial x_m} & \frac{\partial y_2}{\partial x_m} & \cdots & \frac{\partial y_n}{\partial x_m} \end{bmatrix}.$$

Такая матрица называется матрицей Якоби или якобианом. Дифференцирование первого порядка позволяет обнаружить локальное стационарное значение. Чтобы отличить точки максимума от точек минимума, требуются производные второго порядка.

Так, если для

$$y = f(x_1, x_2, \ldots, x_n)$$

положим

$$\frac{\partial y}{\partial x} = 0,$$

то найдем вектор x^0, удовлетворяющий предыдущему уравнению, в котором значение y стационарно. Это стационарное значение будет минимумом, если

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 y}{\partial x_i \partial x_j} dx_i dy_j > 0$$

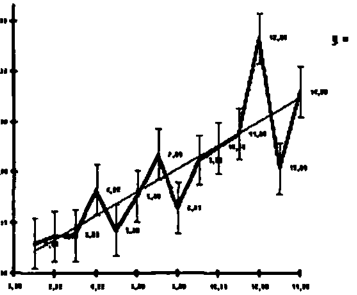
для любого набора элементов dx_i. Выразив данное выражение в матричной форме, получим:

$$\frac{\partial^2 y}{\partial x^2} = \frac{\partial(\partial y/\partial x)}{\partial x} = \begin{bmatrix} \frac{\partial^2 y}{\partial x^2} & \frac{\partial^2 y}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 y}{\partial x_1 \partial x_n} \\ \frac{\partial^2 y}{\partial x_2 \partial x_1} & \frac{\partial^2 y}{\partial x_2^2} & \cdots & \frac{\partial^2 y}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial^2 y}{\partial x_n \partial x_1} & \frac{\partial^2 y}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 y}{\partial x_n^2} \end{bmatrix}.$$

Данная матрица получила название матрицы Гессе или гессиана. Условие в предыдущем выражении (>0) означает, что матрица Гессе должна быть положительно определенной, что говорит о наличии экстремума. Подробней данный вопрос излагается в главе 15.

Глава 11. Корреляционно-регрессионный анализ

11.1. Экономические модели и статистические методы



Закономерности в экономике выражаются в виде связей и свойств экономических показателей, а также в виде математических моделей их поведения. Такие зависимости и модели могут быть получены только путем обработки реальных статистических данных с учетом внутренних механизмов связи и случайных факторов.

Математическая статистика (т. е. теория обработки и анализа данных) и ее применение в экономике – эконометрика – позволяют строить экономические модели и оценивать их параметры, проверять гипотезы о свойствах экономических показателей и формах их связи. В основе методов, с помощью которых решаются данные задачи, лежит **корреляционно-регрессионный анализ.**

Все факторы, неучтенные явно в экономической модели, оказывают на объект некоторое результирующее воздействие, величина которого неизвестна заранее и может быть описана как случайная функция. Для ее описания в модель добавляют случайный параметр ε, интегрирующий в себе влияние всех неучтенных факторов, например:

$$z = f(x, y) + \varepsilon.$$

Введение случайного компонента в экономическую модель приводит к тому, что взаимосвязь остальных ее переменных перестает быть строго детерминированной и становится стохастической, что и наблюдается в реальной действительности.

Работа с эконометрическими моделями требует использования инструментария оценивания и статистической проверки модели, а также решения проблем выбора типа модели, набора объясняющих переменных и вида связей между ними.

11.2. Корреляция между количественными признаками

В математической статистике взаимосвязь явлений и их признаков изучают с помощью метода корреляции (*correlation* (англ.) – соотношение, соответствие, сопоставление).

При обработке и использовании статистических данных с целью получения как научных, так и практических выводов важно проследить, как изменяется один признак при изменении другого, т. е. найти уравнение связи и значения коэффициента и индекса корреляции и детерминации, определяющие степень влияния одного признака на другой.

Связи между явлениями или их признаками, проявляющиеся в изменении статистических характеристик распределения одного признака с изменением значений другого, называются статистическими.

Результативный признак в этих связях не полностью определяется влиянием факториального признака. Это влияние проявляется лишь в среднем, при этом отдельные результаты могут противоречить установленной связи.

В функциональных связях факториальный признак полностью определяет величину результативного признака.

11.3. Корреляционный момент. Коэффициент корреляции

Числовыми характеристиками случайной величины X являются центральные моменты различных порядков. Важнейшие из них: математическое ожидание X и дисперсия σ^2_x.

Особую роль играет второй смешанный центральный момент (т. е. математическое ожидание произведения центрированных величин), называемый корреляционным моментом или моментом связи случайных величин X, Y:

$$\tau = M[(x - X_0)(y - Y_0)].$$

Для дискретных величин момент связи определяют по формуле

$$\tau = \sum\sum(x - X_0)(y - Y_0)\, P_{iu},$$

где $P_{iu} = P(x_i - X)(y_u - Y)$ – вероятность того, что система X, Y примет значения x_i, y_u, а суммирование распространяют по всем возможным значениям случайных величин X, Y.

Для непрерывных величин момент связи определяют по формуле

$$\tau^1 = \int_{-\infty}^{+\infty}\!\!\int (x - X_0)(y - Y_0) f(x,y)\,dx\,dy,$$

где $f(x,y)$ – плотность распределения случайных величин x, y.

Для независимых случайных величин плотность распределения системы равна произведению плотностей распределения отдельных величин, входящих в систему

$$f(x, y) = f_1(x)\, f_2(y).$$

Подставив значение $f(x,y)$ в предыдущую формулу, получим:

$$\tau^1 = \int(x - X_0)\, f_1(x)\, dx \cdot \int(y - Y_0)\, f_2(y)\, dy.$$

Корреляционный момент характеризует не только рассеивание величин X и Y, но и связь между ними.

Каждый из интегралов последней формулы в отдельности есть первый центральный момент случайных величин X и Y. Следовательно, каждый интеграл равен нулю. Таким образом, для независимых случайных величин $\tau^1 = 0$. Если значение корреляционного момента двух случайных величин не равно нулю, то между ними имеется зависимость.

Для характеристики связи между величинами X, Y пользуются безразмерной характеристикой, называемой корреляционным отношением или индексом корреляции:

$$r_{x/y} = \frac{\tau_{x/y}}{\sigma_x \sigma_y}, \text{ аналогично } r_{y/x} = \frac{\tau_{y/x}}{\sigma_x \sigma_y}.$$

Очевидно, для независимых случайных величин $r = 0$.

При прямолинейной зависимости корреляционное отношение называют коэффициентом корреляции. Коэффициент корреляции изменяется в пределах $-1 \le r \le 1$.

Если $r_{x/y} < 0$, то корреляция отрицательная, если $r_{x/y} > 0$, то корреляция положительная. Величина коэффициента корреляции определяет тесноту связи между случайными величинами; чем ближе значение $r_{x/y}$ к единице, тем теснее статистическая связь.

Близкое к нулю значение коэффициента корреляции свидетельствует об отсутствии прямолинейной статистической связи.

Положительная корреляция между случайными величинами характеризует такую зависимость между ними, когда при возрастании одной из них другая также в среднем возрастает.

Отрицательная корреляция характеризует такую зависимость, когда при возрастании одной случайной величины другая в среднем убывает.

Корреляция – это статистическая зависимость между случайными величинами, не имеющими строго функционального характера, при которой изменение одной из случайных величин приводит к изменению математического ожидания другой.

В статистике принято различать следующие варианты зависимостей.

1. Парная корреляция – связь между двумя признаками (результативным и факторным или двумя факторными).

2. Частная корреляция – зависимость между результативным и одним факторным признаками при фиксированном значении других факторных признаков.

3. Множественная корреляция – зависимость результативного и двух или более факторных признаков, включенных в исследование.

Корреляционный анализ имеет своей задачей количественное определение тесноты связи между двумя признаками (при парной связи) и между результативным и множеством факторных признаков (при многофакторной связи).

Теснота связи количественно выражается величиной коэффициентов корреляции. Коэффициенты корреляции, представляя количественную характеристику тесноты связи между признаками, дают возможность определять «полезность» факторных признаков при построении уравнений множественной регрессии. Величина коэффициента корреляции служит также оценкой соответствия уравнения регрессии выявленным причинно-следственным связям.

Корреляция и регрессия тесно связаны между собой: первая оценивает силу (тесноту) статистической связи, вторая исследует ее форму. Та и другая служат для установления соотношения между явлениями, для определения наличия или отсутствия связи.

Корреляционно-регрессионный анализ как общее понятие включает в себя измерение тесноты, направление связи и установление аналитического выражения (формы) связи (регрессионный анализ).

11.4. Уравнение связи

Статистическую связь между признаками выражают с помощью такой математической функции, которая дает наименьшее отклонение от полученных при наблюдении значений признаков. Уравнение такой функции является уравнением связи между результативными и факториальными признаками.

Уравнение связи называют также уравнением регрессии.

Регрессионный анализ заключается в определении аналитического выражения связи, в котором изменение одной величины (называемой зависимой или результативным признаком) обусловлено влиянием одной или нескольких независимых величин (факторов), а множество всех прочих факторов, также оказывающих влияние на зависимую переменную, принимается за постоянные и средние значения. Регрессия может быть однофакторной (парной) и многофакторной (множественной).

По форме зависимости различают:

а) линейную регрессию, которая выражается уравнением прямой (линейной) функцией вида: $y = a + bx$;

б) нелинейную регрессию, которая выражается уравнениями степенной, показательной, экспоненциальной функций, а также уравнениями параболы и гиперболы вида:

парабола $y = a + bx + cx^2$;

гипербола $y = a + b/x$

и т. д.

По направлению связи различают:

а) прямую регрессию (положительную), возникающую при условии, если с увеличением или уменьшением независимой переменной значения зависимой переменной также соответственно увеличиваются или уменьшаются;

б) обратную (отрицательную) регрессию, появляющуюся при условии, что с увеличением или уменьшением независимой переменной зависимая переменная соответственно уменьшается или увеличивается.

11.5. Определение параметров уравнения линейной связи

При линейной связи параметры уравнений регрессии находят по способу наименьших квадратов.

Рассмотрим случай, когда имеется n наблюдений двух переменных x и y. Предположив, что y зависит от x, мы хотим подобрать уравнение

$$\hat{y} = a + bx. \tag{11.1}$$

Расчетное значение зависимой переменной $\hat{y}$ и остаток e_i для наблюдения i определяются уравнениями:

$$\hat{y}_i = a + bx_i;$$

$$e_i = y_i - \hat{y} = y_i - a - bx_i.$$

Сумма квадратов отклонений равна:

$$S = \sum e_i^2 = e_i^2 + \ldots + e_n^2 = \sum (y_i - \hat{y}_i).$$

Тогда требование наименьших квадратов записывается в виде:

$$S = \Sigma (y_i - \hat{y}_i)^2 \to \min, \tag{11.2}$$

т. е. сумма квадратов отклонений фактических ординат точек корреляционного поля от ординат, вычисленных по уравнению (11.1), должна быть наименьшей.

Отыскание параметров прямой, удовлетворяющей этому требованию, называется способом наименьших квадратов.

Для парной линейной регрессии задача заключается в отыскании неизвестных параметров a и b, минимизирующих сумму квадратов отклонений, т. е. функцию S:

$$S = \sum e^2_i = \sum (y_i - a - bx_i)^2.$$

Значения a и b, удовлетворяющие минимуму функции S, находят из уравнений:

$$\frac{\partial S}{\partial a} = 0, \qquad \frac{\partial S}{\partial b} = 0.$$

Выразим квадрат i-го остатка:

$$e_i^2 = (y_i - \hat{y}_i)^2 = (y_i - a - bx_i) = y_i^2 + a^2 + b^2 x_i^2 - 2ay_i + 2abx_i - 2bx_i y_i.$$

Суммируя по всем n наблюдениям, запишем S в виде:

$$S = \sum y_i^2 + na^2 + b^2 \sum x_i^2 - 2a \sum y_i + 2ab \sum x_i - 2b \sum x_i y_i.$$

Тогда условия минимизации функции S принимают вид:

$$\frac{\partial S}{\partial a} = 2an - 2\sum y_i + 2b\sum x_i = 0;$$

$$\frac{\partial S}{\partial b} = 2b\sum x_i^2 + 2a\sum x_i - 2\sum x_i y_i = 0.$$

Подставив в первое уравнение $n\bar{y}$ вместо $\sum y_i$ и $n\bar{x}$ вместо $\sum x_i$, получим:

$$2an - 2n\bar{y} + 2b\bar{x} = 0.$$

Следовательно,

$$a = \bar{y} - b\bar{x}.$$

Подставив полученное выражение для a во второе уравнение минимизации, имеем:

$$2b\sum x_i^2 + 2n\bar{x}\,\bar{y} - 2bn\bar{x}^2 - 2\sum x_i y_i = 0.$$

После преобразования получим:

$$b\left\{\frac{1}{n}\sum x_i^2 - \bar{x}^2\right\} = \frac{1}{n}\sum x_i y_i - \bar{x}\cdot\bar{y} \quad \text{или} \quad b\left\{\frac{1}{n}\sum x_i^2 - \bar{x}^2\right\} = \overline{xy} - \bar{x}\cdot\bar{y}.$$

Откуда

$$b = \frac{\overline{xy} - \bar{x}\cdot\bar{y}}{\frac{1}{n}\sum x_i^2 - \bar{x}^2} \quad \text{или} \quad b = \frac{\sum xy - n\bar{x}\cdot\bar{y}}{\sum x^2 - n\bar{x}^2}.$$

Так как показатель *ковариации* равен:

$$\mathrm{Cov}(x, y) = \frac{1}{n}\sum (x - \bar{x})(y - \bar{y}) = \frac{1}{n}\sum xy - \bar{x}\cdot\bar{y},$$

а *дисперсия* равна:

$$\mathrm{Var}(x) = \frac{1}{n}\sum (x - \bar{x}^2) = \frac{1}{n}\sum x^2 - \bar{x}^2,$$

то параметр b можно выразить как

$$b = \frac{\mathrm{Cov}(x, y)}{\mathrm{Var}(x)}.$$

Коэффициент *корреляции* определяется по формуле:

$$r = \frac{\overline{xy} - \overline{x} \cdot \overline{y}}{\sigma_x \sigma_y},$$

где $\overline{xy}$ – среднее значение произведений x и y; $\overline{x}$, $\overline{y}$ – средние значения соответствующих признаков; σ_x и σ_y – средние квадратические отклонения, найденные по признаку x и по признаку y.

Коэффициент корреляции можно также установить через показатель *ковариации:*

$$r = \frac{\mathrm{Cov}(x,y)}{\sigma_x \sigma_y}.$$

Зная значения r, σ_x, σ_y, можно вычислить коэффициенты корреляционного уравнения по следующим формулам:

$$a = \overline{y} - b\overline{x},$$

$$b = r \cdot \frac{\sigma_y}{\sigma_x}.$$

11.6. Статистический анализ модели

Оценка параметров парной регрессии выполняется из следующих предпосылок. Предположим, что для генеральной совокупности связь между x и y линейна. Наличие случайных отклонений, вызванных воздействием на переменную y множества других, неучтенных в нашем уравнении факторов и ошибок измерения, приведет к тому, что связь наблюдаемых величин x_i и y_i приобретет вид:

$$y_i = \alpha + \beta x_i + \varepsilon_i.$$

Здесь ε_i – случайные ошибки (отклонения, возмущения). Если бы были известны точные значения отклонений ε_i, то можно было бы рассчитать значения параметров α и β. Однако значения случайных отклонений в выборке неизвестны, и по наблюдениям x_i и y_i можно получить оценки параметров α и β, которые сами являются случайными величинами, поскольку соответствуют случайной выборке. Пусть a – оценка параметра α, b – оценка параметра β. Тогда оцененное уравнение регрессии будет иметь вид:

$$y_i = a + bx_i + e_i,$$

где e_i – наблюдаемые значения ошибок ε_i.

Для того чтобы полученные методом наименьших квадратов оценки a и b обладали желательными свойствами, сделаем следующие предпосылки об отклонениях ε_i:

1) величина ε_i является случайной переменной;

2) математическое ожидание ε_i равно нулю;

3) дисперсия ε_i постоянна: $D(\varepsilon_i) = D(\varepsilon_j)$ для всех i, j;

4) значения ε_i независимы между собой.

Известно, что если данные условия выполняются, то оценки, сделанные с помощью метода наименьших квадратов, обладают следующими свойствами:

1. Оценки являются несмещенными, т. е. математическое ожидание оценки каждого параметра равно его истинному значению: $M(a) = \alpha$; $M(b) = \beta$. Это вытекает из того, что $M(\varepsilon_i) = 0$, и говорит об отсутствии систематической ошибки в определении положения линии регрессии.

2. Оценки состоятельны, т. к. дисперсия оценок параметров при возрастании числа наблюдений стремится к нулю:

$$\lim_{n\to\infty} D(a) = 0; \ \lim_{n\to\infty} D(b) = 0.$$

Иначе говоря, надежность оценки при увеличении выборки растет.

3. Оценки эффективны, они имеют наименьшую дисперсию по сравнению с любыми другими оценками данного параметра.

Перечисленные свойства не зависят от конкретного вида распределения величин ε_i, тем не менее обычно предполагается, что они распределены нормально. Эта предпосылка необходима для проверки статистической значимости сделанных оценок и определения для них доверительных интервалов.

Если предположения 3) и 4) нарушены, т. е. дисперсия возмущений непостоянна или значения ε_i связаны друг с другом, то свойства несмещенности и состоятельности сохраняются, но свойства эффективности – нет.

Оценка параметров конкретной регрессии является лишь отдельным этапом длительного и сложного процесса построения эконометрической модели. Первое же оцененное уравнение очень редко является удовлетворительным во всех отношениях. Обычно приходится постепенно подбирать формулу связи и состав объясняющих переменных, анализируя на каждом этапе качество оцененной зависимости. Этот анализ качества включает статистическую и содержательную составляющую. Проверка статистического качества оцененного уравнения состоит из следующих элементов:

- проверка статистической значимости каждого коэффициента уравнения регрессии;
- проверка общего качества уравнения регрессии;
- проверка свойств данных, выполнение которых предполагалось при оценивании уравнения.

Под содержательной составляющей анализа качества понимается рассмотрение экономического смысла оцененного уравнения регрессии: действительно ли значимыми оказались объясняющие факторы, важные с точки зрения теории; положительны или отрицательны коэффициенты, показывающие направление воздействия этих факторов; попали ли оценки коэффициентов регрессии в предполагаемые из теоретических соображений интервалы.

11.7. Проверка гипотез о корреляции случайных величин

Если коэффициент корреляции равен нулю для генеральной совокупности, то это не означает, что он будет равен нулю для выборки. Наоборот, он будет обязательно отклоняться от истинного значения, но чем больше такое отклонение, тем менее оно вероятно при данном объеме выборки. При каждом конкретном значении коэффициента корреляции величин X и Y для генеральной совокупности выборочный коэффициент корреляции является случайной величиной. Следовательно, случайной величиной является также любая его функция, и требуется указать такую функцию, которая имела бы одно из известных распределений, удобное для табличного анализа. Для выборочного коэффициента корреляции и такой функцией является *t–статистика*, рассчитываемая по формуле

$$t = r \cdot \frac{\sqrt{n-2}}{\sqrt{1-r^2}}$$

и имеющая распределение Стьюдента с $(n-2)$ – степенями свободы. Число степеней свободы меньше числа наблюдений на 2, поскольку в формулу выборочного коэффициента корреляции входят средние выборочные значения X и Y, для расчета которых используются две линейные формулы их зависимости от наблюдений случайных величин. Для коэффициента корреляции проверяется нулевая гипотеза, т. е. гипотеза о равенстве его нулю в генеральной совокупности. Эта гипотеза отвергается, если выборочный коэффициент корреляции слишком далеко отклонился от нулевого значения, т. е. произошло событие, которое было бы маловероятным в случае равенства нулю коэффициента корреляции для генеральной совокупности. Для этого задается вероятность такого события, которая называется в статистике «уровень значимости». Чаще всего задается уровень значимости 1% или 5%. Если для некоторого показателя проверяется гипотеза о том, что его истинное значение равно нулю, то данная гипотеза отвергается в том случае, если оценка показателя по данным выборкам такова, что вероятность получения такого или большего (по модулю) ее значения меньше, чем 1% или 5% соответственно.

Сравнивая определенное по выборочным данным значение статистики t с критическими точками, определяемыми по таблицам распределения Стьюдента, можно принять или отвергнуть нулевую гипотезу.

11.8. Анализ статистической значимости параметров линейной регрессии

Надежность оценок коэффициентов регрессии устанавливается по отклонению переменной y от оцененной линии регрессии:

$$e_i = y_i - a - bx_i.$$

Дисперсия ошибки параметра b будет равна:

$$D_b = \sigma_b^2 = \frac{S_e^2}{\sum_i (x_i - \bar{x})^2}.$$

Дисперсия ошибки параметра a будет равна:

$$D_a = \sigma_{ab}^2 = \frac{S_e^2 \cdot \sum_i x_i^2}{n \sum_i (x_i - \bar{x})^2},$$

где $S_e^2 = \frac{\sum_i e_i^2}{n-2}$.

Значимость оцененного параметра регрессии b может быть проверена с помощью анализа его отношения к своей стандартной ошибке. Эта величина имеет t распределение Стьюдента с $(n-2)$-степенями свободы. Она называется *t-статистикой:*

$$t = \frac{b}{\sqrt{D_b}} = \frac{b}{\sigma_b}.$$

Из формулы видно, что t-статистика представляет собой число стандартных отношений. Для t-статистики проверяется нулевая гипотеза. В качестве нулевой гипотезы принимается утверждение о том, что величина y не зависит от x, т. е. $b = 0$. Альтернативная гипотеза заключается в том, что $b \neq 0$, иными словами, что значение x влияет на величину y.

Насколько малой должна быть вероятность для выбора гипотез? В большинстве работ по экономике за критический уровень берется 5% или 1%. Если выбирается уровень 5%, то выбор альтернативной гипотезы происходит в том случае, когда вероятность получения значения b составляет менее 5%. В этом случае говорят, что нулевая гипотеза должна быть отвергнута при 5%-ном уровне значимости (рис. 11.1).

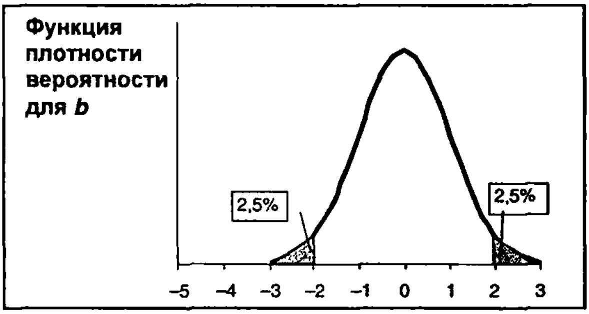


Рис. 11.1

Возникает вопрос: почему исследователи обычно представляют свои результаты при уровнях значимости в 5% и 1%? Почему недостаточно ограничиться только одним уровнем? Причина заключается в том, что обычно делается попытка найти баланс между риском допущения ошибок 1-го и 2-го рода. Ошибка 1-го рода имеет место в том случае, когда вы отвергаете истинную нулевую гипотезу. Ошибка 2-го рода возникает, когда вы не отвергаете ложную гипотезу.

Чем ниже критическая вероятность, тем меньше риск получения ошибок 1-го рода. Если вы используете уровень значимости, равный 5%, то вы отвергаете истинную гипотезу в 5% случаев. Если уровень значимости составляет 1%, вы совершаете ошибку 1-го рода в 1% случаев. Таким образом, в этом отношении однопроцентный уровень значимости более надежен.

В то же время, если нулевая гипотеза ложна, то чем выше уровень значимости, тем шире область принятия гипотезы и тем выше вероятность того, что вы не отвергнете ее, и тем выше риск допущения ошибки 2-го рода. Таким образом, вы оказываетесь перед выбором. Если вы будете настаивать на очень высоком уровне значимости, то столкнетесь с относительно высоким риском допущения ошибки 2-го рода, когда гипотеза окажется ложной. Если вы выбираете низкий уровень значимости, то оказываетесь перед относительно высоким риском допущения ошибки 1-го рода, если гипотеза истинна.

Проверка значимости коэффициента парной линейной регрессии эквивалентна проверке значимости коэффициента корреляции x и y.

Для проверки нулевой гипотезы t-статистика сравнивается с критическим значением t. Значения $t_{крит}$ для некоторых степеней свободы приведены в табл. 11.1. Условие того, что оценка регрессии должна приводить к отказу от нулевой гипотезы, т. е. признать значимым коэффициент b, будет следующим:

$$-t_{крит} > \frac{b}{\sigma_b} > t_{крит} \quad \text{или} \quad t_{ст} > |t_{крит}|.$$

Например, для 22 наблюдений мы получили уравнение регрессии вида:

$$y = 5{,}2 + 3{,}0x.$$

Стандартная ошибка коэффициента b равна 1,0. Тогда t-статистика равна 3,0/1,0 = 3,0. Число степеней свободы равно количеству наблюдений минус количество переменных, т. е. 20. Для данного числа степеней свободы при 5%-ном уровне значимости критическое значение $t_{крит}$ равно 2,086. Сравнение 3,0 > 2,086 позволяет отвергнуть нулевую гипотезу и сделать вывод, что величина b в действительности отличается от нуля и, следовательно, x влияет на y.

Если этот критерий описать словами, то верхний и нижний 2,5%-ные «хвосты» t-распределения начинаются со стандартного отклонения 2,086 вверх и вниз от его математического ожидания, равного нулю. Нулевая гипотеза в этом случае заключается в том, что коэффициент регрессии находится в пределах 2,086 стандартного отклонения и +2,086. В рассматриваемом случае t-статистика эквивалентна 3,0 стандартным отклонениям, и мы приходим к выводу о том, что результат оценивания регрессии противоречит нулевой гипотезе.

Конечно, в том, что мы используем уровень значимости в 5% в качестве основы для проверки гипотезы, существует 5%-ный риск допущения ошибки 1-го рода. В этом случае мы могли бы снизить риск до 1% за счет применения уровня значимости в 1%. Критическое значение для t при 1%-ном уровне значимости с 20 степенями свободы составляет 2,845. Сравнивая это число с $t_{ст}$, мы видим, что можем отказаться от нулевой гипотезы также и при этом уровне значимости.

Таблица 11.1

t-распределение: критические значения t				
Число степеней свободы	Уровень значимости			
	5%		1%	
	Двусторонний тест	Односторонний тест	Двусторонний тест	Односторонний тест
6	2,447	1,943	3,707	3,143
10	2,228	1,812	3,169	2,764

Таблица 11.1 (окончание)

t-распределение: критические значения t				
Число степеней свободы	Уровень значимости			
	5%		1%	
15	2,131	1,753	2,947	2,602
20	2,086	1,725	2,845	2,528
30	2,042	1,697	2,75	2,457
40	2,021	1,684	2,704	2,423

Процедура установления взаимосвязи между зависимой и объясняющей переменными происходит путем формирования, а затем отклонения нулевой гипотезы о том, что $b = 0$, используется очень часто. Соответственно, большая часть программ оценки регрессии в статистических пакетах выводят t-статистику для этого специального случая. Однако если нулевая гипотеза определяет некоторое ненулевое значение величины b, то необходимо использовать более общее выражение:

$$t = \frac{b - b_0}{\sigma_b},$$

где b_0 – некоторое ненулевое значение величины b, принимаемое в качестве нулевой гипотезы.

Односторонние t-тесты

Односторонние t-тесты представляют собой усовершенствованную процедуру проверки в случае, когда альтернативная гипотеза является не простым отрицанием нулевой гипотезы, а сформулирована более конкретно. Формулировка более конкретной гипотезы требует теоретического основания, или предпосылки, вытекающего из анализа связи зависимой и объясняющей переменных.

Рассмотрим три случая: первый случай – весьма частный, когда существует единственное альтернативное истинное значение b, которое обозначим b_1; второй случай – если b не равно b_0, то оно должно быть больше b_0; и третий случай – если величина b не равна b_0, то она должна быть меньше b_0.

В первом случае существует только два возможных истинных значения коэффициента при x – b_0 или b_1. Сначала предположим, что b_1 больше, чем b_0. Допустим, что мы хотим проверить нулевую гипотезу при 5%-ном уровне значимости, и используем для этого обычную процедуру. Найдем границы для верхнего и нижнего 2,5%-ных «хвостов» t-распределения, считая, что нулевая гипотеза верна, и обозначим их как A и B (рис. 11.2). Нулевая гипотеза отклоняется, если коэффициент регрессии b оказывается правее точки B или левее точки A.

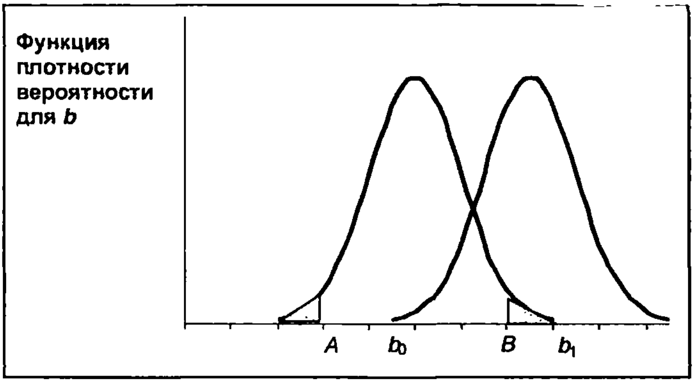


Рис. 11.2

Если значение b находится справа от B, то оно намного лучше совместимо с альтернативной гипотезой, чем с нулевой; вероятность его нахождения справа от B, если истинна альтернативная гипотеза, намного больше, чем при истинности нулевой гипотезы. Здесь нет сомнений в том, чтобы отклонить нулевую гипотезу и принять альтернативную гипотезу.

Если, однако, b находится слева от A, то используемая процедура проверки приведет нас к неверному заключению. Последняя требует отклонить нулевую гипотезу, несмотря на то, что при истинности альтернативной гипотезы вероятность нахождения b слева от A ничтожно мала.

Следовательно, мы отклоним нулевую гипотезу, только если b оказывается в верхнем 2,5%-ном «хвосте» распределения, т. е. справа от B. Это означает, что теперь мы выполняем проверку гипотезы с односторонним критерием, сократив в результате вероятность допущения ошибки 1-го рода до 2,5%. Поскольку уровень значимости определен как вероятность допущения ошибки 1-го рода, то он теперь также составляет 2,5%.

Экономисты обычно предпочитают проверку гипотез с 5%- и 1%-ным уровнем значимости, а не с 2,5%-ным. Если вы хотите провести проверку с 5%-ным уровнем значимости, то следует переместить точку B влево так, чтобы получить 5% вероятности в «хвосте» распределения и увеличить вероятность допущения ошибки первого рода до 5%.

Используя односторонний критерий вместо двустороннего, можно получить большую мощность критерия при любом уровне значимости. При переходе к одностороннему критерию точка B на рис. 11.2 сдвигается влево, и тем самым сокращается вероятность принятия ложной нулевой гипотезы. Нужно, однако, помнить, что выигрыш в мощности будет получен только в условиях, когда использование одностороннего критерия оправданно.

Во втором случае, когда $b_1 > b_0$, желательно исключить левый «хвост» распределения из критической области гипотезы, т. к. низкое значение для b более вероятно получить при нулевой гипотезе $b = b_0$, чем при гипотезе $b > b_0$, а, следовательно, это будет говорить в пользу нулевой гипотезы, а не против нее. Поному вновь односторонний критерий проверки предпочтительней двустороннего, а правый «хвост» распределения мы рассматриваем как область непринятия гипотезы.

Аналогичным образом, если альтернативная гипотеза представлена в виде $b < b_0$, мы предпочтем проверку, основанную на одностороннем критерии, использующим левый «хвост» распределения в качестве области отклонения гипотезы.

Проверки с использованием одностороннего критерия очень важны на практике при решении экономических задач. Как мы уже видели, обычно для установления того, что независимая переменная действительно оказывает влияние на зависимую переменную, формулируется нулевая гипотеза $b = 0$, которую затем пытаются опровергнуть. Очень часто гипотеза оказывается слишком общей, чтобы утверждать, что если x оказывает влияние на y, то это влияние имеет определенную направленность.

Если у нас есть достаточно веские причины считать, что это влияние, скажем, положительно, то можно использовать альтернативную гипотезу $b > 0$ вместо более общей гипотезы $b \neq 0$. Это является преимуществом, поскольку критическое значение t для отклонения нулевой гипотезы при проверке по одностороннему критерию будет меньшим, что облегчает отклонение нулевой гипотезы и установление наличия зависимости.

11.9. Проверка общего качества уравнения регрессии

Квадратическая ошибка, с которой определяется теоретическая линия связи, равна

$$\sigma_{\text{ош}} = \sigma_y \cdot \sqrt{1 - r^2}.$$

Для анализа общего качества оцененной линейной регрессии используют **коэффициент детерминации** R^2, называемый также квадратом множественной корреляции. Для случая парной регрессии это квадрат коэффициента корреляции переменных x и y.

Коэффициент детерминации измеряет долю тех элементов вариации в Y, которые содержатся также и в вариации X. Если известно, что зависимая переменная находится в отношении причинной связи с независимой переменной, то коэффициент детерминации определяет долю, которая может быть рассмотрена как причинно обусловленная изменениями факторного признака.

Коэффициент детерминации рассчитывается по формуле:

$$R^2 = 1 - \frac{\sum_i e_i^2}{\sum_i (y_i - \bar{y})^2}.$$

Рассмотрим вариацию (разброс) $\mathrm{Var}(Y) = \sum (Y_i - \bar{Y})^2$ значений Y_i вокруг среднего значения. Разобьем эту вариацию на две части: объясненную регрессионным уравнением и необъясненную (т. е. связанную с ошибками ε_i).

Обозначим через $\hat{Y}_i = a + bX_i$ предсказанное значение Y_i. Тогда вариация равна:

$$\mathrm{Var}(Y) = \sum(Y_i - \overline{Y})^2 = \sum(Y_i - \hat{Y}_i)^2 + \sum(\hat{Y}_i - \overline{Y})^2.$$

Вариация, деленная на число наблюдений n, представляет собой дисперсию. Обозначим левую часть через TSS (total sum of squared) – вся дисперсия зависимой переменной y, первое слагаемое в правой части, соответствующее необъясненной дисперсии (остаточной) через ESS (error sum of squared), второе слагаемое в правой части – RSS (regression sum of squared) – объясненная часть всей дисперсии:

$$R^2 = 1 - ESS/TSS = RSS/TSS.$$

Коэффициент детерминации характеризует долю дисперсии зависимой переменной, объясненной с помощью регрессии.

Коэффициент детерминации является мерой, позволяющей определить, в какой степени найденная регрессионная прямая дает лучший результат для объяснения поведения зависимой переменной y, чем просто горизонтальная прямая $y = \overline{y}$.

С увеличением объясненной доли разброса коэффициент детерминации приближается к единице. С добавлением еще одной переменной R^2 – обычно увеличивается, однако если переменные x_1 и x_2 сильно коррелируют между собой, то они объясняют одну и ту же часть разброса переменной y, и в этом случае трудно идентифицировать вклад каждой из переменных в объяснение поведения y.

Если существует статистически значимая линейная связь величин x и y, то коэффициент детерминации R^2 близок к единице. Однако он может быть близким к единице просто в силу того, что обе эти величины имеют выраженный временной тренд, не связанный с их причинно-следственной зависимостью.

Если имеются не временные ряды, а перекрестная выборка, т. е. данные об однотипных объектах в один и тот же момент времени, то для оцененного по ним уравнения линейной регрессии величина коэффициента детерминации не превышает обычно уровня 0,6 – 0,7.

11.10. Определение параметров уравнения степенной зависимости

При степенной зависимости параметры уравнения определяются приведением его к линейному виду путем логарифмирования и установления параметров методом наименьших квадратов.

Если существует зависимость вида:

$$y = ax^b,$$

то приведение его к линейному виду осуществляется путем логарифмирования:

$$\lg y = \lg a + b \cdot \lg x,$$

а нахождение параметров a и b осуществляется аналогично, как при линейной связи. Так, если данные сгруппированы, получим:

$$\overline{Y}' = \frac{\sum \lg y_i}{n}; \quad \overline{X}' = \frac{\sum \lg x_i}{n};$$

$$b = \frac{\sum(\lg x \cdot \lg y) - n\overline{X}' \cdot \overline{Y}'}{\sum \lg x^2 - n\overline{X}'^2};$$

$$\lg a = \overline{Y}' - b \cdot \overline{X}'; \quad a = 10^{\lg a}.$$

11.11. Определение параметров гиперболы

Пусть $y = a + b/x$ есть уравнение искомой гиперболы.

Делая замену переменной $1/x_i = z_i$, получим линейное уравнение:

$$y = a + bz,$$

параметры которого определяются по формулам линейной регрессии.

11.12. Определение параметров показательной регрессии

Пусть $y = ab^x$ есть уравнение искомой показательной кривой. Ее можно представить в следующем виде:

$$\lg y = \lg a + x \cdot \lg b.$$

Введя обозначения $\lg a = A$, $\lg b = B$, получим уравнение:

$$\lg y = A + Bx.$$

Отсюда следует, что функция y, представленная на графике с осью ординат, разделенной по логарифмической шкале, и осью абсцисс – по нормальной шкале, дает прямую с угловым коэффициентом B и расстоянием по оси $0y$, равным A.

11.13. Установление параметров параболы

Если связь между признаками Y и X нелинейная и описывается уравнением параболы второго порядка, то

$$\hat{y} = a + bx + cx^2.$$

В данном случае задача сводится к определению неизвестных параметров a, b, c.

Применив метод наименьших квадратов, получим уравнение:

$$S = \sum_{i=1}^{n}(y - a - bx - cx^2)^2 \to \min.$$

Для нахождения значений неизвестных параметров a, b, c, при которых функция была бы минимальной, необходимо приравнять частные производные по этим величинам к нулю, т. е.:

$$\begin{cases} \dfrac{\partial S}{\partial a} = 2\sum(y - a - bx - cx^2) = 0; \\ \dfrac{\partial S}{\partial b} = 2\sum(y - a - bx - cx^2)x = 0; \\ \dfrac{\partial S}{\partial c} = 2\sum(y - a - bx - cx^2)x^2 = 0. \end{cases}$$

Проделав преобразования, получим систему нормальных уравнений:

$$\begin{cases} na + b\sum x + c\sum x^2 = \sum y; \\ n\sum x + b\sum x^2 + c\sum x^3 = \sum yx; \\ a\sum x^2 + b\sum x^3 + c\sum x^4 = \sum yx^2. \end{cases}$$

Решив систему уравнений, найдем значения неизвестных параметров a, b, c и получим уравнение регрессии.

11.14. Множественная линейная регрессия

В экономической практике часто имеет место сложная, многопричинная статистическая связь между признаками. В таком случае зависимость $y = f(x)$ означает, что x – вектор, содержащий m компонентов: $x = (x_1, x_2, ..., x_m)$. Задача оценки статистической взаимосвязи переменных y и $x = (x_1, x_2, ..., x_m)$ формулируется аналогично случаю парной регрессии. Записывается функция $y = f(a,x) + \varepsilon$, где a – вектор параметров, ε – случайная ошибка. Предполагается, что эта функция связывает переменную y с вектором независимых переменных x для данной генеральной совокупности. Как и в случае парной регрессии, предполагается, что ошибки ε_i являются случайными величинами с нулевым математическим ожиданием и постоянной дисперсией; ε_i и ε_j статистически независимы при ij. Кроме того, для проверки статистической значимости оценок a обычно предполагается, что ошибки ε_i нормально распределены. По данным наблюдений выборки размерности n требуется оценить значения параметров a.

Чтобы формально можно было решить задачу, должно быть $n \geq m + 1$. Если число степеней свободы мало, то статистическая надежность оцениваемой формулы невысока. Обычно при оценке множественной регрессии для обеспечения статистической надежности требуется, чтобы число наблюдений по крайней мере в 3 раза превосходило число оцениваемых параметров.

Для оценки параметров применяется, как правило, метод наименьших квадратов. Уравнение регрессии с оцененными параметрами имеет вид:

$$y = a_0 + a_1x_1 + a_2x_2 + ... + a_mx_m + e,$$

и критерием для нахождения вектора (a) является $\min\sum e_j^2$:

$$S = \sum_j e_i^2 = \sum_j (y_j - (a_0 + \sum_i a_i x_{ij}))^2 \to \min.$$

Приравнивая к нулю частные производные по каждому параметру, составляется система нормальных уравнений относительно неизвестных параметров a_i:

$$\frac{\partial S}{\partial a_0} = 0; \quad \frac{\partial S}{\partial a_1} = 0; \quad \frac{\partial S}{\partial a_2} = 0; \quad .... ; \quad \frac{\partial S}{\partial a_m} = 0.$$

Например, по параметру a_1:

$$\frac{\partial S}{\partial a_1} = 2\sum(y - a_0 - a_1x_1 - a_2x_2 - ... - a_mx_m)x_1 = 0.$$

Сделав преобразования, получаем:

$$a_0\sum x_1 + a_1\sum x_1^2 + a_2\sum x_2x_1 + \ldots + a_m\sum x_m x_1 = \sum yx_1 .$$

В результате таких преобразований система нормальных уравнений с m неизвестными (по числу параметров a_i) имеет вид:

$$\begin{cases} na_0 + a_1\sum x_1 + a_2\sum x_2 + \ldots + a_m\sum x_m = \sum y; \\ a_0\sum x_1 + a_1\sum x_1^2 + a_2\sum x_2x_1 + \ldots + a_m\sum x_m x_1 = \sum yx_1; \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_0\sum x_m + a_1\sum x_1x_m + a_2\sum x_2x_m + \ldots + a_m\sum x_m^2 = \sum yx_m. \end{cases}$$

Одним из способов построения множественных уравнений регрессии является построение модели связи в стандартизованном масштабе.

Оценка влияния каждого факторного признака, включенного в уравнение регрессии, на результативный признак может быть затруднена, если факторные признаки различны по своей сущности и имеют различные единицы измерения. В этих случаях для более точной оценки влияния факторных признаков используют множественные модели регрессии в стандартизованном масштабе. Модель регрессии в стандартизованном масштабе предполагает, что все значения исследуемых признаков переводятся в стандарты по формуле:

$$t_{xi} = \frac{x_i - \bar{x}}{\sigma_i},$$

где x_i – значение признака в натуральном масштабе.

Уравнение множественной регрессии в стандартизованном масштабе следующее:

$$\bar{u}_{1,2,\ldots,m} = \beta_1 t_1 + \beta_2 t_2 + \ldots + \beta_m t_m,$$

где $t_1, t_2, \ldots, t_m$ – стандартизированные значения признаков $x_1, x_2, \ldots, x_m$;

$u_{1,2,\ldots,m}$ – среднее значение стандартизованной переменной соответствующего результативного признака, полученного по уравнению регрессии.

Параметры многофакторной модели регрессии в стандартизованном масштабе определяются методом наименьших квадратов.

Система нормальных уравнений будет иметь вид:

$$\begin{cases} \sum ut_1 = \beta_1\sum t_1^2 + \beta_2\sum t_1t_2 + \ldots + \beta_m\sum t_1t_m; \\ \sum ut_2 = \beta_1\sum t_1t_2 + \beta_2\sum t_2^2 + \ldots + \beta_m\sum t_2t_m; \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \sum ut_m = \beta_1\sum t_1t_m + \beta_2\sum t_2t_m + \ldots + \beta_m\sum t_m^2, \end{cases}$$

где u – значение результативного признака в стандартизованном масштабе.

Коэффициенты β_1, β_2, ..., β_m дают возможность провести сравнительную оценку силы влияния изменения каждого факторного признака на изменение результативного (моделируемого) признака.

От уравнения в стандартизированном масштабе легко можно перейти к уравнению в натуральном масштабе. Коэффициенты a_i получают из соотношения

$$a_i = \beta \frac{\sigma_i}{\sigma_y},$$

а свободный член уравнения a_0 – из выражения

$$a_0 = \overline{y} - a_1\overline{x}_1 - a_2\overline{x}_2 - ... - a_m\overline{x}_m.$$

Решение системы нормальных уравнений относительно неизвестных параметров a_i лучше выражать в векторно-матричной форме, иначе оно становится слишком громоздким.

Для линейного уравнения с двумя переменными

$$Y = a + bx + cz$$

система нормальных уравнений имеет вид:

$$\begin{cases} na_0 + b\sum x + c\sum z = \sum y; \\ a\sum x + b\sum x^2 + c\sum xz = \sum xy; \\ a\sum z + b\sum xz + c\sum z^2 = \sum zy. \end{cases}$$

Определить параметры линейного уравнения с двумя переменными можно по частным коэффициентам корреляции. Для этого устанавливают средние значения $\overline{Y}$, $\overline{X}$, $\overline{Z}$, стандартные отклонения σ_y, σ_x, σ_z и парные коэффициенты корреляции нулевого порядка, характеризующие тесноту связи между переменными y и x, y и z, x и z.

Параметры a, b, c устанавливаются по формулам:

$$b = \frac{\sigma_y}{\sigma_x} \cdot \frac{r_{yx} - r_{yz} \cdot r_{xz}}{1 - r_{xz}^2}, \quad c = \frac{\sigma_y}{\sigma_z} \cdot \frac{r_{yz} - r_{yx} \cdot r_{xz}}{1 - r_{xz}^2}, \quad a = \overline{Y} - b\overline{X} - c\overline{Z}.$$

Для вычисления частных корреляций регрессий с большим числом переменных используют формулу:

$$r_{123} = \frac{r_{12} - r_{13} \cdot r_{23}}{\sqrt{(1 - r_{13}^2)(1 - r_{23}^2)}}.$$

Множественная корреляция помогает установить степень влияния каждого фактора на зависимую переменную.

Для анализа статистической значимости полученных коэффициентов множественной линейной регрессии необходимо, как и в случае парной регрессии, оценить дисперсию и стандартные отклонения коэффициентов a_j.

11.15. *F*-тест на качество оценивания

Для определения статистической значимости коэффициента детерминации проверяется нулевая гипотеза для *F-статистики*, рассчитываемой по формуле:

$$F = \frac{R^2}{1 - R^2} \cdot \frac{n - m - 1}{m}.$$

Соответственно, для парной регрессии:

$$F = \frac{R^2(n-2)}{1-R^2}.$$

Смысл проверяемой гипотезы заключается в том, что все коэффициенты линейной регрессии, за исключением свободного параметра, равны нулю (случай отсутствия линейной функциональной связи). Если они действительно равны нулю для генеральной совокупности, то уравнение регрессии должно иметь вид $y = \bar{y}$, а коэффициент детерминации и *F-статистика* Фишера также равны нулю. При этом их оценки для случайной выборки, конечно, отличаются от нуля, но чем больше такое отличие, тем менее оно вероятно. Логика проверки нулевой гипотезы заключается в том, что если произошло событие, которое было бы слишком маловероятным в том случае, если данная гипотеза действительно была бы верна, то эта гипотеза отвергается.

Величина F имеет распределение Фишера с $(m, n - m - 1)$ степенями свободы, где m – число объясняющих переменных, n – число наблюдений. Распределение Фишера – двухпараметрическое распределение неотрицательной случайной величины, являющейся в частном случае при $m = 1$ квадратом случайной величины, распределенной по Стьюденту. Для распределения Фишера имеются таблицы критических значений, зависящих от числа степеней свободы m и $n - m - 1$ при различных уровнях значимости.

Для проверки нулевой гипотезы при заданном уровне значимости по таблицам находится критическое значение $F_{крит}$, и нулевая гипотеза отвергается, если $F > F_{крит}$. Например, при оценке парной регрессии по 15 наблюдениям $R^2 = 0,7$. В этом случае $F = 0,7 \cdot 13/0,3 = 30,3$. По таблицам для распределения Фишера с (1,13) – степенями свободы найдем, что при 5%-ном уровне значимости (доверительная вероятность 95%) критическое значение F равно 4,67, при 1%-ном – 9,07. Поскольку $F = 30,3 > F_{крит}$, нулевая гипотеза в обоих случаях отвергается, т. е. отвергается предположение о незначимости связи.

В случае парной регрессии проверка нулевой гипотезы для *t-статистики* коэффициента регрессии равносильна проверке нулевой гипотезы для *F-статистики* (и, соответственно, R^2). В этом случае *F-статистика* равна квадрату *t-статистики*. В случае парной регрессии статистическая значимость величин R^2 и *t-статистики* коэффициента регрессии определяется коррелированностью переменных x и y. Самостоятельную важность показатель R^2 приобретает в случае множественной линейной регрессии.

Распределение Фишера может быть использовано не только для проверки гипотезы об одновременном равенстве нулю всех коэффициентов линейной регрессии, но и гипотезы о равенстве нулю части этих коэффициентов. Это особенно важно при развитии линейной регрессионной модели, т. к. позволяет оценить обоснованность исключения отдельных переменных или их групп из числа объясняющих переменных, или же, наоборот, включения их в это число.

При анализе адекватности уравнения регрессии исследуемому процессу возможны следующие варианты.

1. Построенная модель на основе ее проверки по F-критерию Фишера в целом адекватна, и все коэффициенты регрессии значимы. Такая модель может быть использована для принятия решений и осуществления прогнозов.

2. Модель по F-критерию адекватна, но часть коэффициентов регрессии не значима. В этом случае модель пригодна для принятия некоторых решений, но не для производства прогнозов.

3. Модель по F-критерию Фишера адекватна, но все коэффициенты регрессии незначимы. В этом случае модель полностью считается неадекватной. На ее основе не принимаются решения и не осуществляются прогнозы.

11.16. Регрессионный анализ в матричном виде

Рассмотрим линейную регрессионную модель с одной независимой переменной. Обозначим

$$X = \begin{bmatrix} x_1 \\ . \\ . \\ . \\ x_n \end{bmatrix}, \ Y = \begin{bmatrix} y_1 \\ . \\ . \\ . \\ y_n \end{bmatrix}, \ s = \begin{bmatrix} 1 \\ . \\ . \\ . \\ 1 \end{bmatrix}, \ e = \begin{bmatrix} e_1 \\ . \\ . \\ . \\ e_n \end{bmatrix},$$

$$\hat{Y} = as + bX, \ e = Y - \hat{Y}$$

где a, b – численные коэффициенты, $\hat{Y}$ – вектор, лежащий в двумерной плоскости π, образуемой векторами s, X.

Задача формулируется следующим образом: найти такие a и b, чтобы вектор e имел наименьшую длину. Другими словами, необходимо наилучшим образом аппроксимировать вектор Y вектором $\hat{Y}$. Решением является такой вектор Y, для которого вектор e перпендикулярен плоскости π. Для этого необходимо и достаточно, чтобы вектор e был ортогонален векторам s и X, порождающим плоскость π:

$$s^T e = 0, \text{ или } \Sigma e_i = 0, \text{ или } \Sigma(Y_i - a - bX_i) = 0,$$
$$X^T e = 0, \Sigma X_i e_i = 0, \Sigma X_i(Y_i - a - bX_i) = 0.$$

Можно отметить, что опять получены необходимые условия экстремума.

Обозначим теперь через X $n{\times}2$ матрицу

$$X = \begin{bmatrix} 1 & x_1 \\ . & . \\ . & . \\ . & . \\ 1 & x_n \end{bmatrix}, \ Y = \begin{bmatrix} y_1 \\ . \\ . \\ . \\ y_n \end{bmatrix}, \ B = \begin{bmatrix} a \\ b \end{bmatrix},$$

где B – матрица (вектор) коэффициентов, $e = Y - XB$.

Условие ортогональности вектора e плоскости π теперь записывается как

$$X^T e = 0, \text{ или } X^T(Y - XB) = X^T Y - X^T XB = 0.$$

Отсюда получаем $X^T XB = X^T Y$, или

$$B = (X^T X)^{-1} X^T Y.$$

Можно показать, что данная формула совпадает с формулами определения параметров по методу наименьших квадратов:

$$B = (X^T X)^{-1} X^T Y = \begin{bmatrix} n & \sum X_i \\ \sum X_i & \sum X_i^2 \end{bmatrix} \begin{bmatrix} \sum Y_i \\ \sum X_i Y_i \end{bmatrix} = \ldots = \begin{bmatrix} a \\ b \end{bmatrix}.$$

Рассмотрим множественную регрессию:

$$Y_i = b_1 + b_2 X_{i2} + \ldots + b_k X_{ik} + e_i,$$

где X_{ik} – значения регрессора X_k в наблюдении i. Если через X_1 обозначить вектор, состоящий из одних единиц (1,..., 1), то можно не различать модели со свободным членом или без свободного члена:

$$Y_i = b_1 X_{i1} + b_2 X_{i2} + \ldots + b_k X_{ik} + e_i.$$

Обозначим через Y $n\times1$ матрицу (вектор-столбец) $Y_1, \ldots, Y_n$; $B = (b_1, \ldots, b_k)'$ – $k\times1$ вектор коэффициентов; $e = (e_1, \ldots, e_n)$ – $n\times1$ вектор ошибок; матрица объясняющих переменных $n\times k$:

$$X = \begin{bmatrix} X_{11} & \ldots & X_{1k} \\ . & \ldots & . \\ X_{n1} & \ldots & X_{nk} \end{bmatrix} = \begin{bmatrix} 1 & X_{21} & \ldots & X_{1k} \\ . & . & \ldots & . \\ 1 & X_{n2} & \ldots & X_{nk} \end{bmatrix}.$$

Как и в случае регрессионного уравнения с одной переменной, задачей является выбор вектора оценок B, минимизирующего сумму квадратов остатков e_i (т. е. квадрат длины вектора остатков e):

$$e = Y - \hat{Y} = Y - XB, \quad ESS = \sum e_i^2 = e^T e \to \min.$$

Выразим $e^T e$ через X и B:

$$e = (Y - XB)^T (Y - XB) = Y^T Y - Y^T XB - B^T X^T Y + B^T X^T XB =$$

$$= Y^T Y - 2B^T X^T Y + B^T X^T XB.$$

Необходимые условия минимума ESS получаются дифференцированием по вектору B:

$$\frac{\partial ESS}{\partial B} = -2X^T Y + 2X^T XB = 0,$$

откуда, учитывая обратимость матрицы $X^T X$, находится оценка метода наименьших квадратов:

$$B_{мн} = (X^T X)^{-1} X^T Y.$$

В дальнейшем изложении мы еще обратимся к вопросам эконометрики: в главе 12 рассмотрим использование Excel и Mathcad для эконометрических расчетов, в главе 13 рассмотрим основные проблемы, возникающие в эконометрических исследованиях, а также вопросы прогнозирования в регрессионных моделях, и в главе 24 – оценку систем одновременных уравнений.

Подробнее вопросы и проблемы эконометрики и вычислительные алгоритмы изложены в работах [7, 8, 11, 12, 18, 22, 26, 30].

Глава 12. Корреляционно-регрессионный анализ в Excel и Mathcad

12.1. Статистические функции в Excel

Статистические функции в Excel находятся в опции **Мастер функций**, категория **Статистические** (рис. 12.1). Здесь представлен довольно обширный список статистических функций. Нас будут интересовать в первую очередь функции, позволяющие выполнять корреляционно-регрессионный анализ. Вторая возможность имеется в опции **Анализ данных**, меню **Сервис**, которая устанавливается, как правило, дополнительно. Рассмотрим вначале возможности, заложенные в Мастере функций.

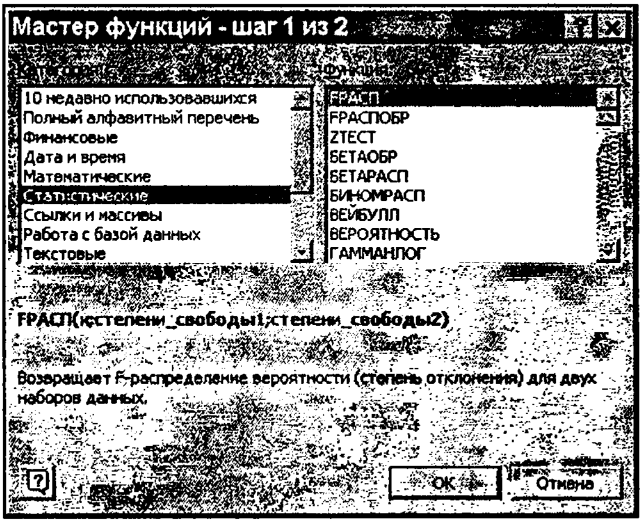


Рис. 12.1. Мастер функций – панель статистических функций

12.2. Установка параметров линейной связи в Excel

Для установки параметров (коэффициентов) линейной связи введем данные рядов X и Y – наблюдаемую статистику независимой и зависимой переменной, в ячейки, адреса которых представлены на рис. 12.2.

Будем определять параметры линейной парной регрессии вида

$$Y = a + bX + e.$$

Рассмотрим разнообразные возможности решения этого вопроса, которые могут потребоваться в зависимости от ситуации.

В ячейке С3 подсчитаем среднее значение ряда X. Для этого вызовем функцию СРЗНАЧ (**Мастер функций – Статистические**) и на появившейся панели функции введем адрес диапазона ряда X – A2:A10 (рис. 12.3).

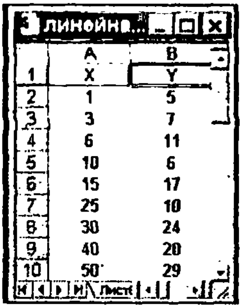


Рис. 12.2. Ввод данных независимой и зависимой переменной

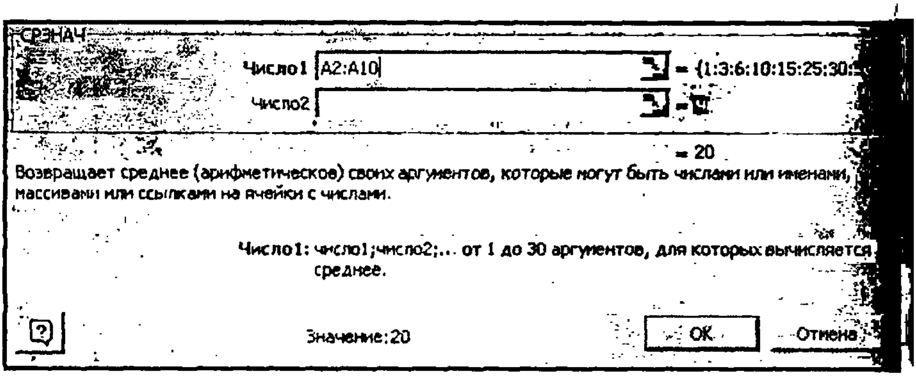


Рис. 12.3. Панель функции СРЗНАЧ

В ячейку D4 введем среднее значение ряда Y, проделав аналогичную операцию, введя на панели функции СРЗНАЧ адрес ряда Y – B2:B10.

В ячейке С6 подсчитаем стандартное отклонение ряда X, вызвав функцию СТАНДОТКЛОН и введя диапазон ряда X, как это показано на рис. 12.4.

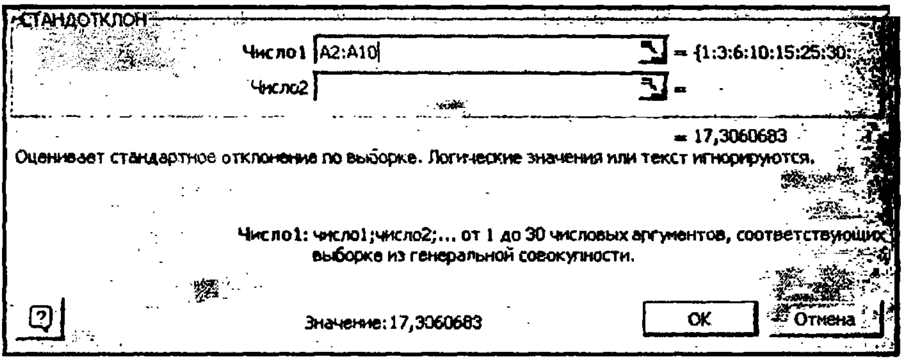


Рис. 12.4. Панель функции СТАНДОТКЛОН

В ячейку D6 введем стандартное отклонение ряда Y, проделав аналогичную операцию.

В ячейке C8 подсчитаем коэффициент корреляции. Для этого существует функция КОРРЕЛ, в окнах панели которой необходимо ввести адреса диапазонов как ряда X, так и ряда Y – рис.12.5.

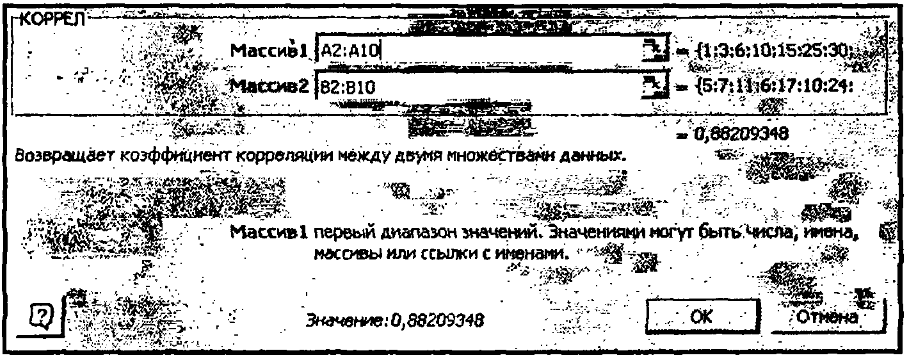


Рис. 12.5. Панель функции КОРРЕЛ

Угловой коэффициент связи (параметр b) равен

$$b = r \cdot \sigma_y / \sigma_x,$$

где r – коэффициент корреляции; σ_y, σ_x – стандартные отклонения соответственно зависимой и независимой переменной.

Параметр a устанавливается по формуле:

$$a = \overline{Y} - b\overline{X},$$

где $\overline{Y}$ и $\overline{X}$ – средние значения зависимой и независимой переменной.

Тогда параметры b и a, учитывая адреса показателей (см. рис. 12.5), установим по формулам:

=C8*D6/C6;
=D3 – C11*C3.

Результат представлен на рис. 12.6.

Параметры линейной парной регрессии можно рассчитать и сразу. Для этого в Excel существуют функции НАКЛОН и ОТРЕЗОК. Функция НАКЛОН служит для определения углового коэффициента связи (параметра b), а функция ОТРЕЗОК – для определения свободного члена уравнения. На панели соответствующих функций необходимо ввести адреса диапазонов Y и X (рис. 12.7 и 12.8).

Кроме перечисленных возможностей, существует еще и следующая. Построим график по имеющимся данным. Чтобы ось X отражала фактические данные, выберем тип диаграммы Точечная – рис. 12.9.

На построенной диаграмме выделим график функции, щелкнув по ней левой кнопкой мыши. Выделение обозначается светлыми маркерами на функции. Нажав правую кнопку мыши, выведем контекстно-зависимое меню (рис. 12.10).

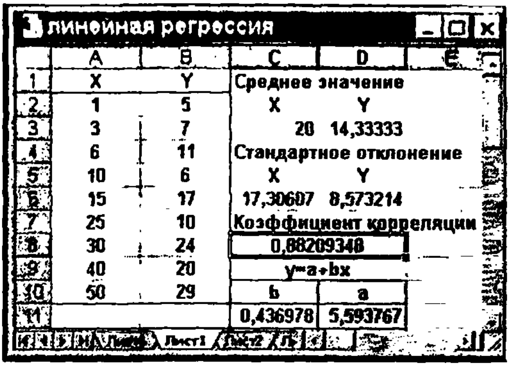


Рис. 12.6. Расчет параметров линейной парной регрессии

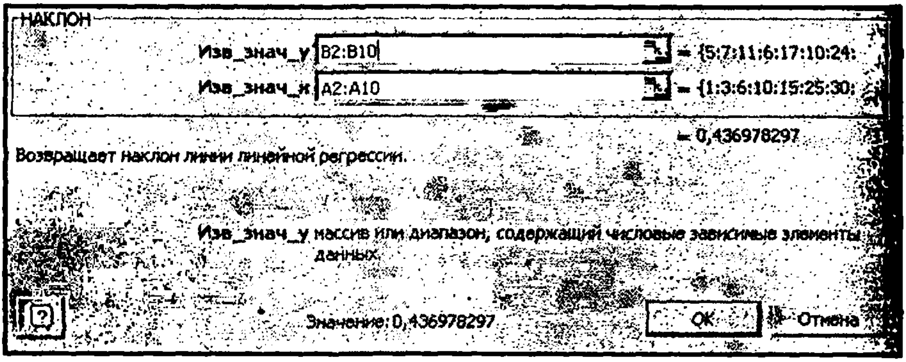


Рис. 12.7. Панель функции НАКЛОН – углового коэффициента связи

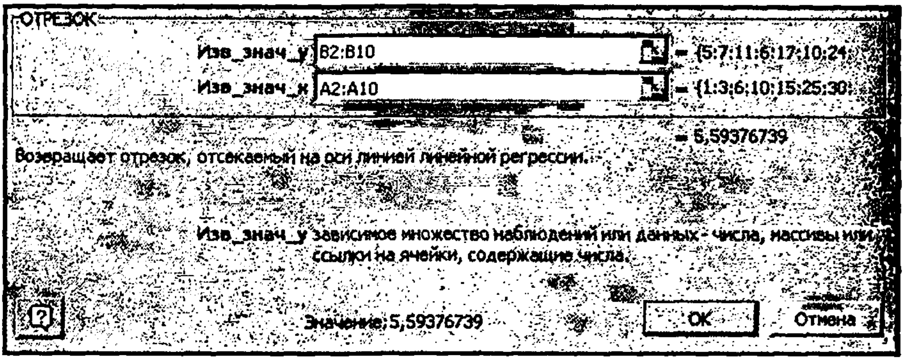


Рис. 12.8. Панель функции ОТРЕЗОК – свободного члена уравнения (параметра а)

Рис. 12.9. Мастер диаграмм – выбор точечной диаграммы

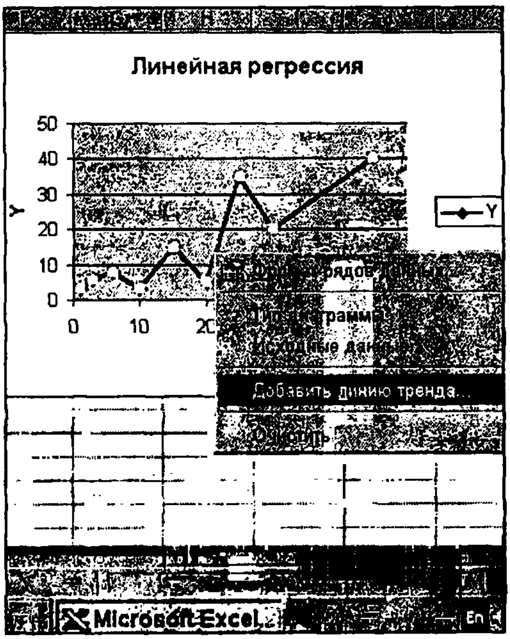


Рис. 12.10. Контекстно-зависимое меню при выделенном ряде данных

Выберем опцию Добавить линию тренда.

В панели линии тренда во вкладке Тип надо выбрать тип функции (по умолчанию выбирается Линейная) – рис. 12.11.

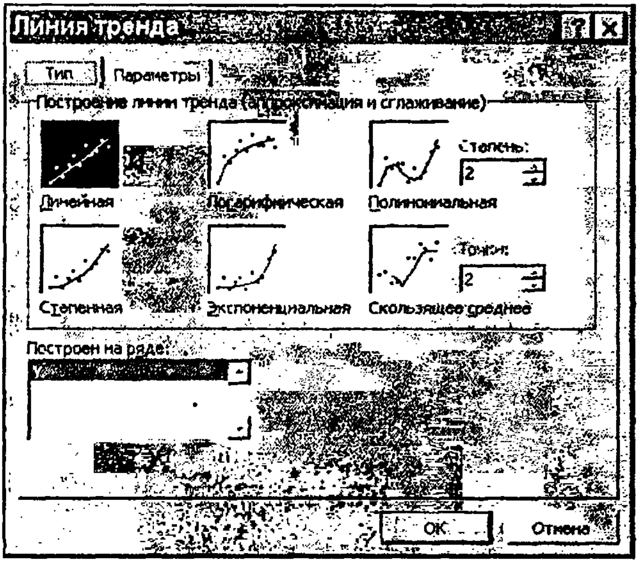


*Рис. 12.11. Панель линии тренда – вкладка **Тип***

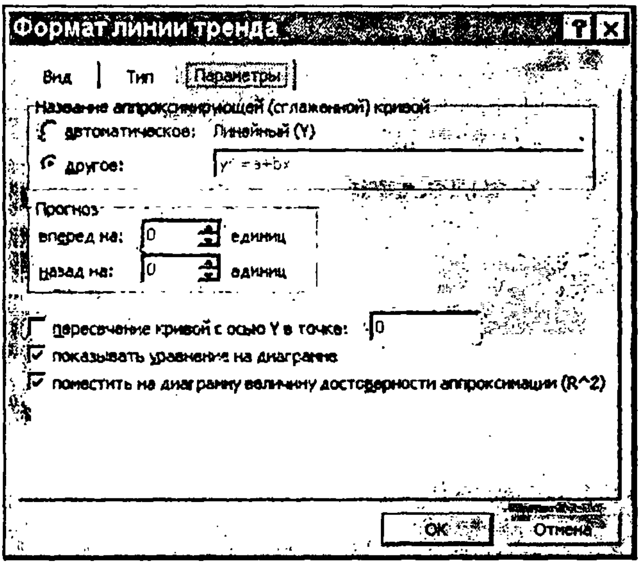


*Рис. 12.12. Панель линии тренда – вкладка **Параметры***

Во вкладке Параметры введем название тренда (теоретической кривой) и установим флажки **показывать уравнение на диаграмме** и **поместить на диаграмму величину достоверности аппроксимации (R^2)**, т. е. коэффициент детерминации, как это показано на рис. 12.12.

В результате на диаграмме появится вид теоретической кривой – тренда, ее параметры и коэффициент детерминации (рис. 12.13).

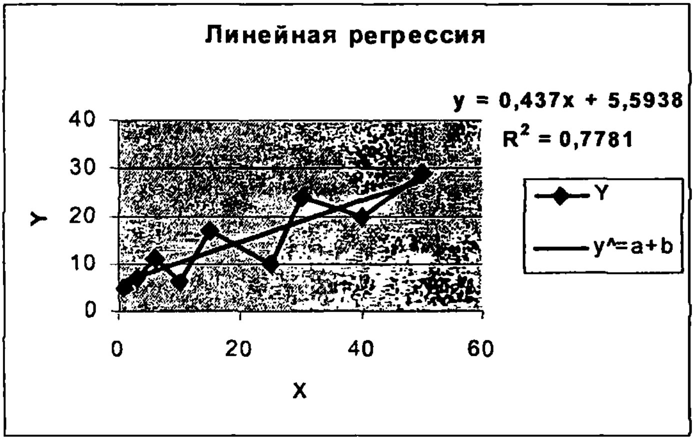


Рис. 12.13. График фактических данных и построенной регрессии с выведенными на диаграмму параметрами уравнения и коэффициентом детерминации

12.3. Установление статистической значимости и общего качества уравнения линейной регрессии в Excel

Рассчитаем сначала коэффициент детерминации и рассмотрим его понятие. Для этого в столбце F рассчитаем вариацию (Yi – Yср)^2 – в ячейку F2 впишем формулу =(B2–$D$3)^2 и скопируем формулу в ячейки F3:F10. В ячейке F11 просуммируем ячейки F2:F10 (рис. 12.14). В столбце G посчитаем вариацию остатков (Yi – a – bx)^2, в ячейке G11 его сумму, а в столбце H – вариацию регрессии (a + bx – Yср)^2, в ячейке H11 – его сумму (рис. 12.14).

Вариация Y_i представляется так:

$$\mathrm{Var}(Y) = \Sigma(Y_i - Y_{ср})^2 = \Sigma(Y_i - a - bx_i)^2 + \Sigma(a + bx_i - Y_{ср})^2.$$

Обозначим $\Sigma(Y_i - Y_{ср})^2$ через *TSS* (total sum of squares) – вся дисперсия;

$\Sigma(Y_i - a - bx_i)^2$ – как *ESS* (erroe sum of squares) – необъясненная дисперсия;

$\Sigma(a + bx_i - Y_{ср})^2$ – как *RSS* (regression sum of squares) – объясненная часть всей дисперсии.

Тогда можно записать:

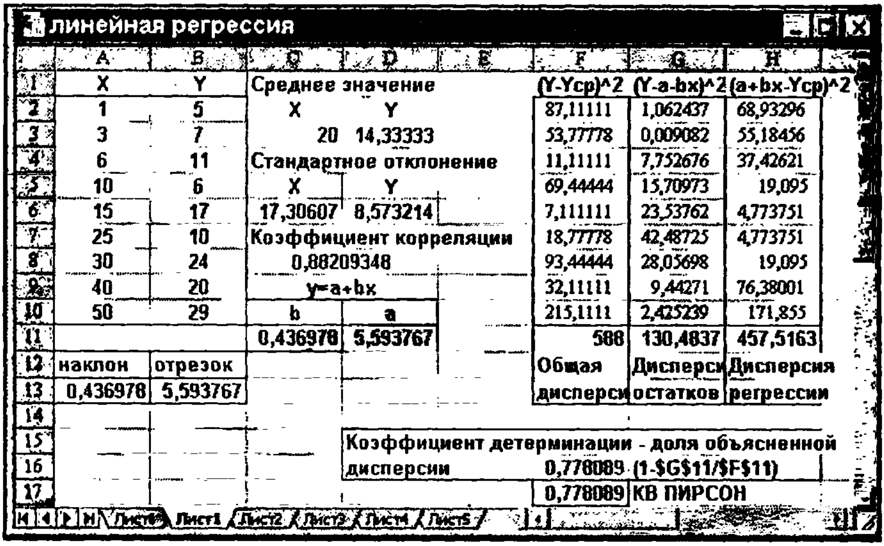
$$\mathrm{Var}(Y) = TSS = ESS + RSS.$$

Коэффициент детерминации будет равен:

$$R^2 = 1 - ESS/TSS = RSS/TSS$$

и представляет собой долю объясненной дисперсии.

Он по данной формуле рассчитан в ячейке F16. А в ячейке F17 коэффициент детерминации определен с использованием функции КВ ПИРСОН (см. рис. 12.14).

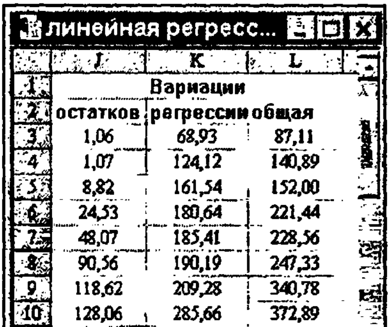

линейная регрессия

	A	B	C	D	E	F	G	H
1	X	Y	Среднее значение			(Y-Ycp)^2	(Y-a-bx)^2	(a+bx-Ycp)^2
2	1	5	X	Y		87,11111	1,062437	68,93296
3	3	7	20	14,33333		53,77778	0,009082	55,18456
4	6	11	Стандартное отклонение			11,11111	7,752676	37,42621
5	10	6	X	Y		69,44444	15,70973	19,095
6	15	17	17,30607	8,573214		7,111111	23,53762	4,773751
7	25	10	Коэффициент корреляции			18,77778	42,48725	4,773751
8	30	24	0,88209348			93,44444	28,05698	19,095
9	40	20	y=a+bx			32,11111	9,44271	76,38001
10	50	29	b	a		215,1111	2,425239	171,855
11			0,436978	5,593767		588	130,4837	457,5163
12	наклон	отрезок				Общая	Дисперси	Дисперсия
13	0,436978	5,593767				дисперси	остатков	регрессии
14								
15				Коэффициент детерминации - доля объясненной				
16				дисперсии		0,778089	(1-$G$11/$F$11)	
17						0,778089	КВ ПИРСОН	

Лист6 Лист1 Лист2 Лист3 Лист4 Лист5

Рис. 12.14. Определение коэффициента детерминации

Представим графически это понятие. Рассчитаем нарастающим итогом вариации остатков, регрессии и общую (рис. 12.15).


линейная регресс...

	J	K	L
1		Вариации	
2	остатков	регрессии	общая
3	1,06	68,93	87,11
4	1,07	124,12	140,89
5	8,82	161,54	152,00
6	24,53	180,64	221,44
7	48,07	185,41	228,56
8	90,56	190,19	247,33
9	118,62	209,28	340,78
10	128,06	285,66	372,89
11	130,48	457,52	588,00
12	ESS	RSS	TSS
13			
14	R² = 1 - ESS/TSS = RSS/TSS		

Лист6 Лист1

Рис. 12.15. Расчет кумулятивных кривых вариации

Построим графики кумулятивных кривых вариации (рис. 12.16). Последние значения этих графиков и представляют собой *TSS*, *RSS* и *ESS* – общую дисперсию, объясненную и необъясненную дисперсию. Отношение величин *RSS*/*TSS* есть доля необъясненной дисперсии, т. е. коэффициент детерминации.

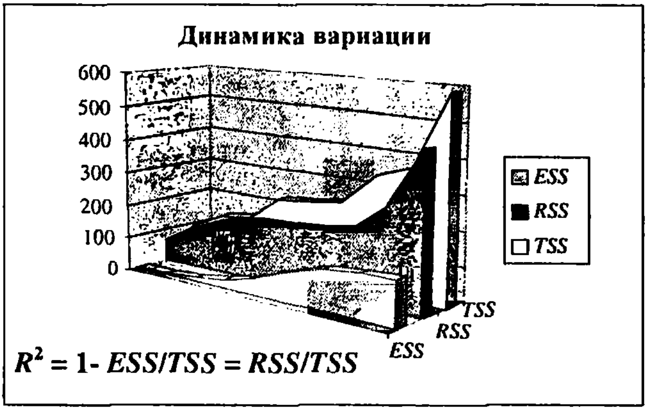


Рис. 12.16. Кумулятивные кривые динамики вариаций: ESS – остатков; RSS – регрессии; TSS – общей

Перейдем к расчету показателей статистической значимости. Проведем предварительные вычисления, показанные на рис. 12.17, которые потребуются для дальнейших вычислений.

линейная регрессия

a+bX	Y-a-bX	(X-Xcp)^2	X^2
6,030746	-1,03075	361	1
6,904702	0,095298	289	9
8,215637	2,784363	196	36
9,96355	-3,96355	100	100
12,14844	4,851558	25	225
16,51822	-6,51822	25	625
18,70312	5,296884	100	900
23,0729	-3,0729	400	1600
27,44268	1,557318	900	2500
		2396	5996

Лист1 Лист2

Рис. 12.17

На рис. 12.18 представлен результат расчета показателей статистической значимости. Стандартное отклонение остатков регрессии определено с помощью функции СТАНДОТКЛОН к ряду остатков регрессии, расчитанной в ячейках K2:K11.

Стандартное отклонение ошибок регрессии определяется по формуле:

$$S = (1/(n-2) \cdot \sum(Y_i - a - bx_i)^2)^{0,5}.$$

Для этого в ячейке M14 введена формула =(1/(9 – 2)*G11)^0,5. Для расчета стандартного отклонения ошибок регрессии применима также функция СТОШYX.

Дисперсия параметра b равна:

$$\sigma_b^2 = \frac{S^2}{\sum_i (x_i - \bar{x})^2}.$$

Для расчета *дисперсии* параметра b в ячейке M16 введена формула G11/L11/(9 – 2), можно также – (M14^2)/L11.

В ячейке M17 рассчитано стандартное отклонение параметра b, как корень квадратный из дисперсии.

В ячейке M17 рассчитана t-статистика параметра b как отношение величины параметра к его стандартному отклонению: =C11/M17.

	J	K	L	M	N	O	P
12				Формула=			
13	Станд отк. остатков регресс			4,03862172	СТАНДОТКЛОН(K2:K11)		
14	Станд отк. ошибок регрессии			4,31746822	(1/(9-2)*$G$11)^0,5	или функ	
15							
16	Дисперсия параметра b			0,00777985	$G$11/$L$11/(9-2)		
17	Стандартное отклонение b			0,08820348	$M$16^0,5		
18	t - статистика параметра b			4,95420673	$C$11/$M$17		
19	Дисперсия параметра а			5,18311208	$G$11*$M$11/(9-2)/$L$11/9		
20	Стандартное отклонение а			2,27664492	$M$19^0,5		
21	t - статистика параметра а			2,45702233	$D$11/$M$20		
22	Критическое значение						
23	t - статистики			2,36462256	СТЬЮДРАСПРОБР(0,05;7)		
24							
25	Вероятность b			0,00164838	СТЬЮДРАСП(M18;7;2)		
26				0,99835162	1 - СТЬЮДРАСП(M18;7;2)		
27	Вероятность а			0,04365546	СТЬЮДРАСП(M21;7;2)		
28				0,95634454	1-СТЬЮДРАСП(M27;7;2)		
29	F - статистика			24,5441644	(9-2)*$F$16/(1-$F$16)		
30	Критическое значение						
31	F - статистики			5,59145974	FРАСПОБР(0,05;1;7)		
32	Значимость F			0,00164838	FРАСП(M29;1;7)		
33	Вероятность F			0,99835162	1-M32		

Лист7 / Лист10 / Лист11 / Лист1 / Лист2 / Лист3

Рис. 12.18. Расчет показателей статистической значимости

В ячейке M19 рассчитана дисперсия параметра a по формуле:

$$\sigma_a^2 = \frac{S^2 \sum x_i^2}{n \sum_i (x_i - \bar{x})^2}.$$

В ячейках M20 и M21 аналогично рассчитаны стандартное отклонение и *t-статистика* параметра *a*.

Критическое значение *t-статистики* рассчитано в ячейке M23 с помощью функции СТЬЮДРАСПОБР, в панели которой (рис. 12.19) вводятся значения пороговой значимости (вероятность) 0,05 и степени свободы ($n - m - 1$), где m – количество независимых переменных.

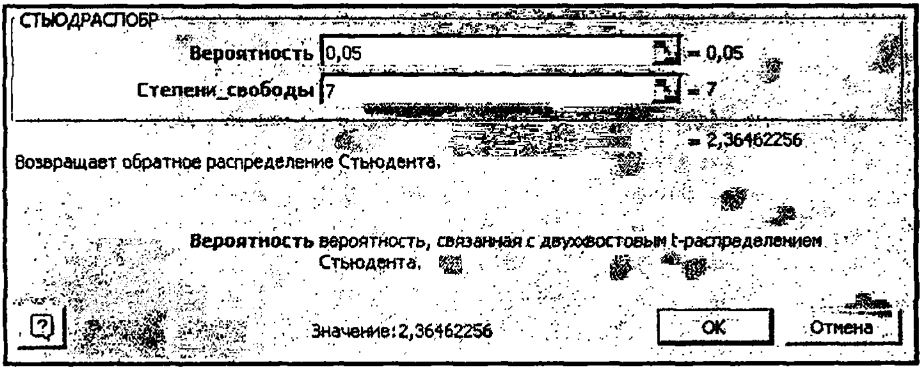


Рис. 12.19. Панель функции обратного распределения Стьюдента

Вероятность распределения *t-статистики* за пределами пороговой значимости можно установить с помощью функции СТЬЮДРАСП, на панели которой в качестве X необходимо установить адрес ячейки с расчетом величины *t-статистики*, степени свободы ($n - m - 1$) и учет порогов (хвостов) распределения с двух сторон – 2.

Тогда вероятность распределения t-статистики в пределах указанной вероятности (0,95) устанавливается как (1-СТЬЮДРАСП), что сделано для параметра *b* в ячейке M26, а для параметра *a* – в ячейке M28.

Необходимо отметить, что проверка значимости коэффициента парной линейной регрессии (параметра *b*) эквивалентна проверке значимости коэффициента корреляции.

Показатель *F-статистики* рассчитывается по формуле:

$$F = \frac{R^2}{1 - R^2} \cdot \frac{n - m - 1}{m},$$

где R^2 – коэффициент детерминации.

Для расчета *F-статистики* в ячейке M29 введена формула =(9–2)F16/(1–F16).

Критическое значение *F-статистики* устанавливается с помощью функции FРАСПОБР, в панели которой устанавливается пороговая значимость – 0,05, степени свободы – ($n - m - 1$), в нашем примере она равна 7, и m равен 1. Вероятность *F-статистики* можно установить с помощью функции FРАСПОБР.

12.4. Статистический пакет анализа данных

Перейдем к возможностям пакета анализа данных меню **Сервис**. Как правило, его необходимо дополнительно установить, если данная команда отсутствует в меню **Сервис**.

Чтобы установить пакет **Анализа данных** в меню **Сервис**, выберите команду **Надстройки** и установите флажок **Пакет анализа**. Диалоговое окно данной опции приведено на рис. 12.20.

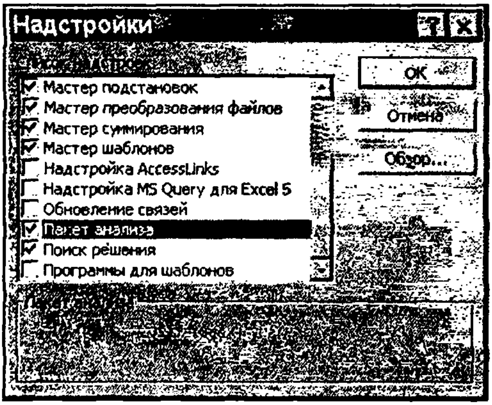


Рис. 12.20. Диалоговое окно ***Надстройки***

Если в списке надстроек нет пакета анализа данных, нажмите кнопку **Обзор** и укажите диск, папку и имя файла для надстройки пакета анализа Analys 32.xll (как правило, папка Library\Analisys) или запустите программу Setup, чтобы установить эту надстройку.

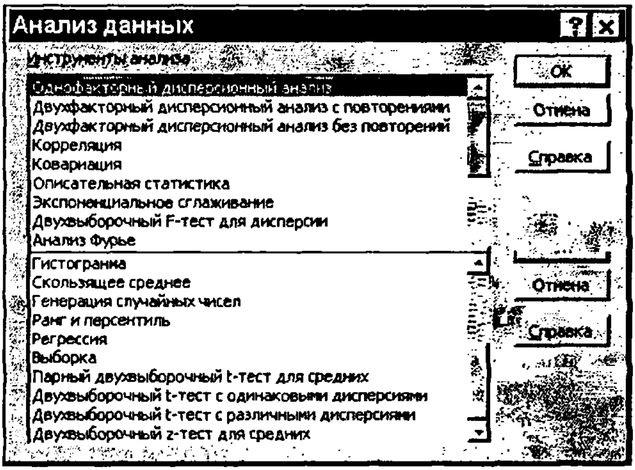


Рис. 12.21. Диалоговое окно ***Анализ данных***

Чтобы запустить пакет анализа в меню **Сервис**, выберите команду **Анализ данных**. В диалоговом окне Анализ данных в списке **Инструменты анализа** выберите нужную строку – рис. 12.21.

Прежде чем мы начнем разбирать анализ регрессии, рассмотрим некоторые возможности пакета Анализ данных. Выберем строку **Описательная статистика**. В диалоговом окне опции введем диапазон ряда *X* (рис. 12.22).

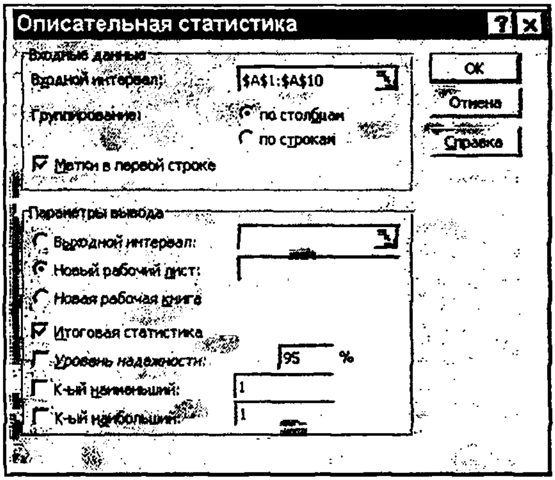


Рис. 12.22. Диалоговое окно опции ***Описательная статистика***

Описательная статистика	X	Y
Среднее	20	14,33333
Стандартная ошибка	5,768689	2,857738
Медиана	15	11
Мода	#Н/Д	#Н/Д
Стандартное отклонение	17,30607	8,573214
Дисперсия выборки	299,5	73,5
Эксцесс	-0,85814	-1,00816
Асимметричность	0,635147	0,600555
Интервал	49	24
Минимум	1	5
Максимум	50	29
Сумма	180	129
Счет	9	9

Рис. 12.23. Вывод показателей опцией ***Описательная статистика***

Данная опция выводит рассчитанные значения многих статистических показателей. На рис. 12.23 показаны значения статистических показателей, полученных с помощью опции **Описательная статистика**, рассчитанные не только для ряда *X*, но и аналогично для ряда *Y*.

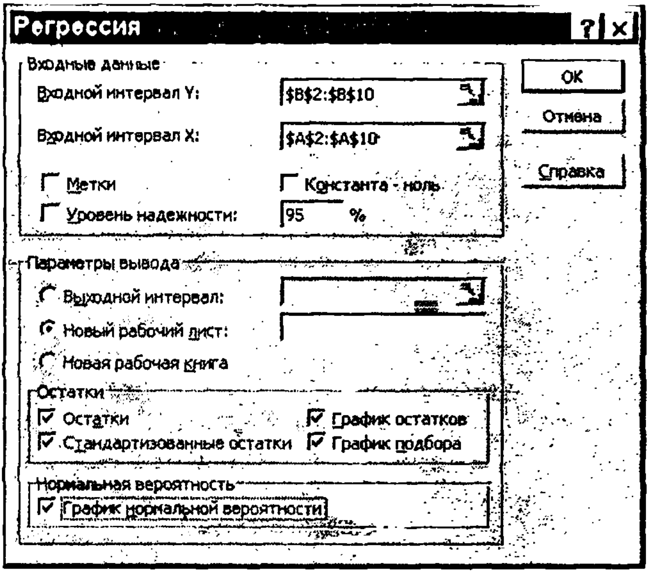
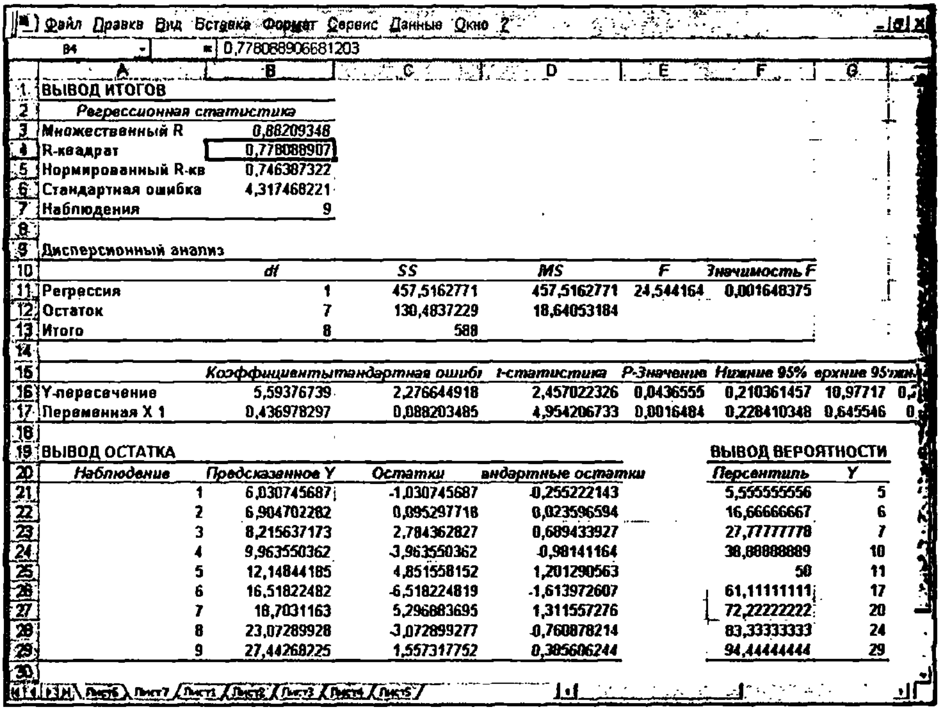


*Рис. 12.24. Диалоговое окно опции **Регрессия***



ВЫВОД ИТОГОВ

Регрессионная статистика	
Множественный R	0,88209348
R-квадрат	0,778088907
Нормированный R-кв	0,746387322
Стандартная ошибка	4,317468221
Наблюдения	9

Дисперсионный анализ

	df	SS	MS	F	Значимость F
Регрессия	1	457,5162771	457,5162771	24,544164	0,001648375
Остаток	7	130,4837229	18,64053184		
Итого	8	588			

	Коэффициенты	тандартная ошибі	t-статистика	P-Значение	Нижние 95%	ерхние 95	ижн
Y-пересечение	5,59376739	2,276644918	2,457022326	0,0436555	0,210361457	10,97717	0,
Переменная X 1	0,436978297	0,088203485	4,954206733	0,0016484	0,228410348	0,645546	0

ВЫВОД ОСТАТКА

Наблюдение	Предсказанное Y	Остатки	андартные остатки
1	6,030745687	-1,030745687	-0,255222143
2	6,904702282	0,095297718	0,023596594
3	8,215637173	2,784362827	0,689433927
4	9,963550362	-3,963550362	-0,98141164
5	12,14844185	4,851558152	1,201290563
6	16,51822482	-6,518224819	-1,613972607
7	18,7031163	5,296883695	1,311557276
8	23,07289928	-3,072899277	-0,760878214
9	27,44268225	1,557317752	0,385606244

ВЫВОД ВЕРОЯТНОСТИ

Персентиль	Y
5,555555556	5
16,66666667	6
27,77777778	7
38,88888889	10
50	11
61,11111111	17
72,22222222	20
83,33333333	24
94,44444444	29

*Рис. 12.25. Вид рабочего листа с выводом показателей опции **Регрессия***

Теперь в окне **Анализ данных** выберем строку **Регрессия**. В окне **Регрессия** вставим входной диапазон Y и X, параметры вывода – новый рабочий лист, и поставим флажки для вывода остатков и графиков (рис. 12.24).

На новом рабочем листе будут выведены показатели, представленные на рис. 12.25, и графики, приведенные на рис. 12.26.

Как видно из рис. 12.25, опция **Регрессия** выводит большинство показателей, которые мы до этого рассчитали с помощью одиночных функций.

*Рис. 12.26. Вид выводимых графиков опции **Регрессия***

12.5. Установка параметров нелинейных регрессий в Excel

Параметры нелинейных регрессий в Excel можно устанавливать несколькими методами.

Во-первых, регрессию можно привести к линейному виду, а затем установить параметры регрессии с помощью функций, перечисленных выше.

Например, степенную функцию можно прологарифмировать, а затем установить параметры. А сейчас рассмотрим возможности Excel устанавливать параметры нелинейной регрессии непосредственно.

Основные возможности заложены в опции **Линия тренда**.

В данной опции, кроме линейной функции, предусмотрены возможности установления параметров логарифмической, полиномиальной, степенной, экспоненциальной функций – рис. 12.11.

Напомню, чтобы воспользоваться этой опцией, необходимо построить диаграмму, активизировать график функции, правой кнопкой мыши открыть контекстно-зависимое меню и выбрать команду **Вставить линию тренда**, в меню **Ти** выбрать вид функции, в меню **Параметры** установить флажки **показывать уравнение на диаграмме и поместить на диаграмму величину достоверности аппроксимации R^2**.

Возьмем данные, приведенные на рис. 12.27, построим диаграмму и установим параметры логарифмической, полиномиальной, степенной и экспоненциальной функций.

	A	B
1	X	Y
2	1	1
3	5	11
4	10	35
5	16	26
6	20	48
7	27	53
8	31	70
9	32	100
10	33	180
11	40	200

Рис. 12.27

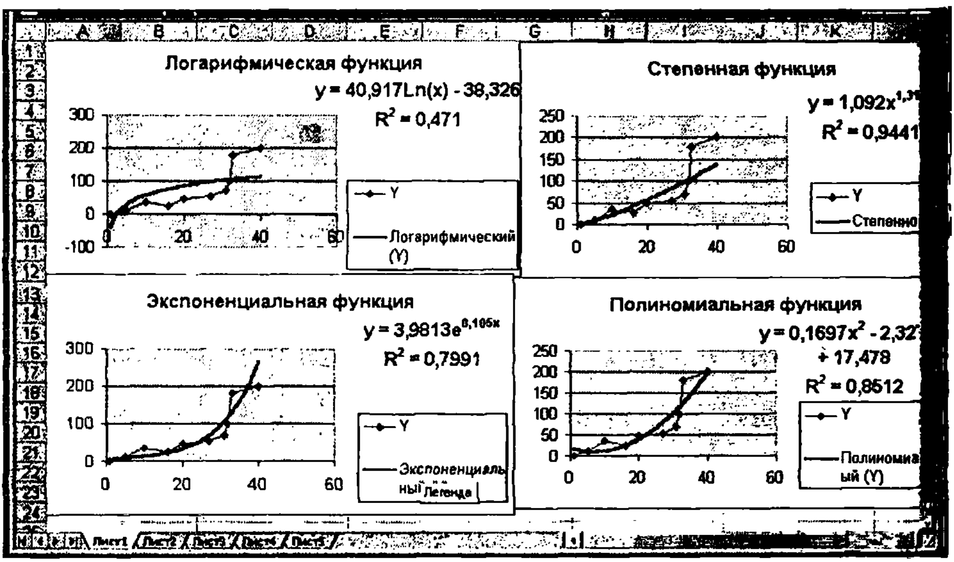


Рис. 12.28. Вид рабочего листа с установленными параметрами нелинейных регрессий

На рис. 12.28 приведены четыре функции с выведенными параметрами и коэффициентами детерминации.

12.6. Установка параметров множественной регрессии в Excel

Параметры множественной регрессии линейного вида в Excel устанавливаются с помощью функции ЛИНЕЙН (рис. 12.29 и 12.30).

Для ввода результата необходимо выделить матрицу в размере: столбцов (m + 1), где m – количество переменных; строк – всегда 5.

В эту функцию необходимо ввести данные в формате:

=ЛИНЕЙН(интервал значений_y; блок значений_xi; константа; статистика).

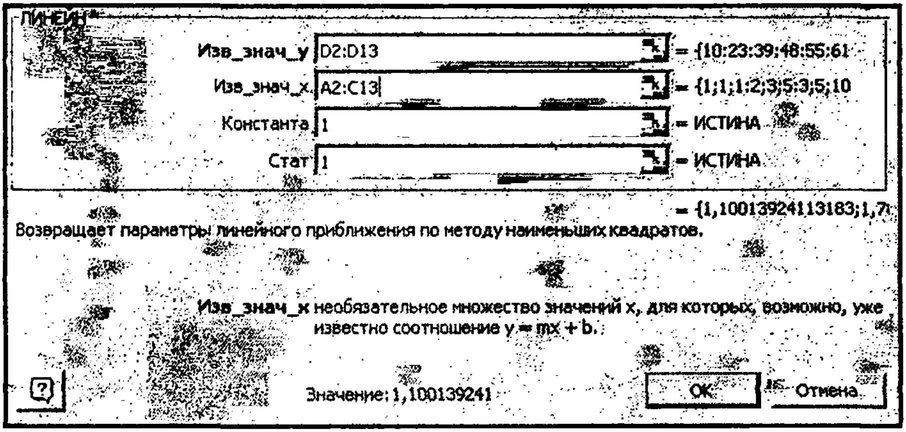


Рис. 12.29. Панель функции ЛИНЕЙН

Поскольку функция ЛИНЕЙН матричная, то необходимо помнить, что независимые переменные (X_1,X_2,...X_n) вводятся в виде единой матрицы.

Ввод константы и статистики обеспечивает представление результатов вычислений в различных вариантах. Константа обозначает свободный член уравнения. Поэтому если константе присвоить значение ИСТИНА (или 1, что равнозначно), то уравнение регрессии будет представлено со свободным членом уравнения. Соответственно, если назначить ЛОЖЬ (или 0), то уравнение будет представлено без свободного члена.

При назначении статистики ИСТИНА будет выведена оценка достоверности, если ЛОЖЬ – то нет.

При работе с массивами данных в Excel необходимо применять команду с одновременным нажатием клавиш: <Shift>+<Ctrl>+<Enter>.

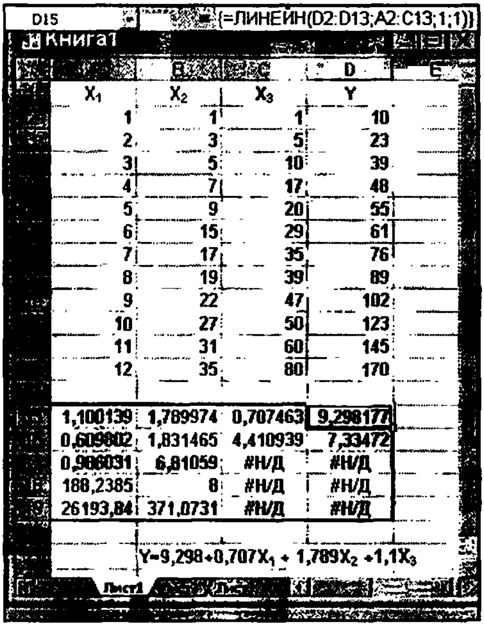
Поэтому вместо команды <закончить> необходимо применять команду <Shift>+<Ctrl>+<Enter>.

В выводимой функцией ЛИНЕЙН таблице результатов представлены следующие параметры (для функции двух независимых переменных):

a_2	a_1	a_0
$\sigma(a_2)$	$\sigma(a_1)$	$\sigma(a_0)$
R^2	$\sigma(y)$	
*F*расч	*df*	
SSreg	*SSresid*	

a_0, a_1, a_2 – параметры регрессии; $\sigma(a_0)$, $\sigma(a_1)$, $\sigma(a_2)$ – стандартные отклонения параметров; $\sigma(y)$ – стандартное отклонение y; R^2 – коэффициент детерминации, *Fрасч* – *F-статистика*; *df* – число степеней свободы, определяемое по формуле $df = n - (m + 1)$, где n – количество исходных данных (число строк в таблице), m – число переменных; *SSreg* – регрессионная сумма квадратов; *SSresid* – остаточная сумма квадратов.

Для функции трех переменных пример вывода показателей представлен на рис. 12.30.


D15 {=ЛИНЕЙН(D2:D13;A2:C13;1;1)}

Книга1

X_1	X_2	X_3	Y
1	1	1	10
2	3	5	23
3	5	10	39
4	7	17	48
5	9	20	55
6	15	29	61
7	17	35	76
8	19	39	89
9	22	47	102
10	27	50	123
11	31	60	145
12	35	80	170

1,100139	1,789974	0,707463	9,298177
0,609802	1,831465	4,410939	7,33472
0,986031	6,81059	#Н/Д	#Н/Д
188,2385	8	#Н/Д	#Н/Д
26193,84	371,0731	#Н/Д	#Н/Д

$Y=9,298+0,707X_1 + 1,789X_2 +1,1X_3$

Лист1

Рис. 12.30. Вывод коэффициентов и показателей статистической значимости при помощи функции ЛИНЕЙН

Функция ЛИНЕЙН применима и к функциям с одной переменной и является очень удобным инструментом с одним, правда, неудобным обстоятельством: коэффициенты регрессий и их стандартные отклонения выводятся справа налево.

Установка параметров нелинейной множественной регрессии (например, такой как степенная регрессия) производится линеаризацией функции (для степенной путем логарифмирования), а затем применением функции ЛИНЕЙН. В случае линеаризации логарифмированием свободный член уравнения, а также его стандартное отклонение, выводятся логарифмированными. Для получения настоящего значения a_0 необходимо основание логарифма возвести в степень полученного значения a_0. Достоверность данного параметра можно соотносить только с линеаризированным видом, т е. устанавливать достоверность параметра $\ln a_0$.

Осуществлять прогнозирование при линейной множественной регрессии можно с помощью функции ТЕНДЕНЦИЯ.

Для установления параметров и статистик нелинейной множественной регрессии в Excel есть только одна функция – экспонента. Функция имеет название ЛГРФПРИБЛ. Применение этой функции имеет одну особенность, которую необходимо помнить для успешного применения. Дело в том, что коэффициенты этой функции выводятся не в обычном для экспоненты виде

$Y = ae^{b_1X_1} \times e^{b_2X_2} \times \ldots \times e^{b_nX_n}$,

а в виде

$Y = ak_1^{X_1} \times k_2^{X_2} \times \ldots \times k_n^{X_n}$,

где $k=e^b$.

То есть функция ЛГРФПРИБЛ выводит следующую матрицу показателей (для двух переменных):

$\exp(b_2)$	$\exp(b_1)$	a
$\sigma(b_2)$	$\sigma(b_1)$	$\ln\sigma(a)$
R^2	$\ln\sigma(y)$	
Fрасч	df	
$SSreg$	$SSresid$	

Это важно для вычисления *t-статистик*. Для этого первоначально необходимо выведенные функцией коэффициенты прологарифмировать.

Осуществлять прогнозирование при экспоненциальной множественной регрессии можно с помощью функции РОСТ.

12.7. Функции регрессионного анализа в Mathcad

Основные функции регрессионного анализа находятся в разделе **Regression and Smooting** в диалоговом окне **Insert Function**, а также **Curve Fitting**. Функции, необходимые для выполнения тестов, находятся в разделах **Probability Density** и **Probability Distribution** (рис. 12.31).

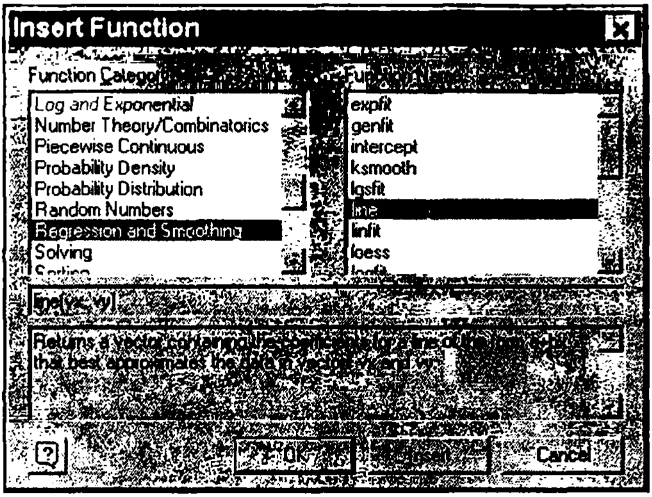


Рис. 12.31. Диалоговое окно ***Insert Function***

12.8. Установка параметров линейной связи в Mathcad

Введем матрицу данных X-У

$$\text{Данные} := \begin{pmatrix} 0 & 1 & 2 & 4 & 6 & 15 & 5 & 8 & 12 & 25 & 22 & 23 & 22 & 27 \\ 0 & 10 & 14 & 20 & 9 & 30 & 14 & 40 & 25 & 56 & 40 & 46 & 56 & 70 \end{pmatrix}$$

$$\text{data} := \text{Данные}^{T}$$

$$X := \text{data}^{\langle 0 \rangle} \qquad Y := \text{data}^{\langle 1 \rangle}$$

Количество данных

$$n := \text{rows(data)} \qquad n = 14$$

Стандартное отклонение

$$SD(x) := \text{stdev}(x) \cdot \sqrt{\frac{n}{n-1}}$$

	X	Y
Среднее	mean(X) = 12.286	mean(Y) = 30.714
Медиана	median(X) = 10	median(Y) = 27.5
Стандартное отклонение	SD(X) = 9.817	SD(Y) = 21.084
Дисперсия	$SD(X)^2 = 96.374$	$SD(Y)^2 = 444.527$

Регресионные статистики

Intercept-свободный член в уравнении регрессии $\quad b_0 := \mathrm{intercept}(X, Y) \quad b_0 = 6.305$

Slope-коэффициент при независимой переменной $\quad b_1 := \mathrm{slope}(X, Y) \quad b_1 = 1.987$

Коэффициент корреляции $\quad \mathrm{corr}(X, Y) = 0.9251$

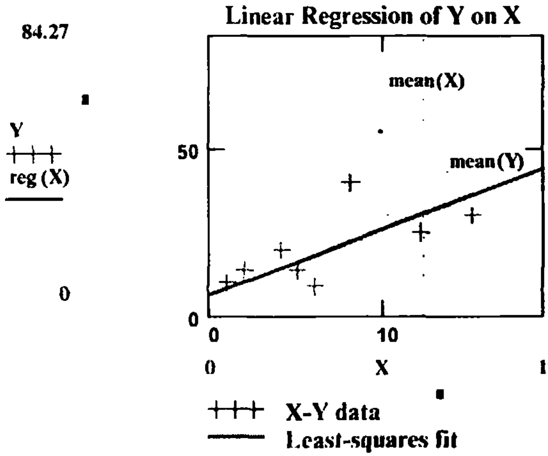
R^2 коэффициент детерминации $\quad \mathrm{corr}(X, Y)^2 = 0.856$

коварнация $\quad \mathrm{cvar}(X, Y) = 177.796$

Стандартная ошибка $\quad \mathrm{stderr}(X, Y) = 8.334$

График $\quad r(x) := b_0 + b_1 \cdot x$

$$\mathrm{scale} := \max\left(\overrightarrow{|r(X) - Y|}\right) \cdot 1.1$$



Введем тестируемую величину коэффициента регрессии

$$b_0 := 0.0$$

Определение коэффициентов регрессии

$$b := \mathrm{slope}(X, Y) \qquad a := \mathrm{intercept}(X, Y)$$

$$b = 0.909 \qquad a = 7.172$$

Стандартная ошибка коэффициента регрессии

$$SEb := \sqrt{\frac{\sum_{i=1}^{n}\left[Y_{i-1} - \left(b \cdot X_{i-1} + a\right)\right]^2}{(n-2)\cdot n \cdot \mathrm{var}(X)}} \qquad SEb = 0.222$$

t-статистика

$$t := \frac{b - b_0}{SEb} \qquad t = 4.1$$

Число степеней свободы

$$df := n - 2 \qquad df = 12$$

Двусторонний тест

$$H_0:\ b = b_0 \qquad H_1:\ b \neq b_0$$

Критическое значение t

$$t_t := qt\left(1 - \frac{\alpha}{2}, df\right) \qquad t_t = 2.179$$

Тест нулевой гипотезы по величине вероятности

$$\frac{\alpha}{2} < pt(t, df) < 1 - \frac{\alpha}{2} = 0 \qquad pt(t, df) = 0.9993$$

Нулевая гипотеза $b = b_0$ отвергается

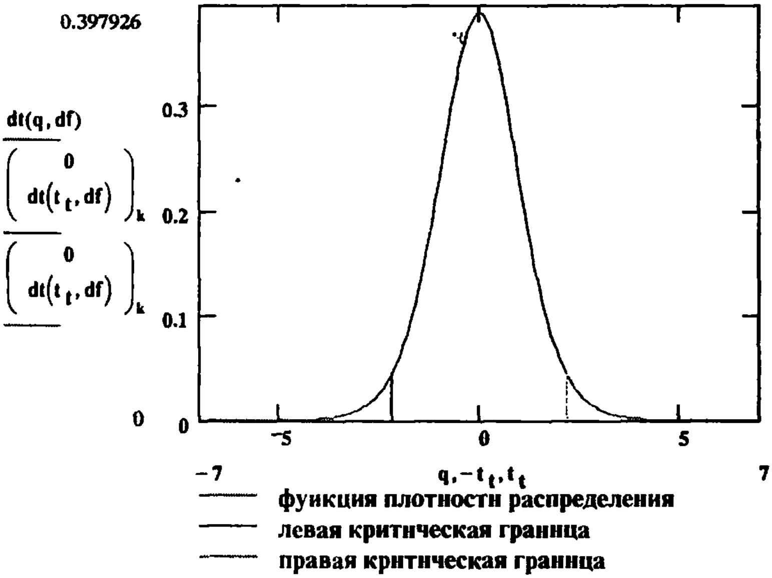
Тест нулевой гипотезы по количеству стандартных отклонений

$|t| < t_t$ **Не выполняется**

Нулевая гипотеза $b = b_0$ **отвергается**

Графическое представление теста

$q := -7, -6.9 .. 7 \qquad k := 0 .. 1$



Левый односторонний тест

H_0: $b \ge b_0$ $\qquad$ H_1: $b < b_0$

Критическое значение t

$t_L := qt(\alpha, df)$ $\qquad t_L = -1.782$

t-статистика

$b_0 := 1.5$

$t = -2.668 \qquad t := \dfrac{b - b_0}{SEb}$

Тест нулевой гипотезы по величине вероятности

$pt(t, df) > \alpha$

Тест нулевой гипотезы по количеству стандартных отклонений

$t > t_L$ **Не выполняется**

$pt(t, df) = 0.010$ $-2.668 < -1.782$

Нулевая гипотеза $b \ge b_0$ отвергается

Графическое представление теста

$q := -7, -6.9 .. 7$ $k := 0 .. 1$

Правый односторонний тест

H_0: $b \le b_0$ H_1: $b > b_0$

Критическое значение t

$t_R := qt(1 - \alpha, df)$ $t_R = 1.782$

Тест нулевой гипотезы по величине вероятности

$$pt(t, df) < 1 - \alpha$$

Тест нулевой гипотезы по количеству стандартных отклонений

$$t < t_R$$

t-статистика

$$b_0 := 1.0 \qquad t := \frac{b - b_0}{SEb}$$

$$t = -0.412$$

$$pt(t, df) = 0.344 \qquad -0.412 < 1.782$$

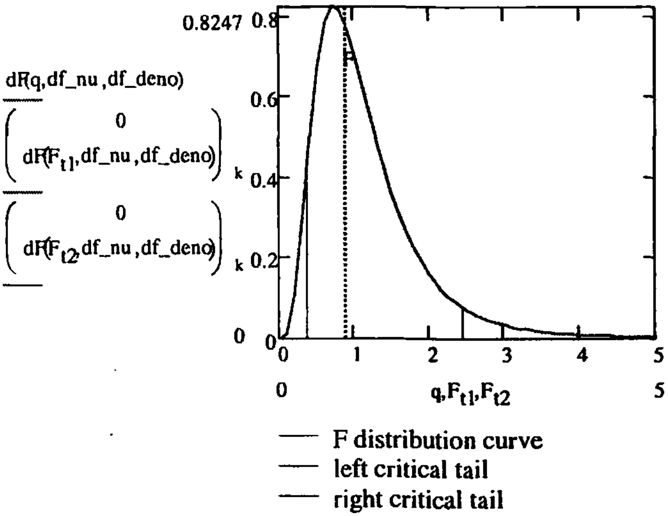
Нулевая гипотеза принимается

Графическое представление теста

$$q := -7, -6.9 .. 7 \qquad k := 0 .. 1$$

Оценка общего качества уравнения связи

F-распределение имеет общий вид:



В данном примере тест будет иметь вид:

$$\mathrm{corr}(X, Y) = 0.764$$

Коэффициент детерминации

$$R2 := \mathrm{corr}(X, Y)^2$$

$$R2 = 0.584$$

F-статистика

$$F := R2 \cdot \frac{df}{(1 - R2)}$$

$$F = 16.814$$

Вероятность

$$pF(F, 1, df) = 0.999$$

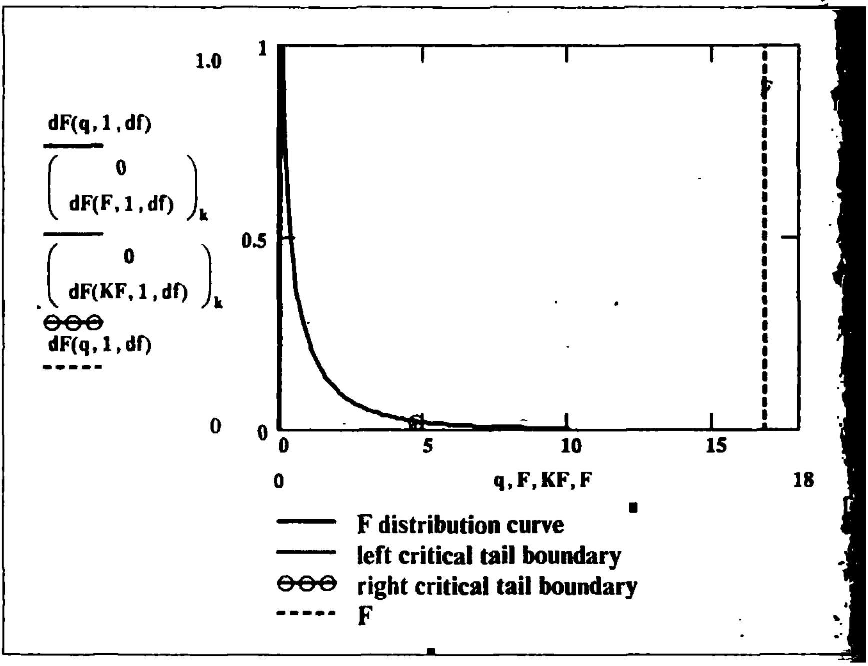
Критическое значение F

$$KF := qF(0.95, 1, 12)$$

$$KF = 4.7472$$

Графическое представление теста

$q := 0, 0.1 .. 10 \qquad k := 0 .. 1$



Коэффициент детерминации $\qquad R2 := 0.3 \qquad df := 8$

F-статистика

$$F := R2 \cdot \frac{df}{(1 - R2)}$$

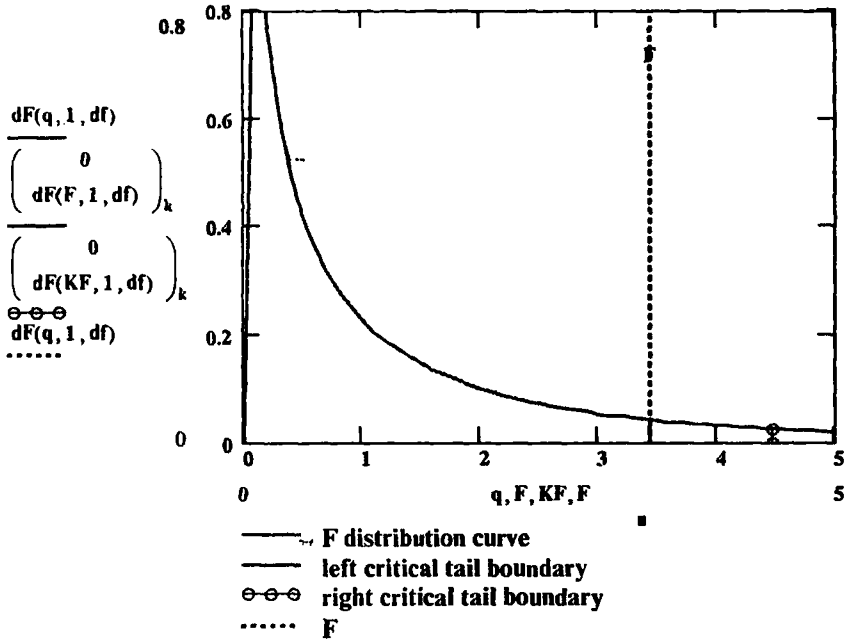
$$F = 3.429$$

Вероятность $\qquad pF(F, 2, df) = 0.916$

Критическое значение F $\qquad KF := qF(0.95, 2, df)$

$q := 0, 0.1 .. 5 \qquad k := 0 .. 1 \qquad KF = 4.459$

Графическое представление теста



12.10. Общие принципы оценки нелинейных регрессий

Мы уже рассматривали подходы к оценке нелинейных регрессий. Это линеаризация исходных функций путем логарифмирования или замены переменных. В этом случае применим метод наименьших квадратов. Однако не все функции можно привести к линейному виду. Например, вы считаете, что переменная y связана с переменной x следующим соотношением:

$$y = a_0 + a_1 x^{a_2} + a_3.$$

Данное уравнение не может быть преобразовано в уравнение линейного вида, поэтому в данном случае невозможно применение обычной процедуры оценивания регрессии.

Тем не менее для получения оценок параметров вы по–прежнему можете применить принцип минимизации суммы квадратов отклонений.

Эта вычислительная процедура состоит из следующих шагов.

Принимаются некоторые исходные значения параметров. Вычисляются значения y по фактическим значениям x с использованием этих значений параметров. Вычисляются остатки для всех наблюдений в выборке и сумма квадратов остатков – S. Далее применяется итерационный алгоритм минимизации S. Как правило, для этого применяются методы нелинейного программирования, которые рассмотрены в соответствующей главе. В частности, для решения данной задачи применяется метод Левенберга–Марквардта.

В Mathcad предусмотрен ряд встроенных функций для установки параметров нелинейных регрессий.

12.11. Установка параметров нелинейных регрессий в Mathcad

Степенная функция

Для установки параметров степенной функции в Mathcad есть несколько возможностей. Во-первых, степенную функцию можно привести к линейному виду путем логарифмирования, и затем применить функции **intercept()**, **slope()** или **line()**. Другие возможности – это использование функций **pwrfit** и **genfit**. Рассмотрим применение данных функций.

Введем исходные данные

DATA :=

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
0	0.1	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	2	6	12	20	28	35	46	60	72	85	92	60	100	90	140

$$\mathbf{data} := \mathbf{DATA}^{\mathrm{T}}$$

$$\mathbf{X} := \mathbf{data}^{\langle 0 \rangle} \qquad \mathbf{Y} := \mathbf{data}^{\langle 1 \rangle}$$

Рассмотрим метод линеаризации логарифмированием и применения функции **line**.

$$q := \overrightarrow{\ln(X)} \qquad s := \overrightarrow{\ln(Y)}$$

$$u := \mathrm{line}(q, s)$$

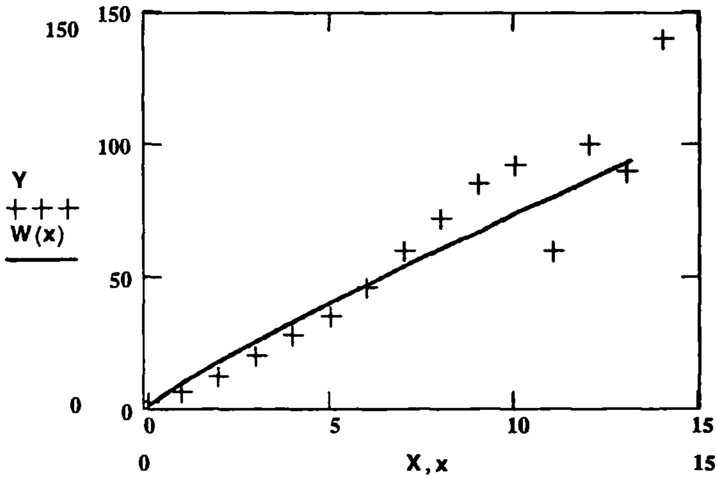
$$a := \exp(u_0) \qquad b := u_1$$

$$a = 9.544 \qquad b = 0.888$$

$$W(x) := a \cdot x^b$$

$$\mathrm{corr}\left(\overrightarrow{W(X)}, Y\right) = 0.948$$

График регрессии



Рассмотрим применение функции **pwrfit**.
Она применяется для регрессий вида $y = a \cdot x^{b} + c$

Зададим первоначальный вектор значений параметров

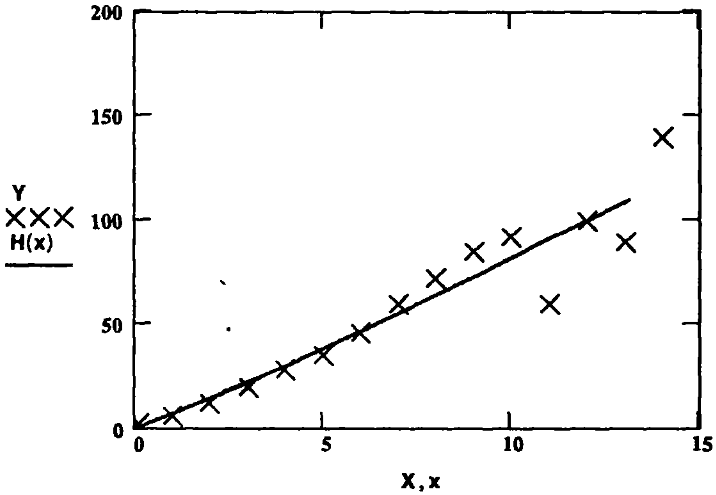
$$\mathbf{Guess} := \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}$$

Используя функцию **pwrfit** найдем параметры регрессии

$$\mathbf{P} := \mathbf{pwrfit}(\mathbf{X}, \mathbf{Y}, \mathbf{Guess}) \qquad \mathbf{P} = \begin{pmatrix} 6.602 \\ 1.094 \\ -0.244 \end{pmatrix}$$

График функции

$$H(x) := P_0 \cdot x^{P_1} + P_2 \qquad x := \min(X) .. \max(X)$$



Теперь рассмотрим применение функции **genfit**.
для установления параметров регрессии вида $y = a \cdot x^b$

$$F(n, a) := \begin{pmatrix} a_0 \cdot n^{a_1} \\ n^{a_1} \\ a_0 \cdot a_1 \cdot n^{a_1 - 1} \end{pmatrix}$$

Задание вектора для функции **genfit**.

$$guess := \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

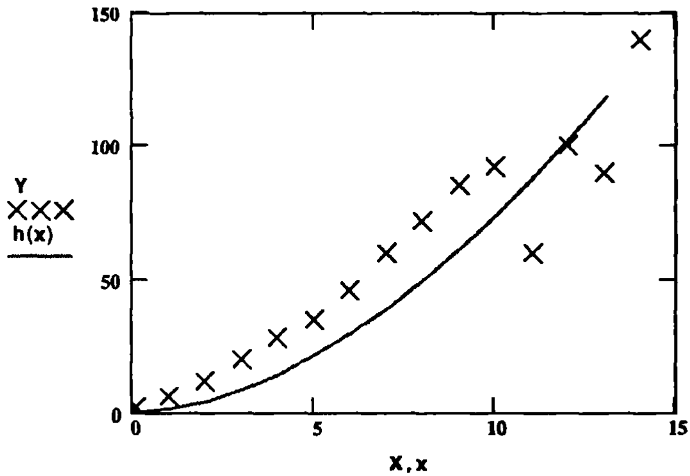
Зададим первоначальный вектор значений параметров.

Используя функцию **genfit** найдем параметры степенной регрессии:

$$G := genfit(X, Y, guess, F)$$

$$G = \begin{pmatrix} 1.189 \\ 1.79 \end{pmatrix}$$

$$h(x) := G_0 \cdot x^{G_1}$$

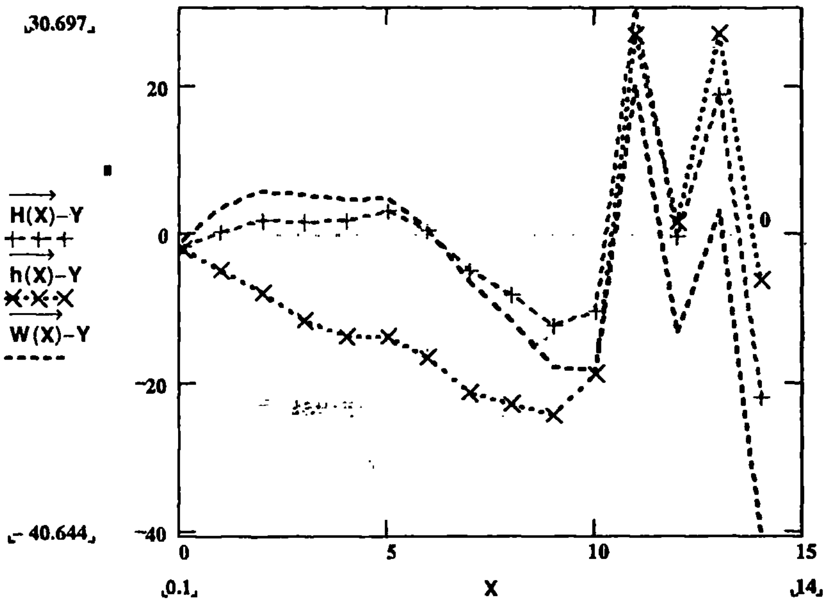


Сравним графики остатков

$$\max\left(\overrightarrow{H(X) - Y}\right) = 30.697 \qquad \max\left(\overrightarrow{W(X) - Y}\right) = 20.208$$

$$\min\left(\overrightarrow{H(X) - Y}\right) = -21.853 \qquad \min\left(\overrightarrow{W(X) - Y}\right) = -40.644$$

$$\max\left(\overrightarrow{h(X) - Y}\right) = 27.215 \qquad \min\left(\overrightarrow{h(X) - Y}\right) = -24.307$$



и коэффициенты корреляции $\text{corr}\left(\overrightarrow{H(X)}, Y\right) = 0.951$

$$\text{corr}\left(\overrightarrow{W(X)}, Y\right) = 0.948$$

$$\text{corr}\left(\overrightarrow{h(X)}, Y\right) = 0.934$$

Экспоненциальная функция

Для установки параметров экспоненциальной функции можно также линеаризовать функцию логарифмированием, а также применить функции **expfit** и **genfit**.

Введем исходные данные

DATA :=

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	0.1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	2	6	11	8	12	20	28	18	34	40	42	50	48	60	80	20

$$data := DATA^T$$

$$X := data^{\langle 0 \rangle} \qquad Y := data^{\langle 1 \rangle}$$

Рассмотрим метод линеаризации логарифмированием и применения функции **line**:

$$s := \overrightarrow{\ln(Y)}$$

$$u := \text{line}(X, s)$$

$$a := \exp(u_0) \qquad b := u_1$$

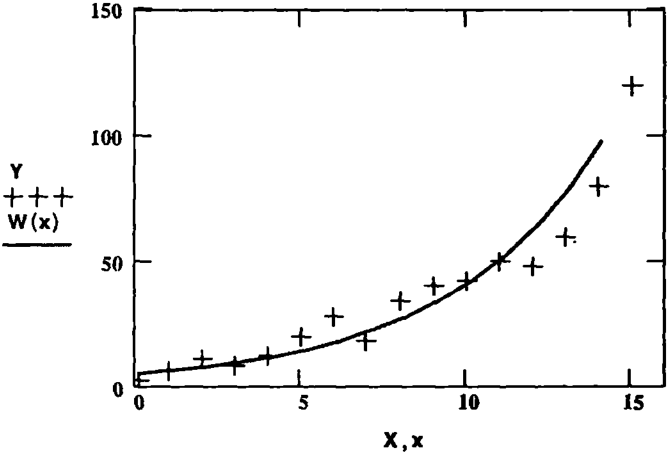
$$a = 4.765 \qquad b = 0.214$$

$$W(x) := a \cdot \exp(b \cdot x)$$

$$\text{corr}\left(\overrightarrow{W(X)}, Y\right) = 0.97411$$

График регрессии

$$x := \min(X) .. \max(X)$$



Рассмотрим применение функции **expfit**.
Она применяется для регрессий вида $y = a \cdot e^{b \cdot x} + c$

$$\text{Guess} := \begin{pmatrix} 3 \\ 0.1 \\ 1 \end{pmatrix}$$

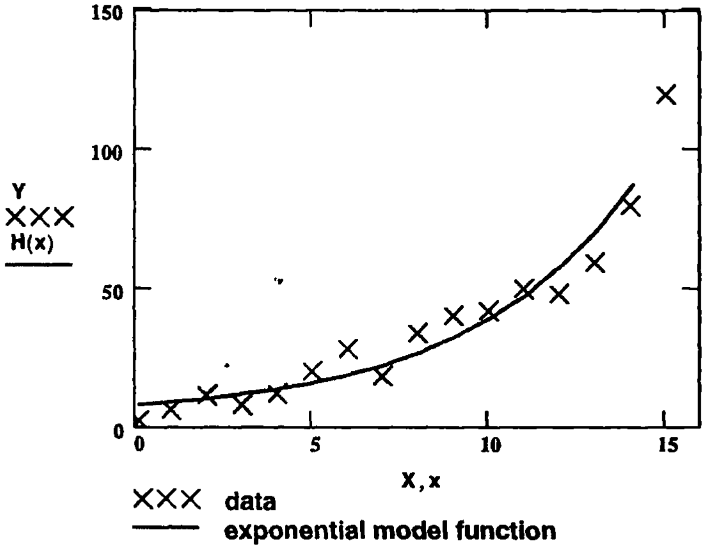
Зададим первоначальный вектор значений параметров.

Используя функцию **expfit** найдем параметры регрессии:

$$E := \text{expfit}(X, Y, \text{Guess}) \qquad E = \begin{pmatrix} 4.03 \\ 0.215 \\ 3.879 \end{pmatrix}$$

График функции

$$H(x) := E_0 \cdot e^{E_1 \cdot x} + E_2 \qquad x := \min(X) .. \max(X)$$



$$\mathrm{corr}\left(\overrightarrow{H(X)}, Y\right) = 0.97413$$

Теперь рассмотрим применение функции **genfit**.
для установления параметров регрессии вида $y = a \cdot e^{b \cdot x}$

$$F(n, a) := \begin{pmatrix} a_0 \cdot e^{a_1 \cdot n} \\ e^{a_1 \cdot n} \\ a_0 \cdot n \cdot e^{a_1 \cdot n} \end{pmatrix}$$

Зададим вектор для функции **genfit**, где первый элемент есть функция, второй элемент есть частная производная по первому параметру, третий элемент есть частная производная по второму параметру.

$$guess := \begin{pmatrix} 3 \\ 0.1 \end{pmatrix}$$

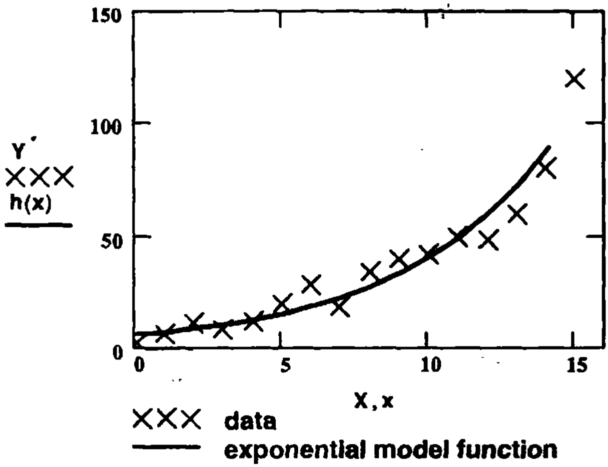
Зададим первоначальный вектор значений параметров.

Используя функцию **genfit** найдем параметры регрессии:

$$G := genfit(X, Y, guess, F)$$

$$G = \begin{pmatrix} 5.75 \\ 0.194 \end{pmatrix}$$

$$h(x) := G_0 \cdot e^{G_1 \cdot x}$$



$$\mathrm{corr}\left(\overrightarrow{h(X)}, Y\right) = 0.9733$$

Логарифмическая функция

Параметры логарифмической функции можно установить с помощью функций **logfit** и **linfit**.

Введем исходные данные

DATA :=

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	2.6	4	4.7	5	6	5.7	5.9	6.1	6.6	7.1	6.8	7	6.9	7.4	7.5

$$data := DATA^T$$

$$X := data^{\langle 0 \rangle} \qquad Y := data^{\langle 1 \rangle}$$

Рассмотрим применение функции **logfit.**
Она применяется для регрессий вида $y = a \cdot \ln(x + b) + c$

$$G := \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

Зададим первоначальный вектор значений параметров

Используя функцию **logfit**, найдем параметры регрессии:

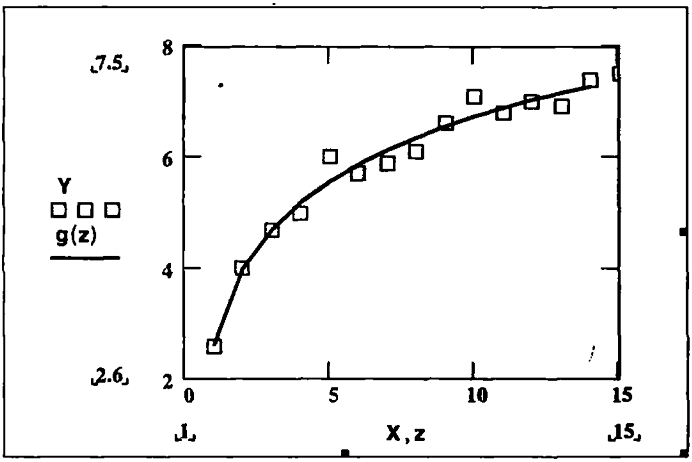
$b := \mathrm{logfit}(X, Y, G)$

$$b = \begin{pmatrix} 1.63 \\ -0.21 \\ 3.002 \end{pmatrix}$$

$g(x) := b_0 \cdot \ln(x + b_1) + b_2$

$$\mathrm{corr}\left(\overrightarrow{g(X)}, Y\right) = 0.9883$$

$z := 0 .. \mathrm{last}(X)$



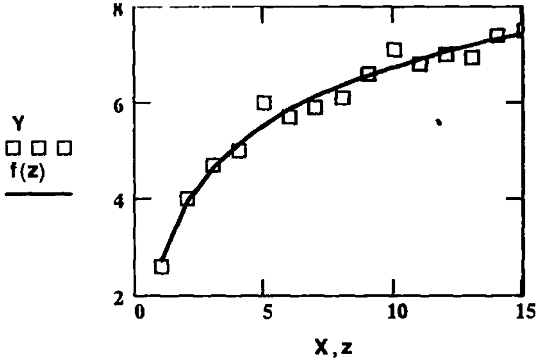
Теперь рассмотрим применение функции **linfit** для установления параметров регрессии вида $y = a \cdot \ln(x) + b$

$F(x) := \begin{pmatrix} \ln(x) \\ 1 \end{pmatrix}$ Задание вектора для функции **linfit**.

$S := \mathrm{linfit}(X, Y, F)$

$$S = \begin{pmatrix} 1.74 \\ 2.717 \end{pmatrix}$$

$f(x) := F(x) \cdot S$ $\quad z := \min(X) .. \max(X)$



$$\text{corr}\left(\overrightarrow{f(X)}, Y\right) = 0.9880$$

Логистическая функция

Параметры логистической регрессии устанавливаются с помощью функций **lgsfit** и **genfit**.

$i := 0..15$

X_i=	Y_i=
1	58
2	56
3	54
4	47
5	30
6	29
7	20
8	13
9	8
10	6
11	3
12	7
13	1.4
14	0.5
15	0.3
16	0.5

Функция **lgsfit** применяется для определения параметров функции

$$y = \frac{a_0}{1 + a_1 \cdot e^{-a_2 \cdot x}}$$

Вектор первоначальных значений устанавливается с помощью выражений

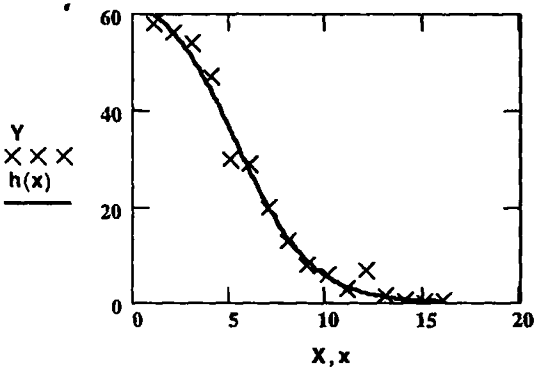
$a := \text{intercept}(X, Y)$ $\qquad b := \frac{\text{mean}(X)}{100}$

$$\text{guess} := \begin{pmatrix} a \\ b \\ -1 \end{pmatrix}.$$

$a := \text{lgsfit}(X, Y, \text{guess})$ $\qquad a = \begin{pmatrix} 66.447 \\ 0.066 \\ -0.512 \end{pmatrix}$

$\text{corr}(X, Y) = -0.93845$

$$h(x) := \frac{a_0}{1 + a_1 \cdot e^{-a_2 \cdot x}} \qquad x := \min(X), \min(X) + 0.1 .. \max(X)$$



Рассмотрим применение функции **genfit**.

$$F(n, a) := \begin{bmatrix} \dfrac{a_0}{1 + a_1 \cdot e^{-a_2 n}} \\ \dfrac{1}{1 + a_1 \cdot e^{-a_2 n}} \\ \dfrac{-a_0 \cdot e^{-a_2 n}}{\left(1 + a_1 \cdot e^{-a_2 n}\right)^2} \\ \dfrac{a_0 \cdot a_1 \cdot n \cdot e^{-a_2 n}}{\left(1 + a_1 \cdot e^{-a_2 n}\right)^2} \end{bmatrix}$$

$$a := \mathrm{intercept}(X, Y)$$

$$b := \frac{\mathrm{mean}(X)}{100}$$

$$guess := \begin{pmatrix} a \\ b \\ -1 \end{pmatrix}$$

$$G := \mathrm{genfit}(X, Y, guess, F)$$

$$G = \begin{pmatrix} 66.447 \\ 0.066 \\ -0.512 \end{pmatrix}$$

$$\mathrm{corr}(X, Y) = -0.93845$$

Полиномиальная функция

data :=

	0	1
0	0	9.1
1	1	7.3
2	2	3.2
3	3	4.6
4	4	4.8
5	5	2.9
6	6	5.7
7	7	7.1
8	8	8.8
9	9	10.2

Установим параметры полинома третьей степени вида

$$y := a_0 + a_1 \cdot x + a_2 \cdot x^2 + a_3 \cdot x^3$$

Введем степень полинома:

$$k := 3$$

$X := data^{\langle 0 \rangle}$ $\quad Y := data^{\langle 1 \rangle}$ $\quad n := rows(data)$

Количество данных $\quad n = 10$

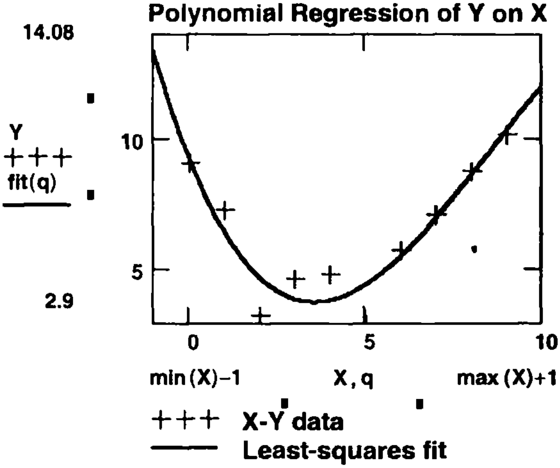
$z := regress(X, Y, k)$ $\quad fit(x) := interp(z, X, Y, x)$

Параметры

$$a := submatrix(z, 3, length(z) - 1, 0, 0)$$

$$a = \begin{pmatrix} 9.298 \\ -3.438 \\ 0.609 \\ -0.024 \end{pmatrix}$$

$$\frac{\sum \overrightarrow{(fit(X) - mean(Y))^2}}{\sum \overrightarrow{(Y - mean(Y))^2}} = 0.88$$ Коэффициент детерминации



12.12. Установка параметров множественной регрессии в Mathcad

Линейная множественная регрессия

Коэффициенты линейной множественной регрессии можно установить как с помощью ввода формулы

$$B = (X^T X)^{-1} X^T Y,$$

так и с помощью встроенных функций **regress** и **interp**:

data :=

	0	1	2
0	20	30	17
1	15	25	14
2	10	18	12
3	7	10	9
4	2	3	6
5	1	2	5

Вид регрессии

$y := a_0 + a_1 \cdot x_1 + a_2 \cdot x_2$

$N := \mathrm{rows}(data)$

$n := \mathrm{cols}(data)$

$Y := data^{\langle 0 \rangle}$

$X := submatrix(data, 0, N-1, 1, n-1)$ то есть $X = \begin{pmatrix} 30 & 17 \\ 25 & 14 \\ 18 & 12 \\ 10 & 9 \\ 3 & 6 \\ 2 & 5 \end{pmatrix}$

$N = 6$ Количество данных

$E := \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$ $W := augment(E, X)$ добавление единичного вектора $W = \begin{pmatrix} 1 & 30 & 17 \\ 1 & 25 & 14 \\ 1 & 18 & 12 \\ 1 & 10 & 9 \\ 1 & 3 & 6 \\ 1 & 2 & 5 \end{pmatrix}$

$$a := \left(W^T \cdot W\right)^{-1} \cdot W^T \cdot Y$$

$$a = \begin{pmatrix} -5.381 \\ 0.178 \\ 1.137 \end{pmatrix} \quad \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix}$$

Теперь рассмотрим применение функций **regress** и **interp** к тем же исходным данным:

$k := 1$ Степень полинома

$z := regress(X, Y, k)$ $\qquad i := 0 .. N-1$

$fit(x) := interp(z, X, Y, x)$

$coeffs := submatrix(z, 3, length(z) - 1, 0, 0)$

$$coeffs = \begin{pmatrix} 0.178 \\ 1.137 \\ -5.381 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_0 \end{pmatrix}$$

$$predY_i := fit\left[\left(X^T\right)^{\langle i \rangle}\right]$$

$$\frac{\sum \overrightarrow{(predY - mean(Y))^2}}{\sum \overrightarrow{(Y - mean(Y))^2}} = 0.989409$$ Коэффициент детерминации

Нелинейная множественная регрессия

Параметры нелинейной множественной регрессии можно устанавливать путем линеаризации и применения метода наименьших квадратов. Встроенные функции **regress(Mxy, VZ,k)**, **interp(VS,Mxy,VZ,V)** и **loes(Mxy,VZ, span)** позволяют выполнять множественную (до 5) полиномиальную регрессию.

$$XY := \begin{pmatrix} 7 & 20 \\ 2 & 12 \\ 3 & 1 \\ 5 & 9 \\ 9 & 25 \\ 4 & 9 \\ 3 & 3 \\ 8 & 15 \\ 15 & 30 \end{pmatrix} \qquad VZ := \begin{pmatrix} 300 \\ 70 \\ 20 \\ 60 \\ 350 \\ 60 \\ 30 \\ 230 \\ 450 \end{pmatrix}$$

$N := 9$

$rows(XY) = 9$

$i := 0 .. N-1$

$k := 2$

$Rg := regress(XY, VZ, k)$

$$Rg = \begin{pmatrix} 3 \\ 3 \\ 2 \\ -5.108 \\ 1.367 \\ 2.766 \\ -144.034 \\ 47.95 \\ 3.82 \end{pmatrix}$$

Параметры полиномиальной регрессии можно установить с помощью вектора, выводимого функцией **regress**.
Первые три величины являются характеристиками вектора:
встроенная функция (3), векторная позиция первого коэффициента (всегда 3), и степень полинома (k).
Сравните с параметрами регрессии, определяемыми
функцией **submatrix**.

$$coeffs := submatrix(Rg, 3, length(Rg) - 1, 0, 0)$$

$$fit(x) := interp(Rg, XY, VZ, x)$$

$$predVZ_i := fit\left[\left(XY^T\right)^{\langle i \rangle}\right] \qquad coeffs = \begin{pmatrix} -5.108 \\ 1.367 \\ 2.766 \\ -144.034 \\ 47.95 \\ 3.82 \end{pmatrix}$$

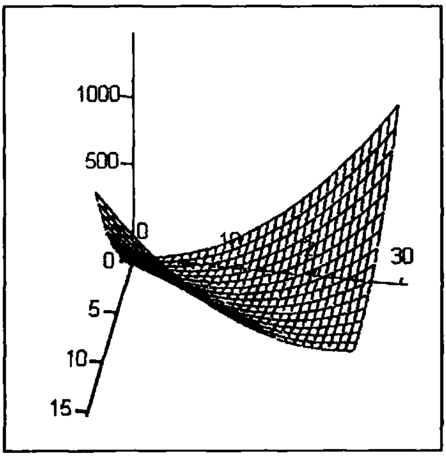
Коэффициент детерминации

$$\frac{\sum \overrightarrow{(predVZ - mean(VZ))^2}}{\sum \overrightarrow{(VZ - mean(VZ))^2}} = 0.982831$$

$$X := XY^{\langle 0 \rangle} \qquad Y := XY^{\langle 1 \rangle}$$

$$x := 0, 1 .. max(X) \qquad y := 0, 1 .. max(Y)$$

$$F_{x,y} := interp\left[Rg, XY, VZ, \begin{pmatrix} x \\ y \end{pmatrix}\right]$$



Глава 13. Специальные методы эконометрики

Наиболее распространенным в практике статистического оценивания параметров уравнений регрессии является метод наименьших квадратов (МНК). Этот метод основан на ряде предпосылок относительно природы данных и результатов построения модели. Основные из них – это некоррелированность факторов, входящих в уравнение, линейность связи, отсутствие автокорреляции остатков, равенство их математических ожиданий нулю и постоянная дисперсия. Эмпирические данные не всегда обладают такими характеристиками, т. е. предпосылки МНК нарушаются. Применение этого метода в чистом виде может привести к таким нежелательным результатам, как смещение оцениваемых параметров, снижение их состоятельности, устойчивости, а в некоторых случаях может совсем не дать решения. Для смягчения нежелательных эффектов при построении регрессионных уравнений, повышения адекватности моделей существует ряд специальных методов, которые применяются в случае нарушения предпосылок применения МНК [8, 11, 14, 22].

13.1. Гетероскедастичность

Здесь мы рассмотрим частный случай регрессионной модели, а именно: модель с гетероскедастичностью. Это означает, что ошибки не коррелированы, но имеют непостоянные дисперсии. Классическая модель с постоянными дисперсиями ошибок называется гомоскедастичной. Возникает гетероскедастичность чаще при анализе неоднородных объектов.

Тесты на гетероскедастичность

Существует несколько тестов на гетероскедастичность. Во всех этих тестах проверяется основная нулевая гипотеза о равенстве дисперсий H_0: $\sigma_1^2 = \sigma_2^2 = ... = \sigma_n^2$, (наличие гомоскедастичности, отсутствие гетероскедастичности) против альтернативной гипотезы H_1: не H_0.

Рассмотрим один из самых распространенных тестов – тест Голдфелда–Квандта. При проведении проверки по этому критерию предполагается, что стандартное отклонение распределения вероятностей ε_i пропорционально значению x в этом наблюдении. Предполагается также, что случайный член распределен нормально и не подвержен автокорреляции.

Все n наблюдений в выборке упорядочиваются по величине x, после чего оцениваются отдельные регрессии для первых n^1 и для последних n^2 наблюдений; средние $(n - n^1 - n^2)$ наблюдений отбрасываются. Если предположение относительно природы гетероскедастичности верно, то дисперсия ε в последних n^2 наблюдениях будет больше, чем в первых n^1, и это будет отражено в сумме квадратов остатков в двух указанных «частных» регрессиях. Обозначая суммы квадратов остатков в регрессиях для первых n^1 и последних n^2 наблюдений соответственно через RSS_1 и RSS_2, рассчитаем отношение RSS_2/RSS_1, которое имеет

F-распределение с (n^1 – m – 1) и (n^2 – m – 1) степенями свободы, где m – число объясняющих переменных в регрессионном уравнении. Мощность критерия зависит от выбора n' по отношению к n. Основываясь на результатах некоторых проведенных ими экспериментов, С. Голфелд и Р. Квандт утверждают, что n' должно составлять порядка 11, когда n = 30, и порядка 22, когда n = 60. Если в модели имеется более одной объясняющей переменной, то наблюдения должны упорядочиваться по той из них, которая, как предполагается, связана с σ_i.

Метод Голфелда – Квандта может также использоваться для проверки на гетероскедастичность при предположении, что σ_i обратно пропорционально ... При этом используется та же процедура, что и описанная выше, но тестовой статистикой теперь является показатель RSS_1/RSS_2, который вновь имеет F-распределение с (n^1 – m – 1) и (n^2 – m – 1) степенями свободы.

Рассмотрим пример выполнения теста в Excel. По 34 странам оценивалась регрессия расходов на образование от валового национального продукта (ВНП). На основе данных, приведенных на рис. 13.1, с помощью функции ЛИНЕЙН были оценены регрессии сначала по наблюдениям 12 стран с наименьшим ВНП, а затем для 12 стран с наибольшим ВНП. Сумма квадратов отклонений в первой регрессии равна 2,68, а во второй – 388,24.

	A	B	C	D	E	F	G	H
1					Продолжение			
2	Страна	Расходы	ВНП	Страна	Расходы	ВНП		
3	1	0,34	5,67	18	5,31	101,65	0,041351	0,082066
4	2	0,22	10,13	19	6,4	115,97	0,012292	0,327561
5	3	0,32	11,34	20	7,15	119,49	0,530905	0,518027
6	4	1,23	18,88	21	11,22	124,15	11,31764	10
7	5	1,81	20,94	22	8,66	140,98	3,037108	2,683617
8	6	1,02	22,16	23	5,56	153,86		
9	7	1,27	23,83	24	13,41	169,38		
10	8	1,07	24,67	25	5,46	186,33	144,6743	4,849142
11	9	0,67	27,56	26	4,79	211,78		
12	10	1,25	27,57	27	8,92	249,72		
13	11	0,76	40,16	28	18,9	261,41		
14	12	2,8	51,62	29	15,95	395,52	0,071109	-8,18725
15	13	4,9	57,71	30	29,9	534,97	0,002738	2,445304
16	14	3,5	63,03	31	33,59	656,29	0,98539	6,230858
17	15	4,45	66,32	32	38,62	815	674,4532	10
18	16	1,6	66,97	33	61,61	1040,45	26184,69	388,2359
19	17	4,26	76,88	34	181,3	2586,4		

Лист8 Лист7 Лист9 Лист6 Лист1

Рис. 13.1. Тест на гетероскедастичность

Соотношение RSS_2/RSS_1, следовательно, составило 144,67. Критическое значение $F(10,10)$ равно 4,849 при однопроцентном уровне значимости, и нулевая гипотеза об отсутствии гетероскедастичности отклоняется.

Метод взвешенных наименьших квадратов

На первом этапе данного метода оценивается линейная регрессионная модель

$$Y = BX + \varepsilon$$

с помощью обычного МНК. Предполагается, что остатки e_i независимы между собой, но имеют разные дисперсии. Поскольку теоретические отклонения ε_i нельзя рассчитать, их обычно заменяют фактическими отклонениями зависимой переменной от линии регрессии e_i. Предполагается, что ковариационная матрица вектора ошибок ε диагональна, $V(e_i) = \sigma_i^2$, $i = 1,...,n$.

Если величины σ_i^2 известны, то делением исходного регрессионного уравнения на σ_i, получаем (выписав каждое уравнение):

$$\frac{Y_i}{\sigma_i} = \sum_{j=1}^{k} b_i \frac{X_{ij}}{\sigma_i} + u_i, \qquad i = 1,...,n,$$

где $u_i = \varepsilon_i/\sigma_i$, причем $V(u_i) = 1$, $\mathrm{Cov}(u_i,u_s) = 0$ при $i \neq s$. Применяя к полученному уравнению стандартный метод наименьших квадратов, оценку получаем минимизацией по $b = (b_1,...,b_k)$ суммы:

$$f(b) = \sum_{i=1}^{n}\left[\frac{1}{\sigma_i}\left(Y_i - \sum_{j=1}^{k} b_j X_{ij}\right)\right]^2 .$$

Содержательный смысл этого преобразования заключается в следующем. Используя обычный метод наименьших квадратов, мы минимизируем сумму квадратов отклонений, в которую, говоря не строго, разные слагаемые вносят разный статистический вклад из-за различных дисперсий, что в конечном итоге и приводит к неэффективности МНК-оценки. «Взвешивая» каждое наблюдение с помощью коэффициента $1/\sigma_i$, мы устраняем такую неоднородность. Применение метода взвешенных наименьших квадратов приводит к уменьшению дисперсий оценок по сравнению с обычным методом наименьших квадратов.

Таким образом, для учета гетероскедастичности в случае пропорциональности дисперсии одному или нескольким регрессорам можно использовать двухшаговую процедуру оценки. Такая двухшаговая процедура дает асимптотически несмещенные оценки стандартных ошибок коэффициентов регрессии.

Предполагается, что дисперсия ошибки есть линейная функция от нескольких регрессоров. Допустим, например:

$$\sigma_i^2 = a_0 + b_1 x_1 + b_2 x_2.$$

На первом шаге процедуры оценивается регрессионное уравнение модели:

$$y = a_0 + \sum b_i x_i + \varepsilon.$$

Устанавливаются остатки e_i, которые подставляются в первое уравнение вместо σ_i. По полученному уравнению находится состоятельная оценка вектора дисперсий σ^2_i.

На втором шаге полученные оценки σ^2_i используются в качестве весовых коэффициентов для взвешенного метода наименьших квадратов (i-е уравнение делится на σ^2_i).

Препятствием для этой процедуры является то, что на практике, как правило, неизвестны фактические значения σ_i. Однако процедура будет применимой, если мы сможем подобрать некоторую величину, пропорциональную, по нашему мнению, σ в каждом наблюдении, и разделим на нее обе части уравнения.

Например, можно предположить, что σ приблизительно пропорциональна x, как в критерии Голдфелда–Квандта. Если после этого мы разделим каждое наблюдение на соответствующее ему значение x, то исходное уравнение примет вид:

$$\frac{y}{x} = a\frac{1}{x} + b + \frac{u}{x},$$

и при этом, возможно, новый случайный член u/x будет иметь постоянную дисперсию. Затем необходимо оценить регрессионную зависимость y/x от $1/x$, включив в уравнение постоянный член. Коэффициент при $1/x$ будет эффективной оценкой a, постоянный член – эффективной оценкой b. В предыдущем примере зависимой переменной будет доля расходов на образование в ВНП, а объясняющей переменной – обратная к ВНП величина. На рис. 13.2 приведены результаты расчетов для этого случая.

	A	B	C	F	G	H	I	J	K
1									
2	Страна	Расход	ВНП	Расх/ВНП	1/ВНП	-0,0657	0,053204		
3	1	0,34	5,67	0,059965	0,176367	0,094411	0,004112	Вся регрессия	
4	2	0,22	10,13	0,021718	0,098717	0,014909	0,019452		
5	3	0,32	11,34	0,028219	0,088183	0,484296	32		
6	4	1,23	18,88	0,065148	0,052966	0,000183	0,012109		
7	5	1,81	20,94	0,086437	0,047755	-0,69591	12,9403		
8	6	1,02	22,16	0,046029	0,045126				
9	7	1,27	23,83	0,053294	0,041964	0,02968	0,043811		
10	8	1,07	24,67	0,043373	0,040535	0,145349	0,010513		
11	9	0,67	27,66	0,024311	0,036284	0,004153	0,021041	Первые	
12	10	1,26	27,67	0,045339	0,036271	0,041698	10	12 наблюдени	
13	11	0,76	40,16	0,01868	0,024907	1,85E-05	0,004427		
14	12	2,8	51,62	0,054243	0,019372				
15	13	4,9	57,71	0,084907	0,017328	-2,62838	0,058455		
16	14	3,5	63,03	0,055529	0,015865	2,586217	0,009839		
17	15	4,45	66,32	0,067099	0,015078	0,093618	0,017933	Последние	
18	16	1,6	66,97	0,023891	0,014932	1,032872	10	12 наблюдени	
19	17	4,26	76,88	0,055411	0,013007	0,000332	0,003216		
20	18	5,31	101,65	0,052238	0,009838				
21	19	6,4	116,97	0,056187	0,008623	1,376711		F-статистика	
22	20	7,15	119,49	0,059838	0,008369		2,97824	F-крит	0,05
23	21	11,22	124,15	0,090375	0,008055		4,849142	F-крит	0,01

Лист8 / Лист7 / Лист9 / Лист6 / Лист1 / Лист2 / Лис

Рис. 13.2. Коррекция на гетероскедастичность при ошибке пропорциональной независимой переменной

Как видно из рис. 13.2, $RSS_1 = 0{,}0044$ больше, чем $RSS_2 = 0{,}0032$; это показывает, что пересчет более чем компенсировал гетероскедастичность. Тестовая статистика в этом случае $RSS_1/RSS_2 = 1{,}376$ невысокая и указывает на статистическую незначимость гетероскедастичности.

Иногда в нашем распоряжении может оказаться несколько переменных, каждую из которых можно использовать для масштабирования уравнения. В рассмотренном примере альтернативной переменной может быть численность населения страны (H). Разделив обе части исходного уравнения на эту величину, получаем:

$$\frac{y}{H} = a\frac{1}{H} + b\frac{x}{H} + \frac{u}{H},$$

и надеемся на то, что случайный член u_i/H_i будет иметь постоянную дисперсию для всех наблюдений. Таки образом, теперь оценивается регрессионная зависимость государственных расходов на образование на душу населения от ВНП на душу населения и обратной величины от численности населения, причем на этот раз без постоянного члена. Статистика численности населения приведена на рис. 13.3.

страна	H	страна	H
1	0,36	18	6,37
2	2,9	19	8,37
3	2,39	20	9,86
4	3,44	21	8,31
5	3,87	22	14,62
6	10,71	23	27,06
7	3,1	24	14,14
8	9,93	25	67,4
9	5,07	26	37,43
10	11,1	27	123,03
11	9,6	28	23,94
12	4,78	29	57,04
13	4,09	30	55,95
14	22,34	31	53,71
15	5,12	32	61,56
16	44,92	33	116,78
17	7,51	34	227,64

Лист7

Рис. 13.3. Данные по численности населения

Данные для новой регрессии необходимо упорядочить по переменной ВНП/Н. Сортировку можно произвести с помощью функции **Данные, Сортировка**. Диалоговое окно этой опции приведено на рис. 13.4.

К отсортированным данным применена указанная схема расчетов, а результаты представлены на рис. 13.5.

В данном случае RSS_2/RSS_1 равняется 4,596, что указывает на то, что нулевая гипотеза о гомоскедастичности должна быть отклонена при уровне значимости в 5% (критическое значение F составляет 2,978).

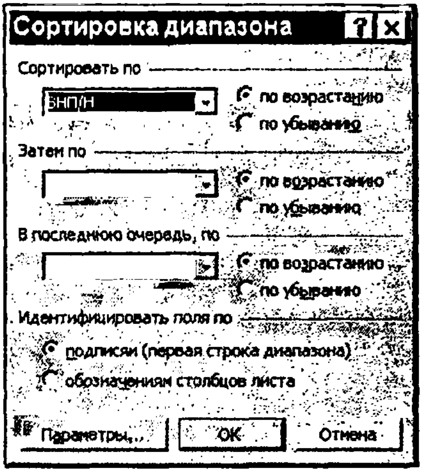
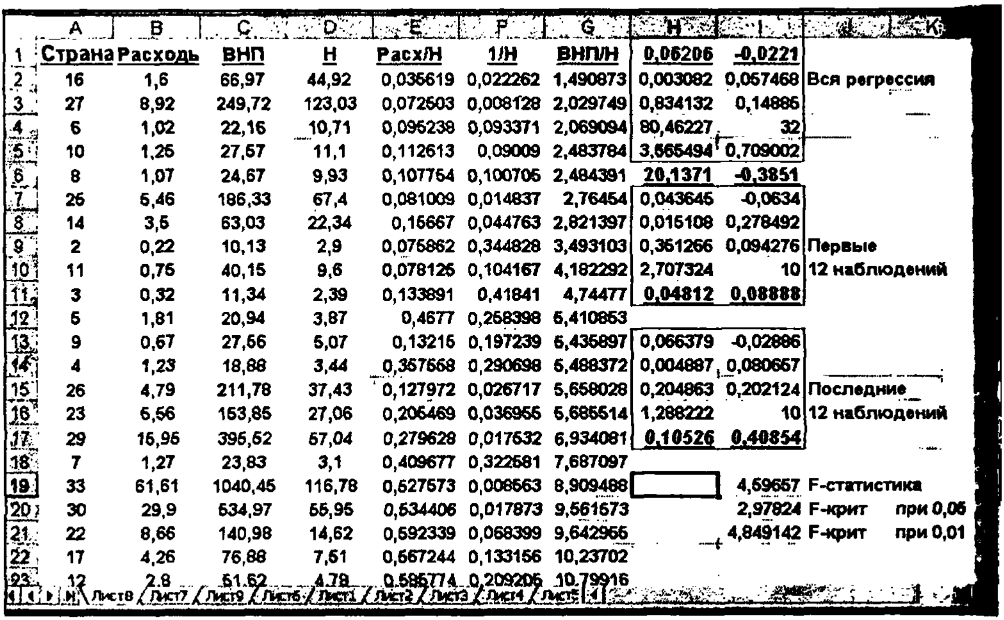


Рис. 13.4. Диалоговое окно
Сортировка диапазона



	A	B	C	D	E	F	G	H	I	J	K
1	Страна	Расходы	ВНП	Н	Расх/Н	1/Н	ВНП/Н	0,06206	-0,0221		
2	16	1,6	66,97	44,92	0,035619	0,022262	1,490873	0,003082	0,057468	Вся регрессия	
3	27	8,92	249,72	123,03	0,072503	0,008128	2,029749	0,834132	0,14886		
4	6	1,02	22,16	10,71	0,096238	0,093371	2,069094	80,46227	32		
5	10	1,25	27,57	11,1	0,112613	0,09009	2,483784	3,665494	0,709002		
6	8	1,07	24,67	9,93	0,107754	0,100705	2,484391	20,1371	-0,3851		
7	26	5,46	186,33	67,4	0,081009	0,014837	2,76454	0,043645	-0,0634		
8	14	3,5	63,03	22,34	0,15667	0,044763	2,821397	0,015108	0,278492		
9	2	0,22	10,13	2,9	0,075862	0,344828	3,493103	0,351266	0,094276	Первые	
10	11	0,75	40,15	9,6	0,078125	0,104167	4,182292	2,707324	10	12 наблюдений	
11	3	0,32	11,34	2,39	0,133891	0,41841	4,74477	0,04812	0,08888		
12	5	1,81	20,94	3,87	0,4677	0,258398	5,410853				
13	9	0,67	27,56	5,07	0,13215	0,197239	5,435897	0,066379	-0,02886		
14	4	1,23	18,88	3,44	0,357558	0,290698	5,488372	0,004887	0,080657		
15	26	4,79	211,78	37,43	0,127972	0,026717	5,658028	0,204863	0,202124	Последние	
16	23	5,56	153,85	27,06	0,205469	0,036955	5,685514	1,288222	10	12 наблюдений	
17	29	15,95	395,52	57,04	0,279628	0,017532	6,934081	0,10526	0,40854		
18	7	1,27	23,83	3,1	0,409677	0,322581	7,687097				
19	33	61,61	1040,45	116,78	0,527573	0,008563	8,909488		4,59657	F-статистика	
20	30	29,9	534,97	55,95	0,534406	0,017873	9,561573		2,97824	F-крит	при 0,05
21	22	8,66	140,98	14,62	0,592339	0,068399	9,642955		4,849142	F-крит	при 0,01
22	17	4,26	76,88	7,51	0,667244	0,133156	10,23702				
23	12	2,8	51,62	4,78	0,585774	0,209205	10,79916				

Лист8 / Лист7 / Лист9 / Лист6 / Лист1 / Лист2 / Лист3 / Лист4 / Лист5

Рис. 13.5. Коррекция на гетероскедастичность
с использованием новой переменной

13.2. Автокорреляция

В значительной части рядов динамики экономических процессов между уровнями, особенно близко расположенными, существует взаимосвязь. Ее удобно представить в виде корреляционной зависимости между рядами $y_1, y_2, y_3, ..., y_n$ и этим же рядом, сдвинутым относительно первоначального положения на h моментов времени $y_{1+h}, y_{2+h}, y_{3+h}, ..., y_{n+h}$. Смещение на время L называется сдвигом, а само явление взаимосвязи – автокорреляцией.

Автокорреляционная зависимость особенно существенна между последующими и предшествующими уровнями ряда динамики. Поскольку классические методы математической статистики применимы лишь в случае независимости отдельных членов ряда между собой, то при анализе нескольких взаимосвязанных рядов динамики важно установить наличие и степень их автокорреляции.

Различаются два вида автокорреляции:

1) автокорреляция в наблюдениях за одной или более переменными;

2) автокорреляция ошибок, или автокорреляция в отклонениях от регрессионной модели.

Наличие последней приводит к искажению величин средних квадратических ошибок коэффициентов регрессии, что затрудняет проверку их значимости.

1. Автокорреляция в наблюдениях за одной или более переменными. Автокорреляцию измеряют при помощи нециклического коэффициента автокорреляции, который может рассчитываться не только между соседними уровнями, т. е. сдвинутыми на один период, но и сдвинутыми на любое число единиц времени (L). Этот сдвиг, именуемый временным лагом, определяет и порядок коэффициентов автокорреляции: первого порядка (при $L = 1$), второго порядка (при $L = 2$) и т. д. Однако наибольший интерес для исследования представляет вычисление нециклического коэффициента первого порядка, т. к. наиболее сильные искажения результатов анализа возникают при корреляции между исходными уровнями ряда (y_t) и теми же уровнями, сдвинутыми на одну единицу времени, т. е. y_{t-1} (y_{t+1}).

Тогда формулу коэффициента автокорреляции можно записать следующим образом:

$$r_a = \frac{\overline{y_t y_{t+1}} - \overline{y_t} \cdot \overline{y_{t+1}}}{\sigma_{y_t} \sigma_{y_{t+1}}},$$

где σ_{yt}, σ_{yt+1} – среднее квадратическое отклонение рядов y_t и y_{t+1}.

Для расчета коэффициента автокорреляции в Excel можно воспользоваться функцией КОРРЕЛ. Предположим, что базовая переменная включает диапазон ячеек A1:A15. Тогда коэффициент автокорреляции равен:

=КОРРЕЛ(A1:A14;A2:A15).

2. Автокорреляция ошибок, или автокорреляция в отклонениях от регрессионной модели. Одним из основных предполагаемых свойств отклонений ε_i от регрессионной модели является их статистическая независимость между собой. Поскольку значения ε_i остаются неизвестными, то проверяется статистическая независимость их аналогов – отклонения e_i. При этом проверяется некоррелированность сдвинутыми на период величинами e_i. Сдвиг производится во времени в случае временных рядов или по возрастанию переменной в случае перекрестных выборок. Для этих величин можно рассчитать коэффициент автокорреляции первого порядка (выборочный коэффициент корреляции между e_i и e_{i-1}):

$$r_e = \frac{\sum_{i=2}^{n} e_i e_{i-1}}{\sqrt{\sum_{i=2}^{n} e_i^2 \cdot \sum_{i=1}^{n-1} e_i^2}} \approx \frac{\sum_{i=2}^{n} e_i e_{i-1}}{\sum_{i=1}^{n} e_i^2}.$$

На практике, однако, в качестве теста используют тесно связанную с коэффициентом автокорреляции r_e статистику Дарбина–Уотсона (*Durbin–Watson*). Тест Дарбина–Уотсона (DW) на наличие или отсутствие автокорреляции ошибок рассчитывается по формуле:

$$DW = \frac{\sum_{i=2}^{n} (e_i - e_{i-1})^2}{\sum_{i=1}^{n} e_i^2}.$$

Нулевая гипотеза состоит в отсутствии автокорреляции. Статистику Дарбина–Уотсона можно выразить через коэффициент автокорреляции:

$$DW \approx 2(1 - r_e).$$

Содержательный смысл теста Дарбина–Уотсона заключается в следующем. Если между e_i и e_{i-1} имеется достаточно высокая корреляция, то e_i и e_{i-1} близки друг к другу и величина статистики DW мала. Это согласуется с последним выражением: если коэффициент r_e близок к единице, то величина DW близка к 2. Другой крайний случай возникает, когда точки наблюдений поочередно отклоняются в разные стороны от линии регрессии, и каждое последующее отклонение e_i имеет, как правило, знак, противоположный предыдущему отклонению e_{i-1}, т. е. $e_i = -e_{i-1}$. В этом случае $(e_i - e_{i-1}) \approx 2e_i$, и $DW = 4$. Это случай отрицательной автокорреляции остатков первого порядка. Во всех других случаях $0 < DW < 4$.

Если бы распределение статистики DW было известно, то для проверки нулевой гипотезы об отсутствии автокорреляции можно было бы для заданного уровня значимости найти такое критическое значение d^*, что если $DW > d^*$, то нулевая гипотеза H_0 принимается, в противном случае она отвергается в пользу наличия автокорреляции. Однако проблема состоит в том, что распределение DW зависит не только от числа наблюдений n и количества регрессоров m, но и от всей матрицы X, и значит практическое применение этой процедуры затруднительно. Тем не менее, Дарбин и Уотсон доказали, что существуют две границы, обычно обозначаемые d_U и d_L, $d_U > d_L$ (u = *upper* – верхняя, l = *low* – нижняя), которые зависят лишь от n, k и уровня значимости (и, следовательно, могут быть затабулированы) и обладают следующим свойством: если $DW > d_U$, то $DW > d^*$, и значит, нулевая гипотеза принимается, а если $DW < d_L$, то $DW < d^*$, и нулевая гипотеза отвергается в пользу альтернативной (табл. 13.1). В случае $d_L < DW < d_U$ ситуация неопределенна, т. е. нельзя высказаться в пользу той или иной гипотезы. Эти результаты можно представить в виде следующей таблицы.

Таблица 13.1

Значение статистики DW	Вывод
$4 - d_L < DW < 4$	Гипотеза H_0 отвергается, есть отрицательная корреляция
$4 - d_U < DW < 4 - d_L$	Неопределенность
$2 < DW < 4 - d_U$	Принимается гипотеза H_0
$d_U < DW < 2$	Принимается гипотеза H_0
$d_L < DW < d_U$	Неопределенность
$0 < DW < d_L$	Гипотеза H_0 отвергается, есть положительная корреляция

Наличие зоны неопределенности, когда нет оснований ни принимать, ни отвергать гипотезу об отсутствии корреляции, представляет определенную трудность при использовании теста Дарбина–Уотсона. В грубом приближении можно сказать, что при числе наблюдений не меньше 12–15 и при 1–3 объясняющих переменных статистика DW должна быть не менее 1 и не более 3. В противном случае признается существование автокорреляции остатков и необходимо улучшить формулу. Если статистика DW находится приблизительно между 1,2–1,3 и 2,7–2,8, можно считать, что статистическая значимая автокорреляция остатков отсутствует. Значения статистики Дарбина–Уотсона приведены в табл. 13.2.

Таблица 13.2. Критерий Дарбина–Уотсона (DW). Значения dL и dU при 5%-ном уровне значимости

n	$m=1$		$m=2$		$m=3$		$m=4$		$m=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	1,08	1,36	0,95	1,54	0,82	1,75	0,69	1,97	0,56	2,21
16	1,10	1,37	0,98	1,54	0,86	1,73	0,74	1,93	0,62	2,15
17	1,13	1,38	1,02	1,54	0,90	1,71	0,78	1,90	0,67	2,10
18	1,16	1,39	1,05	1,53	0,93	1,69	0,82	1,87	0,71	2,06
19	1,18	1,40	1,08	1,53	0,97	1,68	0,86	1,85	0,75	2,02
20	1,20	1,41	1,10	1,54	1,00	1,68	0,9	1,83	0,79	1,99
21	1,22	1,42	1,13	1,54	1,03	1,67	0,93	1,81	0,83	1,96
22	1,24	1,43	1,15	1,54	1,05	1,66	0,96	1,80	0,86	1,94
23	1,26	1,44	1,17	1,54	1,08	1,66	0,99	1,79	0,90	1,92
24	1,27	1,45	1,19	1,55	1,10	1,66	1,01	1,78	0,93	1,90
25	1,29	1,45	1,21	1,55	1,12	1,66	1,04	1,77	0,95	1,89
26	1,30	1,46	1,22	1,56	1,14	1,65	1,06	1,76	0,98	1,88
27	1,32	1,47	1,24	1,56	1,16	1,65	1,08	1,76	1,01	1,86
28	1,33	1,48	1,26	1,56	1,18	1,65	1,10	1,75	1,03	1,85
29	1,34	1,48	1,27	1,57	1,20	1,65	1,12	1,74	1,05	1,84
30	1,35	1,49	1,28	1,57	1,21	1,65	1.14	1,74	1,07	1,83
31	1,36	1,50	1,30	1,57	1,23	1,65	1,16	1,74	1,09	1,83

Таблица 13.2 (окончание)

n	$m=1$		$m=2$		$m=3$		$m=4$		$m=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
32	1,37	1,50	1,31	1,57	1,24	1,65	1,18	1,73	1,11	1,82
33	1,38	1,51	1,32	1,58	1,26	1,65	1,19	1,73	1,13	1,81
34	1,39	1,52	1,33	1,58	1,27	1,65	1,21	1,73	1,15	1,81
35	1,40	1,52	1,34	1,58	1,28	1,65	1,22	1,73	1,16	1,80
36	1,41	1,53	1,35	1,59	1,29	1,65	1,24	1,73	1,18	1,80
37	1,42	1,54	1,36	1,59	1,31	1,66	1,25	1,72	1,19	1,80
38	1,43	1,54	1,37	1,59	1,32	1,66	1,26	1,72	1,21	1,79
39	1,43	1,54	1,38	1,60	1,33	1,66	1,27	1,72	1,22	1,79
40	1,44	1,54	1,39	1,60	1,34	1,66	1,29	1,72	1,23	1,79
45	1,48	1,57	1,43	1,62	1,38	1,67	1,34	1,72	1,29	1,78
50	1,50	1,59	1,46	1,63	1,42	1,67	1,38	1,72	1,34	1,77
55	1,53	1,60	1,49	1,64	1,45	1,68	1,41	1,72	1,38	1,77
60	1,55	1,62	1,51	1,65	1,48	1,69	1,44	1,73	1,41	1,77
65	1,57	1,63	1,54	1,66	1,50	1,70	1,47	1,73	1,44	1,77
70	1,58	1,64	1,55	1,67	1,52	1,70	1,49	1,74	1,46	1,77
75	1,60	1,65	1,57	1,68	1,54	1,71	1,51	1,74	1,49	1,77
80	1,61	1,66	1,59	1,69	1,56	1,72	1,53	1,74	1,51	1,77
85	1,62	1,67	1,60	1,70	1,57	1,72	1,55	1,75	1,52	1,77
90	1,63	1,68	1,61	1,70	1,59	1,73	1,57	1,75	1,54	1,78
95	1,64	1,69	1,62	1,71	1,60	1,73	1,58	1,75	1,56	1,78
100	1,65	1,69	1,63	1,72	1,61	1,74	1,59	1,76	1,57	1,78

Источник : Джонстон Дж. Эконометрические методы. М.: Статистика, 1980, стр. 434 / Biometrika, 1951, vol. 41, p. 173.

В случае наличия автокорреляции остатков полученная формула регрессии считается обычно неудовлетворительной. МНК – оценки в случае наличия авторегрессии первого порядка не смещены, состоятельны, но неэффективны. Оценка дисперсии при использовании МНК является заниженной, что может отрицательно сказаться на проверке гипотез о значимости коэффициентов. Образно говоря, МНК рисует более оптимистическую картину регрессии, чем есть на самом деле. В этом случае можно поискать другую (нелинейную) формулу, включить неучтенные до этого факторы, либо применить к данным уменьшающее автокорреляцию остатков преобразование (например, авторегрессионное преобразование или метод скользящих средних).

13.3. Авторегрессионное преобразование

Есть исходная модель

$$Y = BX + \varepsilon,$$

где i-я компонента вектора Y представляет значение зависимой переменной для наблюдения i, $i = 1,..., n$ (или в момент времени t для временных рядов). Запишем уравнение для наблюдения в момент i:

$$Y_i = a_0 + b_1X_{i1} + ... + b_nX_{ik} + \varepsilon_i.$$

Один из способов учета коррелированности ошибок в разные моменты времени состоит в предположении, что случайная последовательность (ε_i, $i = 1,..., n$) образует авторегрессионный процесс первого порядка. Это означает, что ошибки удовлетворяет соотношению

$$\varepsilon_i = \rho\,\varepsilon_{i-1} + v_i,$$

где (v_i, $i = 1,..., n$) – последовательность независимых нормально распределенных случайных величин с нулевым средним и постоянной дисперсией, а ρ –некоторый параметр, называемый коэффициентом авторегрессии ($|\rho| < 1$).

Запишем исходное уравнение для наблюдения i – 1:

$$Y_{i-1} = a_0 + b_1X_{(i-1)1} + ... + bnX_{(i-1)k} + \varepsilon_{(i-1)}.$$

Умножим обе части уравнения на ρ и вычтем почленно из уравнения для момента i. Тогда с учетом $\varepsilon_i = \rho\,\varepsilon_{i-1} + v_i$, получим:

$$Y_i - \rho Y_{i-1} = a_0(1-\rho) + b_1(X_{i1} - \rho X_{(i-1)2}) + ... + b_k(X_{ik} - \rho X_{(i-1)k}) + v_i. \quad (*).$$

Проблему оценки в модели с авторегрессией рассмотрим на основе **процедуры Кохрейна–Оркатта.**

Начальным шагом этой процедуры является применение обычного метода наименьших квадратов к исходной системе и получение соответствующих остатков $e = (e_1,...,e_n)$.

В качестве приближенного значения ρ_1 берется его МНК–оценка в регрессии:

$$\varepsilon_i = \rho_1\,\varepsilon_{i-1} + v_i.$$

Проводится преобразование при $\rho = \rho_1$:

$$Y_i - \rho_1 Y_{i-1} = a_0(1-\rho_1) + b_1(X_{i1} - \rho_1 X_{(i-1)2}) + ... + b_k(X_{ik} - \rho_1 X_{(i-1)k}) + v_i. \quad (*),$$

и находятся МНК – оценки вектора параметров регрессии B_1.

Строится новый вектор остатков:

$$e_2 = Y - B_1X.$$

Процедура повторяется сначала, т. е. носит итерационный характер. Процесс обычно заканчивается, когда очередное приближение ρ мало отличается от предыдущего.

Процедура Дарбина. Перепишем уравнение (*) в виде:

$$Y_i = a_0(1-\rho) + \rho Y_{i-1} + b_1X_{i1} - b_1\rho X_{(i-1)2}) + ... + b_kX_{ik} - b_k\rho X_{(i-1)k}) + v_i.$$

То есть Y_{i-1} включается в число регрессоров, а ρ – в число оцениваемых параметров. Для этой системы определяются обычные МНК–оценки параметров $a_0(1-\rho)$, ρ, $b_k\rho$. Оценки параметров a_0 и b_k устанавливаются делением соответственно на $(1-\rho)$ и ρ. Полученные оценки подставляются в уравнение (*) и находятся новые МНК–оценки параметров.

В случае когда после применения процедуры авторегрессионного преобразования остатки автокоррелированы, применяется авторегрессионное преобразование более высокого порядка. Так, для авторегрессионного процесса второго порядка ошибки удовлетворяют соотношению:

$$\varepsilon_i = \rho_1\varepsilon_{i-1} + \rho_2\varepsilon_{i-2} + u_i.$$

Однако часто истинной причиной первоначальной автокорреляции остатков может быть нелинейность формулы или неучтенный фактор. Данное обстоятельство является содержательным ограничением для применения авторегрессионного преобразования.

Кроме авторегрессионного преобразования, для устранения автокорреляции остатков и уточнения формулы регрессионной зависимости может использоваться метод скользящих средних (Moving Averages, или MA). В этом случае считается, что отклонения от линии регрессии e_i можно выразить как скользящие средние случайных нормально распределенных ошибок e_i:

$$e_i = \varepsilon_i + \upsilon_1\varepsilon_{i-1} + \ldots + \upsilon_k\varepsilon_{i-k}.$$

Это формула для преобразования MA k – го порядка. Для преобразования первого порядка формула имеет вид:

$$e_i = \varepsilon_i + \upsilon_1\varepsilon_{i-1}.$$

Параметры υ_k, как и в случае авторегрессионного преобразования, могут оцениваться итерационными методами.

Сочетание методов авторегрессионного преобразования (AR) и метода скользящих средних (MA) позволяет решить проблему автокорреляции остатков даже при небольших порядках.

Методы AR и MA могут использоваться в сочетании с переходом от объемных величин в модели к приростным, для которых статистическая взаимосвязь может быть более точной и явной. Модель, сочетающая все эти подходы, называется моделью ARIMA (Autoregressive Integrated Moving Averages).

Преобразования AR, MA и модель ARIMA полезно использовать в тех случаях, когда уже ясен круг объясняющих переменных и общий вид оцениваемой формулы, но в то же время остается существенная автокорреляция остатков. Примером таких функций являются производственные функции, где объясняющими переменными служат используемые объемы или темпы прироста труда и капитала, а требуемой формулой является производственная функция Кобба Дугласа, или CES.

13.4. Прогнозирование в регрессионных моделях

Одна из важнейших целей моделирования заключается в прогнозировании поведения исследуемого объекта. Обычно термин "прогнозирование" используется в тех ситуациях, когда требуется предсказать состояние системы в будущем. Для регрессионных моделей он имеет более широкое значение. Данные могут не иметь временной структуры, но и в этих случаях может возникнуть задача оценки значения зависимой переменной для некоторого набора независимых, объясняющих переменных, которых нет в исходных наблюдениях. Будем понимать прогнозирование как построение оценки зависимой переменной.

Проблема прогнозирования имеет много различных аспектов. Можно различать точечное и интервальное прогнозирование. В первом случае – это конкретное число, во втором – интервал, в котором истинное значение переменной находится с заданным уровнем доверия. Выделяют также безусловное и условное прогнозирование в зависимости от того, известны ли объясняющие переменные точно или приближенно. Кроме того, для временных рядов при нахождении прогноза существенно наличие или отсутствие корреляции во времени между ошибками.

Для регрессионной модели

$$Y = XB + \varepsilon$$

задан дополнительный набор x_{n+1} объясняющих переменных. Зависимая переменная удовлетворяет условию

$$Y_{n+1} = x_{n+1}B + \varepsilon_{n+1}.$$

Требуется по (Y, X, x_{n+1}) оценить Y_{n+1}.

Безусловное прогнозирование. Термин «безусловное прогнозирование» означает, что вектор независимых переменных x_{n+1} известен точно.

Среднеквадратическая ошибка прогноза будет равна:

$$M(\hat{Y} - Y_{n+1})^2 = \sigma^2(1 + x_{n+1}^T(X^TX)^{-1}x_{n+1}),$$

где σ^2 – дисперсия остатков ε.

Обозначим $\delta = \sqrt{\sigma^2(1 + x_{n+1}^T(X^TX)^{-1}x_{n+1})}$. Если ошибки $(\varepsilon, \varepsilon_{n+1})$ имеют совместное нормальное распределение, то случайная величина $(\hat{Y} - Y_{n+1})/\delta$ имеет распределение Стьюдента с $n-k$ степенями свободы. Поэтому доверительным интервалом для Y_{n+1} с уровнем значимости α будет интервал $(\hat{Y} - \delta t_\alpha, \hat{Y} + \delta t_\alpha)$, где t_α двухсторонняя α – квантиль распределения Стьюдента с $n-k$ степенями свободы.

В случае двумерной регрессии среднеквадратическая ошибка прогноза равна:

$$M(\hat{Y} - Y_{n+1})^2 = \sigma^2 \frac{n}{n-2}\left(1 + \frac{1}{n} + \frac{(X_{n+1} - \overline{X})^2}{\sum(X_i - \overline{X})^2}\right)$$

Отсюда следует, что среднеквадратическая ошибка прогноза минимальна при $X_{n+1} = \overline{X}$, и чем дальше X_{n+1} от среднего значения, тем шире доверительный интервал.

Условное прогнозирование. Часто приходится прогнозировать и значения независимых переменных, что неизбежно приводит к отклонениям от истинных значений, и вектор x_{n+1} наблюдается с ошибкой u. В этом случае среднеквадратическая ошибка прогноза будет равна:

$$M(\hat{Y} - Y_{n+1})^2 = \sigma^2\left[1 + x_{n+1}^T(X^TX)^{-1}x_{n+1} + \sigma_u^2 tr((X^TX)^{-1})\right] + \sigma_u B^T B.$$

При наличии ошибок в независимой переменной к ошибке прогнозирования добавляются два новых положительных слагаемых, пропорциональных дисперсии σ_u^2.

Точность прогнозирования. Проблема проверки степени совершенства прогнозов является одной из важнейших в прогнозировании. Степень совершенства прогнозов выражают через различные измерители точности прогнозирования. Точность точечного прогноза в момент t_i определяется разностью между прогнозом P_i и фактическим значением F_i прогнозируемого показателя в этот момент времени. Отдельный точечный прогноз не определяет точность конкретной процедуры прогнозирования в целом, т. е. потребуется некоторая выборка (P_i, F_i), на основе которой рассчитывается значение некоторого измерителя точности прогнозирования.

Разберем несколько важнейших измерителей точности прогнозирования. В первую очередь рассмотрим классический критерий точности прогнозирования – коэффициент корреляции Пирсона:

$$r = \frac{\sum_{i=1}^{n}(P_i - \overline{P})(F_i - \overline{F})}{\sqrt{\sum_{i=1}^{n}(P_i - \overline{P})^2 \sum_{i=1}^{n}(F_i - \overline{F})^2}},$$

где $\overline{P} = \frac{1}{n}\sum_{i=1}^{n} P_i$, $\overline{F} = \frac{1}{n}\sum_{i=1}^{n} F_i$.

Максимальное значение $r = 1$ достигается при наличии линейной связи между P и F, т. е. когда существуют такие a_0 и $a_1 > 0$, что $P = a_0 + a_1 F$.

Проверить значимость коэффициента корреляции Пирсона можно по *t-статистике* с $(n - 2)$-степенями свободы:

$$t = \frac{r\sqrt{n-2}}{1-r^2}.$$

Наиболее распространенными оценками точности прогнозирования являются средняя ошибка аппроксимации:

$$E = \frac{1}{n}\sum_{i=1}^{n}\frac{|P_i - F_i|}{F_i}100(\%),$$

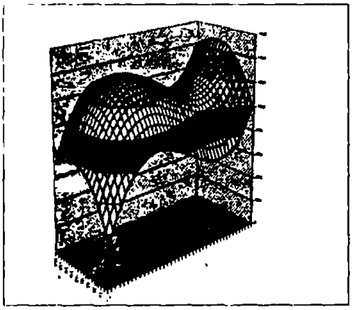
а также средняя квадратическая ошибка прогнозов:

$$S = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_i - F_i)^2}.$$

Точность прогнозирования тем выше, чем меньше значения E или S соответственно. Совершенный прогноз достигается при $E = S = 0$. Средняя ошибка аппроксимации E применяется при постулировании требований к точности прогнозирования показателей. Коэффициент E позволяет сравнивать точность прогнозирования различных процессов. Для обоснованности применения критерия E следует проверить гипотезу о существенности разности $\overline{P} - \overline{F}$ с помощью F-критерия Фишера.

Глава 14. Классические методы оптимизации

14.1. Классический метод безусловной оптимизации



В классических задачах предполагается, что функция $f(x)$ является достаточно гладкой, т. е. она имеет нужное число производных. Необходимым условием оптимума является условие равенства нулю первой производной:

$$f'(x) = 0.$$

Данное условие не является достаточным условием оптимума. Такая точка может быть как точкой локального оптимума функции $f(x)$, так и точкой глобального оптимума. Точки локального оптимума удается исключить из рассмотрения, если проанализировать знак второй производной функции $f''(x)$. В точке, являющейся точкой строгого максимума (локального и глобального), соблюдается условие

$$f''(x) \leq 0.$$

В точке минимума:

$$f''(x) > 0.$$

Это условие второго порядка позволяет не рассматривать часть точек, удовлетворяющих необходимому условию, но не являющихся точками максимума (или минимума). Определив все точки, удовлетворяющие условиям первого и второго порядка, установив значение функции $f(x)$ в этих точках и сравнив их между собой, можно найти ту точку, в которой достигается глобальный максимум (минимум).

Для того чтобы не рассматривать все локальные оптимумы, необходимо проанализировать свойства функции $f(x)$. В экономико-математических методах важную роль играют свойства выпуклости и вогнутости функций. Для выпуклой вверх функции $f(x)$ строгий локальный максимум (если он существует) является единственным и совпадает с глобальным. Аналогично для выпуклой вниз функции локальный минимум является единственным и совпадает с глобальным.

В случае функции многих переменных необходимое условие заключается в равенстве нулю всех частных производных:

$$\partial f / \partial x_i = 0, \ i = 1,\ldots, n.$$

Необходимое условие экстремума можно переформулировать также следующим образом: в точке экстремума градиент равен нулю. Точки, в которых выполнены необходимые условия, называются стационарными.

Для получения достаточных условий следует определить в стационарной точке знак дифференциала второго порядка. Он обозначается как $d^2f(x_i)$ и равен сумме произведений частных производных второго порядка на соответствующие приращения аргументов:

$$d^2 f(X) = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j} \Delta x_i \Delta x_j.$$

В случае, когда в стационарной точке второй дифференциал есть отрицательно определенная квадратичная форма, функция $f(x_i)$ имеет максимум, а в случае, когда этот дифференциал есть положительно определенная квадратичная форма, функция $f(x_i)$ имеет в этой точке минимум.

Квадратичная форма (а также соответствующая матрица A) называется определенной, отрицательно определенной, неотрицательной или неположительной, если соответственно $X^TAX > 0$, $X^TAX < 0$, $X^TAX \geq 0$ или $X^TAX \leq 0$ для каждого набора действительных чисел $\xi_1, \xi_2, \ldots, \xi_n$, не все из которых равны нулю. Все остальные квадратичные формы являются неопределенными (т. е. знак X^TAX зависит от выбора чисел $\xi_1, \xi_2, \ldots, \xi_n$) или тождественно равными нулю.

Квадратичная форма является положительно определенной, отрицательно определенной, неотрицательной, неположительной, неопределенной или тождественно равной нулю в том и только в том случае, если собственные значения λ_i матрицы $A \equiv [a_{ik}]$ соответственно все положительны, все отрицательны, все неотрицательны, все неположительны, имеют различные знаки или все равны нулю.

Для того чтобы квадратичная форма была положительно определена, необходимо и достаточно, чтобы каждый из определителей

$$a_{11}, \quad \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{23} \end{bmatrix}, \quad \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \ldots, \det\left[a_{ik}\right]$$

был положителен.

Для того чтобы квадратичная форма была неотрицательна, необходимо и достаточно, чтобы все главные миноры определителя $\det[a_{ik}]$ (т. е. все миноры, получающиеся из этого определителя вычеркиванием строк и столбцов с одними и теми же номерами или, иначе говоря, симметричные относительно главной диагонали этого определителя) были неотрицательны.

Квадратичная форма отрицательно определена в том случае, если угловые миноры $\Delta_1, \ldots, \Delta_n$ матрицы A знакопеременны:

$$(-1)^i \Delta_i > 0 \quad (i = 1, \ldots, n),$$

значение первого $\Delta_1 < 0$.

Таким образом, условие второго порядка для наличия экстремума заключается в следующем.

Минимум – положительно определенная квадратичная форма, отсюда: собственные значения λ_i матрицы вторых производных все положительны, или определители матрицы положительны.

Максимум – отрицательно определенная квадратичная форма, отсюда: собственные значения λ_i матрицы вторых производных все отрицательны, угловые миноры знакопеременны, начиная с первого отрицательного.

Матрица вторых частных производных (матрица Гессе или гессиан) для функции двух переменных будет:

$$H = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x_1 \partial x_1} & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} \\ \dfrac{\partial^2 f}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f}{\partial x_2 \partial x_2} \end{bmatrix}.$$

Ее определитель, с учетом равенства смешанных производных, выражается как:

$$\det[H] = \frac{\partial^2 f}{\partial x_1^2} \cdot \frac{\partial^2 f}{\partial x_2^2} - \left(\frac{\partial^2 f}{\partial x_1 \partial x_2}\right)^2 > 0.$$

Собственные значения матрицы вторых производных устанавливаются из характеристического уравнения матрицы H:

$$\det(H - \lambda I) = \begin{vmatrix} f''_{xx} - \lambda & f''_{xy} \\ f''_{yx} & f''_{yy} - \lambda \end{vmatrix} = 0,$$

т. е. устанавливаются из уравнения:

$$\lambda^2 - \lambda f''_{xx} - \lambda f''_{yy} + f''_{xx} f''_{yy} - (f''_{xy})^2 = 0.$$

Для функции двух переменных корней собственных значений будет два.

Следовательно, условия минимума формулируются так:

$$\lambda_1 > 0;\ \lambda_2 > 0$$

$$\text{или } f''_{xx} > 0,\quad \det[H] > 0.$$

Условия максимума:

$$\lambda_1 < 0;\ \lambda_2 < 0$$

$$\text{или } f''_{xx} < 0,\quad \det[H] > 0.$$

Для функции трех переменных матрица вторых производных будет равна:

$$H = \begin{bmatrix} f''_{xx} & f''_{xy} & f''_{xz} \\ f''_{yx} & f''_{yy} & f''_{yz} \\ f''_{zx} & f''_{zy} & f''_{zz} \end{bmatrix}.$$

Ее определитель равен:

$$\det(H) = f''_{xx} f''_{yy} f''_{zz} - f''_{xx}(f''_{yz})^2 - f''_{zz}(f''_{xy})^2 + 2f''_{yx} f''_{xz} f''_{zy} - (f''_{xz})^2 f''_{yy}$$

Собственные значения матрицы G определяются из уравнения:

$$-\lambda^3 + b\lambda^2 + c\lambda + d = 0,$$

где $b = f''_{xx} + f''_{yy} + f''_{zz}$

$c = (f''_{xy} + f''_{xz} + (-f''_{zz} - f''_{yy}) f''_{xx} - f''_{zz} f''_{yy} + f''_{yz})$

$d = (-f''_{yz} + f''_{zz} f''_{yy}) f''_{xx} - f''_{xy} f''_{zz} - f''_{xz} f''_{yy} + 2f''_{xy} f''_{xz} f''_{yz}.$

14.2. Классическая оптимизация в Mathcad

Рассмотрим несколько примеров того, как задача классической оптимизации решается в Mathcad. Необходимо отметить, что большее распространение по установлению условий второго порядка получил метод по определителю матрицы и знакам миноров. Однако при наличии Mathcad метод по знакам собственных значений становится удобней в смысле однозначности выводов. Для расчета собственных значений матрицы в Mathcad существует функция **eigenvals[M]**. В первом примере система уравнений первых производных имеет четыре корня, найденных с помощью функции **polyroots**, которые и исследуются на экстремум обоими методами.

Пример 1.

$$f(x,y) := 3\cdot x^3 - x + y^3 - 3\cdot y^2 - 1$$

Находим первые и вторые производные

$$f1(x) := \frac{d}{dx}f(x,y) \qquad f2(x) := \frac{d^2}{dx^2}f(x,y)$$

$$f1(x) \rightarrow 9\cdot x^2 - 1 \qquad f2(x) \rightarrow 18\cdot x$$

$$f1(y) := \frac{d}{dy}f(x,y) \qquad f2(y) := \frac{d^2}{dy^2}f(x,y)$$

$$f1(y) \rightarrow 3\cdot y^2 - 6\cdot y \qquad f2(y) \rightarrow 6\cdot y - 6$$

$$f2(x,y) := \frac{d}{dy}f1(x) \qquad f2(x,y) := \frac{d}{dx}f1(y)$$

$$f2(x,y) \rightarrow 0 \qquad f2(x,y) \rightarrow 0$$

Формируем матрицу вторых производных

$$Hessian(x,y) := \begin{pmatrix} 18\cdot x & 0 \\ 0 & 6\cdot y - 6 \end{pmatrix}$$

Приравниваем нулю первые производные и находим корни

$$9\cdot x^2 - 1 = 0$$

$$3\cdot y^2 - 6\cdot y = 0$$

$$polyroots\left(\begin{pmatrix} -1 \\ 0 \\ 9 \end{pmatrix}\right) = \begin{pmatrix} -0.333 \\ 0.333 \end{pmatrix} \qquad x1 := -0.333 \qquad y1 := 0$$

$$x2 := 0.333 \qquad y2 := 0$$

$$polyroots\left(\begin{pmatrix} 0 \\ -6 \\ 3 \end{pmatrix}\right) = \begin{pmatrix} 0 \\ 2 \end{pmatrix} \qquad x3 := 0.333 \qquad y3 := 2$$

$$x4 := -0.333 \qquad y4 := 2$$

Находим собственные значения матрицы вторых производных

1 $\quad x := \frac{-1}{3} \qquad y := 0 \qquad Hessian(x,y) = \begin{pmatrix} -6 & 0 \\ 0 & -6 \end{pmatrix}$

Находим определитель матрицы вторых производных

$$\det(x,y) := |Hessian(x,y)|$$

$$\det(x,y) = 36$$

$\frac{d^2}{dx^2}(f(x,y) < 0)$ и $\det(x,y) > 0$ максимум

или $CZ := eigenvals(Hessian(x,y))$

$CZ = \begin{pmatrix} -6 \\ -6 \end{pmatrix}$ максимум

2 $x := \frac{1}{3}$ $y := 0$

$$Hessian(x,y) = \begin{pmatrix} 6 & 0 \\ 0 & -6 \end{pmatrix}$$

$$\det(x,y) = -36$$

Определитель матрицы вторых производных отрицательный — нет экстремума

или

$CZ := eigenvals(Hessian(x,y))$

$CZ = \begin{pmatrix} 6 \\ -6 \end{pmatrix}$ нет экстремума

3 $x := \frac{1}{3}$ $y := 2$

$$Hessian(x,y) = \begin{pmatrix} 6 & 0 \\ 0 & 6 \end{pmatrix}$$

$$\det(x,y) = 36$$

$\frac{d^2}{dx^2}(f(x,y) > 0)$ и $\det(x,y) > 0$ минимум

или $CZ := eigenvals(Hessian(x,y))$

$CZ = \begin{pmatrix} 6 \\ 6 \end{pmatrix}$ минимум

4 $x := \frac{-1}{3}$ $y := 2$

$$Hessian(x,y) = \begin{pmatrix} -6 & 0 \\ 0 & 6 \end{pmatrix}$$

$$det(x,y) = -36$$

Определитель матрицы вторых производных отрицательный – нет экстремума

или $CZ := eigenvals(Hessian(x,y))$

$$CZ = \begin{pmatrix} -6 \\ 6 \end{pmatrix}$$ нет экстремума

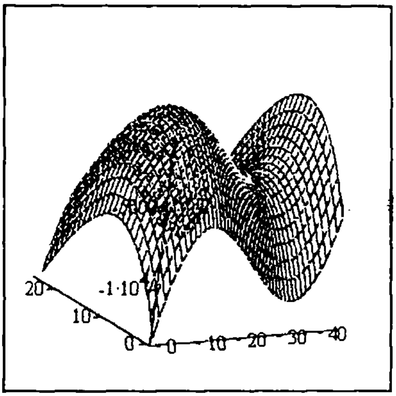
Во втором примере функция двух переменных имеет две стационарные точки. Их поиск, как решение системы уравнений первых производных, осуществлен с помощью функции find(). В этом примере показано, как графическое представление функции помогает анализировать ее на наличие экстремумов. Графики первых производных помогают установить первоначальное приближение для функции find(). Приведен пример построения графика градиентного поля.

Пример 2.

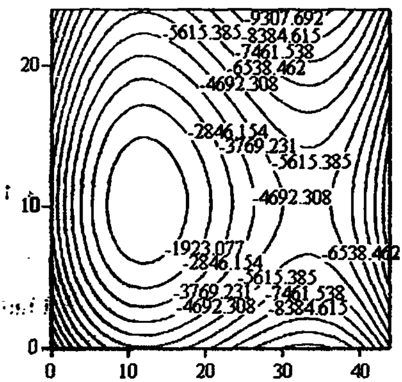
$$f(x,y) := x^3 - 68\cdot x^2 + 1200\cdot x + y^3 - 74\cdot y^2 + 1200\cdot y - 13\,000$$

Построение графиков функции

$m := 45$ $n := 25$
$M := matrix(m,n,f)$



M



M

$$f(x,y) := x^3 - 68\cdot x^2 + 1200\cdot x + y^3 - 74\cdot y^2 + 1200\cdot y - 13\,000$$

Находим первые и вторые производные

$$f1y(x,y) := 3\cdot y^2 - 148\cdot y + 1200 \qquad f1x(x,y) := 3\cdot x^2 - 136\cdot x + 1200$$

$$f2y(x,y) := 6\cdot y - 148 \qquad f2x(x,y) := 6\cdot x - 136$$

Формируем матрицу вторых производных

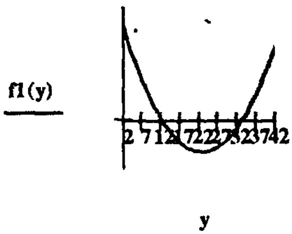
$$G(x,y) := \begin{pmatrix} 6\cdot x - 136 & 0 \\ 0 & 6\cdot y - 148 \end{pmatrix}$$

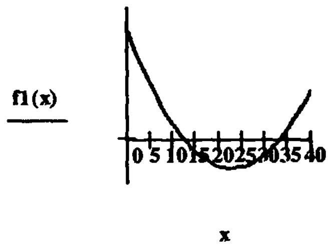
Находим первую стационарную точку

Строим графики первых производных

$$f1(y) := 3\cdot y^2 - 148\cdot y + 1200 \qquad f1(x) := 3\cdot x^2 - 136\cdot x + 1200$$

$$y := 0..42 \qquad x := 0..40$$





$$x := 10 \qquad y := 10$$

Given

$$3\cdot x^2 - 136\cdot x + 1200 = 0$$

$$3\cdot y^2 - 148\cdot y + 1200 = 0$$

$$r1 := \mathrm{Find}(x,y) \qquad r1 = \begin{pmatrix} 12 \\ 10.229 \end{pmatrix}$$

Нахождение определителя

$$G(x,y) = \begin{pmatrix} -76 & 0 \\ 0 & -88 \end{pmatrix}$$

$$\det(G) := |G(x,y)|$$

$$\det(G) = 6.688 \times 10^{3} \quad > 0$$

максимум

$$\frac{d^2}{dx^2} f(x,y) = -76 \qquad < 0$$

Нахождение собственных значений

$$x := r1_0 \qquad y := r1_1$$

$$CZ := \text{eigenvals}(G(x,y))$$

$$CZ = \begin{pmatrix} -64 \\ -86.626 \end{pmatrix}$$

$\lambda_1 < 0 \quad \lambda_2 < 0$ максимум

Находим вторую стационарную точку

$$x := 35 \qquad y := 40$$

Given

$$3 \cdot x^2 - 136 \cdot x + 1200 = 0$$

$$3 \cdot y^2 - 148 \cdot y + 1200 = 0$$

$$r2 := \text{Find}(x,y)$$

$$r2 = \begin{pmatrix} 33.333 \\ 39.104 \end{pmatrix}$$

Нахождение определителя

$$G(x,y) = \begin{pmatrix} 74 & 0 \\ 0 & 92 \end{pmatrix}$$

$$\det(G) := |G(x,y)|$$

$\det(G) = 6.808 \times 10^3$

> 0

минимум

$\frac{d^2}{dx^2} f(x, y) = 74 \qquad > 0$

Нахождение собственных значений

$x := r2_0 \qquad y := r2_1$

$CZ := \mathrm{eigenvals}(G(x, y))$

$$CZ = \begin{pmatrix} 64 \\ 86.626 \end{pmatrix}$$

$\lambda_1 > 0 \quad \lambda_2 > 0$ минимум

Построение графика градиентного поля

$$f(x, y) := x^3 - 68 \cdot x^2 + 1200 \cdot x + y^3 - 74 \cdot y^2 + 1200 \cdot y - 13\,000$$

Пределы изменений $xlow := 0 \quad xhigh := 45 \quad ylow := 0 \quad yhigh := 25$

Число точек $\quad xn := 20 \qquad yn := 20$

Вычисление градиента функции

$$grad(x, y) := \begin{pmatrix} \frac{d}{dx} f(x, y) \\ \frac{d}{dy} f(x, y) \end{pmatrix}$$

Подготовка к построению графика поля

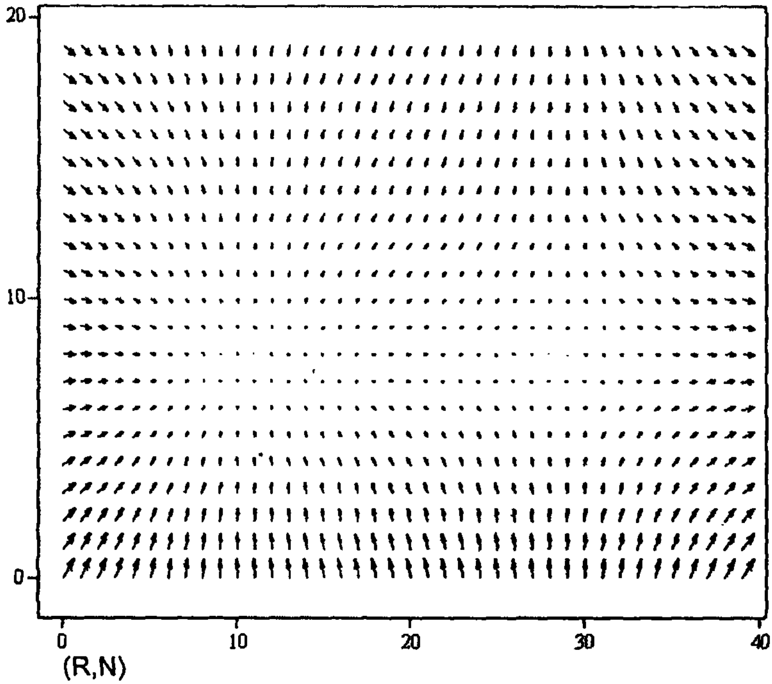
$i := 0 .. xn - 1 \qquad j := 0 .. yn - 1$

$$xind_i := xlow + i \cdot \frac{xhigh - xlow}{xn - 1} \qquad yind_j := ylow + j \cdot \frac{yhigh - ylow}{yn - 1}$$

$$V_{i,j} := grad(xind_i, yind_j)$$

$$R_{i,j} := (V_{i,j})_0 \qquad N_{i,j} := (V_{i,j})_1$$

График градиентного поля функции $f(x,y)$



Стрелки показывают направление роста градиента

В третьем случае приведен пример с функцией трех переменных. Во-первых, рассмотрим разную технику представления матрицы вторых производных. В первом случае для формирования матрицы вторых производных использован якобиан для градиента, формула для расчета которого просто скопирована из справки (Quicksheets, Calculus and Differential Equations, Jacobian Matrix and Determinant).

$$f(x,y,z) := x \cdot y \cdot z - (x-4)^3 - (y-4)^3 - (z-4)^3$$

$$A(x,y,z) := \begin{pmatrix} \frac{d}{dx} f(x,y,z) \\ \frac{d}{dy} f(x,y,z) \\ \frac{d}{dz} f(x,y,z) \end{pmatrix}$$ градиент

$$A(x,y,z) \rightarrow \begin{bmatrix} y \cdot z - 3 \cdot (x-4)^2 \\ x \cdot z - 3 \cdot (y-4)^2 \\ x \cdot y - 3 \cdot (z-4)^2 \end{bmatrix}.$$

$$\mathrm{Jacob}(A,x,y,z) := \begin{bmatrix} \frac{d}{dx}\left(A(x,y,z)_0\right) & \frac{d}{dy}\left(A(x,y,z)_0\right) & \frac{d}{dz}\left(A(x,y,z)_0\right) \\ \frac{d}{dx}\left(A(x,y,z)_1\right) & \frac{d}{dy}\left(A(x,y,z)_1\right) & \frac{d}{dz}\left(A(x,y,z)_1\right) \\ \frac{d}{dx}\left(A(x,y,z)_2\right) & \frac{d}{dy}\left(A(x,y,z)_2\right) & \frac{d}{dz}\left(A(x,y,z)_2\right) \end{bmatrix}$$

$$\mathrm{Jacob}(A,x,y,z) \rightarrow \begin{pmatrix} -6\cdot x+24 & z & y \\ z & -6\cdot y+24 & x \\ y & x & -6\cdot z+24 \end{pmatrix}$$ матрица вторых производных

В дальнейшем матрица вторых производных сформирована и обозначена Hess(x,y,z). При использовании функции **find** необходимо помнить, что она отыскивает правильное решение только вблизи истинного значения корней. Поэтому логичней сформулировать всю задачу целиком, чтобы сразу, после установления первоначальных значений для функции find, было видно значения градиента, и оценить собственные значения матриц (или определителей) на экстремум.

Пример 3.

$$f(x,y,z) := x\cdot y\cdot z-(x-4)^3-(y-4)^3-(z-4)^3$$

$$\mathrm{Grad}(x,y,z) := \begin{pmatrix} \frac{d}{dx}f(x,y,z) \\ \frac{d}{dy}f(x,y,z) \\ \frac{d}{dz}f(x,y,z) \end{pmatrix}$$ градиент

$$\mathrm{Grad}(x,y,z) \rightarrow \begin{bmatrix} y\cdot z-3\cdot(x-4)^2 \\ x\cdot z-3\cdot(y-4)^2 \\ x\cdot y-3\cdot(z-4)^2 \end{bmatrix}$$

$$\mathrm{Hess}(x,y,z) := \begin{pmatrix} \frac{d^2}{dx^2}f(x,y,z) & \frac{d}{dx}\frac{d}{dy}f(x,y,z) & \frac{d}{dx}\frac{d}{dz}f(x,y,z) \\ \frac{d}{dy}\frac{d}{dx}f(x,y,z) & \frac{d^2}{dy^2}f(x,y,z) & \frac{d}{dy}\frac{d}{dz}f(x,y,z) \\ \frac{d}{dz}\frac{d}{dx}f(x,y,z) & \frac{d}{dz}\frac{d}{dy}f(x,y,z) & \frac{d^2}{dz^2}f(x,y,z) \end{pmatrix}$$

$$\text{Hess}(x,y,z) \to \begin{pmatrix} -6\cdot x+24 & z & y \\ z & -6\cdot y+24 & x \\ y & x & -6\cdot z+24 \end{pmatrix}$$ матрица вторых производных

$$\det(\text{Hess},x,y,z) := |(\text{Hess}(x,y,z))|$$

Нахождение первой стационарной точки

$x := 1$ $y := 1$ $z := 1$

Given

$\frac{d}{dx}f(x,y,z) = 0$ $\frac{d}{dy}f(x,y,z) = 0$ $\frac{d}{dz}f(x,y,z) = 0$

$C := \text{Find}(x,y,z)$

$$C = \begin{pmatrix} 2.536 \\ 2.536 \\ 2.536 \end{pmatrix}$$ $x := C_0$ $y := C_1$ $z := C_2$

Проверка

$$\text{Grad}(x,y,z) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \text{Hess}(x,y,z) = \begin{pmatrix} 8.785 & 2.536 & 2.536 \\ 2.536 & 8.785 & 2.536 \\ 2.536 & 2.536 & 8.785 \end{pmatrix}$$

$\det(\text{Hess},x,y,z) = 541.043$

$\lambda := \text{eigenvals}(\text{Hess}(x,y,z))$ $\Delta 2 := \text{submatrix}(\text{Hess}(x,y,z),0,1,0,1)$

$$\lambda = \begin{pmatrix} 13.856 \\ 6.249 \\ 6.249 \end{pmatrix} \qquad \Delta 2 = \begin{pmatrix} 8.785 & 2.536 \\ 2.536 & 8.785 \end{pmatrix}$$

$|\Delta 2| = 70.739$

Собственные значения положительны. Минимум.

Определители положительны. Минимум.

Нахождение второй стационарной точки

$x := 4 \qquad y := 6 \qquad z := 4$

Given

$\frac{d}{dx}f(x,y,z) = 0 \qquad \frac{d}{dy}f(x,y,z) = 0 \qquad \frac{d}{dz}f(x,y,z) = 0$

$C := \mathrm{Find}(x,y,z)$

$$C = \begin{pmatrix} 2.093 \\ 5.209 \\ 2.093 \end{pmatrix} \quad x := C_0 \quad y := C_1 \quad z := C_2$$

Проверка

$$\mathrm{Grad}(x,y,z) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \mathrm{Hess}(x,y,z) = \begin{pmatrix} 11.439 & 2.093 & 5.209 \\ 2.093 & -7.252 & 2.093 \\ 5.209 & 2.093 & 11.439 \end{pmatrix}$$

$\det(\mathrm{Hess}, x, y, z) = -806.808$

$\lambda := \mathrm{eigenvals}(\mathrm{Hess}(x,y,z)) \qquad \Delta2 := \mathrm{submatrix}(\mathrm{Hess}(x,y,z),0,1,0,1)$

$$\lambda = \begin{pmatrix} 17.009 \\ 6.23 \\ -7.613 \end{pmatrix} \qquad \Delta2 = \begin{pmatrix} 11.439 & 2.093 \\ 2.093 & -7.252 \end{pmatrix}$$

$|\Delta2| = -87.339$

нет экстремума нет экстремума

Нахождение третьей стационарной точки

$x := 9 \qquad y := 10 \qquad z := 10$

Given

$\frac{d}{dx}f(x,y,z) = 0 \qquad \frac{d}{dy}f(x,y,z) = 0 \qquad \frac{d}{dz}f(x,y,z) = 0$

$C := \mathrm{Find}(x,y,z)$

$$C = \begin{pmatrix} 9.464 \\ 9.464 \\ 9.464 \end{pmatrix} \quad x := C_0 \quad y := C_1 \quad z := C_2$$

Проверка

$$\mathbf{Grad}(x,y,z) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \mathbf{Hess}(x,y,z) = \begin{pmatrix} -32.785 & 9.464 & 9.464 \\ 9.464 & -32.785 & 9.464 \\ 9.464 & 9.464 & -32.785 \end{pmatrix}$$

$$\det(\mathbf{Hess}, x, y, z) = -2.473 \times 10^4$$

$$\lambda := \mathrm{eigenvals}(\mathrm{Hess}(x,y,z))$$

$$\lambda = \begin{pmatrix} -13.856 \\ -42.249 \\ -42.249 \end{pmatrix}$$

Собственные значения отрицательны. Максимум.

$$\Delta 2 := \mathrm{submatrix}(\mathrm{Hess}(x,y,z), 0, 1, 0, 1)$$

$$\Delta 2 = \begin{pmatrix} -32.785 & 9.464 \\ 9.464 & -32.785 \end{pmatrix}$$

$$|\Delta 2| = 985.261$$

Миноры знакопеременны. Максимум.

14.3. Условная оптимизация – метод множителей Лагранжа

К задаче безусловной оптимизации можно с помощью искусственных приемов свести многие задачи оптимизации в условиях наличия ограничений на переменные, т. е. задачи условной оптимизации. Классическим примером является метод множителей Лагранжа.

Сущность метода состоит в следующем. Если существует целевая функция и ограничение, то можно сформулировать вспомогательную функцию Лагранжа с помощью ограничения и множителя Лагранжа. Поиск максимума целевой функции будет заключаться в нахождении стационарных точек функции Лагранжа.

Рассмотрим следующую задачу. Необходимо найти максимум $y = f(x_1,x_2)$ при ограничении $z(x_1,x_2) = a$.

Введем множитель Лагранжа λ и выразим функцию Лагранжа $L(x_1,x_2,\lambda)$:

$$L(x_1,x_2,\lambda) = f(x_1,x_2) + \lambda(a - z(x_1,x_2)).$$

Сформулируем условия стационарности функции $L(x_1,x_2,\lambda)$. Они имеют вид:

$$\frac{\partial L}{\partial x_1} = \frac{\partial y}{\partial x_1} - \lambda \frac{\partial z}{\partial x_1} = 0;$$

$$\frac{\partial L}{\partial x_2} = \frac{\partial y}{\partial x_2} - \lambda \frac{\partial z}{\partial x_2} = 0;$$

$$\frac{\partial L}{\partial \lambda} = a - z(x_1,x_2) = 0.$$

Выполнение условий стационарности в точке (x_1^*, x_2^*, λ^*) является необходимым условием того, что в точке (x_1^*, x_2^*) достигается решение исходной задачи.

Пусть величина a в ограничении $z(x_1,x_2) = a$ является переменной. Тогда можно выразить множитель Лагранжа λ^* в точке оптимума:

$$\lambda(a)=\frac{dy(x_1^*(a),x_2^*(a))}{da}.$$

Множитель λ^* показывает предельную оценку функции по отношению к правой части ограничения $z(x_1,x_2) = a$.

Сформулированные условия могут быть перенесены на любое число переменных и ограничений. При этом число множителей Лагранжа равно числу ограничений модели.

Пусть решается задача определения условного экстремума функции $y = f(x_1,x_2,\ldots,x_n)$ при ограничениях

$$z(x_1,x_2,\ldots,x_n) = a_j, \quad j = 1,\ldots,m.$$

Необходимым условием решения поставленной задачи является существование вектора $\lambda^* = (\lambda^*_1,\ldots, \lambda^*_m)$, такого, что функция Лагранжа

$$L(X,\lambda)=f(X)+\sum_{i=1}^{m}\lambda_i(a_i-z(X))$$

удовлетворяет в точке оптимума условиям:

$$\frac{\partial L}{\partial x_i}=0, \quad i=1,2,\ldots,n;$$

$$\frac{\partial L}{\partial \lambda_j}=0, \quad j=1,2,\ldots,m.$$

При наличии ограничений типа равенств задача оптимизации оказалась сведенной к задаче поиска стационарной точки без учета ограничений.

Множители Лагранжа приобретают экономический смысл, как только мы составим экономическую модель. Так, если $y = f(x_1,x_2,\ldots, x_n)$ – доход, а функция $z(x_1,x_2,\ldots, x_n)$ – затраты, то λ_j – предельная оценка j-го ресурса, характеризующая изменение экстремального значения целевой функции в зависимости от изменения размера j-го ресурса.

Основные идеи метода множителей Лагранжа перенесены и на задачи с ограничениями типа неравенств, которые являются более подходящими для экономико-математических моделей. Сформулируем необходимое условие максимума функции $y = f(X)$ при наличии ограничений

$$z(x_1,x_2,\ldots,x_n) \le a_j, \quad j = 1,\ldots, m, \quad x \ge 0.$$

Точка $X^*= (x_1^*;x_2^*,\ldots, x_n^*)$ будет решением поставленной задачи только в том случае, если найдется вектор $\lambda^* = (\lambda_1,\ldots, \lambda_m^*)$, такой, что функция Лагранжа

$$L(X,\lambda) = f(X) + \Sigma\lambda_j(a_j - z(X))$$

в точке (X^*, λ^*) удовлетворяет условиям:

$$\frac{\partial L}{\partial x_i} \le 0, \quad i = 1, 2, \ldots, n;$$

$$\frac{\partial L}{\partial \lambda j} \ge 0, \quad j = 1, 2, \ldots, m;$$

$$\sum_{i=1}^{n} \frac{\partial L}{\partial x_i} x_i^* = 0; \qquad \sum_{j=1}^{m} \frac{\partial L}{\partial \lambda_j} \lambda_j^* = 0;$$

$$x^* \ge 0; \qquad \lambda^* \ge 0.$$

Сформулированное утверждение называют теоремой Куна–Таккера. Данная система соотношений отличается тем, что в ней вместо равенств по первым частным производным присутствуют неравенства, дополненные равенствами, которые принято называть условиями дополняющей нежесткости. Смысл первого из равенств состоит в том, что в случае $\partial L/\partial x_i < 0$ выполняется соотношение $x_i^* = 0$, а в случае $x_i^* > 0$ – соотношение $\partial L/\partial x_i = 0$. Поэтому в том случае, когда в точке x^* все неравенства $x \ge 0$ выполняются строго, первое неравенство можно заменить на равенство $\partial L/\partial x_i = 0$ $(i = 1, \ldots, n)$, т. е. необходимое условие оптимальности формулируется точно таким же образом, как и в случае безусловной оптимизации.

Аналогичным образом получаем, что при $\lambda_j^* > 0$ имеем $\partial L/\partial \lambda_j = 0$, т. е. $z(X^*) = a_j$, а при $z(X^*) < 0$ имеем $\lambda_j^* = 0$. При этом в теореме Куна–Таккера множители Лагранжа сохраняют то же свойство, что и в случае ограничений типа равенств: они равны производной критерия по правой части соответствующего ограничения.

Методы, основанные на использовании множителей Лагранжа, позволяют решать задачи оптимизации, если число ограничений не слишком велико, а так же анализировать общие свойства моделей, как это сделано при анализе функций затрат и функций спроса.

Подробнее с вопросами классической оптимизации можно ознакомиться в работах [3, 6, 18, 33].

Глава 15. Оптимизация методом линейного программирования

15.1. Задача линейного программирования

Задача линейного программирования [3, 4, 15, 18, 19, 31] заключается в нахождении r переменных $X_1, X_2,\dots,X_r$, минимизирующих (или максимизирующих) линую линейную функцию (целевую функцию):

$$z = F(X_1,X_2,\dots,X_r) = C_1X_1 + C_2X_2 + \dots + C_rX_r$$

при линейных ограничениях-равенствах

$$a_{i1}X_1 + a_{i2}X_2 + \dots + a_{ir}X_r = A_i, \qquad i = 1, 2,\dots,n$$

и линейных ограничениях-неравенствах

$$A_{j1}X_1 + A_{j2}X_2 + \dots + A_{jr}X_r \geq B_j, \qquad (j = 1, 2,\dots,m).$$

Отметим, что не каждая задача указанного типа имеет решение.

В типичных приложениях это задача о нахождении необходимо положительных величин $X_1, X_2,\dots, X_r$, т. е. r видов сырьевых ресурсов («входные» данные), минимизирующих общую стоимость $F(X_1, X_2, \dots,X_r)$ соответствующих величин выходных продуктов:

$$q_i = A_{j1}X_1 + A_{j2}X_2 + \dots + A_{jr}X_r$$

с нижними уровнями $B_1, B_2, \dots, B_m$. Имея в виду r условий $X_k \geq 0$ (k=1, 2,...,r) и m ограничений $q_i \geq B_i$ (i=1,2,...,m), получаем $n = r + m$ ограничений-неравенств.

Задача линейного программирования в стандартной форме. Допустимые решения. Задача линейного программирования может быть сведена к стандартной форме путем введения в случае необходимости вспомогательных переменных: требуется минимизировать целевую функцию

$$z = f(x_1,x_2, \dots, x_n) = c_1x_1 + c_2x_2 + \dots + c_nx_n$$

при $m < n$ линейных ограничениях – равенствах

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = b_i, \qquad (i = 1, 2, \dots m)$$

и n линейных ограничениях-неравенствах

$$x_k \geq 0, \qquad (k = 1, 2, \dots, n).$$

Допустимое решение задачи линейного программирования, данной в стандартной форме, есть упорядоченное множество чисел ($x_1, x_2, \dots, x_n$), удовлетворяющих ограничениям, которое является точкой в n-мерном пространстве. Допустимое решение, минимизирующее целевую функцию, называется оптимальным решением (оптимальным планом).

Чаще всего оптимальное решение, если оно существует, единственно, однако возможны случаи, когда оптимальных решений бесчисленное множество.

Для существования допустимого решения необходимо (но не достаточно), чтобы система уравнений была совместна. В дальнейшем считается, что это выполняется; ранг матрицы системы $r \leq m < n$. Тогда r линейно независимых уравнений системы определяют некоторые r неизвестных, как линейные функции остальных $n - r$ неизвестных – свободных неизвестных.

Фактически, выражая все переменные через свободные неизвестные, можно прийти к задаче линейного программирования, содержащей $n - r$ переменных и n

линейных ограничений-неравенств, выражающих неотрицательность исходных переменных.

Допустимое решение, соответствующее нулевым значениям свободных неизвестных, называется базисным допустимым решением (опорным планом), невырожденным, если все остальные переменные положительны, и вырожденным, если среди них есть хоть одно нулевое.

Множество всех допустимых решений данной задачи линейного программирования есть выпуклое множество в n-мерном пространстве решений. Более точно, множество допустимых решений есть выпуклый многоугольник или многогранник в плоскости или гиперплоскости, определяемый исходными уравнениями и граничными линиями, плоскостями или гиперплоскостями.

Если существует конечное оптимальное решение, то задача линейного программирования либо имеет единственное оптимальное решение в одной из вершин многогранника решений, либо минимальное значение z достигается на всем выпуклом множестве, порождаемом двумя или более вершинами (вырожденное решение).

15.2. Двойственность в задаче линейного программирования

Задача минимизации

$$z = f(x_1, x_2, \ldots, x_n) = c_1x_1 + c_2x_2 + \ldots + c_nx_n \Rightarrow \min$$

с ограничениями

$$a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n - b_i \geq 0, \qquad (i = 1, 2, \ldots, m)$$

$$x_k \geq 0 \qquad (k = 1, 2, \ldots, r)$$

и соответствующая задача максимизация

$$w = b_1y_1 + b_2y_2 + \ldots + b_ny_n \Rightarrow \max$$

с ограничениями

$$a_{1k}y_1 + a_{2k}y + \ldots + a_{mk}y_m - c_k \leq 0, \qquad (k = 1, 2, \ldots, r)$$

$$y_i \geq 0 \qquad (i = 1, 2, \ldots, m)$$

называются двойственными задачами программирования.

Если одна из задач имеет конечное оптимальное решение, то же имеет место для второй задачи, причем

$$\min z = \max w.$$

Если одна из задач имеет неограниченное решение, то вторая задача не имеет допустимых решений.

Двойственные задачи обладают следующими свойствами:

1. В одной задаче ищут максимум линейной функции, в другой – минимум.
2. Коэффициенты при переменных в линейной функции одной задачи являются свободными членами системы ограничений в другой.
3. Каждая из задач задана в стандартной форме, причем в задаче максимизации все неравенства вида «≤», а в задаче минимизации – все неравенства вида «≥».
4. Матрицы коэффициентов при переменных в системах ограничений обеих задач являются транспонированными друг к другу.
5. Число неравенств в системе ограничений одной задачи совпадает с числом переменных в другой задаче.
6. Условия неотрицательности переменных имеются в обеих задачах.

15.3. Симплекс-метод

Множество, описываемое системой уравнений задачи линейного программирования, является выпуклым и многогранным. В связи с линейностью критерия решение задачи достигается на границе множества допустимых решений. Его выпуклость гарантирует, что найденный локальный максимум будет совпадать с глобальным. Поскольку это множество является многогранным, то из линейности критерия следует, что решение достигается в вершине множества. Если решение задачи не единственно, например, целая грань множества, то среди решений хотя бы одно является вершиной. На этом факте основано большинство методов решения задач линейного программирования.

При большом числе переменных и ограничений процесс перебора вершин многогранника практически неосуществим ввиду крайне большого числа возможных вершин многогранника решений. Впервые эффективный метод решения таких задач, в которых просматриваются не все вершины, был разработан в 1939 г. Л. В. Канторовичем. В 1949 г. американский ученый Дж. Данциг предложил универсальный метод решения задач линейного программирования, названный им симплекс-методом.

Геометрический смысл симплексного метода состоит в последовательном переходе от одной вершины многогранника ограничений (называемой первоначальной) к соседней, в которой линейная функция принимает лучшее значение (по отношению к цели задачи) до тех пор, пока не будет найдено оптимальное решение.

В настоящее время симплекс-метод используется для компьютерных расчетов, однако несложные примеры с применением симплекс-метода можно решать и вручную.

Стратегия поиска оптимального решения симплекс-методом. Оптимальное решение основной задачи линейного программирования следует искать среди допустимых решений системы ограничений. Допустимое базисное решение называется также опорным решением. Допустимые базисные решения могут содержать только неотрицательные переменные. Конечной целью является получение опорного решения с максимальным значением целевой функции.

Следовательно, алгоритм поиска должен отвечать условиям:

- при переходе от одного решения к другому должна сохраняться неотрицательность переменных;
- при переходе от одного опорного решения к другому должен обеспечиваться рост целевой функции.

Такая стратегия направленного поиска предполагает наличие некоторого исходного опорного решения, поэтому весь процесс решения распадается на три этапа – получение исходного базисного решения, поиск опорного и затем оптимального решения.

15.4. Двойственная задача об использовании ресурсов

Необходимо составить такой план производства продукции, при котором прибыль от ее реализации будет максимальной.

Обозначим x_1, x_2, ..., x_n – объем выпуска продукции видов P_1, P_2, ...,P_j, запланированных к производству. Для их изготовления потребуется запас ресурсов $b_1, b_2, \ldots, b_m$ вида $S_1, S_2, \ldots, S_l$. Норма потребления ресурса S_i при производстве единицы продукции P_j составляет a_{ij}. Прибыль (выручка) от реализации единицы продукции P_j составляет c_j.

Тогда задача максимизации прибыли от реализации продукции формулируется следующим образом:

$$F=c_1x_1 + c_2x_2 + \ldots + c_nx_n \to \max$$

при ограничениях:

$$a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \le b_1;$$
$$a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \le b_2;$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \le b_m$$

и условии неотрицательности

$$x_1 \ge 0, \quad x_2 \ge 0, \quad \ldots, \quad x_n \ge 0.$$

Смысл этой задачи состоит в том, чтобы составить такой план выпуска продукции $X = (x_1, x_2, \ldots, x_n)$, при котором прибыль (выручка) от реализации продукции будет максимальной при условии, что потребление ресурсов по каждому виду продукции не превзойдет имеющихся запасов.

Теперь предположим, что некая фирма решила закупить у предприятия ресурсы $S_1, S_2, \ldots, S_m$ и необходимо установить оптимальные цены на эти ресурсы $y_1, y_2, \ldots, y_m$.

Очевидно, что фирма, покупающая ресурсы, заинтересована в том, чтобы затраты на все ресурсы Z в количествах $b_1, b_2, \ldots, b_m$ по ценам соответственно $y_1, y_2, \ldots, y_m$ были минимальны, т. е.

$$Z = b_1y_1 + b_2y_2 + \ldots + b_my_m \to \min.$$

С другой стороны, предприятие, продающее ресурсы, заинтересовано в том, чтобы полученная выручка была не менее той суммы, которую предприятие может получить при переработке ресурсов в готовую продукцию. Поэтому для удовлетворения требований продавца затраты на ресурсы, потребляемые при изготовлении единицы продукции P_n, должны быть не менее ее цены c_n. Тогда экономико-математическая модель двойственной задачи формулируется так:

$$Z = b_1y_1 + b_2y_2 + \ldots + b_my_m \to \min$$

при ограничениях:

$$a_{11}y_1 + a_{12}y_2 + \ldots + a_{1n}y_m \le c_1;$$
$$a_{21}y_1 + a_{22}y_2 + \ldots + a_{2n}y_m \le c_2;$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$a_{1n}y_1 + a_{2n}y_2 + \ldots + a_{mn}y_m \le c_n$$

и условии неотрицательности

$$y_1 \ge 0, \quad y_2 \ge 0, \quad \ldots, \quad y_m \ge 0.$$

Смысловая интерпретация полученной двойственной задачи состоит в следующем. Надо найти такой набор цен (оценок) ресурсов $Y = (y_1, y_2, \ldots, y_m)$, при котором общие затраты на ресурсы будут минимальными при условии, что затраты на ресурсы при производстве каждого вида продукции будут не менее прибыли (выручки) от реализации этой продукции.

Цены ресурсов $y_1, y_2, \ldots, y_m$ в экономической литературе получили различные названия: учетные, неявные, теневые. Смысл этих названий состоит в том, что это условные цены. В отличие от «внешних» цен $c_1, c_2, \ldots, c_n$ на продукцию, известных до начала производства, цены ресурсов $y_1, y_2, \ldots, y_m$ являются внутренними, ибо они задаются не извне, а определяются непосредственно в результате решения задачи, поэтому их чаще называют *оценками* ресурсов.

15.5. Теоремы двойственности

Связь между оптимальными решениями двойственных задач устанавливается с помощью теорем двойственности.

Первая теорема двойственности. Если одна из взаимно двойственных задач имеет оптимальное решение, то его имеет и другая, причем оптимальные значения их линейных функций равны:

$$F_{\max} = Z_{\min} \quad \text{или} \quad F(X^*) = Z(Y^*).$$

Данное равенство является достаточным и необходимым признаком оптимальности решений взаимно двойственных задач.

Экономический смысл первой теоремы двойственности заключается в следующем. План производства $X^* = (x_1^*, x_2^*, \ldots, x_n^*)$ и набор цен (оценок) ресурсов $Y^* = (y_1^*, y_2^*, \ldots, y_m^*)$ оказываются оптимальными только тогда, когда прибыль (выручка) от продукции, найденная при «внешних» (известных заранее) ценах $c_1, c_2, \ldots, c_n$, равна затратам на ресурсы по «внутренним» (определяемым только из решения задачи) ценам $y_1, y_2, \ldots, y_m$.

Вторая теорема двойственности. Пусть даны две взаимно двойственные задачи. Если каждую из этих задач решать симплексным методом, то необходимо привести их к каноническому виду, для чего в систему ограничений задачи 1

$$\sum_{j=1}^{n} a_{ij}x_j \le b_i, \quad i = 1, 2, \ldots, m$$

следует ввести m неотрицательных переменных $x_{n+1}, x_{n+2}, \ldots, x_{n+i}, \ldots, x_{n+m}$, а в систему ограничений задачи 2

$$\sum_{i=1}^{m} a_{ij}y_i \ge c_j, \quad j = 1, 2, \ldots, n$$

n неотрицательных переменных $y_{m+1}, y_{m+2}, \ldots, y_{m+i}, \ldots, y_{m+n}$, где $i(j)$ – номер неравенства, и дополнительная переменная $x_{n+1} \ge 0$ ($y_{m+1} \ge 0$).

Системы ограничений каждой из взаимно двойственных задач примут вид:

$$\sum_{j=1}^{n} a_{ij}x_j + x_{n+1} = b_i, \quad i = 1, 2, \ldots, m,$$

$$\sum_{i=1}^{m} a_{ij}y_i - y_{m+j} = c_j, \quad j = 1, 2, \ldots, n.$$

Положительным (ненулевым) компонентам оптимального решения одной из взаимно двойственных задач соответствуют нулевые компоненты оптимального решения другой задачи, т. е. для любых $i = 1, 2, \ldots, m$ и $j = 1, 2, \ldots, n$: если $x_j^* > 0$, то $y_{m+j} = 0$; если $x_{n+1}^* > 0$, то $y_i^* = 0$, и аналогично, если $y_i^* > 0$, то $x_{n+1}^* = 0$; если $y_{m+j}^* > 0$, то $x_j^* = 0$.

Соответствие между переменными взаимно двойственных задач при достижении оптимума представляет соответствие между основными (как правило, не равными нулю) переменными одной из двойственных задач и неосновны (равными нулю) переменными другой задачи, когда они образуют допустим базисные решения.

Компоненты оптимального решения двойственной задачи равны абсол ным значениям коэффициентов при соответствующих переменных линей функции исходной задачи, выраженной через неосновные переменные ее опт мального решения.

С помощью теорем двойственности можно, решив симплексным методом ходную задачу, найти оптимум и оптимальное решение двойственной задачи

Метод, при котором вначале симплексным методом решается двойствен задача, а затем оптимум и оптимальное решение исходной задачи находят с помощью теорем двойственности, называется двойственным симплексным ме дом. Этот метод бывает выгодно применять, когда первое базисное решение ходной задачи недопустимое или когда число ее ограничений m больше чис переменных n.

15.6. Объективно обусловленные оценки

Компоненты оптимального решения двойственной задачи называются опт мальными (двойственными) оценками исходной задачи. Л. В. Канторович н звал их объективно обусловленными оценками.

Объективно обусловленные оценки ресурсов определяют степень дефицитн сти ресурсов: по оптимальному плану производства дефицитные (т. е. полност используемые ресурсы) получают ненулевые оценки, а недефицитные – нулев оценки.

В оптимальный план производства могут попасть только неубыточные вид продукции.

Компоненты оптимального решения двойственной задачи равны значени частных производных целевой функции $F_{\max}(b_1,b_2,\ldots,b_m)$ по соответствующи аргументам, т. е.

$$\frac{\partial F_{\max}}{\partial b_i}=y_i^* \quad (i=1,2,\ldots,m).$$

Свойства двойственных переменных можно объяснить и следующим обр зом. При описании метода множителей Лагранжа для задач с ограничениям типа равенств было показано, что множитель Лагранжа равен производной кри терия по правой части равенств. Этим же свойством множители Лагранжа обл дают и в задачах линейного программирования.

Значения оптимального решения двойственной задачи характеризуют устой чивость по отношению к изменениям правых частей ограничений. Это определя ет их важную роль в экономическом исследовании при анализе последствий из менения правых частей задачи. Можно сказать, что условные оценки ресурсов показывают, на сколько денежных единиц изменится максимальная прибыль (выручка) от реализации продукции при изменении запаса соответствующего ресурса на одну единицу.

Двойственные оценки могут служить инструментом анализа и принятия пра вильных решений в условиях постоянно меняющегося производства. Так, на пример, с помощью объективно обусловленных оценок ресурсов возможно со поставление оптимальных затрат и результатов производства.

15.7. Линейное программирование в Mathcad

Решение задач линейного программирования в Mathcad осуществляется с помощью функций **maximize** и **minimize** – рис. 15.1.

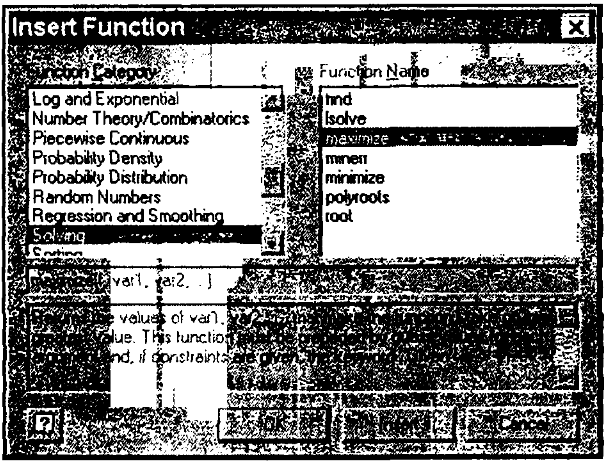


Рис. 15.1. Меню вставки функций

Задача линейного программирования в Mathcad формируется следующим образом: вписывается целевая функция, вводятся значения переменных для начала счета, под словом Given вводятся ограничения и вставляется функция **maximize (minimize)**.

Щелчком правой кнопки мыши по имени функции **maximize (minimize)** вызывается контекстно-зависимое меню (рис. 15.2), с помощью которого можно выбрать метод оптимизации.

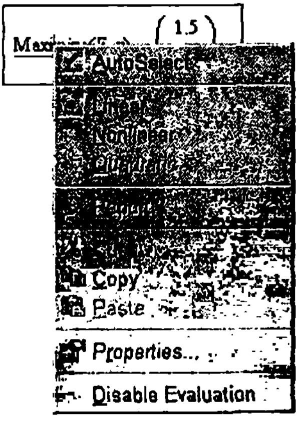


Рис. 15.2. Меню выбора метода оптимизации

Ниже приведены две задачи линейного программирования в Mathcad. Первая задача с двумя переменными, позволяющими представить графическое решение прямой и двойственной задачи.

Первая задача

$F(x) := 2 \cdot x_0 + 4 \cdot x_1$ max **целевая функция**

$x_0 := 1$ $x_1 := 1$ инициализация

Given

$x_0 + x_1 \le 8$

$3 \cdot x_0 + 3 \cdot x_1 \le 30$

$2 \cdot x_0 + x_1 \le 12$ ограничения

$-3 \cdot x_0 + x_1 \le 2$

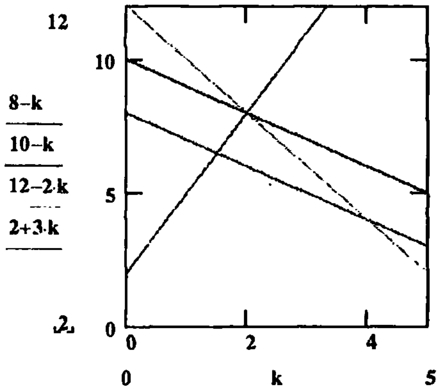
$x_0 \ge 0$ $x_1 \ge 0$

Решение

$$\text{Maximize}(F, x) = \begin{pmatrix} 1.5 \\ 6.5 \end{pmatrix}$$

$$x := \begin{pmatrix} 1.5 \\ 6.5 \end{pmatrix} \qquad F(x) = 29$$

Графическое решение



Двойственная задача

$Z(y) := 8 \cdot y_0 + 30 \cdot y_1 + 12 \cdot y_2 + 2 \cdot y_3$ min — целевая функция

$y_0 := 1 \quad y_1 := 1 \quad y_2 := 1 \quad y_3 := 1$ — инициализация

Given

$y_0 + 3 \cdot y_1 + 2 \cdot y_2 - 3 \cdot y_3 \ge 2$

$y_0 + 3 \cdot y_1 + y_2 + y_3 \ge 4$ — ограничения

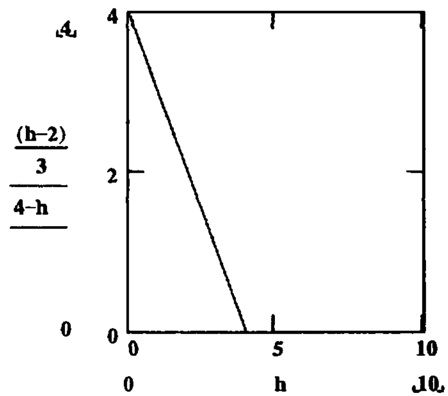
$y_0 \ge 0 \quad y_1 \ge 0 \quad y_2 \ge 0 \quad y_3 \ge 0$

Решение

$$\text{Minimize}(Z, y) = \begin{pmatrix} 3.5 \\ 0 \\ 0 \\ 0.5 \end{pmatrix} \qquad y := \begin{pmatrix} 3.5 \\ 0 \\ 0 \\ 0.5 \end{pmatrix}$$

$Z(y) = 29$

Графическое решение



Во второй задаче четыре переменных. В примере показано, как, сформировав задачу в матричном виде, можно легко сформулировать и решить двойственную задачу.

Вторая задача

$F(x) := 20 \cdot x_0 + 40 \cdot x_1 + 30 \cdot x_2 + 10 \cdot x_3$ max — целевая функция

$x_0 := 1 \quad x_1 := 1 \quad x_2 := 1 \quad x_3 := 1$ — инициализация

Given

$$x_0 + 3 \cdot x_1 + 2 \cdot x_2 + 4 \cdot x_3 \le 20$$

$$10 \cdot x_0 + 4 \cdot x_1 + 5 \cdot x_2 + 2 \cdot x_3 \le 80$$

ограничения

$$5 \cdot x_0 + 8 \cdot x_1 + 4 \cdot x_2 + 10 \cdot x_3 \le 200$$

$$x_0 \ge 0 \quad x_1 \ge 0 \quad x_2 \ge 0 \quad x_3 \ge 0$$

Решение

$$x := \mathrm{Maximize}(F, x)$$

$$x = \begin{pmatrix} 4 \\ 0 \\ 8 \\ 0 \end{pmatrix} \qquad x^T = (4 \quad 0 \quad 8 \quad 0)$$

$$F(x) = 320$$

Представим задачу в матричном виде

$$U(x) := \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix}$$ вектор переменных целевой функции

$$C := \begin{pmatrix} 20 \\ 40 \\ 30 \\ 10 \end{pmatrix}$$ коэффициенты целевой функции

$N(x) := C \cdot U(x)$ — целевая функция

$$x_0 \ge 0 \quad x_1 \ge 0 \quad x_2 \ge 0 \quad x_3 \ge 0$$

$$D := \begin{pmatrix} 1 & 3 & 2 & 4 \\ 10 & 4 & 5 & 2 \\ 5 & 8 & 4 & 10 \end{pmatrix}$$ матрица коэффициентов левых частей ограничений

$$B := \begin{pmatrix} 20 \\ 80 \\ 200 \end{pmatrix}$$ вектор правых частей ограничений

Given

$$D \cdot U(x) \le B$$

$$U(x) \ge 0$$

$$x := \text{Maximize}(N, x)$$

$$x^T = (\ 4\ \ 0\ \ 8\ \ 0\ )$$

$$N(x) = 320$$

Двойственная задача

$$U(z) := \begin{pmatrix} z_0 \\ z_1 \\ z_2 \end{pmatrix}$$ вектор переменных целевой функции

$Q(z) := B \cdot U(z)$ целевая функция (min)

$z_0 := 1$ $z_1 := 1$ $z_2 := 1$

Given

$$D^T \cdot U(z) \ge C$$

$$U(z) \ge 0$$

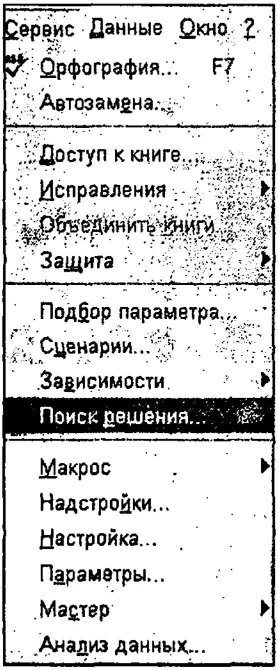
$$z := \text{Minimize}(Q, z)$$

$$z^T = (\ 13.333\ \ 0.667\ \ 0\ )$$

$$Q(z) = 320$$

15.8. Линейное программирование в Excel

Решение задач линейного программирования в Excel производится с помощью решающего блока Solver, вызываемого командой меню **Сервис – Поиск решения**.



Последовательность действий такова. Вводятся исходные данные, лучше в созданную для этого форму. Вводятся зависимости из математической модели. Из меню **Сервис** открывается диалоговое окно **Поиск решения**, в котором вводятся ячейка целевой функции, ее назначение (максимум или минимум), изменяемые ячейки и добавляются ограничения. В опции **Параметры** должен стоять флажок у линейной модели.

Рассмотрим решение той же задачи (второй), которую мы решали в Mathcad.

Ввод исходных данных показан на рис. 15.3.

Теперь необходимо ввести зависимости из математической модели. Эти зависимости представляют собой левые части ограничений и целевую функцию. Данную операцию можно выполнить с помощью функции СУММПРОИЗ, где в первый массив вводятся коэффициенты соответствующего ограничения, а во второй массив переменные x0, x1, x2,x3, точнее ячейки, где мы им присвоили инициирующие значения – ячейки B10:E10. На рис. 15.4 представлены введенные функции, а влияющие ячейки показаны стрелками.

	A	B	C	D	E	F	G	H
1			переменные			левая часть	знак	правая часть
2	наименование	x0	x1	x2	x3			
3	коэф в цел.функции	20	40	30	10			
4								
5	коэф в 1 ограничени	1	3	2	4		<=	20
6	коэф во 2 ограничени	10	4	5	2		<=	80
7	коэф в 3 ограничени	5	8	4	10		<=	200
8								
9		x0'	x1'	x2'	x3'	целевая функция		
10	оптимальные значен	1	1	1	1			max

Отчет по устойчивости 1 / Отчет по пределам 1 / Лист2 / Лист3

Рис. 15.3. Ввод исходных данных

1. Из меню **Сервис** откроем окно **Поиска решения** (рис. 15.5).
2. В поле **Установить целевую ячейку** введем $F$10.
3. Из группы **Равной** выберем переключатель – **максимальному значению**.
4. В поле области **Изменяя ячейки** введем ячейки с первоначальными значениями переменных – $B$10:$E$10.

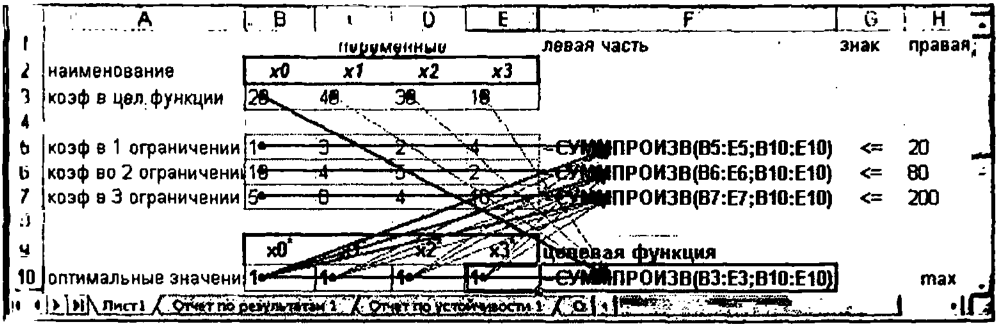


Рис. 15.4. Ввод зависимостей

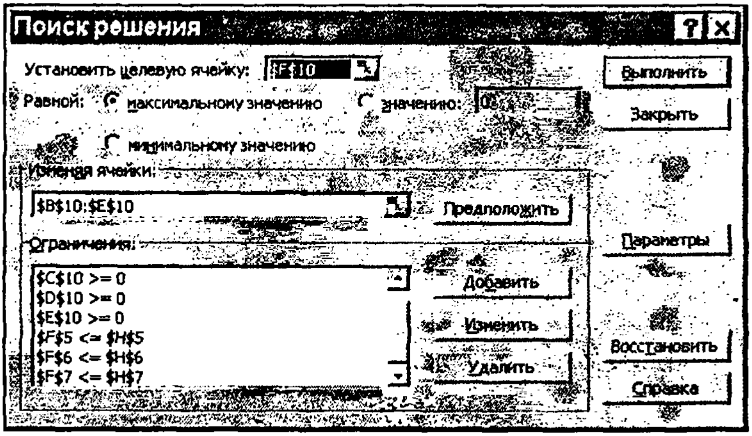


Рис. 15.5. Диалоговое окно **Поиск решения**

5. Нажав кнопку Добавить, откроем диалоговое окно **Добавление ограничения** (рис. 16.6).

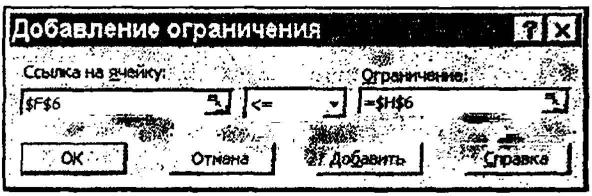


Рис. 15.6. Диалоговое окно **Добавление ограничения**

6. Через данное окно введем ограничения в соответствии со знаком, который принят в модели. В нашей задаче левые части ограничений должны быть меньше или равны правым частям ограничений и переменные должны быть положительны.

Открыв диалоговое окно **Параметры поиска решения** (рис. 15.7) можно изменить параметры **Максимальное время** или **Предельное число итераций** в случае, если за данное количество итераций задача не решена. Если не устраивает погрешность, введенная по умолчанию, ее также можно изменить. Для решения задач линейного программирования должен быть установлен флажок **Линейная модель**.

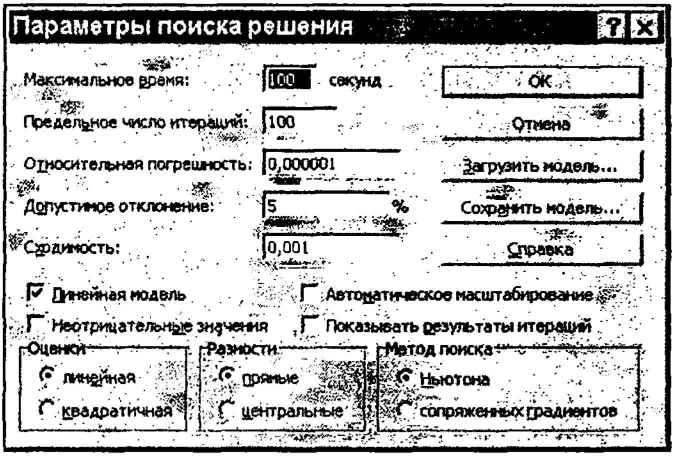


Рис. 15.7. Диалоговое окно ***Параметры поиска решения***

После нажатия кнопки **OK** вновь появится диалоговое окно **Поиск решения**. По нажатии кнопки **Выполнить** на экран выводится окно **Результаты поиска решения** (рис. 15.8).

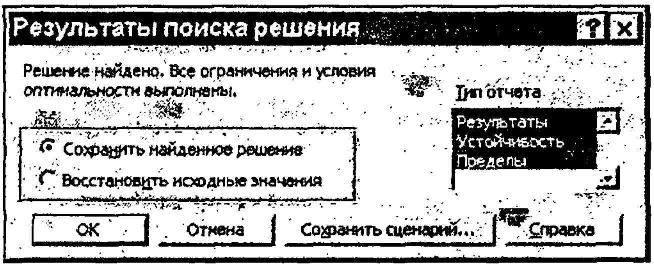


Рис. 15.8. Диалоговое окно ***Результаты поиска решения***

Если решение не найдено, окно выведет соответствующее сообщение.

Если решение найдено, выделим все три типа отчетов, нажмем **OK**, и результат решения задачи – на экране (рис. 15.9).

Для анализа полученного оптимального решения в Excel предусмотрены три типа отчетов: отчет по результатам, устойчивости, пределам.

	A	B	C	D	E	F	G	H	I
1		переменные				левая часть	знак	правая часть	
2	наименование	x0	x1	x2	x3				
3	коэф в цел.функции	20	40	30	10				
4									
5	коэф в 1 ограничении	1	3	2	4	20	<=	20	
6	коэф во 2 ограничени	10	4	5	2	80	<=	80	
7	коэф в 3 ограничении	5	8	4	10	52	<=	200	
8									
9		x0*	x1*	x2*	x3*	целевая функция			
10	оптимальные значени	4	0	8	0	320		max	
11									

Лист1 / Отчет по результатам 1 / Отчет по устойчивост

Рис. 15.9. Результаты решения задачи

В отчете по результатам приведены сведения о целевой функции, значениях искомых переменных и результаты оптимального решения для ограничений (рис. 15.10).

	A	B	C	D	E	F	G
6	Целевая ячейка (Максимум)						
7		Ячейка	Имя	Исходно	Результат		
8		$F$10	оптимальные значения целевая функция	100	320		
9							
10	Изменяемые ячейки						
11		Ячейка	Имя	Исходно	Результат		
12		$B$10	оптимальные значения x0*	1	4		
13		$C$10	оптимальные значения x1*	1	0		
14		$D$10	оптимальные значения x2*	1	8		
15		$E$10	оптимальные значения x3*	1	0		
16							
17	Ограничения						
18		Ячейка	Имя	Значение	формула	Статус	Разница
19		$F$5	коэф в 1 ограничении левая часть	20	$F$5<=$H$5	связанное	0
20		$F$6	коэф во 2 ограничении левая часть	80	$F$6<=$H$6	связанное	0
21		$F$7	коэф в 3 ограничении левая часть	52	$F$7<=$H$7	не связан.	148
22		$E$10	оптимальные значения x3*	0	$E$10>=0	связанное	0
23		$D$10	оптимальные значения x2*	8	$D$10>=0	не связан.	8
24		$C$10	оптимальные значения x1*	0	$C$10>=0	связанное	0
25		$B$10	оптимальные значения x0*	4	$B$10>=0	не связан.	4

Отчет по результатам 3 / Отчет по устойчивости 3 / Отчет по

Рис. 15.10. Отчет по результатам

Для ограничений в столбце **формула** приведены зависимости, которые были введены в диалоговое окно **Поиск решения**; в столбце **Значение** приведены величины использованного ресурса; в столбце **Разница** показано количество неиспользованного ресурса. Если ресурс используется полностью, то в столбце **Ста-**

тус указывается «связанное»; при неполном использовании ресурса в этом столбце указывается «не связан.». Для переменных показывается разность между значением переменных в найденном оптимальным решении и заданным для них граничным условием.

В отчете по устойчивости (рис. 15.11) дан анализ по переменным и ограничениям. В анализе переменных приведены следующие данные:

- результирующие значения переменных;
- нормированная стоимость, т. е. дополнительные двойственные переменные, которые показывают, насколько изменяется целевая функция при принудительном включении единицы этой переменной в оптимальное решение;
- коэффициенты целевой функции;
- допустимые значения приращения коэффициентов целевой функции, при которых сохраняется набор переменных, входящих в оптимальное решение.

Изменяемые ячейки

Ячейка	Имя	Результ. значение	Нормир. стоимость	Целевой Коэффициент	Допустимое Увеличение	Допустимое Уменьшение
$B$10	оптимальные значения x0*	4	0	20	5,714285715	5
$C$10	оптимальные значения x1*	0	-2,666666667	40	2,666666667	1E+30
$D$10	оптимальные значения x2*	8	0	30	10	1,538461539
$E$10	оптимальные значения x3*	0	-44,66666667	10	44,66666667	1E+30

Ограничения

Ячейка	Имя	Результ. значение	Теневая Цена	Ограничение Правая часть	Допустимое Увеличение	Допустимое Уменьшение
$F$5	коэф в 1 ограничении левая часть	20	13,33333333	20	12	12
$F$6	коэф во 2 ограничении левая часть	80	0,666666667	80	120	30
$F$7	коэф в 3 ограничении левая часть	52	0	200	1E+30	148

Отчет по результатам / Отчет по устойчивости / Отчет по пределам

Рис. 15.11. Отчет по устойчивости

В анализе ограничений приведены значения:

- величин использованных ресурсов;
- теневые цены, т. е. двойственные оценки, которые показывают, как изменится целевая функция при изменении ресурсов на единицу;
- значения приращения ресурсов, при которых сохраняется оптимальный набор переменных, входящих в оптимальное решение.

В отчете по пределам (рис. 15.12) показано, в каких пределах может изменяться выпуск продукции, вошедшей в оптимальное решение, при сохранении структуры оптимального решения.

Рассмотрим теперь ввод математической модели в матричном виде, не меняя приготовленной формы (рис. 15.13). Для этого необходимо ввести две матричных функции. В векторе ограничений левой части – ячейки F5:F7 – вводится функция умножения матрицы коэффициентов в ограничениях и транспонированного вектора переменных:

=МУМНОЖ(B5:E7;ТРАНСП(B10:E10)).

Ячейка	Имя	Целевое значение					
$F$10	оптимальные значения целевая ф	320					
Ячейка	Изменяемое Имя	значение	Нижний предел	Целевое результат		Верхний предел	Целевое результат
$B$10	оптимальные значения x0*	4	0	240		4	320
$C$10	оптимальные значения x1*	0	0	320		-7,7686E-12	320
$D$10	оптимальные значения x2*	8	0	80		8	320
$E$10	оптимальные значения x3*	0	0	320		-5,82645E-12	320

Отчет по результатам / Отчет по устойчивости / Отчет по пределам

Рис. 15.12. Отчет по пределам

Целевая функция записывается как функция умножения вектора коэффициентов целевой функции на транспонированный вектор переменных

=МУМНОЖ(B4:E4;ТРАНСП(B10:E10)).

	A	B	C	D	E	F	G	H
1						В матричном виде		
2		переменные				левая часть	знак	правая часть
3	наименование	x0	x1	x2	x3			
4	коэф в цел.функции	20	40	30	10			
5	коэф в 1 ограничении	1	3	2	4	=МУМНОЖ(B5:E7;ТРАНСП(B10:E10))	<=	20
6	коэф во 2 ограничени	10	4	5	2	=МУМНОЖ(B5:E7;ТРАНСП(B10:E10))	<=	80
7	коэф в 3 ограничении	5	8	4	10	=МУМНОЖ(B5:E7;ТРАНСП(B10:E10))	<=	200
8								
9		x0*	x1*	x2*	x3*	целевая функция		
10	оптимальные значени	1	1	1	1	=МУМНОЖ(B4:E4;ТРАНСП(B10:E10))		max

Отчет по пределам / Лист1 / Лист2 / Лист3 / Лист4 / Лист5

Рис. 15.13. Ввод зависимостей в матричном виде

Ввод ограничений показан на рис. 15.14.

Решим двойственную задачу. Схема формирования двойственной задачи приведена на рис. 15.15. Коэффициенты бывшей целевой функции становятся правой частью ограничений. Правая часть ограничений становится коэффициентами новой целевой функции. Матрица коэффициентов ограничений транспонируется.

Ввод зависимостей для двойственной задачи показан на рис. 15.16.

Левая часть ограничений представляет собой произведение матрицы коэффициентов ограничений на вектор переменных. Целевая функция записывается как произведение транспонированного вектора коэффициентов целевой функции на вектор переменных.

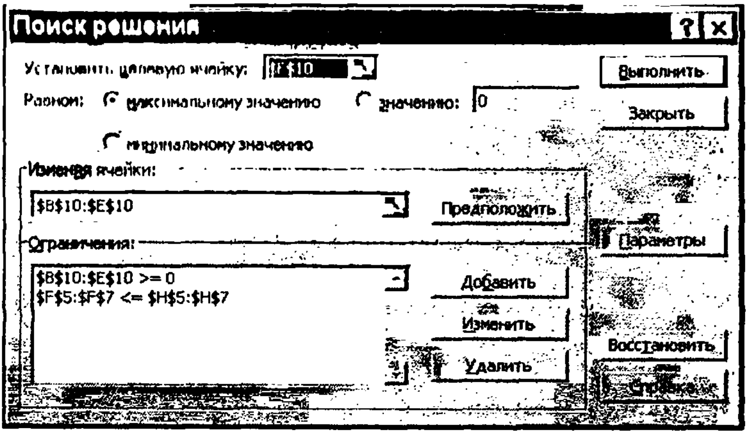
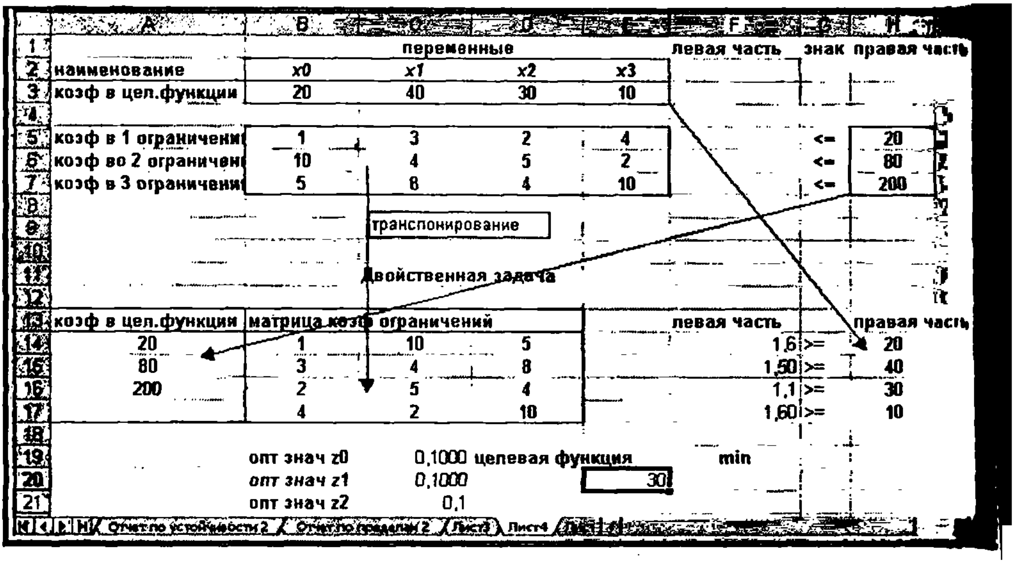


Рис. 15.14. Окно ***Поиск решения*** *с ограничениями для задачи в матричном виде*



	A	B	C	D	E	F	G	H
1		переменные				левая часть	знак	правая часть
2	наименование	x0	x1	x2	x3			
3	коэф в цел.функции	20	40	30	10			
4								
5	коэф в 1 ограничени	1	3	2	4		<=	20
6	коэф во 2 ограничен	10	4	5	2		<=	80
7	коэф в 3 ограничени	5	8	4	10		<=	200
8								
9			транспонирование					
10								
11			Двойственная задача					
12								
13	коэф в цел.функции	матрица коэф ограничений				левая часть		правая часть
14	20	1	10	5		1,6	>=	20
15	80	3	4	8		1,50	>=	40
16	200	2	5	4		1,1	>=	30
17		4	2	10		1,60	>=	10
18								
19		опт знач z0	0,1000	целевая функция		min		
20		опт знач z1	0,1000		30			
21		опт знач z2	0,1					

Рис. 15.15. Схема формирования двойственной задачи

Ограничения приведены на рис. 15.17 в окне **Поиск решения**. Это положительность переменных и то, что вектор левой части ограничений должен быть больше вектора из правой части. Для целевой ячейки устанавливаем флажок минимизации.

Результаты решения двойственной задачи приведены на рис. 15.18.

Рис. 15.16. Ввод зависимостей для двойственной задачи

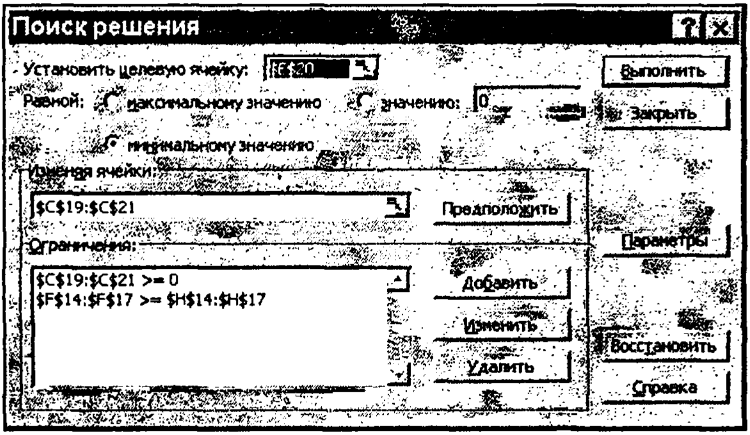


*Рис. 15.17. Окно **Поиск решения** с ограничениями для двойственной задачи*

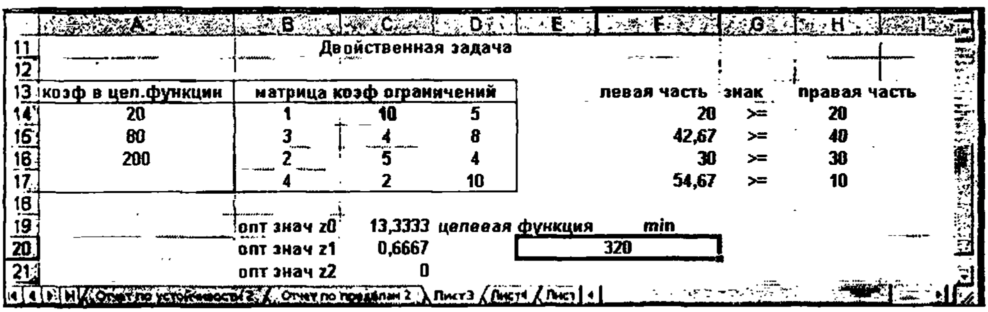


Рис. 15.18. Результаты решения двойственной задачи

Открыв отчет по устойчивости (рис. 15.19), можно увидеть новые двойственные оценки (в столбце Теневая Цена) и убедиться, что значения переменных при решении задачи на максимизацию становятся двойственными оценками при задаче на минимизацию, и наоборот (сравните с рис. 15.11).

Изменяемые ячейки

Ячейка	Имя	Результ. значение	Нормир. стоимость	Целевой Коэффициент	Допустимое Увеличение	Допустимое Уменьшение
$C$20	опт знач z0 Двойственная задача	13,3333	0	20	12	[illegible]
$C$21	опт знач z1 Двойственная задача	0,6667	0	80	120	[illegible]
$C$22	опт знач z2 Двойственная задача	0	148	200	1E+30	140

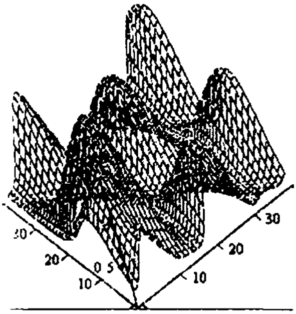
Ограничения

Ячейка	Имя	Результ. значение	Теневая Цена	Ограничение Правая часть	Допустимое Увеличение	Допустимое Уменьшение
$F$15	левая часть	20	4	20	5,71	[illegible]
$F$16	левая часть	42,67	0	40	2,67	1E+30
$F$17	левая часть	30	8	30	10	[illegible]
$F$18	левая часть	54,67	0	10	44,67	1E+30

Отчет по результатам 2 | Отчет по устойчивости 2 | Отчет по пределам 2

Рис. 15.19. Отчет по устойчивости для двойственной задачи

Глава 16. Нелинейное программирование



Если линейная целевая функция или одно или более линейных ограничений в задаче линейного программирования заменены нелинейными относительно переменных x_k, то имеет место задача нелинейного программирования [2, 3, 15, 18]. Задачи нелинейного программирования представляют практический интерес, но, за малым исключением, поддаются лишь численным методам решения.

Задачи нелинейного программирования делятся на задачи с ограничениями и без ограничений, выпуклого и невыпуклого программирования (в зависимости от типа целевой функции).

Задача минимизации функции многих переменных без дополнительных ограничений относится к числу простейших задач нелинейного программирования.

К настоящему времени разработано и исследовано большое число методов минимизации функций многих переменных. Мы остановимся лишь на некоторых наиболее известных и часто используемых на практике методах минимизации.

Значительное внимание уделено классическим методам минимизации – градиентному методу и методу Ньютона. Эти методы имеют большое значение в идейном отношении. Оба они явным образом основаны на идее замены минимизируемой функции в окрестности очередной точки x_k первыми членами ее разложения в ряд Тейлора. В градиентном методе берут линейную часть разложения, в методе Ньютона – квадратическую часть. Многие методы оптимизации базируются на той же идее аппроксимации функций.

16.1. Градиентный метод

Направление наибыстрейшего возрастания функции $f(x)$ в точке x совпадает с направлением градиента $\nabla f(x)$, а направление наибыстрейшего убывания – с направлением антиградиента $(-\nabla f(x))$.

Это замечательное свойство градиента лежит в основе ряда итерационных методов минимизации функций. Одним из таких методов является градиентный метод. Этот метод, как и все итерационные методы, предполагает выбор начального приближения – некоторой точки x_0. Общих правил выбора точки x_0 в градиентном методе, как, впрочем, и в других методах, нет. В тех случаях, когда из геометрических, физических или каких-либо других соображений может быть получена информация об области расположения точки минимума, начальное приближение x_0 стараются выбрать поближе к этой области.

Будем считать, что некоторая начальная точка x_0 уже выбрана. Тогда градиентный метод заключается в построении последовательности $\{x_k\}$ по правилу:

$$x_{k+1} = x_k - a_k \nabla f(x_k),\ a_k > 0,\ k = 0,1\ldots \text{ (если ищется } f_{min}) \tag{16.1}$$

или

$$x_{k+1} = x_k + a_k \nabla f(x_k),\ a_k > 0,\ k = 0,1\ldots \text{ (если ищется } f_{max}). \tag{16.2}$$

Число a_k называют длиной шага или просто шагом градиентного метода. Если $f'(x_k) \neq 0$, то шаг можно выбрать так, чтобы $f(x_{k+1}) < f(x_k)$.

Если $f'(x) = 0$, то x_k – стационарная точка. В этом случае итерационный процесс прекращается, и при необходимости проводится дополнительное исследование поведения функции в окрестности точки x_k для выяснения того, достигается ли в точке x_k минимум функции $f(x)$. В частности, если $f(x)$ – выпуклая функция, то в стационарной точке всегда достигается минимум.

Градиентные методы отличаются способами выбора длины шага и алгоритмами нахождения точки x_{k+1}. Если величина длины шага a выбирается так, чтобы приращение функции Δf при перемещении из точки x_k в точку x_{k+1} было наибольшим (при отыскании f_{max}) или наименьшим (при отыскании f_{min}), то градиентный метод называется методом скорейшего спуска.

По методу скорейшего спуска длина шага выбирается из условия экстремума функции:

$$\Delta f = f(x_{k+1}) - f(x_k).$$

Продифференцировав функцию Δf с учетом выражения x_{k+1} по формуле (16.2) и выражения градиента в точке x_k:

$$\nabla f(x_k) = \left(\frac{\partial f}{\partial x_1}(x_k), \frac{\partial f}{\partial x_2}(x_k), \ldots, \frac{\partial f}{\partial x_n}(x_n) \right),$$

получим, что необходимое условие экстремума $d\Delta f/da = 0$ примет вид:

$$\frac{\partial f}{\partial x_1}(x_{k+1}) \cdot \frac{\partial f}{\partial x_1}(x_k) + \frac{\partial f}{\partial x_2}(x_{k+1}) \cdot \frac{\partial f}{\partial x_2}(x_k) + \ldots + \frac{\partial f}{\partial x_n}(x_{k+1}) \cdot \frac{\partial f}{\partial x_n}(x_k) = 0.$$

Используя скалярное произведение векторов, данное условие можно записать так:

$$\nabla f(x_{k+1}) \nabla f(x_k) = 0 \tag{16.3}$$

Скалярное произведение векторов в прямоугольной системе координат равно сумме произведений их соответствующих координат. Скалярное произведение векторов равно 0 в случае их ортогональности.

Рассмотрим простой пример отыскания максимума функции

$$f = x_1 x_2 - (x_1 - 3)^2 - (x_2 - 4)^2$$

методом скорейшего спуска при ограничениях:

$$x_1 + x_2 \leq 20,$$
$$x_1 > 0,\ x_2 > 0.$$

Функция f является выпуклой, поэтому ее локальный максимум совпадает с глобальным. Для данного примера оптимум достигается внутри области решений системы ограничений.

Градиент этой функции равен:

$$\nabla f = (x_2 - 2x_1 + 6);\ (x_1 - 2x_2 + 8).$$

В качестве исходной точки возьмем точку x_0 (2; 1). Подставляя координаты x_0 в выражение градиента, получим $\nabla f_0 = (3; 8)$. Используя формулу $x_1 = x_0 + a\nabla f$, получим

$$x_1 = (2, 1) + a(3, 8) = (2 + 3a), (1 + 8a).$$

Отсюда

$$\nabla f_1 = (1 + 8a - 2(2 + 3a) + 6), (2 + 3a - 2(1 + 8a) + 8) = (3 + 2a), (8 - 13a).$$

Теперь установим a из выражения

$$\nabla f_0 \nabla f_1 = (3; 8) [(3 + 2a), (8 - 13a)] = 3(3 + 2a) + 8(8 - 13a) = 73 - 98a = 0.$$

Отсюда $a = 0{,}7449$. Подставив величину длины шага, установим координаты следующей точки и градиента:

$$x_1 = (2 + 3 \cdot 0{,}7448), (1 + 8 \cdot 0{,}7449) = (4{,}25,\ 7{,}0);$$

$$\nabla f_1 = (3 + 2 \cdot 0{,}7449), (8 - 13 \cdot 0{,}7449) = (4{,}5,\ -1{,}75).$$

Дальнейшие вычисления представлены в табл. 16.1.

Таблица 16.1. Результаты вычислений по методу скорейшего спуска

k	x_k	$f(x_k)$	∇f_k	x_{k+1}	∇f_{k+1}	$\nabla f_k \nabla f_{k+1}$	a
1	(2; 1)	−8,00	(3; 8)	$(2+3a), (1+8a)$	$(3+2a), (8-13a)$	$73-98a$	0,7449
2	(4,25; 7,0)	19,1875	(4,5; −1,75)	$(4{,}25+4{,}5a)$, $(7-1{,}75a)$	$(4{,}5-10{,}7a)$, $(-1{,}75+8a)$	$23{,}312-62{,}15a$	0,375
3	(5,94, 6,34)	23,5404	(0,4875, 1,25)	$(5{,}94+0{,}48a)$, $(6{,}34+1{,}25a)$	$(0{,}46+0{,}275a)$, $1{,}26-2{,}01a)$	$1{,}8-2{,}379a$	0,7564
4	(6,309, 7,285)	24,220359	(0,67, 0,0546)	$(6{,}309+0{,}67a)$, $(7{,}29+0{,}055a)$	$(0{,}667-1{,}28a)$, $(-0{,}26+0{,}56a)$	$0{,}43-0{,}825a$	0,523
5	(6,658, 7,313)	24,333021	(−0,003, 0,031)	$(6{,}66-0{,}003a)$, $(7{,}31+0{,}031a)$	$(-0{,}003+0{,}037a)$, $(0{,}032-0{,}065a)$	$0{,}001-0{,}0021a$	0,476
6	(6,6566, 7,3278)	24,333257	(0,0146, 0,0009)				

Вычисления остановлены в связи с незначительными изменениями функции.

На практике итерации продолжают до тех пор, пока не выполнится некоторый критерий счета. Здесь часто используются следующие критерии:

$$|x_k - x_{k+1}| \le \varepsilon, \text{ или } |\nabla f_k| \le \gamma, \text{ или } |f(x_k) - f(x_{k+1})| \le \delta,$$

где ε, γ, δ – заданные числа. Иногда заранее задают число итераций; возможны различные сочетания этих критериев.

Следует заметить, что антиградиент $(-\nabla f(x_k))$ указывает направление быстрейшего спуска лишь в достаточно малой окрестности точки x_k. Это означает, что если функция $f(x)$ меняется быстро, то в следующей точке x_{k+1} направление антиградиента $(-\nabla f(x_{k+1}))$ может сильно отличаться от направления $(-\nabla f(x_k))$. Поэтому слишком точное определение a из условий оптимума бывает не всегда целесообразным.

На практике нередко довольствуются нахождением какого-либо $a_k > 0$, обеспечивающего условия монотонности: $f(x_{k+1}) < f(x_k)$. С этой целью задаются какой-либо постоянной $a > 0$ и на каждой итерации берут $a_k = a$. При этом для каждого $k \ge 0$ проверяют условие монотонности, и в случае его нарушения дробят a_k до тех пор, пока не восстановится монотонность метода.

В случае, когда минимизируемая функция f не предполагается выпуклой, градиентной метод может обеспечить лишь сходимость к множеству стационарных точек функции f.

В случае сильно выпуклой минимизируемой функции градиентный метод сходится к точке минимума со скоростью геометрической прогрессии.

Градиентный метод хорошо работает лишь на первых этапах поиска минимума, когда точки x_k не слишком близки к точке минимума x_*, а вблизи точки x_* расстояние $|x_k - x_*|$ часто перестает уменьшаться, сходимость метода ухудшается. Это связано с тем, что в окрестности точки минимума градиент близок к нулю, главная линейная часть приращения $f(x_k) - f(x_*)$, на базе которой выбирается направление спуска, становится малой, учитывается влияние квадратичной части приращения, метод становится слишком чувствительным к неизбежным погрешностям вычислений. Поэтому вблизи точки минимума при необходимости пользуются более точными и, вообще говоря, более трудоемкими методами, лучше учитывающими не только линейные, но и квадратичные части приращения.

16.2. Метод Ньютона

Если минимизируемая функция дважды непрерывно дифференцируема и производные $f'(x)$ и $f''(x)$ вычисляются достаточно просто, то возможно применение методов минимизации второго порядка, которые используют квадратичную часть разложения этой функции в ряд Тейлора. Методы второго порядка сходятся быстрее, чем методы первого порядка, потому что квадратичная часть разложения аппроксимирует функцию гораздо точнее, чем линейная.

Метод Ньютона является методом второго порядка, т. е. использует вычисление вторых производных минимизируемой функции и имеет квадратичную скорость сходимости на классе сильно выпуклых функций. В его модификациях (квазиньютоновских алгоритмах) матрица вторых производных аппроксимируется с помощью информации о значениях градиентов функции, и эти модификации, таким образом, являются методами первого порядка.

Описание метода Ньютона. Предположим, что функция f выпукла и дважды дифференцируема на E^n, причем матрица $f''(x)$ не вырождена на E^n. В методе Ньютона последовательность $x_0, x_1, x_2, \ldots$ генерируется, исходя из следующих соображений.

По определению дважды дифференцируемой функции для очередной точки x_k имеем:

$$f(x) - f(x_k) = \langle f'(x_k), x - x_k\rangle + 1/2\,\langle f''(x_k)\,(x - x_k)\rangle + 0(|x - x_k|^2).$$

Для определения следующей точки x_{k+1} минимизируется функция $f_k(x)$, являющаяся квадратичной частью приращения $f(x) - f(x_k)$, т. е. решается задача:

$$f_k(x) = \langle f'(x_k), x - x_k\rangle + 1/2\langle f''(x_k)\,(x - x_k), x - x_k\rangle \to \min, x \in E^n.$$

Необходимое и достаточное условие минимума имеет вид:

$$f'_k(x_k) = f'(x_k) + f''(x_k)\,(x - x_k) = 0.$$

Решая полученную систему линейных уравнений и принимая найденную точку минимума за x_{k+1}, получаем:

$$x_{k+1} = x_k + h_k,\ h_k = -(f''(x_k))^{-1} f'(x_k).$$

Данное соотношение определяет метод Ньютона минимизации функции f, совпадающий с известным методом Ньютона решения системы уравнений $f'(x)=0$.

Сложность отыскания нужного начального приближения является недостатком метода Ньютона. Еще более существенным недостатком является высокая трудоемкость метода, обусловленная необходимостью вычисления и обращения на каждом шаге матрицы вторых производных минимизируемой функции.

В силу названных причин применение классического метода Ньютона далеко не всегда приводит к успеху. Многочисленные модификации направлены на то, чтобы, сохраняя основное достоинство метода Ньютона – его быструю сходимость, уменьшить трудоемкость и ослабить требования к выбору начального приближения.

Метод Ньютона с регулировкой шага. Рассмотрим метод

$$x_{k+1} = x_k + a_k h_k,\ a_k > 0,\ h_k = -(f''(x_k))^{-1} f'(x_k),$$

называемый методом Ньютона с регулировкой шага. При $a_k \equiv 1$ он совпадает с классическим методом Ньютона.

Выбор коэффициентов a_k производится обычно или из условия минимизации функции вдоль заданного направления, или с помощью метода дробления шага.

Данный метод сходится при любой начальной точке $x_0 \in E^n$. Таким образом, с помощью регулировки шага преодолевается недостаток метода Ньютона, связанный с необходимостью отыскания хорошего начального приближения. Метод Ньютона с регулировкой шага является одним из наиболее популярных в вычислительной практике.

16.3. Квазиньютоновские методы

Рассмотрим метод

$$x_{k+1} = x_k + a_k h_k,\ h = -D_k f'(x_k).$$

Матрицу D_k будем выбирать таким образом, чтобы она в некотором смысле аппроксимировала матрицу $(f''(x_k))^{-1}$. Предполагая невырожденной матрицу $f''(x_{k+1})$, отсюда с точностью до членов более высокого порядка малости по сравнению с $\|x_k - x_{k+1}\|$ имеем:

$$(f''(x_{k+1}))^{-1}(f'(x_{k+1}) - f'(x_k)) \approx x_{k+1} - x_k.$$

При этом, если $f(x) = \langle Ax, x\rangle/2 + \langle b, x\rangle$ – квадратичная функция, A – симметрическая положительно определенная матрица, то $f'(x) = Ax + b, f''(x) = A$ и приближенное равенство обращается в точное:

$$(f''(x_{k+1}))^{-1} \Delta f'_k = \Delta x_k,$$

где $\Delta x_k = x_{k+1} - x_k$, $\Delta f'_k = f'(x_{k+1}) - f'(x_k)$.

Поэтому для матрицы D_{k+1}, приближающей $(f''(x_{k+1}))^{-1}$, можно записать условие:

$$D_{k+1} \Delta f'_k = \Delta x_k.$$

Это условие носит название *квазиньютоновского*. Оно лежит в основе целого ряда методов аппроксимации $(f'')^{-1}$. Соответствующие методы минимизации, для которых на каждом шаге выполняется квазиньютоновское условие, также называются квазиньютоновскими.

Пусть приближения к $(f'')^{-1}$ пересчитываются от шага к шагу по формуле:

$$D_{k+1} = D_k + \Delta D_k.$$

Представим данное выражение в виде:

$$\Delta D_k \, \Delta f'_k = \Delta x_k - D_k \, \Delta f'_k.$$

Этому равенству удовлетворяет матрица ранга 1, заданная формулой:

$$\Delta D_k = (1/\langle z_k, \Delta f'_k\rangle)(\Delta x_k - D_k \, \Delta f'_k)\, z_k,$$

где z_k – произвольный вектор, такой, что $\langle z_k, \Delta f'_k\rangle \neq 0$.

Выбрав $z_k = \Delta x_k - D_k \, \Delta f'_k$, получаем при $\langle \Delta x_k - D_k \, \Delta f'_k, \Delta f'_k\rangle \neq 0$ следующую формулу для пересчета матриц D_k:

$$D_{k+1} = D_k + \frac{(\Delta x_k - D_k \Delta f'_k)\cdot(\Delta x_k - D_k \Delta f'_k)}{\langle \Delta x_k - D_k \Delta f'_k, \Delta f'_k\rangle}.$$

Часто используется также формула:

$$D_{k+1} = D_k + \frac{\Delta x_k \Delta x_k}{\langle \Delta x_k, \Delta f'_k\rangle} - \frac{D_k(\Delta f'_k \cdot \Delta f'_k) D_k}{\langle D_k \Delta f'_k, \Delta f'_k\rangle}.$$

В качестве D_0 можно выбрать любую положительно определенную симметрическую матрицу. На практике часто выбирается единичная матрица.

Длина шага в квазиньютоновских методах чаще всего выбирается из условия минимизации функции вдоль заданного направления:

$$f(x_k + a_k h_k) = \min f(x_k + a h_k).$$

Итак, в квазиньютоновской процедуре направления поиска задаются в виде $-D_k \nabla f(x_k)$. Направление градиента является, таким образом, отклоненным в результате умножения на D_k, где D_k – положительно определенная симметрическая матрица порядка $n \times n$, аппроксимирующая обратную матрицу Гессе. На следующем шаге матрица D_{k+1} представляется в виде суммы D_k и двух симметрических матриц ранга 1 каждая.

Квазиньютоновские методы являются эффективным средством решения задач безусловной оптимизации. Их отличает высокая скорость сходимости, в то же время при реализации квазиньютоновских алгоритмов не приходится выполнять такие трудоемкие операции, как вычисление матрицы вторых производных или обращение матрицы. Однако при большой размерности пространства необходимость хранения и пересчета на каждом шаге матриц D_k обусловливает высокие требования к объему занимаемой памяти ЭВМ. Этот недостаток не присущ методу сопряженных градиентов.

16.4. Метод сопряженных градиентов

В методах сопряженных направлений требуется найти направления $h_0, h_1, \ldots$ $\ldots, h_{n-1}$ – такие, что последовательность n одномерных минимизаций вдоль этих направлений приводит к отысканию минимума функции $f(x_n) = \min f(x)$ при любом $x_0 \in E^n$, где

$$x_{k+1} = x_k + a_k h_k, f(x_k + a_k h_k) = \min f(x_k + a h_k), k = 0,1,\ldots$$

Оказывается, что указанным свойством обладает система взаимно сопряженных относительно матрицы A направлений.

Понятие сопряженных направлений и их свойства. Пусть A – симметрическая положительно определенная матрица размера $n\times n$. Векторы (направления) h_I и h_{II} называются сопряженными относительно матрицы A или A-ортогональными, если они отличны от нуля и $\langle Ah_I, h_{II}\rangle = 0$. Векторы (направлеиия) $h_0, h_1, \dots, h_n$ называются взаимно сопряженными (относительно матрицы A), если все они отличны от нуля и $\langle Ah_i, h_j\rangle = 0, i \neq j, 0 \leq i, j \leq k$.

Взаимно сопряженные векторы линейно независимы.

Рассмотрим задачу минимизации квадратичной функции $f(x) = \langle Ax, x\rangle/2 + \langle b, x\rangle$ с помощью метода сопряженных градиентов. Построим систему взаимно сопряженных направлений по правилу:

$$h_0 = -f'(x_0),\ h_k = -f'(x_k) + \beta_{k-1} h_{k-1},\ k \geq 1.$$

Из условия сопряженности векторов h_{k-1} и h_k имеем:

$$0 = \langle h_k, Ah_{k-1}\rangle = \langle -f'(x_k), Ah_{k-1}\rangle + \beta_{k-1}\langle h_{k-1}, Ah_{k-1}\rangle,$$

откуда:

$$\beta_{k-1} = \frac{\langle f'(x_k), Ah_{k-1}\rangle}{\langle h_{k-1}, Ah_{k-1}\rangle}.$$

Метод обеспечивает отыскание точки минимума квадратичной функции не более чем за n шагов.

Сформулируем теперь метод сопряжениых градиентов для миниимизации неквадратичной функции. Для этого преобразуем формулу определения β_{k-1} так, чтобы в ней не фигурировала матрица A:

$$\beta_{k-1} = \frac{\langle f'(x_k), Ah_{k-1}\rangle}{\langle h_{k-1}, Ah_{k-1}\rangle} = \frac{\langle f'(x_k), f'(x_k) - f'(x_{k-1})\rangle}{\langle h_{k-1}, f'(x_k) - f'(x_{k-1})\rangle} =$$

$$= \frac{\langle f'(x_k), f'(x_k) - f'(x_{k-1})\rangle}{\langle -f'(x_{k-1}) - \beta_{k-2} f'(x_{k-2}) - \dots, f'(x_k) - f'(x_{k-1})\rangle} =$$

$$= \frac{\langle f'(x_k), f'(x_k) - f'(x_{k-1})\rangle}{\|f'(x_{k-1})\|^2} = \frac{\|f'(x_k)\|^2}{\|f'(x_{k-1})\|^2}.$$

Поскольку в данной задаче квадратичность функции не предполагается, то нельзя ожидать, что описанный метод сопряженных градиентов за конечное число итераций приведет к точке минимума функции. Далее, точное определение величины a_k возможно лишь в редких случаях, поэтому реализация каждой итерации метода будет сопровождаться неизбежными погрешиостями. Как показывает практика, эти погрешности, накапливаясь, могут привести к тому, что векторы h_k перестанут указывать направление убывания фуикции, и сходимость метода может нарушиться. Чтобы бороться с этим явлением, метод сопряженных градиентов время от времени обновляют, полагая $\beta_k = 0$. Номера $k \geq 1$, при которых принимается $\beta_k = 0$, называются моментами обновления метода. На практике часто берут моменты обновления $\{n, 2n, 3n, \dots\}$, где n – размериость рассматриваемого пространства. В этом случае вариант метода сопряжениых градиеитов принимает вид:

$$x_{k+1} = x_k + a_k h_k, f(x_k + a_k h_k) = \min f(x_k + a h_k),\ k = 0,1,\dots$$

$$h_0 = -f'(x_0),\ h_k = -f(x_k) + \beta_{k-1} h_{k-1},\ k \geq 1;$$

$$\beta_{k-1}=\begin{cases}\dfrac{\langle f'(x_k),f'(x_k)-f'(x_{k-1})\rangle}{\|f'(x_{k-1})\|^2}, & k\notin\{n,2n,3n,\ldots\},\\ 0, & k\in\{n,2n,3n,\ldots\}.\end{cases}$$

Метод сопряженных градиентов обладает высокой скоростью сходимости. В то же время его трудоемкость сравнительно невелика. Все это позволяет отнести метод сопряженных градиентов к числу наиболее эффективных алгоритмов первого порядка. Вычислительная практика показывает, что этот метод незначительно уступает по эффективности квазиньютоновским методам; в то же время он предъявляет меньшие требования к объему занимаемой памяти ЭВМ. Следует отметить, что в настоящее время построено и применяется много различных вариантов метода сопряженных градиентов.

Коротко еще о двух методах.

Квадратичное программирование. В задаче квадратичного программирования целевая функция является квадратичной, а ограничения – линейными функциями. Квадратичной является функция, представляемая в виде:

$$f(x_1,\ldots,x_n)=\sum_{i=1}^{n}\sum_{j=1}^{m}a_{ij}x_ix_j+\sum_{i=1}^{n}b_ix_i+c.$$

Метод Левенберга–Марквардта. Данный метод совмещает в себе методы Ньютона и градиентного спуска. Левенберг и Марквардт, добавив неотрицательный член к диагональным элементам матрицы ($f''(x_k)$) в процедуре Ньютона, предложили искать следующую точку из уравнения:

$$x_{k+1} -= x_k-(f''(x_k)+\nu I)^{-1}f'(x_k),$$

где I – единичная матрица размера $n\times n$; а ν – некоторая величина, называемая числом Марквардта.

При $\nu=0$ приходим к методу Ньютона, а при достаточно большом ν получаем метод градиентного спуска. Основная идея метода заключается в том, чтобы при высокой степени нелинейности, пока «расстояние» между итеративным и искомым решением значительно, использовать большие ν (метод градиентного спуска), а при приближении к искомому решению постоянно уменьшать значения ν, что позволяет достаточно быстро достичь желаемого минимума.

Метод Левенберга–Марквардта позволяет найти локальный минимум суммы квадратов регрессионных остатков.

16.5. Нелинейное программирование в Excel

Задачи нелинейного программирования в Excel решаются с помощью знакомых уже опций **Сервис – Поиск решения**. Однако в диалоговом окне **Параметры поиска решения** не надо устанавливать флажок **Линейная модель**.

Схема работы следующая. Вводятся начальные значения искомых переменных. Через адреса ячеек, в которых присвоены начальные значения искомых переменных, вводится целевая функция и ограничения. Целевая функция в начальной точке не должна быть равной нулю. Из меню **Сервис** необходимо вызвать диалоговое окно **Поиск решения** (рис. 16.1). В нем в поле **Установить целевую** ячейку ввести адрес ячейки, где записана целевая функция, далее установить флажок по цели решения на максимум или минимум.

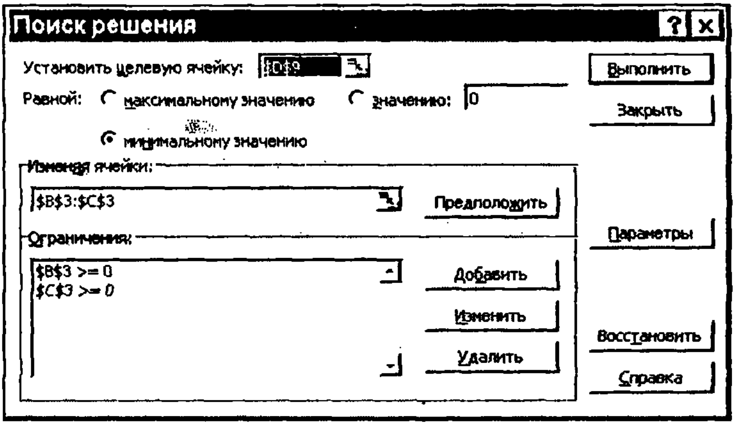


Рис. 16.1. Диалоговое окно ***Поиск решения***

В поле области **Изменяя ячейки** ввести адреса ячеек, где установлены начальные значения переменных, *и в поле области* **Ограничения** ввести ограничения модели. Нажать кнопку **Параметры**. В открывшемся диалоговом окне **Параметры поиска решения** (рис. 16.2) снять флажок **Линейная модель**, если он установлен.

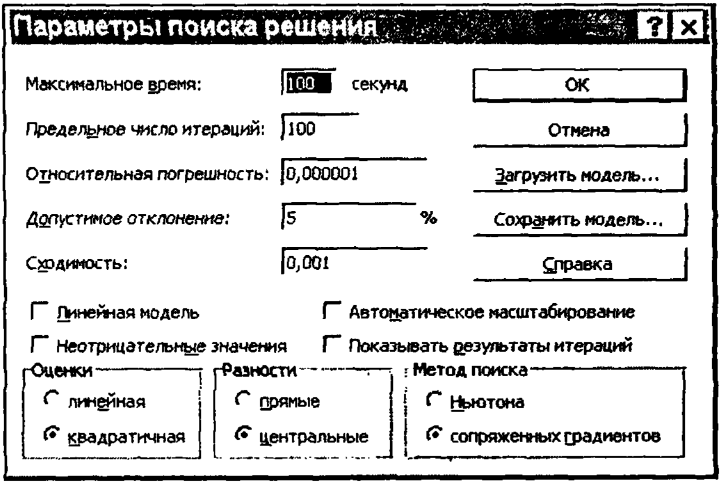


Рис. 16.2. Диалоговое окно ***Параметры поиска решения***

Выбрать метод решения. В Excel существует два метода – **квазиньютоновский** метод и метод **сопряженных градиентов**. Под параметром **Производные** понимается *метод численного дифференцирования, аппроксимация производных с по-*

мощью центральных разностей предпочтительней. Под командой **Оценка** понимается характер шага метода. Смысл остальных параметров очевиден.

Рассмотрим пример. Минимизировать функцию

$$f(x) = (x_1 - 2)^4 + (x_1 - 2x_2)^2$$

при ограничениях $x_1 > 0$, $x_2 > 0$.

Введем в ячейки B3 и C3 начальные значения (рис. 16.3), в ячейку D3 – целевую функцию, в ячейки E3 и F3 – производные (для контроля).

	B	C	D	E	F
1			Метод сопряженных градиентов		
2	x1	x2	f(x)	f/(x1)	f/(x2)
3	0	3	=(B3-2)^4+(B3-2*C3)^2	=4*(B3-2)^3+2*B3-4*C3	=-4*B3+8*C3
4					

Рис. 16.3. Ввод начальных значений и формул

Тогда получим вид листа, приведенный на рис. 16.4.

	B	C	D	E	F
1			Метод сопряженных градиентов		
2	x1	x2	f(x)	f/(x1)	f/(x2)
3	0,000000	3,000000	52,00000000	-44,00000	24,00[illegible]
4					
5					

Рис. 16.4. Вид листа после ввода формул

Далее выполним действия, приведенные выше. Вид диалоговых окон для данного примера показан на рис. 16.1 и 16.2. После нажатия кнопок **ОК** и **Выполнить** получим результат минимизации функции, приведенный на рис. 16.5.

	B	C	D	E	F
1			Метод сопряженных градиентов		
2	x1	x2	f(x)	f/(x1)	f/(x2)
3	2,039907301	1,020363253	0,000003207	-0,001384187	0,003276822
4					
5					

Рис. 16.5. Результат минимизации функции

При желании увеличить точность можно уменьшить относительную погрешность и сходимость, и получить результат, как на рис. 16.6.

B	C	D	E	F
		Метод сопряженных градиентов		
x1	x2	f(x)	f/(x1)	f/(x2)
2,001485	1,000742	0,000000000	0,000000082	-0,000000137

Рис. 16.6. Результат вычислений после уменьшения относительной погрешности и сходимости

Рассмотрим другой пример — функцию Розенброка, часто используемую как тестовую. Необходимо минимизировать функцию

$$f(x) = 100(y - x^2)^2 + (1 - x)^2.$$

Функция Розенброка имеет минимум в точке $x = 1$, $y = 1$. Введем начальные значения и целевую функцию, как показано на рис. 16.7.

	A	B	C	D	E
1			Функция Розенброка		
2	x	y	f(x)	f/(x)	f/(y)
3	2	-2	=100*(B3-A3^2)^2+(1	-400*(B3-A3^2)*A3-2+2*A3	=200*B3-200*A3^2
4					
5					

Рис. 16.7. Ввод начальных значений и целевой функции

После первого запуска решающего блока получим результат, представленный на рис. 16.8.

	A	B	C	D	E
1			Функция Розенброка		
2	x	y	f(x)	f/(x)	f/(y)
3	0,5649483747	0,3171974953	0,1896576800	-0,4251113161	-0,3938341575
4					

Рис. 16.8. Результат после первого запуска

Не меняя погрешности и сходимости, сделаем еще два запуска решающего блока. Получим удовлетворительный результат, представленный на рис. 16.9.

	A	B	C	D	E
1			Функция Розенброка		
2	x	y	f(x)	f/(x)	f/(y)
3	0,9999118584	0,9998233669	0,0000000078	-0,0000332125	-0,0000715417
4					
5					

Рис. 16.9. Результат минимизации функции Розенброка

Этого результата можно также добиться, изменяя значения погрешности и сходимости.

16.6. Нелинейное программирование в Mathcad

В среде Mathcad щелчок правой кнопкой мыши по имени «решающих» функций (**minimize**, **maximize**, **minerr** и **find**) вызывает дополнительное меню со списком трех возможных методов численного решения задачи нелинейного программирования (рис. 16.10):

- сопряженных градиентов;
- Левенберга–Марквардта;
- квазиньютоновский.

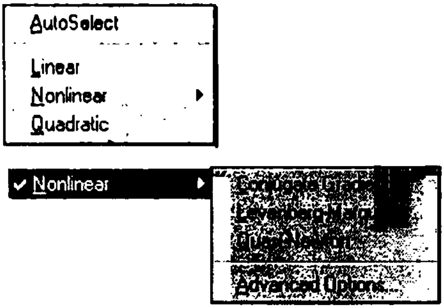


Рис. 16.10. Меню выбора метода

Выбор команды **AutoSelect** поможет квалифицировать задачу и выбрать метод между линейным, нелинейным и квадратичным программированием (рис. 16.11). Также решающий блок даст подсказку по выбору метода нелинейного программирования, что не мешает выбрать их самостоятельно или скомбинировать.

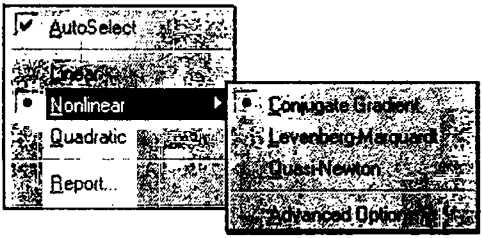


Рис. 16.11. Применение опции ***AutoSelect***

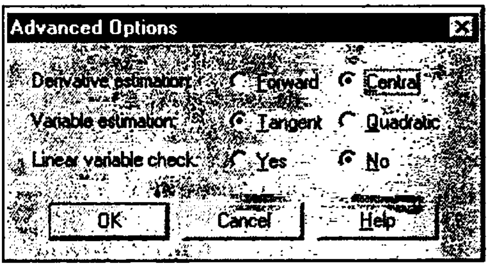


Рис. 16.12. Диалоговое окно ***Advanced Options***

В окне **Advanced Options** (Дополнительные параметры) содержатся некоторые средства управления блока для решения проблем особо тонкого характера. Он-ции окна **Advanced Options** связаны только с нелинейными оптимизаторами. Командой **Advanced Options** открывается одноименное диалоговое окно, показанное на рис. 16.12.

Диалоговое окно вносит в список три параметра, каждый из которых имеет два взаимно исключающих выбора. Заданные по умолчанию параметры настройки приведены на рисунке. Эти параметры входят в порядок алгоритма. Заданные по умолчанию параметры решают большинство проблем.

Derivative estimation – аппроксимация производных. Панель для выбора типа численного дифференцирования состоит из двух опций:

- **Forward** – аппроксимация с помощью разностей вперед (правых разностей);
- **Central** – аппроксимация с помощью центральных разностей.

Аппроксимация с помощью центральных разностей предпочтительней. Однако не всегда. Например, рассмотрим задачу:

$$x := 0$$

Given

$$\sqrt{x} = 10$$

Find(x) = ▮

Эта задача при использовании центральных разностей не решается, поскольку при начальном значении 0 в подкоренном выражении получаем комплексное число. С изменением установки типа аппроксимации решение будет найдено.

Variable estimation – изменение типа шага метода при перемещении от одной точки до следующей:

- **Tangent** (tangential approximation) – аппроксимация, направленная по касательной к данной кривой, т. е. по направлению вектора градиента;
- **Quadratic** – квадратичная аппроксимация шага.

Linear variable check – Разложение математического выражения в простые линейные функции для увеличения быстродействия вычислений.

Схема работы следующая. Ввод начальных значений переменных, целевой функции и ограничений. Ввод слова Given. Вставка «решающей» функции, щелчком правой кнопки мыши вызывается меню выбора метода. При необходимости изменяется точность – TOL.

Рассмотрим пример с минимизацией уже знакомой функцией Розенброка.

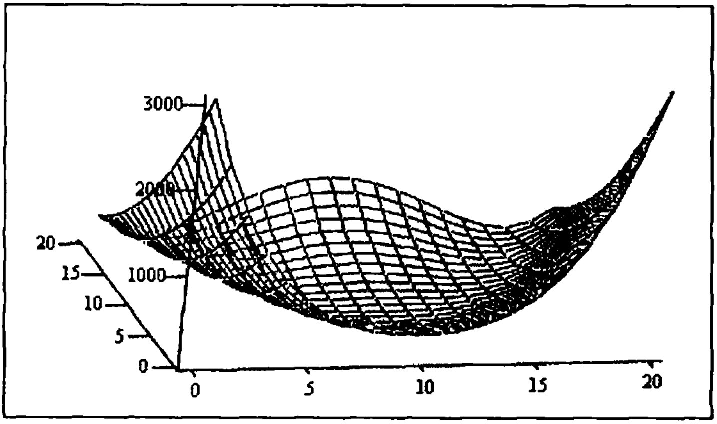
Минимум функции Розенброка трудно увидеть на графике. Представим ее график (рис. 16.13).

$$f(x, y) := 100 \cdot \left(y - x^2\right)^2 + (1 - x)^2$$

$$i := 0 .. 20 \qquad j := 0 .. 20$$

$$x_i := -2 + 4 \cdot \frac{1 \cdot i}{20} \qquad y_j := -1.5 + 4 \cdot \frac{1 \cdot j}{20}$$

$$M_{i,j} := f\left(x_i, y_j\right)$$



M

Рис. 16.13. График функции Розенброка

Теперь найдем ее минимум методами сопряженных градиентов и квазиньютоновского.

$$f(x, y) := 100 \cdot \left(y - x^2\right)^2 + (1 - x)^2$$

$$x := 2 \qquad y := -2$$

Given

1 $\quad \text{Minimize}(f, x, y) = \begin{pmatrix} 0.979 \\ 0.957 \end{pmatrix}$ Метод сопряженных градиентов

$$\begin{pmatrix} x \\ y \end{pmatrix} := \begin{pmatrix} 0.979 \\ 0.957 \end{pmatrix}$$

2 $\quad \text{Minimize}(f, x, y) = \begin{pmatrix} 0.99999956 \\ 0.99999911 \end{pmatrix}$ Повторно Метод сопряженных градиентов

$x := 2 \qquad y := 2$

Given

$$1 \qquad \text{Minimize}(f, x, y) = \begin{pmatrix} 0.863 \\ 0.75 \end{pmatrix}$$

Квазиньютоновский метод

$$\begin{pmatrix} x \\ y \end{pmatrix} := \begin{pmatrix} 0.863 \\ 0.75 \end{pmatrix}$$

$$2 \qquad \text{Minimize}(f, x, y) = \begin{pmatrix} 0.999998825 \\ 0.999997763 \end{pmatrix}$$

Повторно Квазиньютоновский метод

$$\begin{pmatrix} x \\ y \end{pmatrix} := \begin{pmatrix} 0.999998825 \\ 0.999997763 \end{pmatrix}$$

$$3 \quad \text{Minimize}(f, x, y) = \begin{pmatrix} 0.999998868 \\ 0.999997732 \end{pmatrix}$$

3 Квазиньютоновский метод

Увеличим точность

$$f(x, y) := 100 \cdot \left(y - x^2\right)^2 + (1 - x)^2 \qquad TOL := 1 \cdot 10^{-9}$$

$x := 2 \qquad y := -2$

Given

$$1 \qquad \text{Minimize}(f, x, y) = \begin{pmatrix} 1.00000122 \\ 1.000002445 \end{pmatrix}$$

Метод сопряженных градиентов при увеличении точности

Как видим, результат большой точности получается при запуске решающего устройства два или три раза в зависимости от метода или увеличения точности.

Визуальная демонстрация поиска минимума нелинейным программированием функции Розенброка прекрасно выполнена в демонстрационной версии пакета Matlab (рис. 16.14 – 16.17).

Так, на рис. 16.14 приведен график функции Розенброка и путь решения от исходной точки (2,0; – 1,9) квазиньютоновским методом по алгоритму Дэвидона–Флетчера–Пауэлла. Количество итераций для достижения решения – 86.

На рис. 16.15 приведен путь решения методом наискорейшего спуска. Количество итераций – 301.

На рис. 16.16 приведены результаты минимизации функции методом Ньютона. Количество итераций – 32.

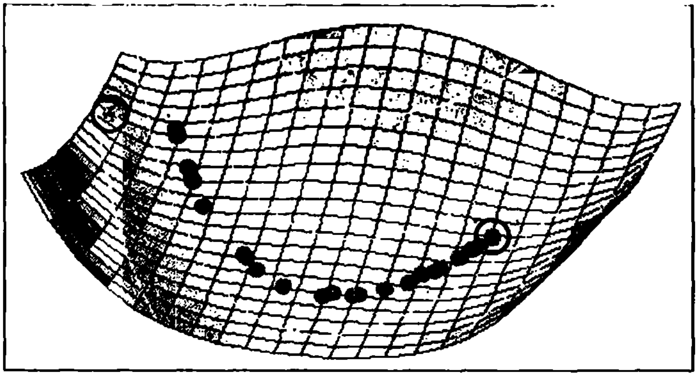


Рис. 16.14. Минимизация функции Розенброка квазиньютоновским методом

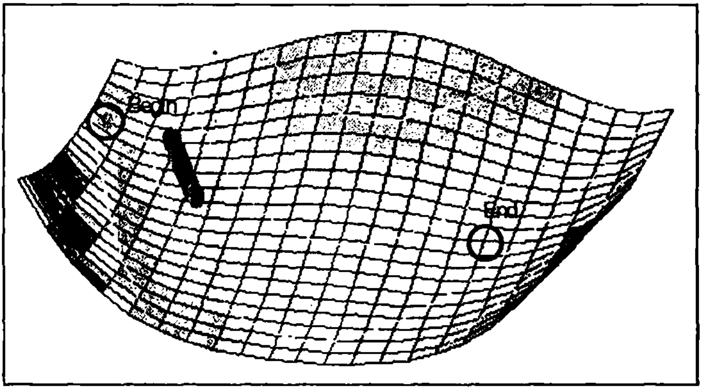


Рис. 16.15. Минимизация функции Розенброка методом наискорейшего спуска

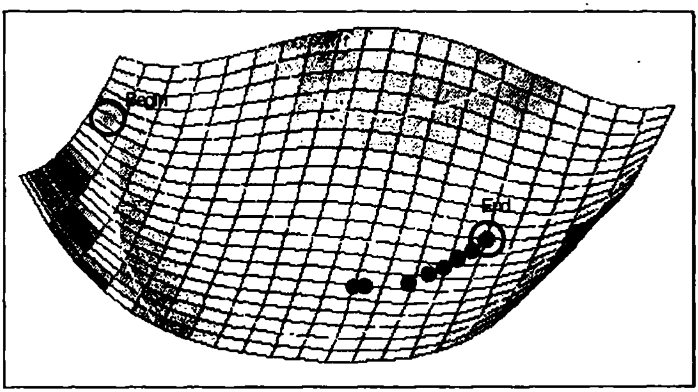


Рис. 16.16. Минимизация функции методом Ньютона

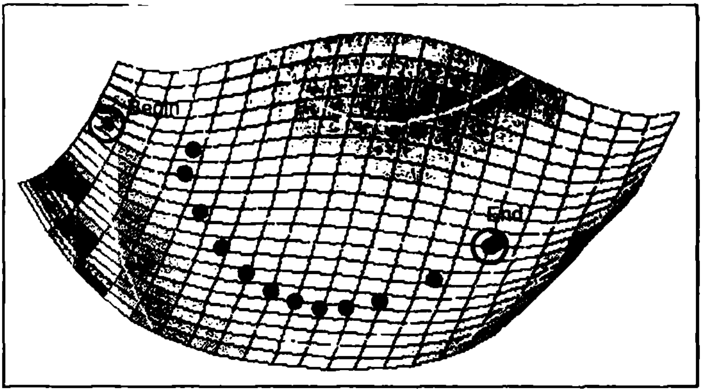


Рис. 16.17. Минимизация функции методом Левенберга–Марквардта

На рис. 16.17 приведен путь решения минимизации функции методом Левенберга–Марквардта. Количество итераций – 60.

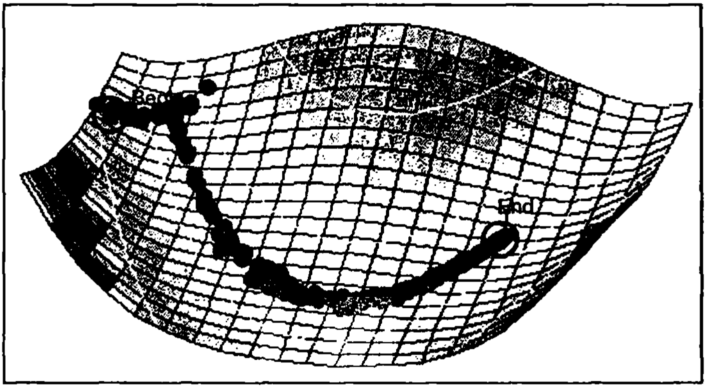


Рис. 16.18. Минимизация функции методом покоординатного спуска

На рис. 16.18 приведены результаты использования метода покоординатного спуска с использованием алгоритма симплексного поиска (метода многогранника). Количество итераций – 210.

Рассмотрим задачу *нелинейного программирования с ограничениями* в Mathcad.

Введем целевую функцию и ограничения, зададим инициализацию и в решающем блоке под словом Given сформулируем задачу и применим метод сопряженных градиентов:

Целевая функция **Ограничения**

$f(x,y) := 12x + 5y - 0.7x^2$ max $5 \cdot y + x^2 - 10 \cdot x \le 30$

$x \ge 0 \qquad y \ge 0$

Решение

$x := 1 \qquad y := 1$

Given

$5 \cdot y + x^2 - 10 \cdot x \le 30$

$x \ge 0 \qquad y \ge 0$

$$\begin{pmatrix} x0 \\ y0 \end{pmatrix} := \text{Maximize}(f, x, y)$$

$$\begin{pmatrix} x0 \\ y0 \end{pmatrix} = \begin{pmatrix} 6.471 \\ 10.567 \end{pmatrix} \qquad f(x0, y0) = 101.177$$

Построим графики целевой функции и ограничений в осях x–y, целевую функцию для значений оптимума $k(x,y)$ – рис. 16.19.

$$12 \cdot x + 5 \cdot y - 0.7 \cdot x^2 = k$$

$$k := 101.76$$

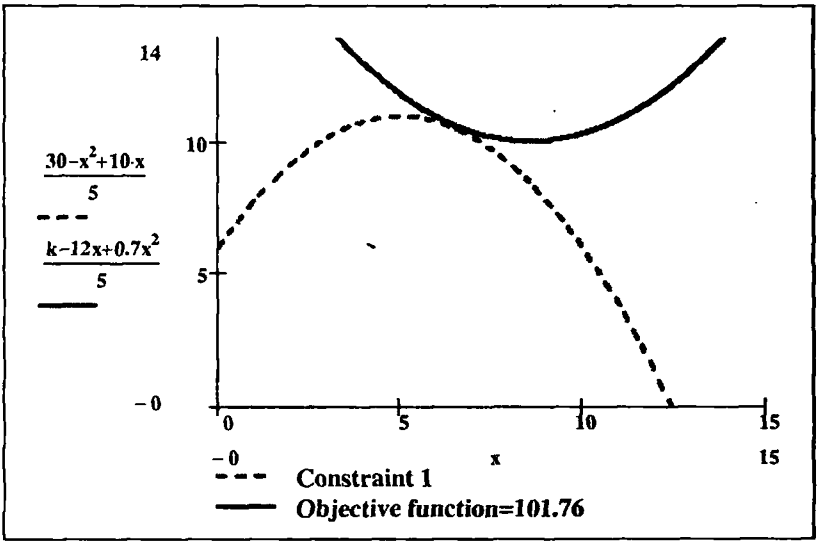


Рис. 16.19

Ограничение можно представить как область допустимых значений, в которой ищется максимум целевой функции (рис. 16.20).

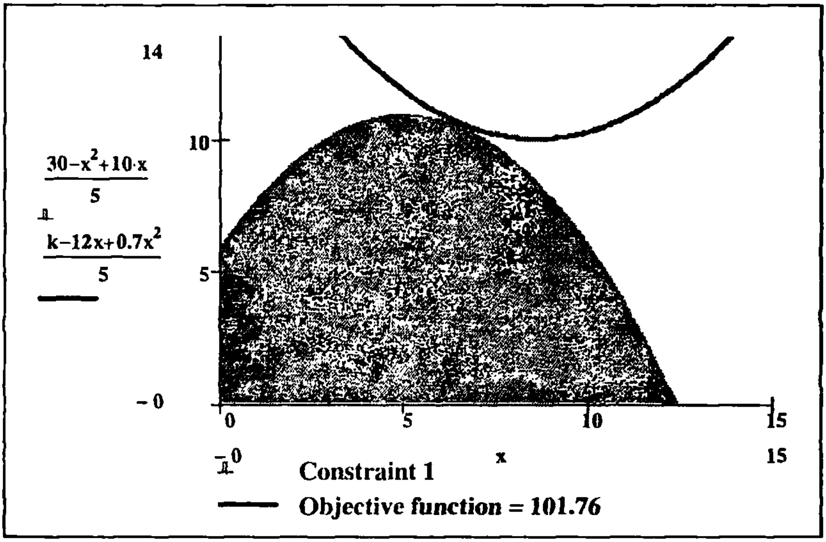


Рис. 16.20

Введем еще одно ограничение

$x := 1 \qquad y := 1$

Given

$$5 \cdot y + x^2 - 10 \cdot x \le 30$$

$$2 \cdot y + 8 \cdot x \le 32$$

$$x \ge 0 \qquad y \ge 0$$

$$\begin{pmatrix} x1 \\ y1 \end{pmatrix} := \text{Maximize}(f, x, y)$$

$$\begin{pmatrix} x1 \\ y1 \end{pmatrix} = \begin{pmatrix} 1.771 \\ 8.915 \end{pmatrix}$$

$$f(x1, y1) = 63.634$$

Представим решение графически (рис. 16.21):

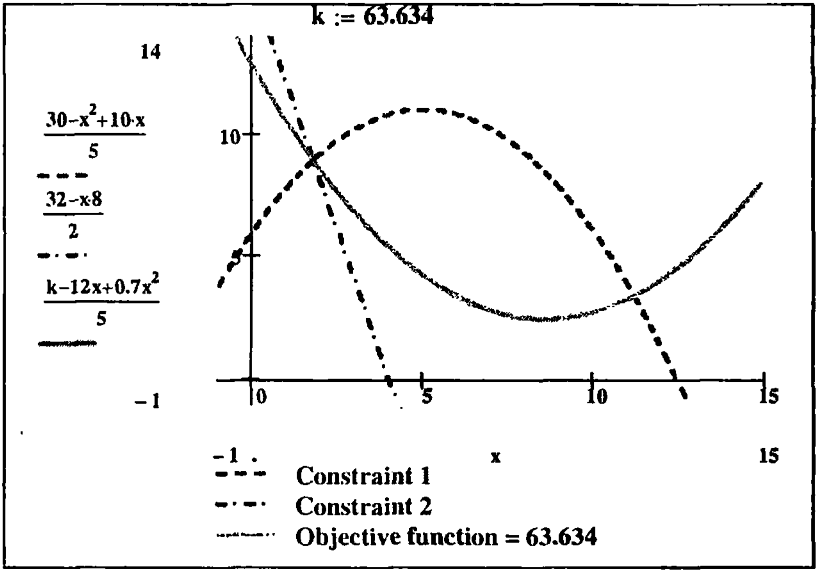


Рис. 16.21

Или выделив допустимую область значений – рис.16.22.

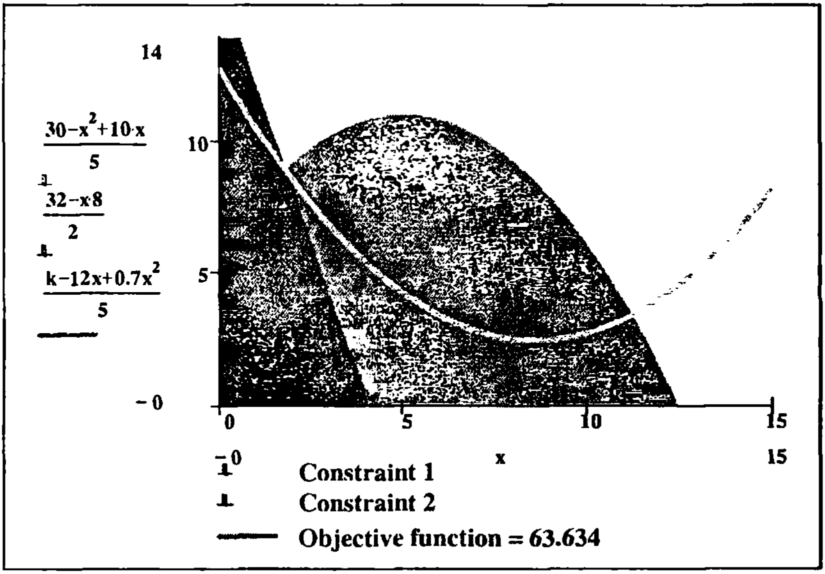


Рис. 16.22

Расширим область ограничения и установим оптимум:

$$x := 1 \qquad y := 1$$

Given

$$5 \cdot y + x^2 - 10 \cdot x \le 30$$

$$2 \cdot y + 8 \cdot x \le 90$$

$$x \ge 0 \qquad y \ge 0$$

$$\begin{pmatrix} x2 \\ y2 \end{pmatrix} := \mathrm{Maximize}(f, x, y)$$

$$\begin{pmatrix} x2 \\ y2 \end{pmatrix} = \begin{pmatrix} 6.471 \\ 10.567 \end{pmatrix}$$

$$f(x2, y2) = 101.177$$

Представим результат графически – рис.16.23.

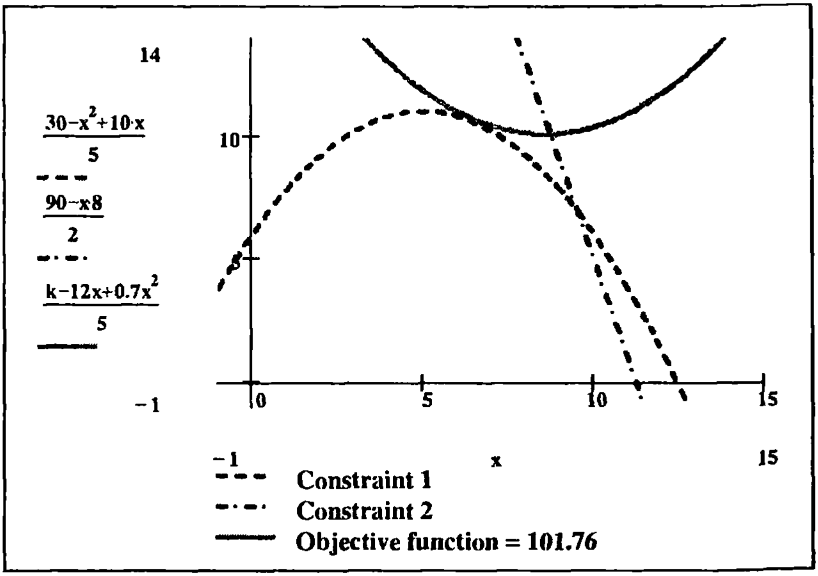


Рис. 16.23

Как видно из результата решения и графика, это тот случай, когда ограничение не задействовано, т. е. оптимальное решение лежит в его области допустимых значений.

Методы нелинейного программирования широко применяются в экономике. Это задачи оптимизации прибыли фирмы, оптимизации портфеля ценных бумаг, установки параметров нелинейных функций и множество других. Пример установки параметров производственной функции с постоянной эластичностью замещения методом Левенберга–Марквардта приведен в главе 20.

Много примеров приведено в справочной системе Mathcad.

Приведенными встроенными функциями и алгоритмами не ограничиваются возможности Mathcad. Во-первых, можно самостоятельно моделировать алгоритмы. Во-вторых – дополнительно инсталлировать пакет *Expert Solver*, что важно при практической работе по оптимизации с функциями многих переменных, использовании квадратического программирования, например оптимизации портфеля ценных бумаг.

The ***Solving and Optimization Extension Pack*** (***Expert Solver***) – Пакет Расширения Оптимизации (Экспертное Решающее устройство) – расширяет возможности Mathcad, включая большее количество подпрограмм оптимизации для решения проблем более высокой сложности, включая смешанное целочисленное программирование и генерацию отчетов. Этот пакет расширения имеет наиболее мощную технологию решающего устройства, позволяя вычислять сложные модели.

Пакет Расширения Оптимизации (Экспертное Решающее устройство) добавляет к уже имеющемуся сложному решающему устройству, представленному в Mathcad 2000:

- увеличение числа переменных, позволенных в системе, до 1000;
- решения подпрограмм для нелинейных и квадратичных систем:
- предоставление возможности генерировать итеративные отчеты для линейных и квадратичных решений программирования;
- обеспечение смешанного целочисленного программирования.

Обратите внимание: если на вашу машину не установлен пакет Экспертного Решающего устройства, то вместо квадратичного программирования будете видеть в диалоговом окне выбора метода метод сопряженного градиента.

Глава 17. Основные представления об экономико-математическом моделировании

17.1. Методологические основы экономико-математического моделирования

Экономическая наука включает в себя – как необходимые инструментальные средства – математические методы и модели. Их использование позволяет формализовать важнейшие связи экономических систем и на этой основе производить анализ экономических систем, осуществлять прогнозирование и оптимизацию. Математические и эконометрические методы позволяют получать новые знания об экономическом объекте и его поведении, оценивать форму и параметры зависимостей его переменных [14, 21, 36].

Любое экономическое исследование всегда предполагает объединение теории (математической модели) с практикой (экспериментом и статистическими данными). Примерами экономических моделей являются модели предприятия, модели конкурентного равновесия на товарных и финансовых рынках, модели экономического роста и многие другие.

Формализация основных особенностей функционирования экономических объектов позволяет оценить возможные последствия и использовать такие оценки в управлении.

Модель - это материальный или мысленно представляемый объект, который в процессе исследования замещает объект-оригинал, так что его непосредственное изучение дает новые знания об объекте-оригинале.

Под моделированием понимают процесс построения, изучения и применения моделей. Оно является методом познания с помощью объектов-заместителей. Необходимость использования этого метода определяется тем, что многие объекты или проблемы непосредственно исследовать или совсем невозможно, когда объект недосягаем либо реально не существует (будущее состояние экономики), или же это исследование требует много времени и средств.

Основным способом моделирования в экономике является метод ***математического моделирования***, который представляет собой описание основных особенностей реального процесса с помощью системы математических формул. Математическая модель объекта, процесса или явления, как определенная математическая задача, включает две группы элементов:

- характеристики объектов, которые нужно определить (неизвестные величины), – компоненты вектора $Y = (y_i)$);
- характеристики внешних (по отношению к моделируемому объекту) изменяющихся условий – компоненты вектора $X = (x_i)$.

Моделируемая задача может также включать совокупность внутренних параметров объектов A. Условия и параметры, описываемые X и A, рассматриваются как экзогенные, определяемые вне модели, а величины, составляющие вектор Y, – как эндогенные, определенные с помощью модели (рис. 17.1).

11 Зак 1030

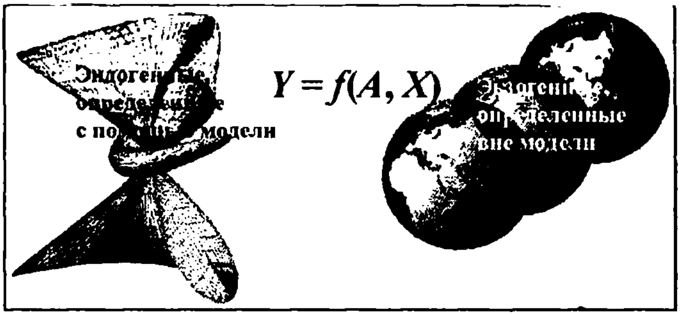


Рис. 17.1

Математическую модель можно интерпретировать как особый преобразователь внешних условий объекта («входа») X в искомой характеристике объекта («выхода») Y.

По способам выражения соотношений между внешними условиями, внутренними параметрами и искомыми характеристиками математические модели делятся на два основных типа: функциональные и структурные.

Основная идея функциональных моделей – познание сущности объекта через важнейшие проявления этой сущности: деятельности, функционирования, поведения. Внутренняя структура при этом не изучается, а информация о структуре не используется. Функциональная модель имитирует поведение объекта так, что, задавая значения «входа» X, можно получать значения «выхода» Y. Построить функциональную модель – значит отыскать оператор D, связывающий X и Y:

$$Y = D(X).$$

Структурные модели отражают внутреннюю организацию объекта: его составные части, внутренние параметры и их связи с «входом» и «выходом».

Наиболее распространены два вида структурной модели:

- все неизвестные выражаются в виде функции от внешних условий и внутренних условий объекта:

$$y_j = f_j(A, x),\ j \in J;$$

- неизвестные определяются совместно, на основе системы отношений этого вида, т. е. уравнений, неравенств и т. д.:

$$\varphi_i(A, x, y) = 0,\ i \in I.$$

Функциональные и структурные модели дополняют друг друга. С одной стороны, при изучении функциональных моделей возникают гипотезы о внутренней структуре объекта, объясняющей его функционирование, и тем самым открывается путь для структурного моделирования. С другой стороны, анализ структурных моделей дает ценную информацию о том, как объект реагирует на изменение внешних условий.

По степени сложности математические модели экономических объектов и явлений подразделяются на несколько основных типов.

К первому типу относятся функциональные модели, которые выражают прямые зависимости между известными (экзогенными) или неизвестными (эндогенными) величинами. Необходимые для построения модели параметры определяются на основе нормативных данных или статистическим способом.

Модели второго типа — это модели, выраженные с помощью систем уравнений относительно эндогенных величин. Они выражают обычно балансовое соотношение между различными экономическими показателями и используются для нахождения сбалансированных плановых решений (например, модель межотраслевого баланса).

Третья группа – модели оптимизационного типа. Основную часть такой модели составляют системы уравнений или неравенств относительно эндогенных величин. При этом необходимо найти такое решение этой системы, которое давало бы оптимальное (т. е. максимальное или минимальное в зависимости от постановки задачи) значение некоторого экономического показателя. Основными моделями этого типа, применяемыми в планировании, являются задачи линейного программирования.

К четвертой группе относятся имитационные модели, которые используются в составе человеко-машинных или игровых имитационных систем.

Пятую группу составляют более сложные образования – такие, как системы и комплексы взаимосвязанных моделей, относящихся к перечисленным выше типам.

Развитие систем моделей позволяет точнее отразить самые различные аспекты планирования и функционирования экономических объектов, в частности, находить оптимальные решения, сочетающие рациональное соотношение между отраслевым и региональным аспектами планирования и управления.

17.2. Основные классы экономико-математических моделей

Экономико-математические модели можно подразделить на классы по ряду признаков: по цели моделирования; по характеристике моделируемого объекта; по исследуемым экономическим процессам и содержательной проблематике; по используемому инструментарию; по характеру описания случайных процессов; по способу отражения фактора времени.

По целевому назначению экономико-математические модели делятся на:

- теоретические, используемые при исследовании общих свойств и закономерностей экономических процессов;
- прикладные, используемые для решения конкретных экономических задач (модели экономического анализа, прогнозирование, управление).

Экономико-математические модели могут предназначаться для исследования разных сторон народного хозяйства (производственно-технологической, территориальной) и его отдельных частей.

При классификации моделей по исследуемым экономическим процессам и содержательной проблематике выделяются модели народного хозяйства в целом и его отдельных подсистем – отраслей, регионов и т. д., комплексы моделей производства, потребления, формирования и распределения доходов, трудовых ресурсов, ценообразования, финансовых связей...

В исследованиях на народнохозяйственном уровне чаще применяются структурные или структурно-функциональные модели, поскольку для планирования и управления большое значение имеют взаимосвязи подсистем. Функциональные модели широко применяются в экономическом регулировании.

Макроэкономические модели описывают экономику страны как единое целое, связывая между собой укрупненные материальные и финансовые показатели: ВНП, потребление, инвестиции, занятость, ставку процента, денежные агрегаты и др.

Микроэкономические модели описывают взаимодействие структурных и функциональных составляющих экономики либо поведение отдельной такой составляющей в рыночной среде. Вследствие разнообразия типов экономических элементов и форм их взаимодействия на рынке микроэкономическое моделирование занимает основную часть экономико-математической теории. Основной целью микроэкономического моделирования является анализ одновременного установления как цен, так и количества произведенных, обмененных и потребленных продуктов в условиях определенных институциональных структур и процессов.

Различают дескриптивные и нормативные модели. Дескриптивные модели объясняют наблюдаемые факты или дают вероятностный прогноз. Нормативные отвечают на вопрос: как это должно быть – т. е. предполагают целенаправленную деятельность. Примером нормативных моделей являются модели оптимального планирования, формализующие тем или иным способом цели экономического развития, возможности и средства их достижения.

Дескриптивный подход применяется для установления статистических закономерностей экономических процессов, изучения вероятных путей развития каких-либо процессов при неизменяющихся условиях или протекающих без внешних воздействий. Примерами дескриптивных моделей являются производственные функции и функции потребительского спроса, построенные на основе обработки статистических данных.

По характеру отражения причинно-следственных связей различают модели жестко детерминистские и модели, учитывающие случайность и неопределенность. В результате накопления опыта использования жестко детерминистских моделей были созданы реальные возможности для успешного применения более совершенной методологии моделирования экономических процессов, учитывающих стохастику и неопределенность; проведения многовариантных расчетов и модельных экспериментов с вариацией конструкции модели и ее исходных данных; изучения устойчивости и надежности получаемых решений; выделения зоны неопределенности; включения в модель резервов; применения приемов, повышающих приспособляемость (адаптивность) экономических решений к вероятным и непредвиденным ситуациям. Получают распространение модели, непосредственно отражающие стохастику и неопределенность экономических процессов и использующие соответствующий математический аппарат: теорию вероятностей и математическую статистику, теорию игр и статистических решений, теорию массового обслуживания, теорию случайных процессов.

По способам отражения фактора времени экономико-математические модели делятся на статические и динамические. В статических моделях все зависимости относятся к одному моменту времени. Динамические модели характеризуют изменение экономических процессов во времени.

В моделировании рыночной экономики особое место занимают равновесные модели. На их основе на микроэкономическом уровне моделируются оптимизационные модели. На макроуровне равновесные статические модели помогают исследовать состояние экономических систем. Экономическая динамика на макроуровне описывается с помощью моделей роста (модель магистрального типа, модель Харрода–Домара, модель Солоу и др.).

17.3. Основные представления о математических моделях для прикладных экономических исследований

Математическая модель – это описание объекта на математическом языке. Для формулировки модели необходимо указать список переменных модели, т. е. нефиксированных заранее величин, описывающих ту или иную сторону моделируемого явления. При этом надо указать, какие значения могут принимать переменные, какие преобразования можно проводить с ними. В некоторых случаях переменные могут принимать только целые неотрицательные значения. В экономико-математических исследованиях часто встречаются переменные, являющиеся функциями других переменных.

После формулировки списка переменных модели, необходимо указать, какие значения переменных могут реализоваться, т. е. указать **множество допустимых значений переменных**. Такое множество часто представляется с помощью системы ограничений на значения переменных. Эти ограничения выделяют среди всевозможных значений переменных **допустимые значения**.

Рассмотрим некоторые основные типы математических моделей, встречающихся в экономико-математических исследованиях.

Линейные статические модели

В них рассматривается конечное число переменных – n. Переменные модели мы обозначим как $x_1, x_2,... x_n$. Предполагается, что эти переменные принимают вещественные значения. Связи в линейной модели в соответствии с ее названием имеют вид системы линейных равенств и неравенств.

$$\sum_{j=1}^{n} a_{ij}x_j = b_i, \qquad i = 1,...,k; \tag{17.1}$$

$$\sum_{j=1}^{n} a_{ij}x_j \le b_i, \qquad i = k+1,...,1, \tag{17.2}$$

где a_{ij} и b_i – заданные числа.

Каждое равенство системы (17.1) можно представить в виде неравенств, т. е. система эквивалентна совокупности неравенств:

$$\sum_{j=1}^{n} a_{ij}x_j \le b_i, \qquad i = 1,...,k,$$

$$-\sum_{j=1}^{n} a_{ij}x_j \le -b_i.$$

Поэтому линейную систему часто представляют в виде:

$$\sum_{j=1}^{n} a_{pj}x_j \le b_p, \qquad p = 1,...,m. \tag{17.3}$$

Здесь числа a_p и b_p не совпадают с коэффициентами систем (17.1) и (17.2).

Модели типа (17.3) наиболее простые среди экономико-математических моделей. Часто их записывают в сокращенном векторном виде. Для этого вместо n переменных в модели используют единственную – вектор x, имеющий n составляющих:

$$\vec{x} = (x_1, x_2, ..., x_n).$$

Чтобы подчеркнуть векторную природу переменной $\vec{x}$, применяют запись вида $\vec{x} \in E^n$, где E^n – n-мерное евклидово пространство. Принадлежность вектора $\vec{x}$ пространству E^n означает, что вектор $\vec{x}$ имеет n вещественных составляющих x_j, причем векторы $\vec{x} \in E^n$ и $\vec{y} \in E^n$, можно складывать по правилу:

$$(\vec{x} + \vec{y}) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n),$$

и умножать на вещественное число λ:

$$\lambda \vec{x} = (\lambda x_1, \lambda x_2, \ldots, \lambda x_n).$$

Кроме того, определено произведение двух векторов:

$$(\vec{x}, \vec{y}) = \sum_{j=1}^{n} x_j y_j.$$

На основе понятия скалярного произведения модель (17.3) можно представить в сокращенном виде:

$$(a_p, x) \le b_p, p = 1, \ldots, m, \tag{17.4}$$

где $a_p = (a_{p1}, a_{p2}, \ldots, a_{pn}) \in E^n$ – векторы, состоящие из коэффициентов системы (17.3).

Обычно векторная запись имеет еще более сокращенный вид. Для этого из коэффициентов a_{pj} системы (17.3) образуют прямоугольную матрицу:

$$A = \begin{bmatrix} a_{11} & a_{12} & . & a_{1n} \\ a_{21} & a_{22} & . & a_{2n} \\ . & . & . & . \\ a_{m1} & a_{m2} & . & a_{mn} \end{bmatrix},$$

а из коэффициентов b_p составляют вектор

$$b = (b_1, b_2, \ldots, b_m) \in E^n.$$

Тогда соотношение (17.3) переписывается в виде:

$$Ax \leqq b. \tag{17.5}$$

В записи (17.5) использован знак неравенства $\leqq$. Для двух векторов a и b, принадлежащих пространству E^n, запись $a \leqq b$ означает, что выполняются все неравенства

$$a_i \le b_i, i=1,\ldots,m, \tag{17.6}$$

причем все они могут одновременно быть равенствами. Запись $\le$ для векторов означает, что все неравенства (17.6) одновременно в равенство обращаться не должны.

Множество допустимых значений переменной x, которое мы обозначим через X, для модели (17.3) является многогранным. Можно сказать, что рассматриваемые линейные статические модели имеют общий вид:

$$x \in X \in E^n, \tag{17.7}$$

где X – многогранное множество.

Для того чтобы описать конкретное множество X, его представляют в одном из видов (17.3), (17.4) или (17.5).

Нелинейные статические модели

Пусть $x \in E^n$. Рассмотрим представление линейной модели в виде (17.4). Вместо линейных связей

$$(a_p, x) \leq b_p \; (p = 1,..., m)$$

на переменные модели наложим нелинейные связи:

$$g_p(x) \leq 0, p = 1,..., m. \tag{17.8}$$

Модель (17.8) – общее представление нелинейной статической модели с конечным числом ограничений.

Так же как и в случае линейных систем, модель (17.8) можно представить в виде (17.7), только теперь множество X уже не является многогранным.

Среди нелинейных статических моделей, используемых в экономико-математических моделях, наиболее важную роль играют модели, для которых множество допустимых значений X является **выпуклым**, точнее говоря, вместе с любыми двумя векторами $x^* \in X$ и $x^{**} \in X$ этому множеству принадлежит весь отрезок, равный

$$x = ax^* + (1 - a)\, x^{**}, \text{ где } a \in (0; 1).$$

Для рассмотренных линейных статических моделей множество X всегда выпукло.

Для нелинейных моделей это не всегда так.

Модели с выпуклым множеством X более удобны для исследования, поэтому модель стараются сформулировать так, чтобы множество X было выпукло.

Рассмотрим вопрос об условиях, достаточных для того, чтобы множество допустимых значений X, описываемое соотношением (17.8), было выпуклым. Этот вопрос решается на основе введения понятия выпуклой функции. Функция $g(x)$, где $x \in E^n$, называется выпуклой вниз, если для любых значений x^* и x^{**} и при любом числе a, изменяющимся от **0** до **1**, выполнено неравенство:

$$g(ax^* + (1 - a)\, x^{**}) \leq ag(x^*) + (1 - a)\, g(x^{**}). \tag{17.9}$$

Если же для некоторой функции выполнено обратное условие, то ее называют вогнутой (или выпуклой вниз).

В том случае, когда все нелинейные функции $g_p(x)$, ($p = 1,..., m$) выпуклы, множество x тоже выпукло.

Понятие выпуклости функции играет важную роль в экономико-математическом моделировании, поскольку позволяет получить интересные качественные результаты.

Как в линейных, так и в нелинейных экономико-математических моделях множество X обычно содержит более одного допустимого вектора. Это означает, что имеется некоторая свобода выбора: соотношение модели не определяет единственным образом то, что произойдет с изучаемой экономической системой. Это позволяет ввести понятие внешнего воздействия или управления, определяющего судьбу моделируемой системы.

В статических моделях типа (17.3) или (17.8) управлением является вектор $x \in E^n$, причем необходимо выбирать такие его значения, которые удовлетворяют всем ограничениям (17.3) или (17.8).

При исследовании некоторых моделей типа (17.3) или (17.8) фактически надо выбирать не n значений переменных, а меньшее число. Например, если в системе (17.3) часть неравенств объединяется в равенство (17.1), то, разрешив эту систему равенств, часть переменных можно выразить через другие, т. е. представить модель (17.3) в виде:

$$x^{**} = Bx^{*};$$
$$Ax^{*} \leq b, \qquad (17.10)$$

где A и B – матрицы, а b – вектор. Тогда управлениями будут переменные, вошедшие в вектор x^{*}, а переменные из вектора x^{**} будут определяться внутри модели.

Представление моделей вида (17.3) или (17.5) в виде (17.10) позволяет ввести понятие экзогенных и эндогенных переменных. В модели (17.10) экзогенными (т. е. происходящими извне) являются переменные, составляющие вектор x^{*}, а эндогенными – переменные, вошедшие в вектор x^{**}.

Общий вид математической модели

Процесс построения моделей имеет следующие общие черты. Прежде всего устанавливается, какие переменные рассматриваются в модели: либо вещественные векторы, либо целочисленные переменные, либо функции времени. В результате записывается пространство переменных моделей. Затем формулируются связи, накладываемые на переменные модели. Эти связи позволяют выделить среди всевозможных сочётаний переменных те, которые соответствуют нашим представлениям об изучаемой системе. В процессе построения математической модели постепенно формулируются соотношения между переменными, делающие множества допустимых сочетаний все уже. Модель в общем виде можно представить:

$$x \in X \in Q_x,$$

где: x – переменные модели, X – множество допустимых значений переменных, Q_x – пространство переменных.

В дальнейшем изложении мы рассмотрим два подхода к моделированию экономических систем. Первый из них базируется на теории межотраслевого анализа, и его основы в статике изложены в главе 18. Второй подход описывает методы моделирования на основе микроэкономической теории – моделирование выпуска, затрат, спроса, равновесия и анализ одновременного установления как цен, так и количеств произведенных, обменных и потребленных продуктов. Объем книги не позволяет рассматривать методы моделирования динамики и неопределенности.

Глава 18. Межотраслевой анализ

18.1. Метод межотраслевого анализа

Создатель теории межотраслевого анализа экономических систем – лауреат Нобелевской премии по экономике Василий Васильевич Леонтьев [20]. По определению академика А. Г. Гранберга, сущность и сила межотраслевого анализа В. В. Леонтьева состоит в соединении теории функционирования экономических систем, метода математического моделирования, приемов систематизации и обработки экономической информации. Типичный продукт и вместе с тем предмет межотраслевого анализа – межотраслевой баланс экономики. Это и система показателей, характеризующих соотношения, структуру, связи экономики, и математическая модель, позволяющая не только изучать взаимовлияние множества экономических величин, но и конструировать возможные (альтернативные) состояния экономики.

Экономическая система, для исследования которой применяется метод межотраслевого анализа, может быть большой, как народное хозяйство страны или даже вся мировая экономика, или малой, такой как экономика региона или даже одного предприятия.

В любом случае подход в основном один и тот же. Структура производственного процесса в каждом секторе представляется определенным вектором структурных коэффициентов, который количественно характеризует связь между затратами этого сектора и результатами его деятельности. Взаимозависимость между секторами рассматриваемой экономики описывается системой линейных уравнений, выражающих балансы между совокупными затратами и агрегированным выпуском каждого продукта и услуг, производимых и используемых в течение одного или нескольких промежутков времени.

Соответственно, технологическая структура системы в целом может быть представлена матрицей технологических коэффициентов «затраты–выпуск» всех ее секторов. В то же время эта матрица содержит множество параметров, на которых основываются балансовые соотношения.

Таблица межотраслевого баланса

Таблица межотраслевого баланса описывает потоки товаров и услуг между всеми секторами народного хозяйства в течение фиксированного периода времени, например года. Упрощенный пример такой таблицы, отражающий трехсекторную экономику приведен в табл. 18.1.

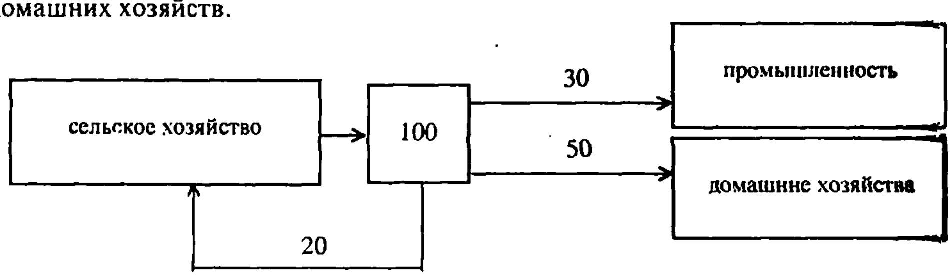
Таблица 18.1

Из ⇓	**В ⇒**	**Сектор 1 Сельское хозяйство**	**Сектор 2 Промышленность**	**Сектор 3 Домашнее хозяйство**	**Общий выпуск**
Сектор 1 Сельское хозяйство		20	30	50	100 т зерна

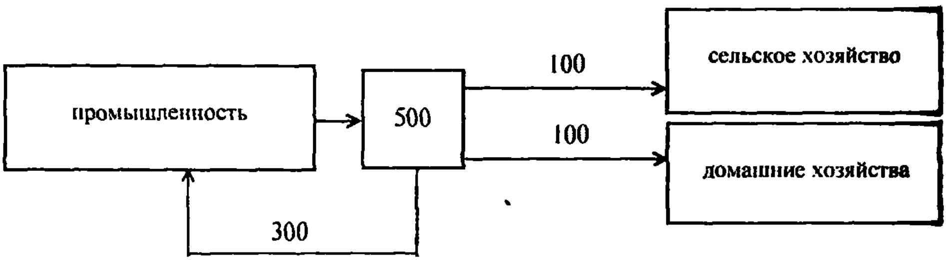
Таблица 18.1 (окончание)

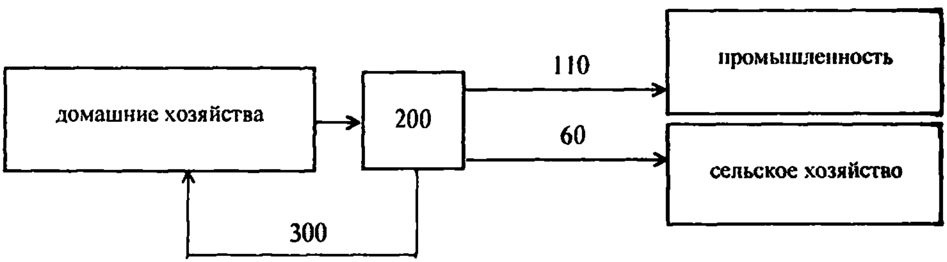
Из ⇓	В ⇒	Сектор 1 Сельское хозяйство	Сектор 2 Промышленность	Сектор 3 Домашнее хозяйство	Общий выпуск
Сектор 2 Промышленность		100	300	100	500 т стали
Сектор 3 Домашнее хозяйство		60	110	30	200 человеко лет труда

Этими тремя секторами являются сельское хозяйство, совокупный годовой продукт которого составляет 100 т зерна, промышленность, производящая 500 т стали, и сектор домашних хозяйств, который предлагает 300 человеко – лет труда. Девять чисел, составляющих основное содержание таблицы, характеризуют межсекторные потоки. Из 100 т зерна, производимых сельскохозяйственным сектором, 20 т используются внутри него самого, 30 поставляются промышленности и поглощаются ею в качестве одного из ресурсов, и 50 т потребляются в секторе домашних хозяйств.



Вторая и третья строка табл.18.1 аналогичным образом описывают распределение продукции двух других секторов.





В столбцах числа описывают структуру затрат соответствующего сектора. Чтобы произвести 100 т своего совокупного продукта, сельское хозяйство потребляет 20 т своего собственного продукта, 100 т продукции промышленности и 60 человеко – лет труда из сектора домашних хозяйств.

Чтобы произвести 500 т своего совокупного продукта, сектор промышленности должен получить и использовать 30 т сельскохозяйственной и 300 т собственной продукции, а также 180 человеко-лет труда из сектора домашних хозяйств.

В свою очередь, сектор домашних хозяйств расходует доход, полученный за предложение 200 человеко-лет труда, для оплаты потребления 50 т сельскохозяйственных и 100 т промышленных товаров, а также 30 человеко-лет непосредственных затрат труда.

Предполагается, что все числа в табл. 18.1 представляют количества или, по крайней мере, физические индексы количеств определенных товаров или услуг. Более детализированная таблица межотраслевого баланса позволяет получить более определенную характеристику каждого отдельного числа.

Балансовые уравнения

Всю производимую отраслями продукцию удобно разделить на две части: промежуточный продукт и конечный продукт.

Промежуточный продукт – это та часть совокупного продукта, которой производители обмениваются между собой или используют для собственных нужд.

Конечный продукт – это продукция, предназначенная для потребителей.

Сектор конечного спроса – это сектор, где потребляется конечный продукт – домашние хозяйства, экспорт, правительственные закупки.

Применение данных показателей представлено в табл. 18.2, сделанной на основе табл. 18.1.

Таблица 18.2

Отрасль	Межотраслевые потоки в экономической системе			Общий выпуск
	1	2	Конечный спрос	
1	промежуточный продукт из 1 → в 1	промежуточный продукт из 1 → в 2	конечный продукт из 1	Объем выпуска в 1
2	промежуточный продукт из 2 → в 1	промежуточный продукт из 2 → в 2	конечный продукт из 2	Объем выпуска в 2
Конечный спрос	промежуточный продукт из 3 → в 1	промежуточный продукт из 3 → в 2	конечный продукт из 3	Объем выпуска в 3
	В столбцах – структура затрат отрасли			

Введем обозначения: U – общий выпуск; v – промежуточный продукт; k – конечный продукт. Представим табл. 18.2 в принятых обозначениях в табл. 18.3.

Таблица 18.3

Отрасль	Межотраслевые потоки в экономической системе			Общий выпуск
	1	2	Конечный спрос	
1	v_{11}	v_{12}	k_{13}	U_1
2	v_{21}	v_{22}	k_{23}	U_2
Конечный спрос	v_{31}	v_{32}	k_{33}	U_3

Объем продукции данной отрасли равняется сумме потоков продукции этой отрасли в другие отрасли, продукции, потребляемой в данной отрасли, и конечного продукта данной отрасли. Следовательно, баланс между совокупным выпуском и суммарными затратами продукции каждого сектора, показанный в нашем примере, может быть описан следующей системой уравнений:

$$U_1 = v_{11} + v_{12} + k_{13};$$
$$U_2 = v_{21} + v_{22} + k_{23};$$
$$U_3 = v_{31} + v_{32} + k_{33}.$$

Эти уравнения называются балансовыми уравнениями производства. Для экономики с n отраслями балансовые уравнения будут иметь вид:

$$U_1 = v_{11} + v_{12} + \ldots + v_{1n} + k_1;$$
$$U_2 = v_{21} + v_{22} + \ldots + v_{2n} + k_2;$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$U_n = v_{n1} + v_{n2} + \ldots + v_{nn} + k_n.$$

Изложенная модель получила название модели «затраты–выпуск», или модели межотраслевого анализа (англ.: Input – Output Analysis).

Технологические коэффициенты

Объем выпуска сектора i, используемого сектором j при производстве единицы его совокупного выпуска j, обозначается символом a_{ij} и называется технологическим коэффициентом затрат продукта i в секторе j:

$$a_{ij} = \frac{v_{ij}}{U_j}.$$

Представим вычисление технологических коэффициентов для примера трехсекторной экономики в табличном виде (табл. 18.4).

Таблица 18.4

Отрасль	Межотраслевые потоки в экономической системе			Общий выпуск
	1	2	Конечный спрос	
1	$a_{11} = v_{11}/U_1 = 0{,}2$	$a_{12} = v_{12}/U_2 = 0{,}06$	k_{13}	U_1

Таблица 18.4 (окончание)

Отрасль	Межотраслевые потоки в экономической системе			Общий выпуск
	1	2	Конечный спрос	
2	$a_{21} = v_{21}/U_1 = 1{,}0$	$a_{22} = v_{22}/U_2 = 0{,}60$	k_{23}	U_2
Конечный спрос	$a_{31} = v_{31}/U_1 = 0{,}6$	$a_{32} = v_{32}/U_2 = 0{,}22$	k_{33}	U_3

Множество всех коэффициентов затрат всех секторов рассматриваемой экономики, представленной в форме прямоугольной таблицы, соответствующей таблице межотраслевого баланса, называется структурной матрицей этой экономики. Технологические коэффициенты образуют следующую квадратную матрицу n-го порядка:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ . & . & . & . \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}.$$

На практике структурные матрицы обычно вычисляются на основе межотраслевого баланса в стоимостном выражении. Но во всех случаях коэффициенты затрат должны интерпретироваться как отношения двух количеств, измеренных в физических единицах.

Решение системы балансовых уравнений

Из определения технологических коэффициентов вытекает:

$$v_{ij} = a_{ij}\,U_j.$$

Следовательно, балансовые уравнения можно записать так:

$$U_1 = a_{11}U_1 + a_{12}U_2 + \dots + a_{1n}U_n + k_1;$$
$$U_2 = a_{21}U_1 + a_{22}U_2 + \dots + a_{2n}U_n + k_2;$$
$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$
$$U_n = a_{n1}U_1 + a_{n2}U_2 + \dots + a_{nn}U_n + k_n.$$

Или в матричной форме:

$$\begin{bmatrix} U_1 \\ U_2 \\ \dots \\ U_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ . & . & . & . \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \cdot \begin{bmatrix} U_1 \\ U_2 \\ \dots \\ U_n \end{bmatrix} + \begin{bmatrix} k_1 \\ k_2 \\ \dots \\ k_n \end{bmatrix}.$$

Введя обозначения:

$$\begin{bmatrix} U_1 \\ U_2 \\ \dots \\ U_n \end{bmatrix} = U, \quad \begin{bmatrix} k_1 \\ k_2 \\ \dots \\ k_n \end{bmatrix} = K,$$

получим матричное уравнение

$$U = A\,U + K,$$

откуда:

$$(I - A)\,U = K$$

Умножим полученное уравнение на $(I - A)^{-1}$:

$$(I - A)^{-1}(I - A)\,U = (I - A)^{-1} K,$$

откуда:

$$U = (I - A)^{-1} K.$$

Планирование материального производства начинается с определения размеров и структуры общественного продукта. Решение матричного уравнения позволяет определить плановый объем производства отдельных продуктов таким образом, чтобы получить необходимые количества конечных продуктов. Полученное выражение позволяет быстро разработать разные варианты плана материального производства в соответствии с вариантами заданного конечного общественного продукта.

Введем обозначение:

$$(I - A)^{-1} = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ . & . & . & . \\ A_{n1} & A_{n2} & \dots & A_{nn} \end{bmatrix}.$$

Тогда можем записать:

$$\begin{bmatrix} U_1 \\ U_2 \\ \dots \\ U_n \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ . & . & . & . \\ A_{n1} & A_{n2} & \dots & A_{nn} \end{bmatrix} \cdot \begin{bmatrix} k_1 \\ k_2 \\ \dots \\ k_n \end{bmatrix}.$$

Умножив, получим:

$$\begin{bmatrix} U_1 \\ U_2 \\ \dots \\ U_n \end{bmatrix} = \begin{bmatrix} A_{11}k_1 & A_{12}k_2 & \dots & A_{1n}k_n \\ A_{21}k_1 & A_{22}k_2 & \dots & A_{2n}k_n \\ . & . & . & . \\ A_{n1}k_1 & A_{n2}k_2 & \dots & A_{nn}k_n \end{bmatrix} \cdot \begin{bmatrix} k_1 \\ k_2 \\ \dots \\ k_n \end{bmatrix},$$

откуда:

$$U_1 = A_{11}k_1 + A_{12}k_2 + \dots + A_{1n}k_n,$$
$$U_2 = A_{21}k_1 + A_{22}k_2 + \dots + A_{2n}k_n,$$
$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$
$$U_n = A_{n1}k_1 + A_{n2}k_2 + \dots + A_{nn}k_n.$$

Данное уравнение показывает, что элементы матрицы $(I - A)^{-1}$ есть величины, определяющие количественные соотношения между конечными продуктами всех отраслей и продуктами каждой отрасли. Постоянные A_{ij} показывают, насколько увеличится выпуск U_i сектора i при увеличении k_j, т. е. количества товара j, потребляемого домашними хозяйствами (или любым другим потребителем этого сектора) на единицу.

В качестве примера рассмотрим нашу трехсекторную экономику, где матрица технологических коэффициентов равна:

$$A=\begin{bmatrix}0{,}2 & 0{,}06\\ 1 & 0{,}6\end{bmatrix}.$$

Примем два варианта потребности в конечном продукте:

1. $k_1 = 50$; $k_2 = 100$.
2. $k_1 = 75$; $k_2 = 110$.

Тогда

$$(I-A)=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}-\begin{bmatrix}0{,}2 & 0{,}06\\ 1 & 0{,}6\end{bmatrix}=\begin{bmatrix}0{,}8 & -0{,}06\\ -1 & 0{,}4\end{bmatrix};$$

$$(I-A)^{-1}=\begin{bmatrix}1{,}538 & 0{,}231\\ 3{,}846 & 3{,}076\end{bmatrix};$$

$$(I-K)^{-1}\cdot K=\begin{bmatrix}1{,}538 & 0{,}231\\ 3{,}846 & 3{,}076\end{bmatrix}\cdot\begin{bmatrix}50\\ 100\end{bmatrix}=\begin{bmatrix}100\\ 500\end{bmatrix}.$$

Для второго варианта имеем:

$$(I-A)^{-1}\cdot K=\begin{bmatrix}1{,}538 & 0{,}231\\ 3{,}846 & 3{,}076\end{bmatrix}\cdot\begin{bmatrix}75\\ 110\end{bmatrix}=\begin{bmatrix}140{,}77\\ 629{,}92\end{bmatrix}.$$

Кроме приведенной задачи, с помощью модели межотраслевого баланса решаются и другие задачи: определение занятости в производстве; определение совокупных затрат труда; распределение совокупного общественного продукта; определение цен продуктов; определение размеров капитальных вложений. Рассмотрим одну из них.

Определение цен продуктов

Цены в системе межотраслевых связей определяются из системы уравнений, каждое из которых устанавливает, что цена единицы выпуска соответствующего производственного сектора должна быть равна совокупным издержкам в процессе производства этой продукции (в расчете на единицу выпуска). В эти издержки входит не только оплата затрачиваемых ресурсов, но и добавленная стоимость, которая представляет собой в основном платежи секторам конечного спроса (d_i). Эти платежи состоят обычно из зарплаты, процента на капитал, предпринимательской прибыли, налогов, выплачиваемых правительству и другим секторам конечного спроса.

Обозначим через p_i цену единицы i-го продукта. Тогда балансовые уравнения можно записать так:

$$U_1 p_1 = a_{11}U_1p_1 + a_{12}U_2p_2 + \ldots + a_{1n}U_n p_n + d_1U_1;$$
$$U_2 p_2 = a_{21}U_1p_1 + a_{22}U_2p_2 + \ldots + a_{2n}U_n p_n + d_2U_2;$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$U_np_n = a_{n1}U_1p_1 + a_{n2}U_2p_2 + \ldots + a_{nn}U_np_n + d_nU_n.$$

Сократив в обеих частях уравнений U_i, получим систему уравнений:

$$p_1 = a_{11}p_1 + a_{12}p_2 + \ldots + a_{1n}p_n + d_1;$$
$$p_2 = a_{21}p_1 + a_{22}p_2 + \ldots + a_{2n}p_n + d_2;$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$p_n = a_{n1}p_1 + a_{n2}p_2 + \ldots + a_{nn}p_n + d_n$$

или в матричной форме:

$$p = A^T p + D,$$

где A^T – транспонированная матрица технологических коэффициентов, а p и D вектора цен и платежей секторам конечного спроса соответственно:

$$p=\begin{bmatrix} p_1 \\ p_2 \\ \dots \\ p_n \end{bmatrix}, \quad D=\begin{bmatrix} d_1 \\ d_2 \\ \dots \\ d_n \end{bmatrix}.$$

Матричное уравнение можно представить в виде:

$$(I - A^T)p = D,$$

откуда получим в окончательном виде:

$$p = (I - A^T)^{-1} D.$$

Данное уравнение позволяет определить соответствующую цену продукта отрасли. Элементы матрицы $(I - A^T)^{-1}$ измеряют зависимость цены p_j продукции сектора j добавленной стоимости d_i, полученной в секторе i в расчете на единицу продукции этого сектора.

В применявшемся выше примере добавленная стоимость, выплаченная в сельском хозяйстве и промышленности (т. е. зарплата), в расчете на единицу выпуска составляет 0,6 и 0,22 соответственно. Транспонированная матрица технологических коэффициентов равна:

$$A^T=\begin{bmatrix} 0{,}2 & 1 \\ 0{,}06 & 0{,}6 \end{bmatrix}.$$

Далее рассчитываем:

$$(I - A^T)=\begin{bmatrix} 0{,}8 & -1 \\ -0{,}06 & 0{,}4 \end{bmatrix}, \quad (I - A^T)^{-1}=\begin{bmatrix} 1{,}538 & 3{,}846 \\ 0{,}231 & 3{,}077 \end{bmatrix}.$$

Тогда цены равны:

$$p=(I - A^T)^{-1}\cdot D=\begin{bmatrix} 1{,}538 & 3{,}846 \\ 0{,}231 & 3{,}077 \end{bmatrix}\cdot\begin{bmatrix} 0{,}6 \\ 0{,}22 \end{bmatrix}=\begin{bmatrix} 1{,}769 \\ 0{,}815 \end{bmatrix},$$

т. е. цены на сельскохозяйственную и промышленную продукцию, используемые при расчете стоимостных показателей межотраслевых потоков.

Внутреннее единство стоимостных и физических взаимосвязей в рамках открытой системы межотраслевых связей подтверждается следующим тождеством:

$$v_1 d_1 + v_2 d_2 + \dots + v_n d_n = k_1 p_1 + k_2 p_2 + \dots + k_n p_n.$$

В левой части соотношения находится общая сумма добавленных стоимостей, выплаченная секторами системы секторам конечного спроса; в правой части – сумма стоимостей продуктов, доставленных всеми секторами секторам конечного спроса.

18.2. Разработка плана предприятия методом межотраслевого анализа

Метод межотраслевого анализа применим и для такой экономической системы, как предприятие. В этом случае место отраслей займут цеха, а конечного продукта – товарная продукция предприятия. Допустим, что предприятие состоит из n производственных цехов, производящих однородные продукты 1, 2, ..., n. Основа технико-экономического плана промышленного предприятия есть система технико-экономических норм. В эту систему входят:

1. Нормы затрат продуктов собственного производства в отдельных цехах; эти нормы можно представить в виде матрицы:

$$H_Z = [z_{ij}] \ (i, j = 1, 2, ..., n).$$

2. Нормы расхода сырья, основных материалов, топлива и электроэнергии на единицу продукта, произведенного в соответствующем цехе; эти нормы можно записать в виде матрицы:

$$H_S = [s_{ij}] \ (i = 1,2, ..., g; j = 1,2, ..., n).$$

3. Нормы времени работы машин и оборудования; эти нормы можно представить в виде матрицы:

$$H_{BM} = [m_{ij}] \ (i = 1,2, ..., r; j = 1,2, ..., n).$$

4. Нормы, определяющие время работы отдельных групп персонала, необходимое для производства единицы продукта в соответствующем цехе; эти нормы можно представить в виде матрицы:

$$H_{BL} = [l_{ij}] \ (i = 1,2, ..., f; j = 1,2, ..., n).$$

Обозначим через U_i совокупную продукцию i-го цеха, а через k_i – товарную продукцию этого цеха, т. е. ту часть совокупной продукции, которая остается после обеспечения производственных цехов и предназначается для сбыта. Поскольку затраты продукции i-го цеха на единицы продукта j-го цеха определяются по матрице $H_Z = [z'_{ij}]$ $(i, j = 1, 2, ..., n)$, можно записать следующую систему уравнений:

$$U_1 = z_{11}U_1 + z_{12}U_2 + ... + z_{1n}U_n + k_1;$$
$$U_2 = z_{21}U_1 + z_{22}U_2 + ... + z_{2n}U_n + k_2;$$
$$\dots\dots\dots\dots\dots\dots\dots\dots$$
$$U_n = z_{n1}U_1 + z_{n2}U_2 + ... + z_{nn}U_n + k_n$$

или в матричной форме:

$$U = H_Z U + K.$$

Решение данного уравнения есть матрица:

$$U = (I - H_Z)^{-1} K.$$

Отсюда следует, что матрица продукции есть произведение матрицы норм полных затрат продуктов, произведенных отдельными цехами, и матрицы плановой товарной продукции предприятия.

Матрица H_S норм расхода сырья, материалов, топлива и электроэнергии есть основа плана материально – технического снабжения. Из матрицы H_S следует, что расходы отдельных видов сырья и материалов составляют:

$$R_1 = s_{11}U_1 + s_{12}U_2 + ... + s_{1n}U_n;$$
$$R_2 = s_{21}U_1 + s_{22}U_2 + ... + s_{2n}U_n;$$
$$\dots\dots\dots\dots\dots\dots\dots\dots$$
$$R_n = s_{n1}U_1 + s_{n2}U_2 + ... + s_{nn}U_n$$

или в матричной форме:

$$R = H_S U.$$

Подставляя выражение для определения матрицы товарной продукции, получаем:

$$R = H_S (I - H_Z)^{-1} K.$$

Элементы произведения $S(I - Hz)^{-1}$ можно назвать коэффициентами полных затрат сырья и материалов. Матрицу потребности в сырье и материалах можно получить, умножив матрицу коэффициентов полных затрат сырья и материалов на матрицу товарной продукции.

Матрица H_{BM} – основа плана использования машин и оборудования. Использование машин и оборудования в производстве составляет:

$$U_1 = m_{11}U_1 + m_{12}U_2 + \ldots + m_{1n}U_n;$$
$$U_2 = m_{21}U_1 + m_{22}U_2 + \ldots + m_{2n}U_n;$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$U_n = m_{n1}U_1 + m_{n2}U_2 + \ldots + m_{nn}U_n$$

или в матричной форме:

$$U = H_{BM}\, U = H_{BM}\,(I - Hz)^{-1}\, K.$$

Элементы матрицы $H_{BM}\,(I - Hz)^{-1}$ называются коэффициентами полного использования машин и оборудования. Матрицу плана использования машин и оборудования можно получить, умножив матрицу коэффициентов полного использования машин и оборудования на матрицу товарной продукции.

Матрица H_{BL} – основа плана по труду. Матрица рабочей силы есть:

$$L = H_{BL}\, U = H_{BL}\,(I - Hz)^{-1}\, K.$$

Элементы произведения $H_{BL}\,(I - Hz)^{-1}$ называются коэффициентами полных затрат рабочей силы. Матрица плановых затрат рабочей силы представляет собой произведение матрицы коэффициентов полных затрат рабочей силы и матрицы товарной продукции.

Матричная форма технико-экономического плана в значительной мере упрощает планирование и уменьшает его трудоемкость: она позволяет быстро разработать различные варианты технико-экономического плана.

Пример. Промышленное предприятие состоит из шести цехов. Продукт одного цеха перерабатывается в следующих цехах (производственные услуги и полуфабрикаты). Продукты цехов также могут быть товарной продукцией (продукт основного цеха, услуги автотранспортного цеха, ремонтного цеха и др.). Плановая деятельность предприятия основывается на следующей системе технико-экономических норм (табл. 18.5).

Таблица 18.5

Наименование показателей	Единица измерения	Цеха				Матрица
		1	2	3	4	
1. Производственные услуги и полуфабрикаты						Hz
1	т	0,1	0,1	0,3	0,1	
2	м³	0,1	0,2	0,4	0,2	
3	т	0,2	0,2	0,5	0,15	
4	т	0,1	0,7	0,2	0,3	

Таблица 18.5 (окончание)

Наименование показателей	Единица измерения	Цеха				Матрица
		1	2	3	4	
2. Сырье, материалы, электроэнергия:						H_S
s_1	т	0,1	0	0	0,2	
s_2	т	0,15	0,05	0	0,07	
s_3	т	0,2	0,03	0,06	0,1	
s_4	т	0	0	0,5	0,3	
s_5	т	0,11	0,07	0,14	0,08	
3. Труд:						H_{BL}
персонал категории 8	чв-ч	1,2	0,3	0,6	0,4	
персонал категории 9	чв-ч	0,9	0,2	0,4	0,5	
персонал категории 10	чв-ч	0,6	0,1	0,8	0,7	

Предприятие разрабатывает проект производственного плана в двух вариантах:

1) для планового задания по товарной продукции цехов соответственно:

$k_1 = 0$; $k_2 = 100$; $k_3 = 200$; $k_4 = 30\,000$;

2) для планового задания по товарной продукции цехов соответственно:

$k_1 = 400$; $k_2 = 600$; $k_3 = 800$; $k_4 = 30\,000$.

Необходимо разработать два варианта плана выпуска продукции и потребности в сырье, материалах и электроэнергии, а также два плана по труду.

Матрица норм использования производственных услуг и потребления продуктов, произведенных соответствующими цехами, есть:

$$H_Z = \begin{bmatrix} 0{,}1 & 0{,}1 & 0{,}3 & 0{,}1 \\ 0{,}1 & 0{,}2 & 0{,}4 & 0{,}1 \\ 0{,}2 & 0{,}2 & 0{,}5 & 0{,}1 \\ 0{,}1 & 0{,}7 & 0{,}2 & 0{,}1 \end{bmatrix}.$$

Матрица $I - H_Z$ будет равна:

$$I - H_Z = \begin{bmatrix} 0{,}9 & -0{,}1 & -0{,}3 & -0{,}1 \\ -0{,}1 & 0{,}8 & -0{,}4 & -0{,}1 \\ -0{,}2 & -0{,}2 & 0{,}5 & -0{,}1 \\ -0{,}1 & -0{,}7 & -0{,}2 & 0{,}9 \end{bmatrix},$$

а обратная ей матрица:

$$(I - H_Z)^{-1} = \begin{bmatrix} 1{,}929 & 1{,}477 & 2{,}606 & 0{,}668 \\ 1{,}204 & 3{,}020 & 3{,}481 & 0{,}856 \\ 1{,}552 & 2{,}408 & 5{,}268 & 1{,}025 \\ 1{,}496 & 3{,}048 & 4{,}167 & 2{,}079 \end{bmatrix}.$$

Согласно формуле $U = (I - Z)^{-1} K$, первый вариант матрицы выпуска продукции будет равен:

$$U = \begin{bmatrix} 10{,}6 & 19{,}1 & 26{,}7 & 12{,}7 \\ 14{,}8 & 30{,}8 & 41{,}4 & 19{,}8 \\ 16{,}5 & 32{,}8 & 46{,}9 & 21{,}8 \\ 21{,}1 & 42{,}9 & 58{,}7 & 29{,}3 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 100 \\ 200 \\ 30000 \end{bmatrix} = \begin{bmatrix} 388722 \\ 605410 \\ 666305 \\ 894172 \end{bmatrix}.$$

То есть продукция цехов по первому варианту плана должна составить:

$U_1 = 388\,722$ т, $U_2 = 605\,410$ т, $U_3 = 666\,305$ т, $U_4 = 894\,172$ т.

Второй вариант матрицы плана выпуска продукции равен:

$$U = \begin{bmatrix} 1{,}929 & 1{,}477 & 2{,}606 & 0{,}668 \\ 1{,}204 & 3{,}020 & 3{,}481 & 0{,}856 \\ 1{,}552 & 2{,}408 & 5{,}268 & 1{,}025 \\ 1{,}496 & 3{,}048 & 4{,}167 & 2{,}079 \end{bmatrix} \cdot \begin{bmatrix} 400 \\ 600 \\ 800 \\ 30000 \end{bmatrix} = \begin{bmatrix} 418582 \\ 651616 \\ 717457 \\ 959258 \end{bmatrix},$$

т. е. продукция цехов должна составить:

$U_1 = 418\,582$ т, $U_2 = 651\,616$ т, $U_3 = 717\,457$ т, $U_4 = 959\,258$ т.

Объем поставок i-го продукта j-му цеху рассчитывается по формуле:

$$v_{ij} = z_{ij}\, U_j \; (i, j = 1, 2, 3, 4).$$

Рассчитав v_{ij}, получим два варианта плана выпуска продукции в натуральном выражении.

Таблица 18.6. Первый вариант

Продукт	Единица измерения	Объем выпуска продукции	Поставки в цеха				Товарная продукция
			1	2	3	4	
1	тыс. т	388,7	38,9	60,5	199,9	89,4	0
2	тыс. м^3	605,4	38,9	121,1	266,5	178,8	0,10
3	тыс. т	666,3	77,7	121,1	333,1	134,1	0,20
4	тыс. т	894,1	38,8	423,8	133,2	268,2	30,0

Таблица 18.7. Второй вариант

Продукт	Единица измерения	Объем выпуска продукции	Поставки в цеха				Товарная продукция
			1	2	3	4	
1	тыс. т	418,5	41,8	65,1	215,2	95,9	0,4
2	тыс. м^3	651,6	41,8	130,3	286,9	191,8	0,6
3	тыс. т	717,4	83,7	130,3	358,7	143,9	0,8
4	тыс. т	959,2	41,9	456,1	143,5	287,8	30,0

Основа разработки плана затрат сырья, материалов и электроэнергии – это матрица:

$$H_S = \begin{bmatrix} 0{,}1 & 0 & 0 & 0{,}2 \\ 0{,}15 & 0{,}05 & 0 & 0{,}07 \\ 0{,}2 & 0{,}03 & 0{,}06 & 0{,}1 \\ 0 & 0 & 0{,}5 & 0{,}3 \\ 0{,}11 & 0{,}07 & 0{,}14 & 0{,}08 \end{bmatrix}.$$

Матрица материальных затрат в первом варианте плана выпуска продукции это матрица:

$$R_1 = \begin{bmatrix} 0{,}1 & 0 & 0 & 0{,}2 \\ 0{,}15 & 0{,}05 & 0 & 0{,}07 \\ 0{,}2 & 0{,}03 & 0{,}06 & 0{,}1 \\ 0 & 0 & 0{,}5 & 0{,}3 \\ 0{,}11 & 0{,}07 & 0{,}14 & 0{,}08 \end{bmatrix} \cdot \begin{bmatrix} 388722 \\ 605410 \\ 666305 \\ 894172 \end{bmatrix} = \begin{bmatrix} 217706 \\ 151170 \\ 225302 \\ 601404 \\ 249954 \end{bmatrix}.$$

Во втором варианте:

$$R_2 = \begin{bmatrix} 0{,}1 & 0 & 0 & 0{,}2 \\ 0{,}15 & 0{,}05 & 0 & 0{,}07 \\ 0{,}2 & 0{,}03 & 0{,}06 & 0{,}1 \\ 0 & 0 & 0{,}5 & 0{,}3 \\ 0{,}11 & 0{,}07 & 0{,}14 & 0{,}08 \end{bmatrix} \cdot \begin{bmatrix} 418582 \\ 651616 \\ 717457 \\ 959258 \end{bmatrix} = \begin{bmatrix} 233710 \\ 162516 \\ 242238 \\ 646506 \\ 268842 \end{bmatrix}.$$

Основа разработки плана по труду есть матрица:

$$H_{BL} = \begin{bmatrix} 1{,}2 & 0{,}3 & 0{,}6 & 0{,}4 \\ 0{,}9 & 0{,}2 & 0{,}4 & 0{,}5 \\ 0{,}6 & 0{,}1 & 0{,}8 & 0{,}7 \end{bmatrix}.$$

Матрицу затрат рабочей силы в первом варианте получаем перемножением матрицы H_{BL} и матрицы выпуска продукции:

$$L_1 = \begin{bmatrix} 1{,}2 & 0{,}3 & 0{,}6 & 0{,}4 \\ 0{,}9 & 0{,}2 & 0{,}4 & 0{,}5 \\ 0{,}6 & 0{,}1 & 0{,}8 & 0{,}7 \end{bmatrix} \cdot \begin{bmatrix} 388722 \\ 605410 \\ 666305 \\ 894172 \end{bmatrix} = \begin{bmatrix} 1405541 \\ 1184540 \\ 1452738 \end{bmatrix},$$

во втором варианте:

$$L_2 = \begin{bmatrix} 1{,}2 & 0{,}3 & 0{,}6 & 0{,}4 \\ 0{,}9 & 0{,}2 & 0{,}4 & 0{,}5 \\ 0{,}6 & 0{,}1 & 0{,}8 & 0{,}7 \end{bmatrix} \cdot \begin{bmatrix} 418582 \\ 651616 \\ 717457 \\ 959258 \end{bmatrix} = \begin{bmatrix} 1511962 \\ 1273660 \\ 1561758 \end{bmatrix}.$$

Данные матрицы составляют необходимые затраты рабочей силы (в человеко-часах).

18.3. Разработка плана предприятия в Excel

Рассмотрим разработку плана предприятия в Excel, используя цифры приведенного примера. Введем матрицу норм потребления продуктов в диапазон ячеек A2:D5, матрицу товарной продукции в F2:F5, единичную матрицу в A8:D11, матрицу норм затрат сырья в A21:D25 и матрицу плана по труду в A28:D30.

Выделим ячейки F8:F11 для ввода расчетной матрицы выпуска продукции. Решение состоит в умножении обратной матрицы разницы между единичной матрицей и матрицей норм потребления продуктов на матрицу товарной продукции.

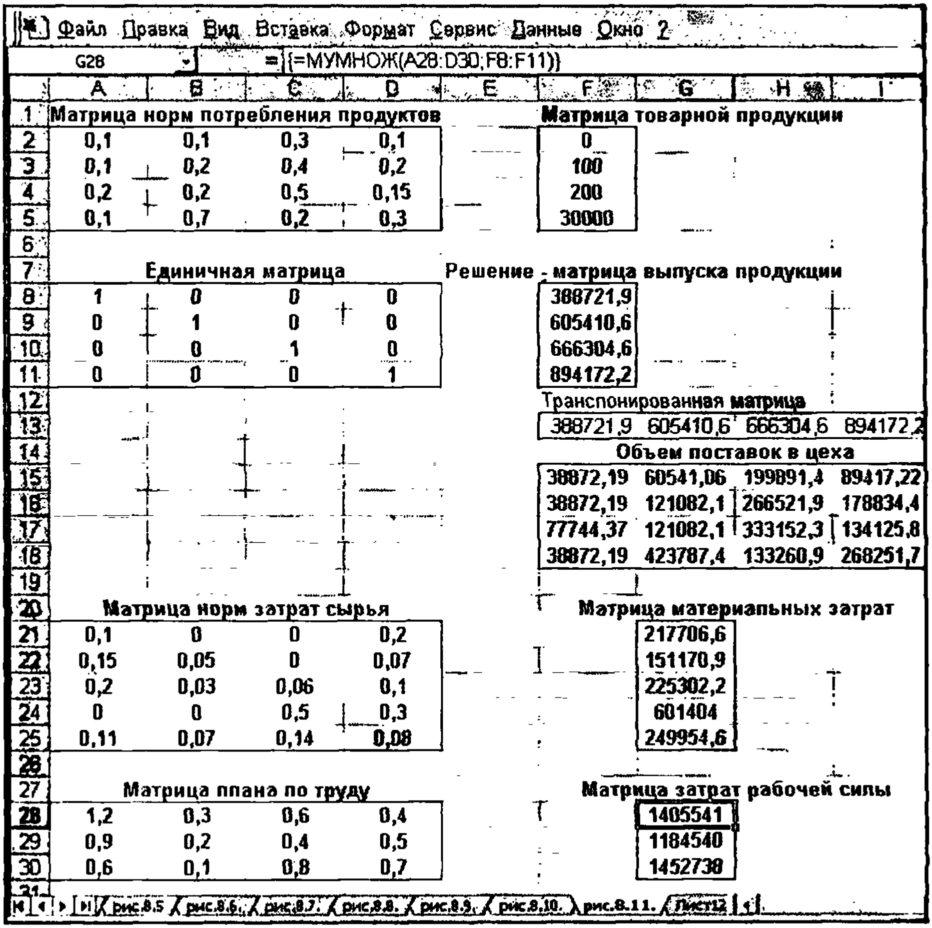


Рис. 18.1. Расчет технико-экономического плана предприятия

Введем функцию МУМНОЖ, затем в опции массива 1 введем встроенную функцию МОБР, в диалоговом окне которой запишем разницу между единичной матрицей и матрицей норм потребления продуктов (A8:D11 – A2:D5), закроем МОБР, в диалоговом окне функции МУМНОЖ в массив 2 введем диапазон матрицы товарной продукции (F2:F5). Закончим ввод комбинацией клавиш <Shift>+ +<Ctrl>+<Enter>. В строке формул (рис. 18.1) должна быть записана формула:

{=МУМНОЖ(МОБР(A8:D11 – A2:D5);(F2:F5))}.

Полученную матрицу выпуска продукции скопируем и через специальную вставку транспонируем значения в ячейки F13:I13. Данная операция не обязательна, но нужна для ускорения расчета объема поставок в цеха.

Для расчета объема поставок в цеха первоначально выделим ячейки F15:F18, куда введем формулу умножения первого столбца матрицы норм потребления продуктов на первое число транспонированной матрицы товарной продукции F13 и закончим ввод комбинацией клавиш <Shift>+<Ctrl>+<Enter>. То есть команда записывается следующим образом:

{=A2:A5*F13}.

Не убирая выделения, скопируем мышью (перетащим) формулу в столбцы

G5:G18; H15:H18; I15:I18,

т. е. получим формулы

{=B2:B5*G13; {=C2:C5*H13}; {=D2:D5*I13}.

Для расчета матрицы материальных затрат выделим диапазон ячеек G21:G25 и введем формулу умножения матрицы норм затрат сырья на матрицу выпуска продукции:

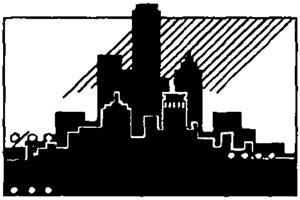
{=МУМНОЖ(A21:D25; F8:F11)}.

Матрица затрат рабочей силы устанавливается перемножением матрицы плана по труду на матрицу выпуска продукции.

{=МУМНОЖ(A28:D30; F8:F11)}.

Глава 19. Основные принципы построения экономико-математических моделей производства

19.1. Общее представление об экономических моделях производства



Моделируемая экономическая система представляется в виде совокупности некоторого числа «элементарных» экономических единиц, каждая из которых имеет в экономической системе определенную функцию, связанную с производством, потреблением, распределением или хранением материальных благ [14, 21, 23].

Экономико-математическая модель производственно-технологического уровня экономической системы состоит в описании:

- потоков материальных благ и трудовых ресурсов между элементарными экономическими единицами;
- закономерностей преобразования ресурсов и продуктов в этих элементарных единицах.

Потоки материальных благ между экономическими единицами должны удовлетворять физическим законам сохранения вещества. Эти законы выражаются в виде балансовых соотношений. Если в модели имеются две производственные единицы, причем первая из них выпускает за заданный период времени некоторый продукт в количестве y, который может быть сырьем для второй производственной единицы либо может вывозиться за пределы моделируемой системы в количестве q, то в отсутствие запасов и ввоза этого продукта можно сформулировать соотношение:

$$q + x = y,$$

где x – потребление второй производственной единицы.

Если дополнительно предположить, что этот продукт используется полностью, то неравенство следует заменить на равенство:

$$q + x = y.$$

Данные соотношения дают представление о балансовых соотношениях, общий принцип построения которых можно сформулировать так: суммарное потребление любого продукта не превышает или равно сумме его исходных запасов производства в системе и поставок извне.

Кроме законов сохранения, в описание потоков включают также различные ограничения на величины этих потоков. Такие ограничения могут иметь под собой чисто технологическую основу – например, ограниченную пропускную способность транспортной сети.

После описания потоков между элементарными единицами модели необходимо сформулировать в математической форме закономерности преобразования ресурсов и продуктов в этих единицах.

Соотношение, описывающее закономерности выпуска новых продуктов в производственных элементарных единицах моделей, принято называть производственными функциями.

Выделяются основные виды производственных функций – функции выпуска, функции затрат и производственные способы.

19.2. Производственные функции как основа описания закономерностей производства

Производственными функциями называют соотношения между используемыми в производстве материальными благами и трудовыми ресурсами (называемыми в совокупности производственными ресурсами), а также выпускаемой продукцией [1, 14, 17, 21, 23, 34, 36].

Пусть в модели рассматривается n производственных ресурсов. Количество i-го ресурса, используемого (или потребляемого) в течение некоторой единицы времени обозначим через x_i. Пусть выпускается m продуктов, причем объем выпуска j-го продукта мы обозначим через y_j. Производственная функция связывает значение вектора продукции y со значениями вектора ресурсов x:

$$F(x, y, a) = 0, \tag{19.1}$$

где $a = a_1, a_2,\ldots,a_p$ – вектор параметров производственной функции.

Соотношение (19.1) может быть векторным, т. е. оно может состоять из нескольких равенств и может быть задано не только в аналитическом виде, но и в виде таблицы.

Описание связи между использованием ресурсов и выпуском продукции в виде (19.1) подразумевает, что не учитываются эффекты, связанные с продолжительностью производственного цикла, т. е. с периодом между затратами ресурсов и выпуском продукции. Далее в этом параграфе будем предполагать, что это возможно в силу того, что единица времени, для которой строится производственная функция, значительно превосходит продолжительность производственного цикла.

Описание элементарной производственной единицы начинается с формулировки списка ресурсов и номенклатуры продукции с указанием характерных значений и пределов изменения этих величин.

Материальные производственные ресурсы необходимо различать по способам их расходования в производственных процессах. Обычно выделяют материальные ресурсы двух типов: предмет труда (сырье) и основные фонды (здания, оборудование и т. д.). Ресурсы первого типа в процессе производства в течение одного производственного цикла (периода выпуска продукции) расходуются полностью. Ресурсы второго типа используются в течение значительного числа производственных циклов.

Вместо общего представления производственных функций в виде (19.1) часто используют два частных случая.

1. **Функция выпуска**, в которой в качестве независимых переменных берутся затраты ресурсов, а функцией является выпуск:

$$y = f(x,a). \tag{19.2}$$

2. Функция производственных затрат, в которой независимой переменной является выпуск, а функцией – затраты:

$$x = f(y,a). \qquad (19.3)$$

В соотношениях (19.2) и (19.3) величины *x*, *y*, и *a* могут быть многокомпонентными или векторными.

В том случае, когда вектор ресурсов *x* является многокомпонентным, между функциями выпуска и функциями затрат возникает принципиальное различие. В функции выпуска (19.2) возможны различные сочетания количеств производственных ресурсов, что приводит к тому, что один и тот же объем продукции может быть произведен при разных сочетаниях производственных ресурсов.

В функции затрат (19.3) задание выпуска продукции полностью определяет затраты ресурсов. Поэтому функция затрат используется в том случае, когда в описываемой элементарной экономической единице отсутствует возможность замещения одного ресурса другим.

Функция выпуска используется тогда, когда такая замена допустима. Как правило, в экономической литературе под производственной функцией подразумевают функцию выпуска.

С понятием производственной функции тесно связано понятие **множество производственных возможностей**, которое определяется как множество всех возможных сочетаний затрат трудовых материальных ресурсов и выпусков продукции:

$$[x,y] \in G(a), \qquad (19.4)$$

где $G(a)$ – некоторое множество G в пространстве ресурсов и продуктов, зависящее от вектора параметра (a).

Множество производственных возможностей задается соотношением:

$$0 \le a \le x^a,\ x > 0, \qquad (19.5)$$

где a – параметр удовлетворяет соотношению $0 \le a \le 1$.

Множество производственных возможностей можно представить графически (рис. 19.1).

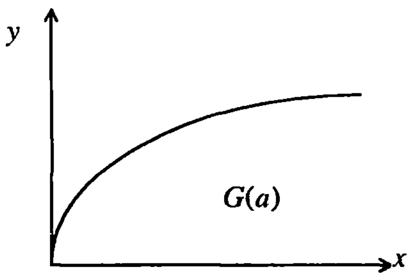


Рис. 19.1

Связь между производственной функцией и множеством производственных возможностей устанавливается следующим образом. Предположим, что производство ведется эффективно, т. е. при данном количестве ресурсов выпускается максимально возможное количество продукции. Тогда имеет смысл рассматривать не все множество производственных возможностей, а только его границу:

$$y = x^a;\ x > 0. \qquad (19.6)$$

Таким образом, мы получаем производственную функцию в виде функции выпуска. Если же с самого начала задана производственная функция в виде (19.6), то множество производственных возможностей можно получить, предполагая, что с помощью тех же ресурсов можно выпустить и меньшее количество продукции. Тогда из (19.6) с учетом того, что выпуск неотрицателен, сразу же получается (19.5).

Аналогичным образом от описания множества производственных возможностей (19.5) можно перейти к функции затрат. Для этого необходимо предположить, что производство данного неотрицательного объема продукции y достигается при минимальных затратах ресурса x. Из (19.5) получаем функцию

$$x = y^{1/a},\ y > 0, \qquad (19.7)$$

которая имеет вид функции затрат (19.3).

Заметим, что соотношения (19.6) и (19.7) описывают одну и ту же зависимость.

На рис. 19.2 изображены функция затрат (19.7) и множества производственных возможностей (19.5).

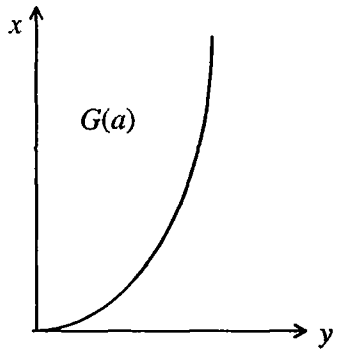


Рис. 19.2

Такая простая связь между множеством производственных возможностей и производственными функциями устанавливается лишь в случае одного продукта и одного ресурса. В общем случае переход от множества производственных возможностей (19.4) к производственной функции (19.1) требует построения множества всех эффективных точек множества (19.4), т. е. всех таких объемов затрачиваемых ресурсов и выпускаемой продукции, что при данных затратах нельзя выпустить большее количество хотя бы одного вида продукции, не уменьшив производства других видов продукции.

Кроме того, данного выпуска нельзя добиться при меньших затратах хотя бы одного ресурса, не увеличив при этом затраты других ресурсов.

19.3. Свойства производственных функций

Обратимся к некоторым наиболее общим свойствам производственных функций, имеющих форму $y = f(x, a)$, т. е. функций выпуска, допускающих замещение одного ресурса другим [21]. Рассмотрим в данном разделе функции с одним продуктом и несколькими ресурсами – трудовыми и материальными.

Вектор параметров a в данном соотношении будем опускать, считая, что параметры уже определены и их влияние нас не интересует. Тогда функция выпуска приобретает вид:

$$y = f(x), \tag{19.8}$$

где: $x = (x_1,...,x_n)$ – вектор.

Соотношение (19.8) задано при неотрицательных значениях компонентов вектора x.

Обычно относительно производственной функции (19.8) делают предположение, очень удобное с математической точки зрения, – о непрерывном изменении переменных x и достаточно плавном изменении выпуска при изменении затрат ресурсов. В математической форме эти предположения имеют следующий вид: функция (19.8) задана при всех неотрицательных значениях составляющих вектора x и является непрерывной или нужное число раз дифференцируемой функцией своих аргументов.

Перейдем к формулировке предположений (свойств), имеющих под собой экономическое обоснование. Для этого нам потребуются показатели предельного анализа.

Частная производная производственной функции по одному из ресурсов является предельной производительностью (эффективностью) данного ресурса $\partial f/\partial x_i$. Она характеризует скорость изменения функции выпуска по отношению к изменению затрат ресурса. Если предельная производительность ресурса положительна, то, следовательно, выпуск растет при росте затрат ресурса. Если предельная производительность ресурса отрицательна, то выпуск уменьшается при росте затрат ресурса.

Средней производительностью ресурса будет показатель $f(x)/x_1$.

Относительной характеристикой изменения выпуска продукции при увеличении затрат ресурсов будет показатель эластичности выпуска по отношению к изменению затрат i-го ресурса:

$$\varepsilon_i = \frac{x_i}{f(x)}\frac{\partial f}{\partial x_i}.$$

Эластичность выпуска по отношению к изменению затрат ресурса показывает, на сколько процентов возрастет объем продукции при увеличении затрат ресурсов на 1%.

Величину $\varepsilon_i(x)$ можно вычислить по другой, эквивалентной формуле:

$$\varepsilon_i(x) = \partial(\ln f(x)) / \partial(\ln x_i).$$

Определим данные показатели для производственной функции $y = x^a$ при $x > 0$. Предельная эффективность ресурса равна:

$$\partial f/\partial x = ax^{a-1} > 0;$$

Средняя эффективность ресурса равна:

$$f(x)/x_i = x^{a-1}.$$

В силу того, что $0 < a < 1$, для этой производственной функции предельная эффективность меньше средней.

Эластичность выпуска по ресурсу будет равна:

$$\varepsilon(x) = \partial(\ln f(x)) / \partial(\ln x) = \partial(a \ln x) / \partial(\ln x) = a.$$

Эта производственная функция характеризуется постоянной эластичностью выпуска по отношению к изменению ресурса.

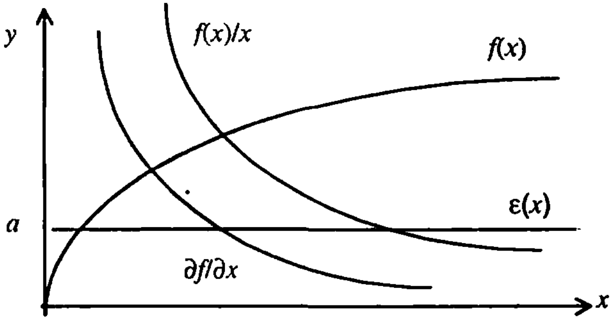


Рис. 19.3

На рис. 19.3 изображен график производственной функции $y = x^a$, ее предельной и средней эффективностей, а также эластичности выпуска по ресурсу.

Теперь сформулируем экономические предположения.

Первое предположение. Производство невозможно при отсутствии хотя бы одного ресурса (точнее незаменимого ресурса), т. е.

$$\begin{aligned} f(0, x_2, \ldots, x_n) &= 0; \\ f(x_1, 0, \ldots, x_n) &= 0; \\ f(x_1, x_2, \ldots, 0) &= 0. \end{aligned} \tag{19.9}$$

Это означает, что каждый из ресурсов необходим хотя бы в малых количествах. Полное его отсутствие не может быть компенсировано другими ресурсами.

Второе предположение. При увеличении затрат производственных ресурсов выпуск продукции не уменьшается. Это означает, что предельные эффективности ресурсов положительны. В математической форме:

$$\partial f / \partial x_i \geq 0, i = 1, \ldots, n. \tag{19.10}$$

Предположение (19.10), являющееся на первый взгляд очевидным, выполняется не всегда. Например, при возрастании количества удобрений, приходящихся на единицу площади, производство зерна сначала растет, а затем начинает снижаться. Поэтому для производственных функций, не удовлетворяющих соотношению (19.10), вводится понятие **экономической области**. Использование ресурсов в сочетаниях, не попадающих в экономическую область, бессмысленно с экономической точки зрения.

Для функций (19.8), имеющих непрерывные производные, границами экономической области являются поверхности $\partial f / \partial x_i = 0$, которые называют разделяющими поверхностями.

Третье предположение. По мере увеличения количества одного ресурса при постоянных количествах других предельная эффективность использования этого ресурса не возрастает. Математически это требование для дважды дифференцируемых функций выглядит следующим образом:

$$\frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_i}\right) \le 0, \; i = 1, \ldots, n. \tag{19.11}$$

Для производственной функции вида $y = x^a$ это условие выполняется. Оно означает, что рост вооруженности средствами производства приводит к росту выпуска продукции, но темп роста выпуска продукции все время падает. В случае экстенсивного роста производства, т. е. роста только за счет количества ресурсов без повышения эффективности их использования на основе достижений научно-технического прогресса, соотношение (19.11) имеет разумную интерпретацию: поскольку каждая следующая единица производственного ресурса, количество которого возрастает, должна соединиться со все меньшим приходящимся на нее количеством других ресурсов, эффективность использования этого ресурса уменьшается.

Часто вместо условия (19.11) формулируется более сильное математическое требование, близкое к (19.11) по смыслу. Если $f(x)$ – выпуклая вверх функция своих аргументов, на неотрицательном ортанте для любых двух неотрицательных векторов x^* и x^{**} и любого числа $a \in [0,1]$ справедливо неравенство:

$$f(ax^* + (1-a)x^{**}) \ge af(x^*) + (1-a)f(x^{**}). \tag{19.12}$$

Если функция $f(x)$ дважды непрерывно дифференцируема, условие выпуклости эквивалентно требованию неположительной определенности квадратичной формы при всех положительных значениях вектора ресурсов x:

$$(\xi, H\xi) = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}\xi_i \xi_j \le 0. \tag{19.13}$$

Если используется единственный ресурс, а функция $f(x)$ достаточно гладкая, то требования (19.11), (19.12) и (19.13) равносильны. Если же ресурсов несколько, то (19.11) не эквивалентно (19.12) или (19.13), т. е. не эквивалентно выпуклости вверх функции $f(x)$.

Четвертое предположение. Производственная функция характеризуется определенной отдачей от расширения масштабов производства. Последняя характеризует изменение выпуска продукции при пропорциональном изменении затрат ресурсов и математически выражается в умножении всех компонентов вектора x на положительный скаляр t. Скалярная функция $f(x)$ является однородной функцией степени δ, если для любого вектора x и любого скаляра t она удовлетворяет соотношению:

$$f(tx) = t^{\delta} f(x). \tag{19.14}$$

Математически четвертое предположение состоит в требовании однородности производственной функции. Если $\delta > 1$, то производственная функция характеризуется возрастающей отдачей от расширения масштабов производства; если

$\delta = 1$ – постоянной отдачей; при $\delta < 1$ – убывающей отдачей. Естественно, что выполняется предположение $\delta \geq 1$, ибо в противном случае нарушалось бы условие (19.10) во всех точках положительного ортанта и отсутствовала бы экономическая область. Данное предположение выполняется далеко не для всех производственных функций, используемых в экономических исследованиях. Для характеристики последствий изменения масштаба производства вводят показатель $\varepsilon(x)$, называемый **эластичностью производства** и определяемый следующим образом:

$$\varepsilon(x) = \lim_{t \to 1} \frac{t}{f(tx)} \frac{\partial f(tx)}{\partial t}. \qquad (19.15)$$

Этот показатель характеризует процентное изменение выпуска продукции при изменении масштаба производства на 1% при данной структуре ресурсов. Для производственных функций, удовлетворяющих соотношению (19.14), получаем $\varepsilon(x) = \delta$.

Можно установить связь между эластичностью производства и эластичностью выпуска по отношению к изменению затрат ресурсов $\varepsilon_i(x)$. Учитывая, что

$$\frac{\partial f(tx)}{\partial t} = \sum_{i=1}^{n} \frac{\partial f(tx)}{\partial (tx_i)} x_i, \qquad (19.16)$$

тогда

$$\varepsilon(x) = \lim_{t \to 1} \frac{t}{f(tx)} \frac{\partial f(tx)}{\partial t} = \lim_{t \to 1} \frac{t}{f(tx)} \sum_{i=1}^{n} \frac{\partial f(tx)}{\partial (tx_i)} x_i = \frac{1}{f(x)} \sum_{i=1}^{n} \frac{\partial f}{\partial x} x_i = \sum_{i=1}^{n} \varepsilon_i(x). \qquad (19.17)$$

Таким образом, эластичность производства в некоторой точке пространства ресурсов равна сумме эластичности выпуска по отношению к затратам производственных ресурсов в этой точке.

В случае единственного ресурса, например в функции (19.6), эластичность производства совпадает с эластичностью выпуска по отношению к изменению затрат ресурса. Для производственных функций с постоянной отдачей от расширения масштабов производства (19.14) связь между эластичностями выпусков и эластичностью производства приобретает вид:

$$\sum_{i=1}^{n} \varepsilon_i(x) = \delta. \qquad (19.18)$$

Рассмотрим производственные функции, удовлетворяющие четырем сформулированным выше предположениям, а именно: (19.9), (19.11). (19.14), (19.18). Возьмем t, удовлетворяющее условиям $0 < t < 1$. Из условия (19.15) получаем:

$$t f(x) + (1 - t) f(0) \leq f(tx + (1 - t) \times 0) = f(tx).$$

Поскольку в силу (19.9) имеем $f(0) = 0$, то $tf(x) \leq f(tx)$. Из соотношения (19.14) получаем $\delta \leq 1$, т. е. для выпуклых вверх производственных функций имеет место невозрастающая отдача от увеличения масштаба производства. Если производственная функция является строго выпуклой, условие (19.12) выполняется со знаком строгого неравенства ($\delta < 1$). Это означает, что отдача от увеличения масштаба может быть только убывающей. Таким образом, для производственных функций, удовлетворяющих четырем соотношениям, в силу (19.18) и неотрицательности эластичности выпуска по ресурсам существует ограничение по эластичности выпуска:

$$0 \le \varepsilon_i(x) \le 1. \qquad (19.19)$$

Таким образом, в основе производственных функций лежат предположения, приведенные на рис. 19.4.

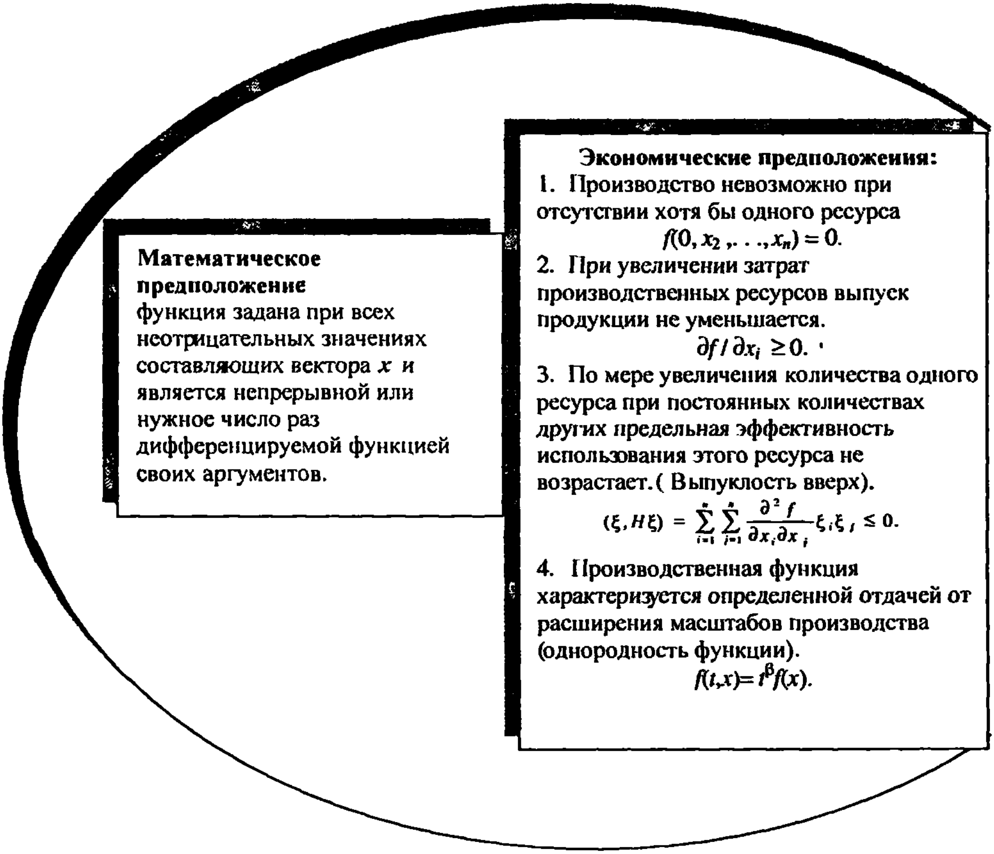


Рис. 19.4

19.4. Возможности замещения ресурсов

Возьмем производственную функцию с двумя ресурсами:

$$y = x_1^{a} x_2^{1-a}, \; 0 < a < 1. \qquad (19.20)$$

Функции такого типа часто используются при описании народного хозяйства или его структурных единиц. В таких производственных функциях величина y имеет смысл конечной продукции народного хозяйства, x_1 – общего количества основных фондов, x_2 – общего количества трудовых ресурсов в стране.

Функция (19.20) удовлетворяет всем предположениям предыдущего раздела, причем для нее $\delta = 1$. Поэтому можно построить функцию $\varphi(x)$, которая в данном случае будет показывать объем продукции на 1 трудящегося и имеет вид:

$$\varphi(\tilde{x}) = \tilde{x}^a,$$

где $\tilde{x}$ – отношение количества основных фондов к численности трудящихся, т. е. фондовооруженность.

График функции $\varphi(\tilde{x})$ совпадает в этом случае с графиком производственной функции с одним ресурсом. Возможность взаимного замещения ресурсов означает, что одно и то же количество продукта y может быть произведено при различных сочетаниях ресурсов. Совокупность таких сочетаний ресурсов, т. е. точек в пространстве ресурсов, при котором может быть произведено определенное количество продукции y, называется **изоквантой** и обозначается:

$$Q(y_0) = \{x : f(x) = y_0\}. \tag{19.21}$$

Рассмотрим произвольный луч в пространстве ресурсов, исходящий из начала координат и лежащий в положительном ортанте. Математически этот луч описывается как множество:

$$L = \{x : x = tx^0, t \geq 0\}, x^0 \geq 0.$$

Согласно соотношению (19.16), получается, что для точек луча L имеет место соотношение:

$$f(x) = t^\delta f(x^0).$$

Если $f(x^0) > 0$ и $\delta > 0$, то при достаточно больших t выпуск продукции на луче может достичь любых предварительно заданных величин, в том числе и y_0.

Пусть $y_0 = t_0^\delta f(x^0)$, тогда в точке $x^* = t_0 x^0$ луч L пересекается с изоквантой $Q(y_0)$. В точках луча, лежащих ближе к началу координат, т. е. $t < t_0$, выполняется соотношение $y < y_0$. А в точках луча, лежащих от начала координат дальше, чем точка x^*, имеем $y > y_0$. Поскольку данное утверждение верно для любого луча с положительным направляющим вектором x^0 – таким, что $(f(x^0) > 0$, то можно сделать следующие выводы о свойствах изоквант:

- изокванты не пересекаются друг с другом;
- изокванта $Q(y_0)$ разбивает неотрицательный ортант пространства ресурсов на 2 множества: в одном из которых $y < y_0$, в другом $y > y_0$, причем граница между этими множествами проходит по изокванте $Q(y_0)$;
- большему выпуску продукции соответствует изокванта, более удаленная от начала координат;
- изокванты не имеют общих точек с осями координат.

Одна из изоквант производственной функции $y = x_1^{a_1} x_2^{1-a_1}$ изображена на рис. 19.5. Луч L представлен на рисунке отрезком OA. Изокванта $Q(y_0)$ представляет собой зависимость $x_2(x_1)$.

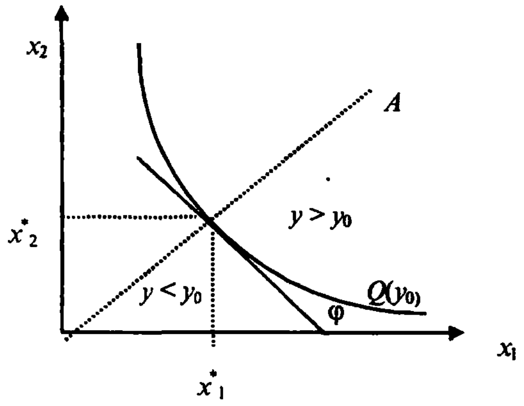


Рис. 19.5

Уравнение изокванты (19.21) задает эту зависимость неявно:

$$y_0 = x_1^a x_2^{1-a}.$$

В явном виде получаем:

$$x_2(x_1) = \left(\frac{y_0}{x_1^a}\right)^{\frac{1}{1-a}}. \tag{19.22}$$

Функция $x_2(x_1)$, имеющая смысл количества трудовых ресурсов, необходимых для получения заданного конечного продукта в зависимости от использующегося объема основных фондов, является монотонно убывающей функцией. При $\partial f/\partial x_2 > 0$ вдоль изокванты выполняется соотношение:

$$\gamma = \frac{\partial x_2}{\partial x_1} = \frac{\partial f/\partial x_1}{\partial f/\partial x_2}. \tag{19.23}$$

Из условия (19.11) получается, что $\gamma \le 0$, а при строгой положительности предельных эффективностей ресурсов $\gamma < 0$. Величину γ принято называть **предельной нормой замещения** одного ресурса другим. Она показывает, сколько второго ресурса может быть высвобождено при увеличении затрат первого ресурса, если выпуск продукции остается неизменным.

Предельная норма замещения γ имеет отрицательную величину, т. к. при уменьшении использования одного из ресурсов для сохранения выпуска продукции использование другого ресурса надо увеличить. На рис. 19.5 предельная норма замещения $\tilde{a}$ совпадает по величине с тангенсом угла φ. Можно заметить, что $\operatorname{tg}\varphi < 0$, а угол φ и величина γ меняются при движении вдоль изокванты $Q(y_0)$.

Для производственной функции $y = x_1^{a_1} x_2^{1-a_1}$ имеем:

$$\gamma = -\frac{\partial y/\partial x_1}{\partial y/\partial x_2} = -\frac{a}{(1-a)}\frac{x_2}{x_1}. \tag{19.24}$$

Из формулы следует, что для функции (19.20) абсолютная величина предельной нормы замещения труда основными фондами обратно пропорциональна фондовооруженности x_1/x_2. Этот факт можно легко объяснить: увеличение фондовооруженности приводит к уменьшению количества трудовых ресурсов, высвобождаемых каждой новой единицей основных фондов. Такой результат тесно связан со свойством (19.14) функции (19.20).

Линии $\gamma(x_1,x_2) = \gamma_0$ называют **изоклиналями** производственных функций с двумя ресурсами. Для функции (19.20) изоклинали имеют вид:

$$x_2 = -\gamma_0 \frac{1-a}{a} x_1.$$

На рис. 19.6 представлены две изокванты, $Q\ (y_0)$ и $Q\ (y_1)$, и три изоклинали, соответствующие значениям нормы замещения, γ_1, γ_2, и γ_3, где $(-\gamma_1) > (-\gamma_2) > (\ \gamma_1)$ для производственной функции (19.20).

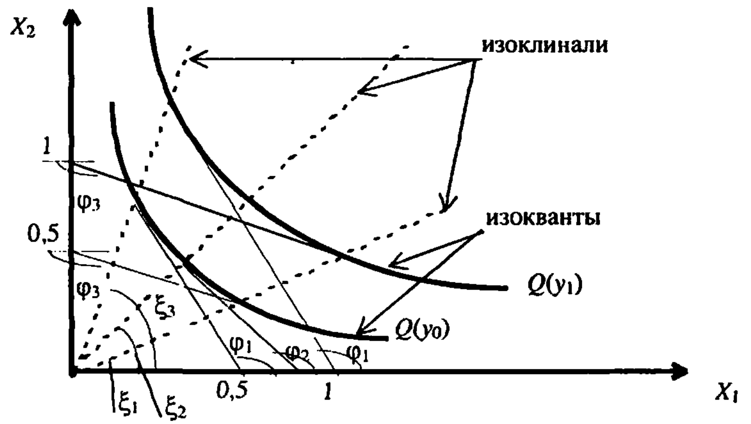


Рис. 19.6

Величины углов φ_1 , φ_2 и φ_3 удовлетворяют соотношению:

$$\gamma_i = \mathrm{tg}\varphi_i\,, \quad i = 1, 2, 3,$$

а уравнения изоклиналей выглядят так:

$$x_2 = -\gamma_i \frac{1-a}{a} x_1, \quad i = 1, 2, 3.$$

В данном случае изоклинали имеют особенно простой вид – они являются лучами, исходящими из начала координат.

Такое свойство имеют изоклинали для важного класса производственных функций – однородных функций.

Рассмотрим пример построения изоквант и изоклиналей в Excel. Для примера возьмем производственную функцию $y = x_1^{0,4}\ x_2^{0,6}$. Используя окно **Таблица подстановки**, рассчитаем матрицу для производственной функции – рис. 19.7.

Рис. 19.7

По данным рассчитанной таблицы построим график производственной функции «Поверхность» – рис. 19.8.

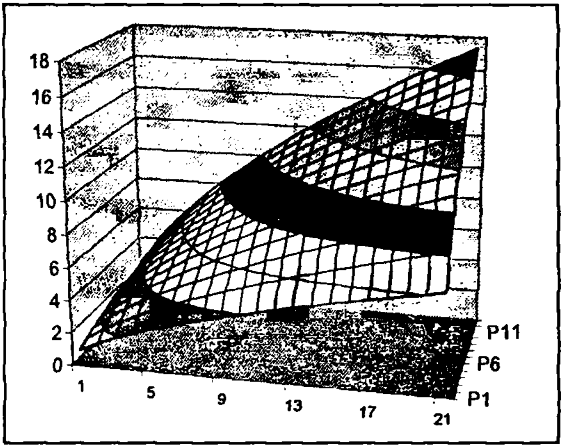


Рис. 19.8

Изолинии на графике производственной функции являются изоквантами, правда, построенными в соответствии с ценой деления. Их легко представить в осях ресурсов, задав контурный тип диаграммы, – рис. 19.9.

Построим изокванты с необходимыми нам уровнями выпуска, допустим: 4, 6, 8. Для этого рассчитаем изокванты для данных уровней, задав формулу изокванты, – рис. 19.10.

По рассчитанным данным построим графики трех изоквант – рис. 19.11.

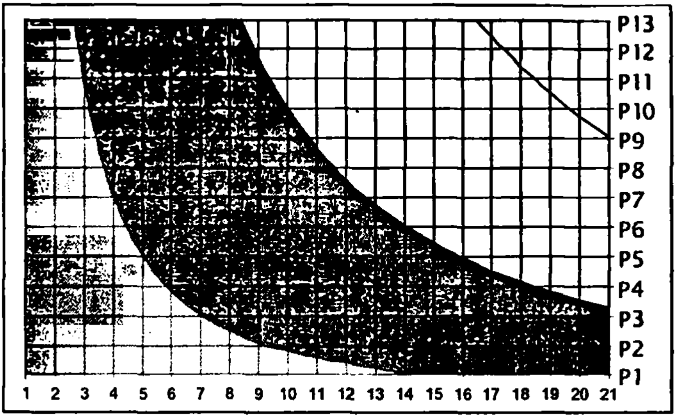
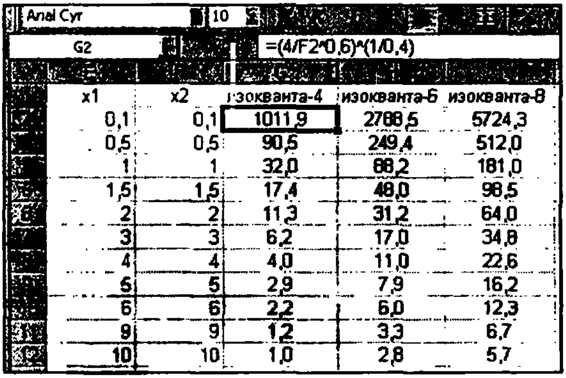


Рис. 19.9


Arial Cyr 10

G2 =(4/F2^0,6)^(1/0,4)

x1	x2	изокванта-4	изокванта-6	изокванта-8
0,1	0,1	1011,9	2788,5	5724,3
0,5	0,5	90,5	249,4	512,0
1	1	32,0	88,2	181,0
1,5	1,5	17,4	48,0	98,5
2	2	11,3	31,2	64,0
3	3	6,2	17,0	34,8
4	4	4,0	11,0	22,6
5	5	2,9	7,9	16,2
6	6	2,2	6,0	12,3
9	9	1,2	3,3	6,7
10	10	1,0	2,8	5,7

Рис. 19.10

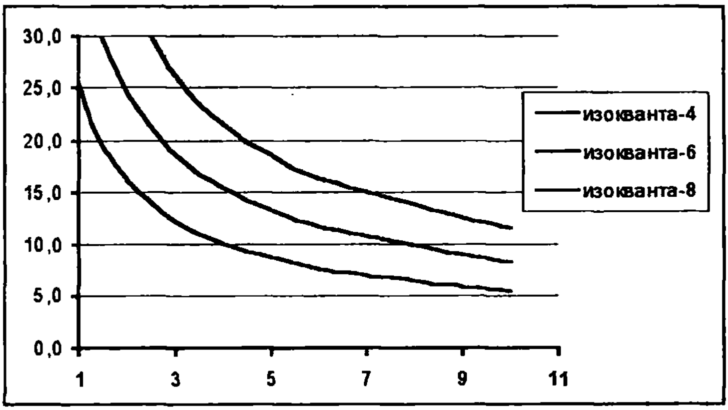


Рис. 19.11

Зададимся тремя соотношениями ресурсов, установим предельные нормы замещения и по ним рассчитаем изоклинали – рис. 19.12.

изокванта-6	изокванта-8	изоклиналь-2	изоклиналь-1	изоклиналь-0,5
180,7	252,5	0,3	0,15	0,075
61,8	86,4	1,5	0,75	0,375
38,9	54,4	3	1,5	0,75
29,7	41,5	4,5	2,25	1,125
24,5	34,3	6	3	1,5
18,7	26,2	9	4,5	2,25
15,5	21,6	12	6	3
13,3	18,6	15	7,5	3,75
11,8	16,5	18	9	4,5
9,0	12,6	27	13,5	6,75
8,4	11,7	30	15	7,5

изокванты

Рис. 19.12

Вид изоклиналей по этим расчетам показан на рис. 19.13.

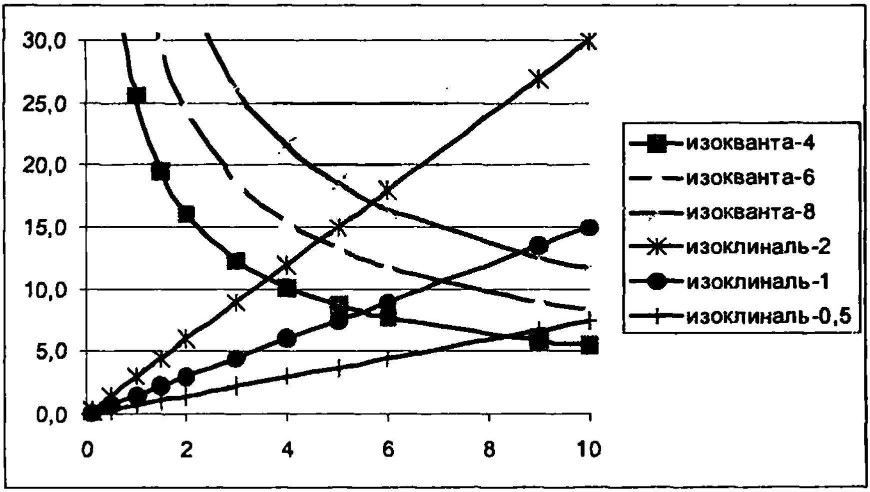


Рис. 19.13

Для количественной характеристики скорости изменения предельной нормы замещения вдоль изокванты используется понятие **эластичности замещения ресурсов** $\sigma(x_1, x_2)$:

$$\sigma(x_1,x_2)=\frac{\gamma}{x_2/x_1}\cdot\frac{d(x_2/x_1)}{d\gamma}. \qquad (19.25)$$

Эластичность замещения ресурсов имеет следующий экономический смысл: она приближенно показывает, на сколько процентов должно измениться отношение ресурсов при движении вдоль изокванты, чтобы при этом предельная норма замещения γ изменилась на 1%.

Для производственной функции (19.20) эластичность замещения ресурсов имеет простую геометрическую интерпретацию: поскольку изоклинали этой функции – прямые линии, то отношение x_2/x_1 характеризуется тангенсом угла ξ наклона изоклинали (см. рис.19.5). Поэтому величина δ показывает, на сколько процентов необходимо повернуть изоклиналь (т. е. изменить $\mathrm{tg}\xi$), чтобы $\mathrm{tg}\varphi$ изменился на 1%.

Как и в случае эластичности выпуска по ресурсу, эластичность замещения ресурсов также может быть представлена в более удобной форме:

$$\sigma(x_1,x_2)=\frac{d(\ln(x_2/x_1))}{d(\ln(-\gamma))}.$$

Для производственной функции (21.20), учитывая соотношение

$$\ln(-\gamma)=\ln(a/(1-a))+\ln(x_2/x_1),$$

получаем:

$$\sigma(x_1,x_2)=\frac{d(\ln(x_2/x_1))}{d(\ln(-\gamma))}=\frac{d(\ln(-\gamma)-\ln(a/(1-a)))}{d(\ln(-\gamma)}=1.$$

Постоянство эластичности замещения ресурсов σ многих производственных функций позволяет охарактеризовать с ее помощью возможность замещения ресурсов в целом (а не при каком-то конкретном соотношении ресурсов, как это удается на основе предельной нормы замещения γ). Чем больше σ, тем в более широких пределах производственные ресурсы могут замещать друг друга.

Все изложенные понятия, относящиеся к анализу замещения ресурсов в производственных функциях с двумя ресурсами, могут быть обобщены и на случай произвольного числа ресурсов. Понятие изокванты (19.21) с самого начала введено для произвольного числа ресурсов. Продифференцировав функцию $f(x)$ вдоль изокванты, получим:

$$\sum_{i=1}^{k}\frac{\partial f}{\partial x_i}dx_i=0. \qquad (19.26)$$

Зафиксируем затраты всех ресурсов, кроме i-го и j-го. Получаем соотношение

$$\frac{\partial f}{\partial x_i}dx_i+\frac{\partial f}{\partial x_j}dx_j=0,$$

которое полностью совпадает с соотношением (19.22) для производственной функции с двумя ресурсами. Это дает возможность ввести предельную норму замещения для ресурсов i и j:

$$\gamma_{ij}(x_1,\dots,x_n)=\frac{\partial f/\partial x_j}{\partial f/\partial x_i}, \qquad (19.27)$$

где $\partial f/\partial x_i>0$. Величина γ_{ij} характеризует отношения между малыми изменениями количеств этих ресурсов при сохранении выпуска на прежнем уровне.

Можно ввести понятие эластичности замещения ресурсов i и j:

$$\sigma_{ij}(x_1,...,x_n) = \frac{\gamma_{ij}}{x_i/x_j}\frac{d(x_i/x_j)}{d\gamma_{ij}} = \frac{d(\ln(x_i/x_j))}{d(\ln(-\gamma_{ij}))}, \qquad (19.28)$$

где по-прежнему меняются объемы только двух ресурсов, i-го и j-го, а производная берется вдоль изокванты.

Эластичность замещения ресурсов i и j приближенно показывает, на сколько процентов должно измениться отношение ресурсов i и j, чтобы при этом предельная норма замещения этих ресурсов изменялась на 1%.

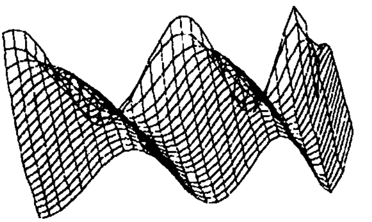
Итак, основными показателями анализа замещения ресурсов являются параметры, приведенные в табл. 19.1.

Таблица 19.1

Изокванта	Функция одного ресурса от остальных при постоянном выпуске	$y_0 = x_1^a x_2^{1-a}$ $x_2(x_1) = \left(\frac{y_0}{x_1^a}\right)^{\frac{1}{1-a}}$	Показывает, при каких раз-личных сочетаниях ресурсов может быть обеспечен один и тот же выпуск
Предельная норма замещения одного ресурса другим	Отношение частных производных по ресурсам, т. е. отношение предельных производительностей ресурсов	$\gamma_{ij}(x_1,...,x_n) = $ $= \frac{\partial f/\partial x_j}{\partial f/\partial x_i}$	Показывает, сколько второго ресурса может быть высвобождено при увеличении затрат первого ресурса, если выпуск продукции остается неизменным
Изоклиналь	Функция одного ресурса от остальных при постоянной предельной норме замещения	$x_2 = -\gamma_i \frac{1-a}{a} x_1$	Показывает, при каких раз-личных сочетаниях ресурсов может быть обеспечена одинаковая предельная норма замещения
Эластичность замещения ресурсов	Эластичность отношения ресурсов по предельной норме замещения	$\sigma_{ij}(x_1,...,x_n) = $ $= \frac{\gamma_{ij}}{x_i/x_j}\frac{d(x_i/x_j)}{d\gamma_{ij}}$	Показывает процентное изменение отношения ресурсов i и j, при изменении предельной нормы замещения этих ресурсов на 1%

Глава 20. Основные типы функций выпуска

20.1. Степенные производственные функции



Степенная производственная функция с (n) ресурсами имеет следующий общий вид:

$$y = a\prod_{i=1}^{n} x_i^{a_i} = a x_1^{a_1} \cdot x_2^{a_2} \dots x_n^{a_n}, \tag{20.1}$$

где y – объем продукции, который является скалярной величиной; $a, a_1, a_2, \dots, a_n$ – положительные параметры. Обычно предполагается, что $a_i < 1$ для всех $i = 1, \dots, n$.

Степенную производственную функцию (20.1) часто представляют в более удобном логарифмическом виде, эквивалентном (20.1) при $x_i > 0$ ($i = 1, \dots, n$):

$$\ln y = \ln a + \sum_{i=1}^{n} a_i \cdot \ln x_i.$$

Степенные производственные функции были предложены в двадцатых годах двадцатого столетия Ч. Коббом и П. Дугласом для описания связи между объемом общественного продукта и двумя важнейшими ресурсами – трудовыми ресурсами и основными производственными фондами. В настоящее время степенные производственные функции используются для широкого класса экономических систем.

Для функции (20.1) выпуск продукции невозможен при отсутствии хотя бы одного ресурса.

Предельная эффективность j-го ресурса имеет вид:

$$\frac{\partial y}{\partial x_j} = a \cdot a_j \cdot x_j^{a_{(j-1)}} \cdot \prod_{i \neq j} x_i^{a_i} = \frac{a_j}{x_j} \cdot y(x). \tag{20.2}$$

При $x_i > 0$ ($i = 1, \dots, n$) имеем $\partial y / \partial x_j > 0$, причем в силу $0 < a_i < 1$ при стремлении x_j к нулю предельная эффективность j-го ресурса стремится к бесконечности, а при стремлении x_j к бесконечности предельная эффективность стремится к нулю (конечно, при постоянных объемах других ресурсов).

Эластичность выпуска по j-му ресурсу легко получить из логарифмического представления степенной функции:

$$\varepsilon_j = \frac{x_j}{y} \frac{\partial y}{\partial x_j} = \frac{\partial(\ln y)}{\partial(\ln x_j)} = a_j. \tag{20.3}$$

Соотношение (20.3) показывает роль параметров a_j – это эластичность выпуска по соответствующему ресурсу. Учитывая экономический смысл эластичности выпуска по ресурсу, получаем, что показатели степени производственной функции (20.1) характеризуют отношение предельной и средней эффективностей использования производственных ресурсов. Таким образом, отношение этих двух эффективностей в степенной производственной функции не зависит от количества используемых ресурсов.

Эластичность производства равна:

$$\varepsilon = \sum_{i=1}^{n} \varepsilon_i = \sum_{i=1}^{n} a_i. \qquad (20.4)$$

В степенных производственных функциях выполняется предположение о наличии определенной отдачи от изменения масштабов производства. При $\sum a_i > 1$ имеем возрастающую, при $\sum a_i = 1$ – постоянную, а при $\sum a_i < 1$ – убывающую отдачу от увеличения масштабов производства.

Итак, степенная производственная функция удовлетворяет всем четырем предположениям о производственных функциях, сформулированных ранее.

Рассмотрим вопрос о замещаемости производственных ресурсов в степенной производственной функции (20.1). Предельные нормы замещения имеют вид:

$$\gamma_{ij} = -\frac{\partial y/\partial x_j}{\partial y/\partial x_i} = -\frac{a_j \cdot x_i}{a_i \cdot x_j}. \qquad (20.5)$$

Предельные нормы замещения являются линейными функциями отношения объема ресурсов, поэтому изоклинали степенной производственной функции плоскости (или линии при $n = 2$). При пропорциональном росте объемов производственных ресурсов предельная норма замещения не изменяется. При стремлении количества замещаемого ресурса к нулю предельная норма замещения падает, но остается положительной, т. е. возможность замещения сохраняется при любых малых (но не нулевых) количествах замещаемого ресурса.

Подсчитаем эластичность замещения σ_{ij}. Прежде всего заметим, что из (20.5) следует:

$$\ln(x_i/x_j) = \ln(a_i/a_j) + \ln(-\gamma_{ij}).$$

Поэтому:

$$\sigma_{ij} = \frac{d(\ln(x_i/x_j))}{d(\ln(-\gamma_{ij}))} = 1. \qquad (20.6)$$

Таким образом, свойства производственной функции с двумя производственными ресурсами переносятся на степенные производственные функции с любым числом ресурсов. Равенство единице эластичности замещения ресурсов в степенных производственных функциях вне зависимости от коэффициентов a, $a_1, \ldots, a_n$ является одним из важнейших свойств производственных функций этого типа. Она показывает, что характеристика замещения одного ресурса другим при выборе степенной производственной функции задана заранее – вне зависимости от желания исследователя. Это является одним из недостатков степенной производственной функции вида (20.1).

Изокванты степенной производственной функции (20.1) описываются соотношением:

$$a x_1^{a_1} \ldots x_n^{a_n} = y_0.$$

Для двух ресурсов характерный вид изоквант изображен на рис. 20.5. На рисунке изокванта приближается к оси координат при стремлении одного из ресурсов к бесконечности. Это можно показать аналитически из уравнения изокванты:

$$y_0 = a x_1^{a_1} x_2^{a_2}.$$

Получаем ее уравнение как функцию $x_2(x_1)$:

$$x_2 = (y_0/ax_1^{a_1})^{1/a_2}.$$

Отсюда в силу $0 < a_i < 1$ следует, что на изокванте $x_2(x_1)$

$$\lim_{x_1 \to +\infty} x_2(x_1) = 0,$$

т. е. в соответствии с рис. 20.5 изокванта неограниченно приближается к оси x_1. Аналогичным образом получаем, что при $x_2 \to \infty$ изокванта приближается к оси x_2.

Это стремление изоквант к координатным осям означает, что любое заранее заданное количество продукции может быть выпущено при сколь угодно малом количестве одного из ресурсов, если имеется в достаточном количестве другой ресурс. Такое свойство изоквант степенной ПФ с двумя ресурсами переносится на степенные ПФ с любым числом переменных (20.1): одним производственным ресурсом можно компенсировать недостаток всех остальных ресурсов.

Возможность замещения одного ресурса другим (равенство единице эластичности замещения ресурсов и неограниченная возможность компенсации недостатка одних ресурсов другими) часто вступает в противоречие со свойствами моделируемых производственных единиц. В связи с этим все чаще используются ПФ, близкие к степенной, но отличающиеся от нее возможностями замещения ресурсов. Такие функции характеризуются показателем эластичности замещения ресурсов, не равным единице.

20.2. Оценка параметров степенной производственной функции в Excel

Рассмотрим вопрос установления параметров a_0, a_1, a_2 двухфакторной производственной функции

$$Y = a_0 K^{a_1} L^{a_2},\ a_1 > 0,\ a_2 > 0.$$

Параметры этой функции можно установить множественным регрессионным анализом, предварительно приведя функцию к линейному виду путем ее логарифмирования:

$$\ln Y = \ln a + a_1 \ln K + a_2 \ln L.$$

В 1928 г. американский экономист Ч. Кобб в сотрудничестве с математиком П. Дугласом (Cobb C. W., Douglas P. H.) оценили значения a_0, a_1, a_2, используя данные по американской обрабатывающей промышленности за период с 1899 по 1922 год, – индекс производства Y, индекс основного капитала K и индекс труда L. В результате исследователи пришли к выводу о том, что $a_1 + a_2 = 1$ (т. е. имеет место неизменный эффект масштаба). С тех пор производственную функцию

$$Y = a_0 K^{a_1} L^{a_2},$$

для которой $a_1 + a_2 = 1$, называют функцией Кобба–Дугласа.

Повторим вслед за Ч. Коббом и П. Дугласом вычисления, используя Excel. Введем данные (индексы Y, K, L) – рис. 20.1.

Прологарифмируем введенные данные (см. рис. 20.1) и применим к ним функцию ЛИНЕЙН. Для этого выделим область ячеек 3×5, вызовем функцию ЛИНЕЙН и введем диапазоны данных (рис. 20.2).

	A	B	C	D	E	F	G
1	Выпуск	Капитал	Труд		lnY	lnK	lnL
2	100	100	100,0		4,60517	4,60517	4,60517
3	101	107	104,8		4,615121	4,672829	4,652054
4	112	114	110,0		4,718499	4,736198	4,70048
5	122	122	117,2		4,804021	4,804021	4,763882
6	124	131	121,9		4,820282	4,875197	4,803201
7	122	138	115,6		4,804021	4,927254	4,750136
8	143	149	125,0		4,962845	5,003946	4,828314
9	152	163	134,2		5,023881	5,09375	4,899331
10	151	176	139,9		5,01728	5,170484	4,940928
11	126	185	123,2		4,836282	5,220356	4,813809
12	155	198	142,7		5,043425	5,288267	4,960745
13	159	208	147,0		5,068904	5,337538	4,990433
14	153	216	148,1		5,030438	5,375278	4,997888
15	177	226	155,0		5,17615	5,420535	5,043425
16	184	236	156,2		5,214936	5,463832	5,051137
17	169	244	152,2		5,129899	5,497168	5,025195
18	189	266	155,8		5,241747	5,583496	5,048573
19	225	298	183,0		5,4161	5,697093	5,209486
20	227	335	197,5		5,42495	5,814131	5,285739
21	223	366	201,1		5,407172	5,902633	5,303802
22	218	387	195,9		5,384495	5,958425	5,277604
23	231	407	194,4		5,442418	6,008813	5,269918
24	179	417	146,4		5,187386	6,033086	4,986343
25	240	431	160,5		5,480639	6,066108	5,078294

Лист1 / Лист2 / Лист3 / Лист4 / Лист5

Рис. 20.1

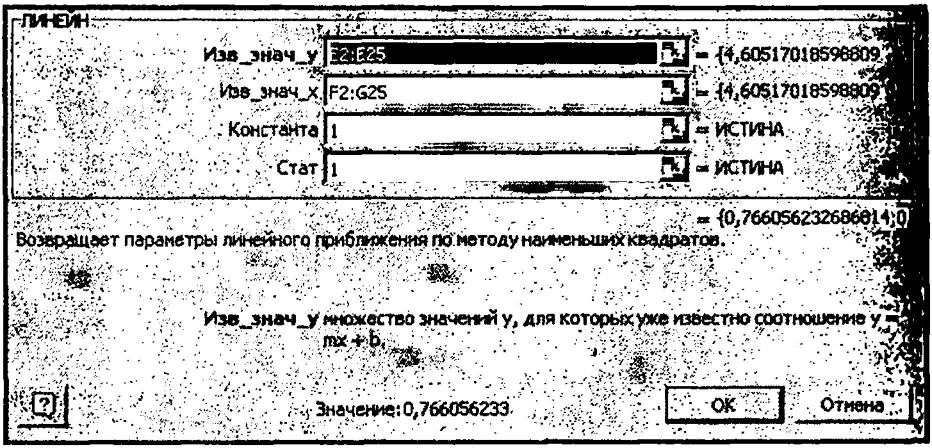
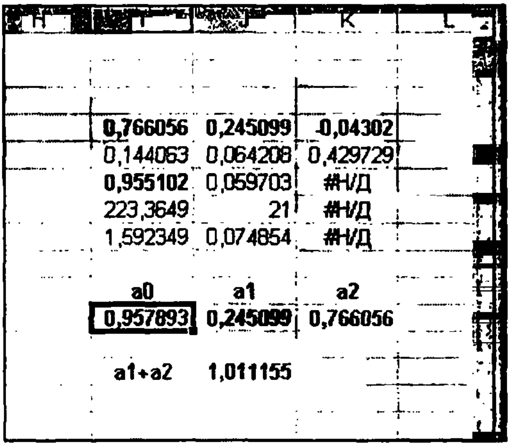


Рис. 20.2

Закончив ввод данных одновременным нажатием <Shift>+<Ctrl>+<Enter>, получим результат, показанный на рис. 20.3.



H	I	J	K	L
	0,766056	0,245099	-0,04302	
	0,144063	0,064208	0,429729	
	0,955102	0,059703	#Н/Д	
	223,3649	21	#Н/Д	
	1,592349	0,074854	#Н/Д	
	a0	a1	a2	
	0,957893	0,245099	0,766056	
	a1+a2	1,011155		

Рис. 20.3

Как видно из рис. 20.3, сумма a_1 и a_2 равна 1,011155.

20.3. Производственные функции с постоянной эластичностью замещения ресурсов

Они имеют следующий вид:

$$y = b\left(\sum_{i=1}^{n} \beta_i x_i^{-\rho}\right)^{-\delta/\rho}, \qquad (20.7)$$

где параметры b, δ, ρ, $\beta_1, \ldots, \beta_n$ положительны. Функцию можно представить в более удобной форме, эквивалентной исходной при $x_i > 0$:

$$\ln y = \ln b - \frac{\delta}{\rho} \ln \sum_{i=1}^{n} \beta_i x_i^{-\rho}. \qquad (20.8)$$

Кроме того, функцию (20.7) иногда записывают в форме, удобной для запоминания (при $\delta = 1$; $b = 1$):

$$y = \sum_{i=1}^{n} \beta_i \cdot x_i^{-\rho}.$$

В качестве примера функции типа (20.7) рассмотрим функцию с двумя производственными ресурсами и $\delta = 1$:

$$y = b(\beta_1 x_1^{-\rho} + \beta_2 x_2^{-\rho})^{-1/\rho}. \qquad (20.9)$$

Переменные y, x_1, x_2 будем интерпретировать так же, как и для степенной функции: y – конечная продукция народного хозяйства; x_1 – общее количество основных фондов, x_2 – количество трудовых ресурсов. Функция (20.9) характеризуется постоянной отдачей от увеличения масштаба производства, поэтому для нее можно рассмотреть функцию одной переменной, которая для (20.9) выглядит так:

$$\tilde{y} = b(\beta_1\tilde{x}^{-\rho} + \beta_2)^{-1/\rho}, \qquad (20.10)$$

где $\tilde{y} = y/x_2$, $\tilde{x} = x_1/x_2$.

Функция (20.10) изображена на рис. 20.4. Относительно параметров β_1 и β_2 сделано предположение, что $\beta_1 = \beta_2 = 1$. Для сравнения показана функция:

$$\tilde{y} = b\tilde{x}^{\alpha}; \ \alpha = \beta_1.$$

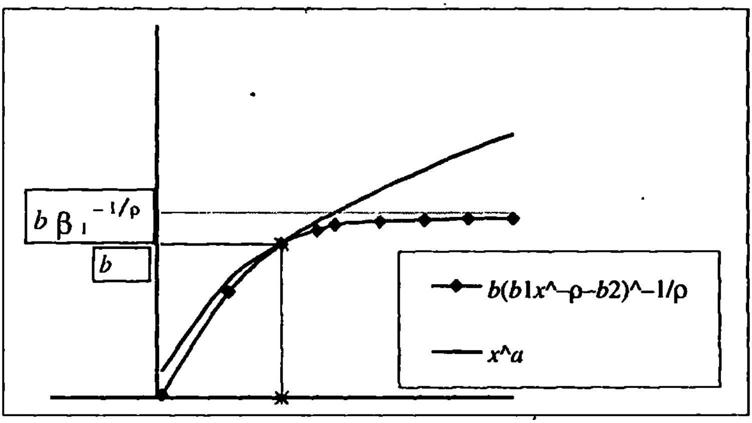


Рис. 20.4

Эти функции имеют много общего. Они равны нулю при $\tilde{x} = 0$, монотонно возрастают и вогнуты. Но есть два важных различия: предельная эффективность $\tilde{y}'(\tilde{x})$ функции (20.10) при $\tilde{x} \to 0$ не стремится к бесконечности, а при $\tilde{x} \to +\infty$ не стремится к бесконечности сама функция, т. к. ограничена асимптотой

$$\tilde{y} = b\beta_2^{-1/\rho}.$$

Поскольку $\tilde{x} = x_1/x_2$, то сильное отличие функции (20.10) от степенной функции при больших и малых значениях $\tilde{x}$ означает, что функция (20.9) отличается от степенной свойствами замещения одного ресурса другим. Для анализа свойств замещения построим изокванту функции (20.9), которая описывается соотношением:

$$y_0 = b\left(\beta_1 x_1^{-\rho} + \beta_2^{-\rho}\right)^{-1/\rho}.$$

Преобразовав это выражение, получим:

$$(y_0/b)^{-\rho} = \beta_1 x_1^{-\rho} + \beta_2 x_2^{-\rho}.$$

То есть уравнение изокванты имеет вид (рис. 20.5):

$$x_2 = \left[\frac{1}{\beta_2}\left(\left(\frac{y_0}{b}\right)^{-\rho} - \beta_1 x_1^{-\rho}\right)\right]^{-1/\rho}.$$

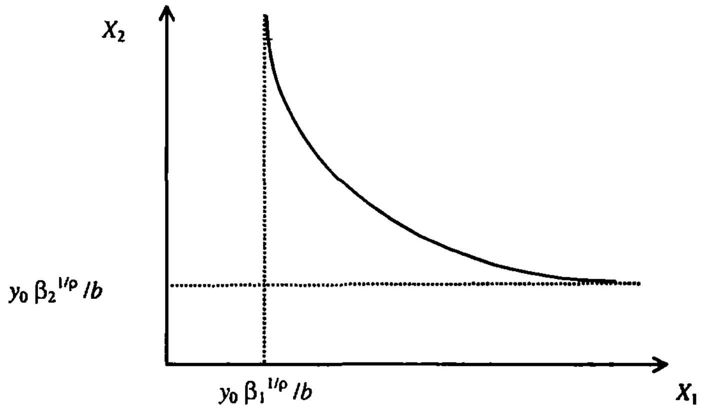


Рис. 20.5

При x → ∞ изокванта имеет асимптоту:

$$x_2 = \left(\frac{y_0}{b}\right)\beta_2^{1/\rho}.$$

Это означает, что даже при неограниченном росте количества первого ресурса для выпуска продукции в количестве y_0 необходим второй ресурс в количестве, большем $(y_0/b)\beta_2^{1/\rho}$.

Аналогичная ситуация складывается при увеличении количества второго ресурса. Таким образом, полное замещение ресурсов, характерное для степенных производственных функций, отсутствует. Рассмотрим свойства. Подсчитаем эффективность j-го ресурса. Для этого продифференцируем (20.8) по x_j:

$$\frac{1}{y}\cdot\frac{\partial y}{\partial x_j} = -\frac{\delta}{\rho}\cdot\frac{1}{\sum_{i=1}^{n}\beta_i\cdot x_i^{-\rho}}\cdot\beta_j(-\rho)\cdot x_j^{-(1+\rho)}.$$

Отсюда следует:

$$\frac{\partial y}{\partial x_j} = \frac{\delta\cdot\beta_j\cdot x_j^{-(1+\rho)}}{\sum_{i=1}^{n}\beta_i\cdot x_i^{-\rho}}\cdot y(x) > 0. \qquad (20.11)$$

На основе (22.11) можно показать, что предельная эффективность ресурса падает с ростом его объема при постоянных количествах других ресурсов.

Эластичность выпуска по j-му ресурсу имеет вид:

$$\varepsilon_j = \frac{x_j}{y} \cdot \frac{\partial y}{\partial x_j} = \frac{\delta \cdot \beta_j \cdot x_j^{-\rho}}{\sum_{j=1}^{n} \beta_i \cdot x_i^{-\rho}}. \tag{20.12}$$

Таким образом, для функции (20.7) эластичности выпусков по ресурсам, в отличие от степенных функций, уже непостоянны, т. е. отношение предельной эффективности ресурса к его средней эффективности изменяется.

При фиксированных затратах остальных ресурсов уменьшение количества j-го ресурса приводит к увеличению эластичности выпуска до величины δ, неограниченное увеличение – к падению эластичности выпуска по этому ресурсу до нуля. Поэтому отношение предельной эффективности ресурса к средней эффективности падает с ростом используемого объема ресурса.

Эластичность производства, согласно (20.12), имеет вид:

$$\varepsilon = \sum_{j=1}^{n} \varepsilon_j = \frac{\sum_{j=1}^{n} \delta\beta_j x_j^{-\rho}}{\sum_{i=1}^{n} \beta_i x_i^{-\rho}} = \delta. \tag{20.13}$$

Эластичность производства не зависит от соотношения ресурсов, как и в случае степенной функции. Из (20.13) вытекает экономический смысл параметра δ: он характеризует отдачу от увеличения масштабов производства.

Итак, четыре предположения выполняются для функций (20.7).

Рассмотрим вопрос о замещении ресурсов. Предельные нормы замещения имеют вид:

$$\gamma_{ij} = -\frac{\partial y / \partial x_j}{\partial y / \partial x_i} = -\frac{\beta_j}{\beta_i} \cdot \left(\frac{x_i}{x_j} \right)^{1+\rho}. \tag{20.14}$$

Таким образом, предельные нормы замещения зависят от отношения ресурсов, причем так же, как и в случае степенных функций, изоклинали являются плоскостями, а при пропорциональном увеличении количеств обоих ресурсов предельная норма замещения не изменяется.

Эластичность замещения будет равна:

$$\sigma_{ij} = \frac{d(\ln(x_i / x_j))}{d \ln(-\gamma_{ij})} = \frac{1}{1+\rho}. \tag{20.15}$$

Таким образом, хотя функция типа (20.7) по – прежнему имеет постоянную эластичность замещения ресурсов, эта эластичность, в отличие от степенных производственных функций, не равна 1 и меняется при изменении параметра ρ от 1 (при $\rho = 0$) и до 0 (при $\rho \to \infty$).

Из-за этого свойства производственная функция (20.7) получила название производственной функции с постоянной эластичностью замещения, или сокращенно – ПЕЗ-функции. Распространено также название CES-функция, от английского названия Constant Elastisity of Substitution.

При $\rho \to 0$ все характеристики функций с постоянной эластичностью замещения ε_j, ε_i, ε_{ij}, δ_{ij} стремятся к соответствующим характеристикам степенной производственной функции (20.1), причем между параметрами обеих производственных функций устанавливается следующее соответствие:

$$a = b,\ a_j = \frac{1}{\sum_{i=1}^{n} \beta_i} \delta\beta_j,\ j = 1,\dots,n.$$

Если $\sum\beta_i = 1$ и $\delta = 1$, то $a = b$ и $a_j = b_j$, и сама функция (20.7) стремится к функции (20.1) при $\rho \to 0$.

Таким образом, степенная производственная функция оказывается предельным случаем производственной функции с постоянной эластичностью замещения, и которая (CES), в свою очередь, является обобщением степенной производственной функции.

20.4. Оценка параметров производственной функции с постоянной эластичностью замещения ресурсов в Mathcad

Для установления параметров функций данного типа применяются методы нелинейного программирования. Рассмотрим пример установления параметров производственной функции с постоянной эластичностью замещения (CES) вида:

$$Y = ae^{\lambda t}(bK^{-\rho} + (1 - b)L^{-\rho})^{-1/\rho}.$$

Данная функция остается нелинейной даже после логарифмирования. Поэтому алгоритм решения данной задачи может быть следующим. Сначала устанавливаем параметры аппроксимационного уравнения данной функции. Для этого выражаем аппроксимационное уравнение, формируем целевую функцию как минимум суммы регрессионных остатков, задаем начальные значения параметров и применяем сначала метод Левенберга–Марквардта, а затем метод сопряженных градиентов. Далее представляем производственную функцию в логарифмическом виде, формируем целевую функцию как минимум суммы регрессионных остатков. В качестве начальных значений используем значения параметров, полученных решением аппроксимационного уравнения. Применяем метод Левенберга–Марквардта и затем метод сопряженных градиентов.

1 этап. Ввод исходных данных.

$$Y := \begin{pmatrix} 33.15 \\ 38.2 \\ 42.56 \\ 47.35 \\ 53.46 \\ 60.34 \\ 66.33 \\ 73.01 \\ 81.29 \\ 90.56 \\ 100 \\ 109.56 \\ 120.27 \\ 131.21 \\ 141.78 \\ 153.41 \\ 167.09 \\ 183.65 \\ 199.91 \\ 215.2 \end{pmatrix} \quad K := \begin{pmatrix} 33.77 \\ 37.59 \\ 42.35 \\ 46.96 \\ 52.18 \\ 58.64 \\ 65.64 \\ 72.5 \\ 80.45 \\ 89.67 \\ 100 \\ 111.59 \\ 123.93 \\ 138.1 \\ 154.31 \\ 170.42 \\ 186.13 \\ 201.73 \\ 217.68 \\ 235.1 \end{pmatrix} \quad L := \begin{pmatrix} 79.92 \\ 84.52 \\ 87.6 \\ 91.41 \\ 95.83 \\ 97.61 \\ 96.32 \\ 96.6 \\ 99.14 \\ 100.66 \\ 100 \\ 99.54 \\ 102.81 \\ 106.44 \\ 110.91 \\ 115.65 \\ 118.66 \\ 122.55 \\ 125.62 \\ 130.6 \end{pmatrix} \quad t := \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 10 \\ 11 \\ 12 \\ 13 \\ 14 \\ 15 \\ 16 \\ 17 \\ 18 \\ 19 \\ 20 \end{pmatrix}$$

2 этап. Решение аппроксимационного уравнения.

$n := \mathrm{length}(Y)$ $\qquad$ $\mathrm{ORIGIN} = 1$

$i := 1..n$ $\qquad$ $\mathrm{last}(Y) = 20$

$$Q(K,L,t,a,b,p,\lambda) := \ln(a) + \lambda \cdot t + b \cdot \ln(K) + (1-b)\cdot \ln(L) - 0.5 \cdot p \cdot b \cdot (1-b) \cdot (\ln(K) - \ln(L))^2$$

$$SSE(a,b,p,\lambda) := \sum_i \left(\ln(Y_i) - Q(K_i, L_i, t_i, a, b, p, \lambda)\right)^2$$

$a := 1$ $\quad b := 0.5$ $\quad p := 1.5$ $\quad \lambda := 0.05$

Given

$SSE(a,b,p,\lambda) = 0$ $\quad TOL = 1 \cdot 10^{-9}$ $\quad CTOL = 1 \cdot 10^{-9}$

Метод Левенберга–Марквардта

$CTOL = 1 \cdot 10^{-9}$

$$\begin{pmatrix} a \\ b \\ p \\ \lambda \end{pmatrix} := Minerr(a,b,p,\lambda)$$

$$\begin{pmatrix} a \\ b \\ p \\ \lambda \end{pmatrix} = \begin{pmatrix} 0.969 \\ 0.865 \\ 2.046 \\ 1.902 \times 10^{-3} \end{pmatrix}$$

$$SSE(a,b,p,\lambda) = 4.379 \times 10^{-3}$$

$$\begin{pmatrix} a \\ b \\ p \\ \lambda \end{pmatrix} := Minimize(SSE, a, b, p, \lambda)$$

Метод сопряженных градиентов

$$\begin{pmatrix} a \\ b \\ p \\ \lambda \end{pmatrix} = \begin{pmatrix} 0.864 \\ 0.738 \\ 1.42 \\ 0.012 \end{pmatrix}$$

$$SSE(a,b,p,\lambda) = 3.434 \times 10^{-3}$$

3 этап. Решение в логарифмированном виде.

$$H(K,L,t,a,b,p,\lambda) := \ln\left[(a)\cdot e^{\lambda\cdot t}\cdot\left[b\cdot K^{-p}+(1-b)\cdot L^{-p}\right]^{\frac{-1}{p}}\right]$$

$$SSE(a,b,p,\lambda) := \sum_i \left(\ln(Y_i) - H(K_i,L_i,t_i,a,b,p,\lambda)\right)^2$$

$$\begin{pmatrix} a \\ b \\ p \\ \lambda \end{pmatrix} := \begin{pmatrix} 0.864 \\ 0.738 \\ 1.42 \\ 0.012 \end{pmatrix}$$

Метод Левенберга–Марквардта

Given

$$SSE(a,b,p,\lambda) = 0$$

$$\begin{pmatrix} a \\ b \\ p \\ \lambda \end{pmatrix} := \mathrm{Minerr}(a,b,p,\lambda)$$

$$\begin{pmatrix} a \\ b \\ p \\ \lambda \end{pmatrix} = \begin{pmatrix} 0.852 \\ 0.733 \\ 1.559 \\ 0.013 \end{pmatrix}$$

$$SSE(a,b,p,\lambda) = 3.561\times 10^{-3}$$

$$\begin{pmatrix} a \\ b \\ p \\ \lambda \end{pmatrix} := \mathrm{Minimize}(SSE,a,b,p,\lambda)$$

Метод сопряженных градиентов

$$\begin{pmatrix} a \\ b \\ p \\ \lambda \end{pmatrix} = \begin{pmatrix} 0.78650 \\ 0.63852 \\ 1.46115 \\ 0.02063 \end{pmatrix}$$

$$SSE(a,b,p,\lambda) = 3.207\times 10^{-3}$$

4 этап. Устанавливаем эластичность замещения, коэффициент детерминации, статистику Дарбина Уотсона.

Эластичность замещения $\sigma := \dfrac{1}{p+1}$ $\sigma = 0.4063$

$$TSS(Y) := \sum_{i=1}^{n} \left(\ln\left(Y_i\right) - mean(\ln(Y))\right)^2$$

$$R2 := 1 - \frac{SSE(a,b,p,\lambda)}{TSS(Y)}$$

Коэффициент детерминации $R2 = 0.999494$

$$DW := \frac{\left[\sum_{i=2}^{n} \left\lfloor\left(\ln\left(Y_i\right) - H\left(K_i, L_i, t_i, a, b, p, \lambda\right)\right) - \left(\ln\left(Y_{i-1}\right) - H\left(K_{i-1}, L_{i-1}, t_{i-1}, a, b, p, \lambda\right)\right)\right\rfloor^2\right]}{SSE(a,b,p,\lambda)}$$

$DW = 0.819$

20.5. Производственные функции с постоянными пропорциями

Рассмотрим, что произойдет с производственной функцией (20.7) в том случае, когда параметр ρ стремится к своему другому пределу, т. е. $\rho \to \infty$.

Очевидно, что, согласно (20.15), в этом случае $\sigma \to 0$, т. е. эластичность замещения также стремится к своему другому крайнему значению.

Построим функцию

$$f_0(x) = \lim_{\rho \to \infty} f_\sigma(x),$$

где $f_\sigma(x)$ – (20.7) . $f_0(x)$ – искомая функция, где индекс 0 – подчеркивает равенство нулю ее эластичности замещения.

При $\delta = 1$ для случая двух производственных ресурсов, т. е. $x = (x_1, x_2)$:

$$f_0(x_1, x_2) = \lim_{\rho \to \infty} b(\beta_1 x_1^{-\rho} + \beta_2 x_2^{-\rho})^{-1/\rho}.$$

После преобразования получим:

при $x_1 \geq x_2$ $f_0(x_1, x_2) = bx_2$;

при $x_1 \leq x_2$ $f_0(x_1, x_2) = bx_1$.

В целом функция $f_0(x_1, x_2)$ может быть представлена в виде:

$$f_0(x_1, x_2) = b \min\{x_1, x_2\}. \qquad (20.16)$$

Для функции (20.16) построим соответствующую ей функцию $\bar{y} = b \cdot \min\{\bar{x}, 1\}$ (рис. 20.6).

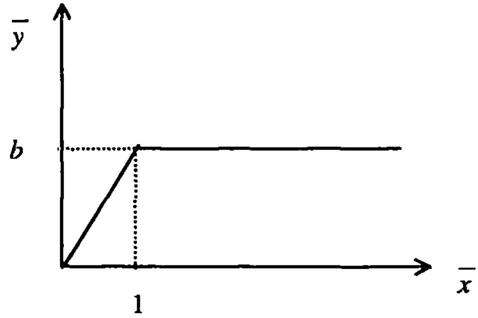


Рис. 20.6

Эта функция имеет излом в точке $x = 1$, т. е. для функции не выполняется предположение о дифференцируемости при всех положительных значениях переменных. Рассмотрим изокванты функции (20.16). Уравнение изоквант имеет вид:

$$y_0 = b \cdot \min\{x_1, x_2\}.$$

Из этого соотношения следует, что для выпуска продукции в объеме, равном y_0, достаточно использовать ресурсы в объеме:

$$x_1 = y_0/b,\ x_2 = y_0/b.$$

Увеличение количества одного из ресурсов сверх этой величины без изменения количества другого приводит к выпуску того же самого объема продукции y_0, поэтому изокванты имеют вид, показанный на рис. 20.7.

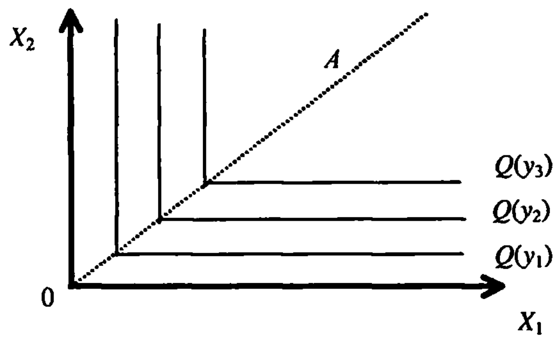


Рис. 20.7

В точках, лежащих на луче OA, на котором выполнено условие $x_1 = x_2$, находятся изломы изоквант. Особенностью полученной функции является наличие рациональных пропорций между ресурсами, задаваемыми соотношением $x_1 = x_2$. Когда количество одного ресурса превышает количество другого, избыток ресурсов пользы принести не может. Таким образом, замена одного ресурса другим здесь отсутствует. Производственные функции такого типа принято называть производственными функциями с постоянными пропорциями. Они имеют вид:

$$y = b \cdot \min\left(\frac{x_1}{a_1}, \frac{x_2}{a_2}, \dots, \frac{x_n}{a_n}\right), \tag{20.17}$$

где b, a_1, a_2 – положительные параметры.

Функция характеризуется постоянной отдачей от расширения масштабов производства. Предельная норма замещения γ при $x_1 > x_2$ равна $-\infty$, а при $x_1 < x_2$ равна нулю.

Функция имеет нулевую эластичность замещения. Функция предназначена для моделирования строго определенных технологий, не допускающих отклонений от норм используемых ресурсов и от технологических норм использования ресурсов на единицу продукции.

20.6. Некоторые виды функций выпуска

Кроме перечисленных функций выпуска, существуют и другие функции [17, 21]. Наиболее распространенными являются линейная функция, функции Аллена, функции с линейной эластичностью замены факторов, функция Солоу, ограниченная функция CES, многорежимная функция, функция линейного программирования.

Линейная функция имеет вид:

$$y = a_1x_1 + a_2x_2.$$

У данной функции предельные производительности факторов постоянны, эластичность замены факторов – бесконечна. Функция может использоваться в тех случаях, когда вклад каждого ресурса независим, например: производственная система состоит из отдельных производственных единиц, каждая из которых использует свой собственный производственный ресурс, подходящий только для этого производства.

Функция Аллена имеет вид:

$$y = a_0x_1x_2 - a_1x_1^2 - a_2x_2^2.$$

Такая функция предназначена для описания производственных процессов, в которых чрезмерный рост любого из факторов оказывает отрицательное воздействие на объем выпуска. Обычно такая функция используется для описания мелкомасштабных систем с ограниченными возможностями переработки ресурсов.

Функция с линейной эластичностью замены факторов (функция LES):

$$y = x_1^{a_0}(a_1x_1 + a_2x_2)^{a_3}.$$

Функция LES применяется для описания производственных процессов, у которых (в отличие от описываемых функцией CES) возможность замещения вовлекаемых факторов существенно зависит от их пропорций, причем при низком уровне отношений x_1/x_2 близка к единице, а с ростом отношения x_1/x_2 неограниченно возрастает. Такая ситуация возможна, например, если рост ресурсов x_1 связан с общим расширением производства, появлением множественных технологических процессов с широкими возможностями комбинирования.

Функция Солоу имеет вид:

$$y = (a_1 x_1^{a_3} + a_2 x_2^{a_4})^{a_5}.$$

Характеризуется тем, что величина процентного изменения предельной нормы замещения факторов, вызванного увеличением любого фактора на один процент, не зависит от начального уровня фактора. Эта функция может использоваться, когда влияние на объем выпуска увеличения каждого из факторов проявляется различным образом.

Ограниченная функция CES:

$$y = \min(x_1/a_1, x_2/a_2, (a_3 x_1^{a_5} + a_4 x_2^{a_5})^{a_6}).$$

Функция предназначена для выражения двухрежимного производственного процесса, в котором один из режимов характеризуется отсутствием заменяемости факторов, другой – ненулевой постоянной величиной эластичности замены. При этом переход от одного режима к другому осуществляется в зависимости от уровня лимитирующего первый режим фактора.

Многорежимная функция:

$$y = (a_{11} x_1^{a_0} + a_{21} x_2^{a_0})^{a_1} \ldots (a_{1n} x_1^{a_0} + a_{2n} x_2^{a_0})^{a_n}.$$

Одна из наиболее общих форм производственных функций. Она используется при описании процессов, в которых уровень отдачи каждой новой единицы ресурса скачкообразно меняется в зависимости от соотношения факторов. Функцию целесообразно применять при наличии информации о числе режимов и о ширине «переходной» области между режимами (чем выше a_0, тем более отчетливо выделяются режимы).

Функция линейного программирования:

$$y = \min(x_1/a_{11}, x_2/a_{12}) + \ldots + \min(x_1/a_{k1}, x_2/a_{k2}).$$

Функцию имеет смысл использовать в тех случаях, когда выпуск продукции является результатом одновременного функционирования k-фиксированных технологий, использующих одни и те же ресурсы.

Описание технического прогресса

При построении производственных функций научно-технический прогресс может быть учтен с помощью множителя $e^{\lambda t}$, где параметр λ характеризует темп прироста выпуска под влиянием научно-технического прогресса:

$$y(t) = e^{\lambda t} f(x_1(t), x_2(t)),\ t = 0, 1, \ldots, T.$$

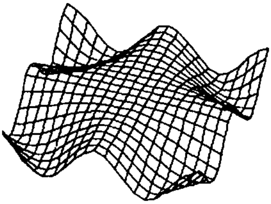
Данная производственная функция является примером динамической производственной функции. Она включает нейтральный, то есть не материализованный в одном из факторов технический прогресс. Другим подходом является выражение технического прогресса от прироста основных фондов в году t или от инвестиций в научные исследования, что эконометрически предпочтительней. В более сложных случаях технический прогресс может воздействовать непосредственно на производительность труда или капиталоотдачу:

$$y(t) = f(A(t)K, B(t)L),$$

где K – основные фонды; L – трудовые ресурсы; $A(t)$ и $B(t)$ – заданные функции времени, причем $A(t)$ описывает повышение эффективности использования основных фондов; $B(t)$ – трудовых ресурсов.

Глава 21. Функции затрат

21.1. Функции затрат и их свойства



Рассмотрим функцию выпуска $y = f(x)$ с одним продуктом и единственным ресурсом. Пусть эта функция – непрерывно дифференцируемая и удовлетворяет условиям:

$$f(0) = 0; f'(x) > 0 \text{ при } x > 0. \qquad (21.1)$$

В этом случае существует непрерывно дифференцируемая обратная функция: $x = h(y)$. Это – функция затрат [1, 21, 23, 34]. В качестве примера функции $f(x)$ можно рассмотреть функцию выпуска

$$y = x^a, x \geq 0,$$

представленную в виде функции затрат

$$x = y^{1/a}, y \geq 0.$$

Рассмотрим свойства функции затрат $x = c(y)$. Из условия $f(0) = 0$ следует, что

$$c(0) = 0, \qquad (21.2)$$

т. е. в случае отсутствия выпуска продукции тратить ресурс нет необходимости. Из условия $f'(x) > 0$ при $x > 0$ следует, что

$$c'(y) = 1/f'(x) > 0, \qquad (21.3)$$

это означает, что с ростом выпуска продукции затраты ресурса растут. Функцию $h'(y)$ принято называть предельными затратами ресурсов. Как видно из (21.3), предельные затраты ресурса обратно пропорциональны предельной эффективности ресурсов.

Предположим, что для функции $f(x)$ выполнено предположение об убывании предельной эффективности ресурса, т. е. $f''(x) < 0$. Тогда из (21.3) получаем, что функция $c'(y)$ монотонно возрастает и

$$c''(y) > 0. \qquad (21.4)$$

Введем понятие средних удельных затрат ресурса: $g(y) = x/y$. Отношение предельных затрат ресурса к средним удовлетворяет соотношению:

$$c'(y)/g(y) = (\partial x/\partial y)\,(y/x) = 1/a(x),$$

где $a(x)$ – эластичность выпуска по ресурсу для $f(x)$.

При выполнении предположения о том, что $f''(x) < 0$, получаем, что $a(x) < 1$. Поэтому в таком случае предельные затраты ресурса больше средних. Для функции затрат $x = y^{1/a}$, порождаемой функцией выпуска $y = x^a$, получаем:

$$c(0) = 0; g(y) = y^{(1-a)/a},$$

$$c'(y) = 1/a\, y^{(1-a)/a} > 0,$$

$$c''(y) = (1-a)/a\, y^{(1-2a)/a} > 0.$$

Графики перечисленных функций для функции затрат вида $x = y^{1/a}$ при $a = 0{,}5$ приведены на рис. 21.1.

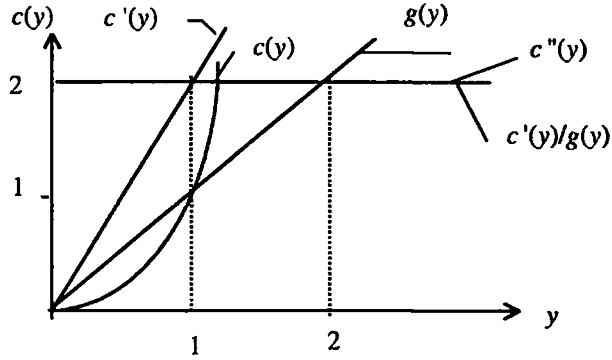


Рис. 21.1

Функция выпуска с одним продуктом и единственным ресурсом и соответствующая ей функция затрат эквивалентны: замена одной из них на другую не может привести к новым представлениям или дать преимущество при моделировании производственных единиц. Иное дело в случае нескольких ресурсов. Функция затрат для нескольких ресурсов и одного продукта имеет следующий вид:

$$x_i = c_i(y), \text{ где } i = 1,\dots, n. \qquad (21.5)$$

Потребление каждого из ресурсов задается однозначной функцией количества выпускаемой продукции. Замещение ресурсов здесь невозможно. Ресурсы в функции затрат являются взаимодополняющими, т. е. объемы потребления ресурсов определяются жесткими технологическими условиями, и нехватка хотя бы одного из ресурсов не позволяет полностью использовать остальные ресурсы. Таким образом, описание производства с помощью функции затрат принципиально отличается от его описания с помощью функции выпуска, где замещение ресурсов допустимо.

Свойства функции затрат

Относительно функции затрат (21.5) формулируются предположения, близкие по характеру к свойствам функции затрат с одним ресурсом (21.2). Прежде всего для простоты предполагается, что функция затрат является дважды непрерывно дифференцируемой. По аналогии с (21.2) считается, что, во-первых:

$$c_i(0) = 0, \quad i = 1,\dots, n, \qquad (21.6)$$

т. е. при отсутствии производства ресурсы не нужны, и, во-вторых:

$$c'_i(y) \geq 0, \quad i = 1,\dots, n, \qquad (21.7)$$

т. е. рост производства требует увеличения количества используемых ресурсов.

Иногда делается следующее предположение:

$$c''_i(y) > 0, \quad i = 1, \dots, n, \qquad (21.8)$$

т. е. предельные затраты с ростом производства растут.

Часто за счет концентрации производства имеется противоположный эффект: с ростом объема производства предельные затраты падают; в таких случаях вместо предположения (21.8) используется противоположное:

$$c''_i(y) < 0, \quad i = 1,\dots, n, \qquad (21.9)$$

означающее невозрастание предельных затрат с ростом производства.

В этом случае предельные затраты $c'(y)$ оказываются не больше средних, причем средние затраты также убывают с ростом выпуска продукции. Встречаются также функции затрат, для которых в некоторых диапазонах затрат выполняется соотношение (21.8), в других – соотношение (21.9). Такая ситуация может возникнуть, если при росте выпуска (y) сначала основное влияние оказывает экономия ресурсов за счет концентрации производства, а при слишком большом выпуске эффективность начинает падать.

21.2. Некоторые виды функции затрат

Наиболее простая функция затрат – это линейная однородная функция:

$$x_i = a_i y;\ i = 1,\dots, n, \qquad (21.10)$$

где a – неотрицательный параметр.

Для функции (21.10) выполняются предположения (21.6) и (21.7), при этом средние затраты $g_i(y)$ и предельные затраты $c_i'(y)$ совпадают и равны a_i.

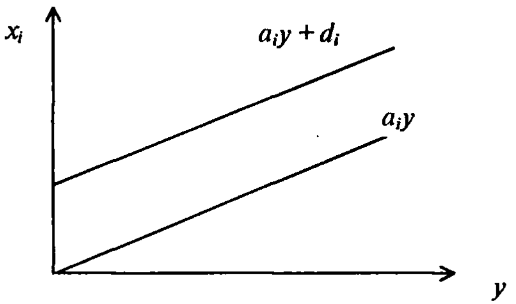


Рис. 21.2

Близка к линейной функции затрат линейная неоднородная функция (рис. 21.2):

$$x_i = a_i y + d_i, \qquad (21.11)$$

где d_i – неотрицательный параметр.

Если все $d_i = 0$, то функция (21.11) совпадает с (21.10), в противном случае затраты не равны нулю даже тогда, когда продукция не выпускается. Эта функция может быть использована, когда приходится заранее делать капиталовложения, объем которых не зависит от масштабов производства. Так как для функции (21.11), по крайней мере для некоторых ресурсов, имеем: $c_i(0) = d_i > 0$, то предположение (21.6) здесь не выполняется. Поскольку $c_i'(y) = a_i \geq 0$, то предположение (21.7) выполняется по-прежнему.

При анализе функции (21.11) представляет интерес соотношение между предельными и средними затратами. Здесь средние затраты имеют вид:

$$g_i(y) = x_i/y = a_i + d_i/y, \text{ поэтому } g_i(y) > c_i'(y) \text{ при } d_i > 0$$

$$\text{и } \lim_{y\to\infty} g_i(y) = a_i = c_i'(y),$$

т. е. средние затраты для функции (21.11) превосходят предельные и стремятся к ним при стремлении выпуска к бесконечности.

Для того чтобы не нарушать предположение (21.6) о нулевых затратах при нулевом выпуске, иногда вместо функции (21.11) используют близкую к ней функцию затрат вида:

$$x_i = \begin{cases} d_i + a_i y & \Rightarrow \text{при } y > 0; \\ 0 & \Rightarrow \text{при } y = 0. \end{cases} \qquad (21.12)$$

Эта функция обладает существенным недостатком: она имеет разрыв в точке 0, что затрудняет исследование моделей.

Функции (21.11) и (21.12) применяются на практике достаточно часто благодаря тому, что они хорошо выражают закономерности производства во многих экономических исследованиях.

В качестве функции затрат, характеризующейся возрастающими или убывающими предельными затратами ресурсов, можно использовать степенную функцию затрат:

$$x_i = a_i y^{\alpha_i},\ i = 1,.., n, \qquad (21.13)$$

где a и α – положительные параметры. Представим две такие функции в Mathcad (рис. 21.3).

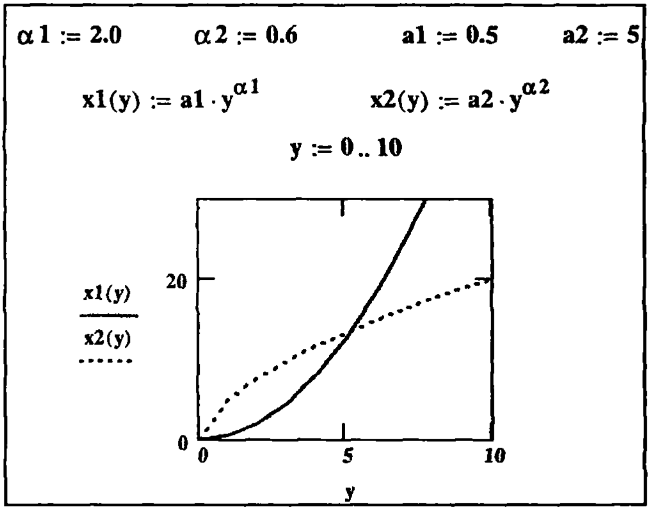


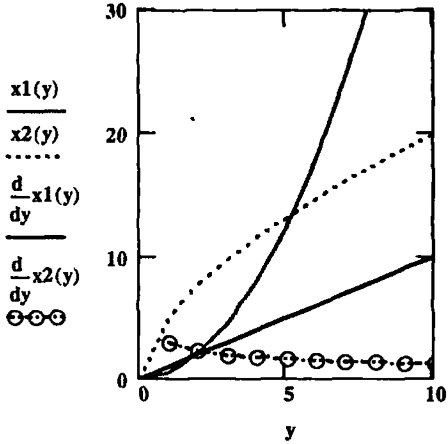
Рис. 21.3

Для данной функции:

$$c_i(0) = 0;$$
$$c_i'(y) = a_i \alpha_i y^{\alpha - 1} > 0;$$
$$c_i''(y) = a_i \alpha_i (\alpha_i - 1)\, y^{\alpha_i - 2};$$
$$g_i(y) = a_i y^{\alpha_i - 1}.$$

Если $\alpha_i < 1$, то эта функция с убывающими предельными затратами. Для нее: $c_i'(y) < g_i(y)$, т. е. предельные затраты меньше средних.

Если $\alpha_i > 1$, то эта функция с возрастающими затратами. Для нее: $c_i'(y) > g_i(y)$ (рис. 21.4)



$$g1(y) := \frac{x1(y)}{y} \qquad g2(y) := \frac{x2(y)}{y}$$

$$y := 1..10$$

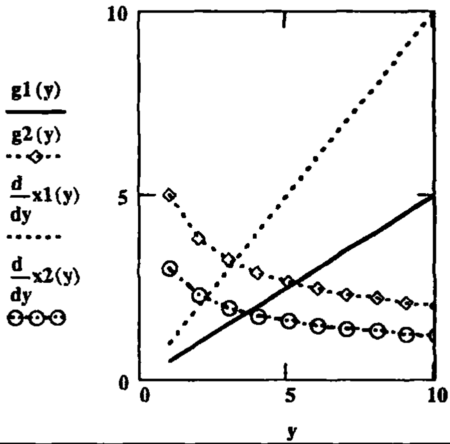


Рис. 21.4

В зависимости от свойств моделируемой производственной единицы может быть выбрана та или иная величина α_i.

В отличие от функции выпуска, которая обычно используется для описания сложных производственных единиц, функция затрат чаще всего применяется для описания производства в относительно простых экономических системах. Разнообразие производственных объектов такого типа приводит к тому, что встречается большее число различных типов функции затрат. Более того, при описании одной и той же производственной единицы могут использоваться различные функции затрат для ресурсов разных типов. Так, в некоторых моделях затраты сырьевых ресурсов выражаются линейными функциями типа (21.10), а затраты трудовых ресурсов и основных фондов – степенными функциями, характеризующими экономию затрат, связанную с увеличением масштабов производства.

Теория предприятия часто рассматривается с позиции функции затрат, принятой в качестве первичного понятия. Такой подход значительно упрощает проведение анализа. Однако он может быть подвергнут критике с двух позиций.

С одной стороны, соотношение между стоимостью затрачиваемых ресурсов и произведенным количеством зависит от цен p_i различных ресурсов, так что функция затрат изменяется при изменении этих цен. Производственная функция представляет собой, таким образом, более фундаментальное понятие, т. к. отражает технологические ограничения независимо от системы цен.

С другой стороны, теория предприятия, построенная на основе анализа затрат, плохо вписывается в теорию общего равновесия, в которой цены рассматриваются как эндогенные, а не являются определенными заранее.

Тем не менее, рассмотрение функций затрат позволяет вскрыть ряд интересных классических свойств.

Рассмотрим степенную производственную функцию с двумя ресурсами (x_1 и x_2) вида:

$$f = a_0 x_1^{a_1} x_2^{a_2},$$

при этом затраты определяются как:

$$C(x_1, x_2) = p_1 x_1 + p_2 x_2,$$

и будем считать, что рынки ресурсов подчиняются законам совершенной конкуренции, так что цены p_1 и p_2 заданы для предприятия экзогенно.

Для того чтобы определить функцию затрат, мы должны сначала отыскать комбинацию ресурсов, которые позволяют осуществить производство заданного количества f_h продукта с наименьшими затратами, т. е. максимизировать прибыль при ограничении $f = f_h$.

Применим метод Лагранжа. Функция Лагранжа будет иметь вид:

$$L(x_1, x_2, \lambda) = p_1 x_1 + p_2 x_2 + \lambda(f(x_1, x_2) - f(x_1^h, x_2^h)).$$

Возьмем первые производные и приравняем их к нулю:

$$\frac{\partial L}{\partial x_1} = p_1 + \lambda \frac{\partial f}{\partial x_1};$$

$$\frac{\partial L}{\partial x_2} = p_2 + \lambda \frac{\partial f}{\partial x_2}; \qquad (21.14)$$

$$\frac{\partial L}{\partial \lambda} = f(x_1, x_2) - f(x_1^h, x_2^h).$$

Из первых двух уравнений получаем, что предельные нормы замещения между ресурсами равны отношениям между ценами этих ресурсов:

$$\frac{\partial f}{\partial x_1} \bigg/ \frac{\partial f}{\partial x_2} = \frac{p_1}{p_2}.$$

Предельные нормы замещения для данной производственной функции равны:

$$\frac{\partial f}{\partial x_1} \Big/ \frac{\partial f}{\partial x_2} = \frac{a_1 x_2}{a_2 x_1}.$$

Откуда

$$x_1 = \frac{a_1}{a_2}\frac{p_2}{p_1}x_2.$$

Тогда

$$C = p_2 x_2(a_1 + a_2)/a_2.$$

Из функции выпуска будем иметь:

$$x_2 = \left(\frac{a_2 p_1}{a_1 p_2}\right)^{\frac{a_1}{a_1+a_2}}\left(\frac{f}{a_0}\right)^{\frac{1}{a_1+a_2}}.$$

Теперь получаем окончательное выражение для функции затрат от выпуска:

$$C(f) = p_2\left(\frac{a_1 + a_2}{a_2}\right)\left(\frac{a_2 p_1}{a_1 p_2}\right)^{\frac{a_1}{a_1+a_2}}\left(\frac{f}{a_0}\right)^{\frac{1}{a_1+a_2}}.$$

Предельные затраты будут равны:

$$\frac{dC}{df} = \frac{p_2}{a_2}\frac{\left[\left(\frac{a_2 p_1}{a_1 p_2}\right)^{a_1}\frac{f}{a_0}\right]^{\frac{1}{a_1+a_2}}}{f}$$

или после преобразований:

$$\frac{dC}{df} = \left(\frac{p_1}{a_1}\right)^{\frac{a_1}{a_1+a_2}}\left(\frac{p_2}{a_2}\right)^{\frac{a_2}{a_1+a_2}} a_0^{\left(-\frac{1}{a_1+a_2}\right)} f^{\frac{1-a_1-a_2}{a_1+a_2}}.$$

Если выразить прибыль как

$$PR = P_0 f - C(f),$$

то условием первого порядка нахождения максимума прибыли в условиях конкурентного равновесия будет равенство предельных затрат цене выпускаемой продукции:

$$P_0 = C'f.$$

Для выполнения условий второго порядка необходимо, чтобы вторая производная прибыли была отрицательная или равна нулю, т. е. чтобы предельные затраты были возрастающими или постоянными.

В точке равновесия множитель Лагранжа λ равен предельным затратам.

Предположение о невозрастании предельной выручки влечет за собой постоянство или возрастание предельных затрат.

Кривая затрат, соответствующая производственной функции, предполагающей невозрастание предельной выработки, вогнута вверх. Классическая кривая, представляющая функцию затрат, вначале вогнута вниз. Эта часть кривой соответствует области значений выпуска, для которых неделимость играет существенную роль, а предельная выработка возрастает.

13 Зак. 1010

Заметим также, что предельные затраты постоянны, если производственная функция удовлетворяет гипотезе постоянства от изменения масштаба.

Функция предложения может быть получена приравниванием предельных затрат к цене продукта, что после преобразования дает:

$$f(P_0)=a_0^{\frac{1}{1-a_1-a_2}}\left(\frac{a_1}{p_1}\right)^{\frac{a_1}{1-a_1-a_2}}\left(\frac{a_2}{p_2}\right)^{\frac{a_2}{1-a_1-a_2}}P_0^{\frac{a_1+a_2}{1-a_1-a_2}}.$$

21.3. Краткосрочные и долгосрочные решения

В некоторых случаях предприятие выбирает объемы использования не всех своих ресурсов, а только некоторых из них; использование других предопределено ранее. Таким образом, часто для одного и того же предприятия различают решения на длительный период, касающиеся всей организации производства, выбора оборудования и технологических процессов, а также решения на короткий период, касающиеся использования уже имеющихся в распоряжении производственных мощностей. Для решения на короткий период ресурсы задаются в соответствии с имеющимся оборудованием.

Рассмотрим вопрос на примере функции затрат, полученной из степенной производственной функции. Долгосрочную функцию затрат, в которой капитал также является переменной величиной, мы уже установили. Вид удельных средних и приростных затрат, т. е. на единицу выпуска, показан на рис. 21.5.

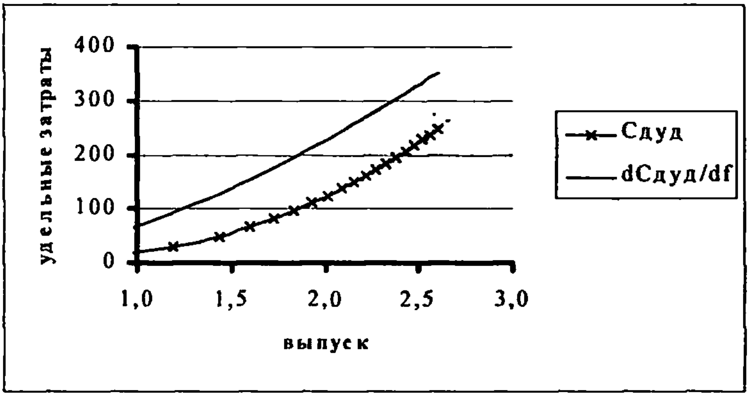


Рис. 21.5

Используя аналогичный метод, мы можем получить функции затрат для краткосрочных решений. В этом случае капитал будем рассматривать как постоянную величину. Тогда функция затрат для короткого периода из степенной производственной функции будет равна:

$$C_k(f)=p_1K+\left(\frac{p_2K^{-a_1}}{a_0}\right)^{\frac{1}{a_2}}f_k^{\frac{1}{a_2}},$$

где f_k – соответствующая функция выпуска:

$$f_k = a_0 x_1^{a_1} x_2^{a_2} \text{ при } x_1 < K;$$

$$f_k = a_0 K^{a_1} x_2^{a_2} \text{ при } x_1 > K.$$

Характерный вид функций затрат приведен на рис. 21.6.

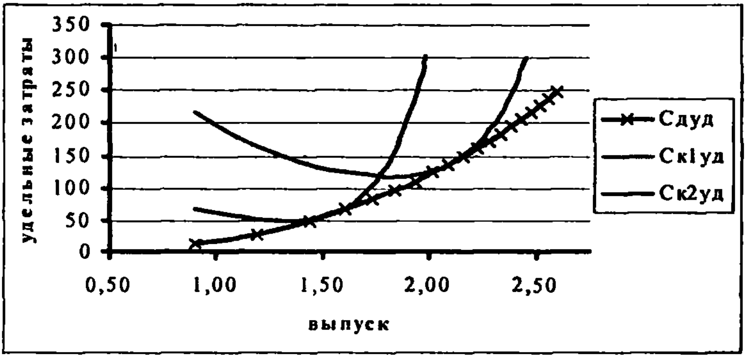


Рис. 21.6

Предельные затраты для короткого периода будут равны:

$$\frac{dC}{df} = \frac{1}{a_2}\left(\frac{p_2 K^{-a_1}}{a_0}\right)^{\frac{1}{a_2}} f_k^{\left(\frac{1}{a_2}-1\right)}.$$

Вид предельных затрат представлен на рис. 21.7.

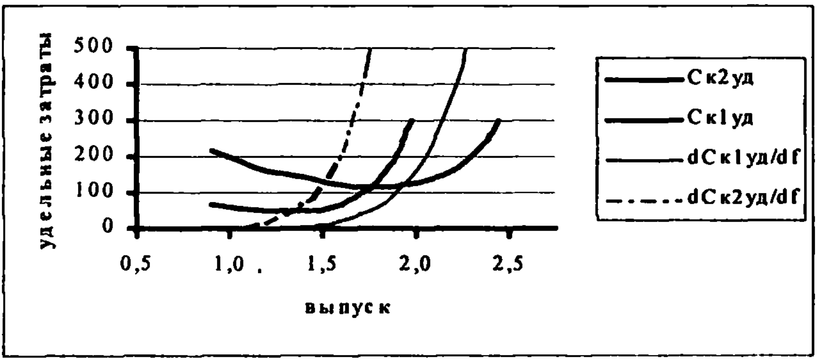


Рис. 21.7

Предельные затраты на короткий период равны также равновесному значению множителя Лагранжа λ^*.

Теория равновесия предприятия показана на рис. 21.8, где представлены различные функции затрат как функции аргумента f.

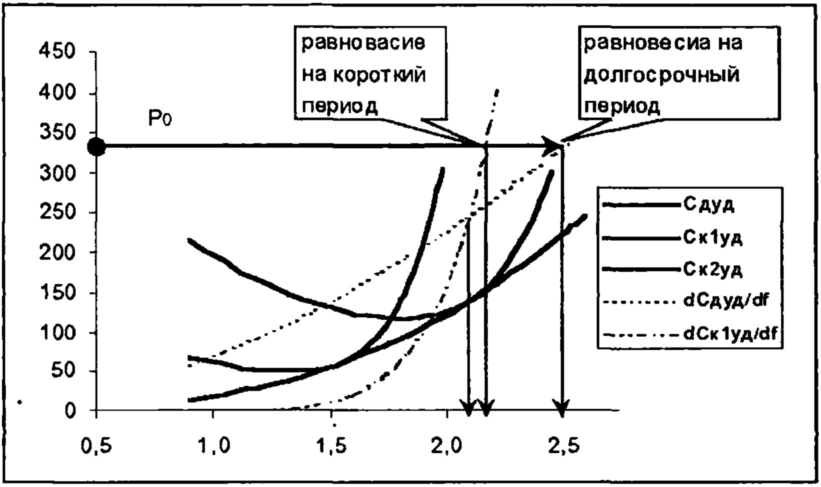


Рис. 21.8

Значение долгосрочного равновесия при цене продукции P_0 определяется как абсцисса точки пересечения цены и функции предельных затрат долгосрочного периода. Значение равновесия на короткий период определяется как абсцисса точки пересечения функции предельных затрат на короткий период и цены продукции P_0.

Кривые долгосрочных и краткосрочных предельных затрат обычно имеют общую точку, соответствующую значению f_h, для которой решение системы (21.14), определяющее долгосрочные затраты, дает для f значение f_h. В этой точке также удовлетворяются решения аналогичной системы для краткосрочных затрат. Поэтому в ней краткосрочные и долгосрочные предельные затраты равны, и они равны равновесному множителю Лагранжа. Также в этой точке кривые краткосрочных и долгосрочных средних затрат касаются друг друга. Этот результат может оказаться очевидным. Так как существующее оборудование совпадает с тем, которое предприятие выбирает на длительный период при тех же ценах, в этой точке долгосрочное и краткосрочное равновесие должны совпадать.

Кривая средних затрат за длительный период является огибающей кривых средних затрат за короткий период. Ни в каком случае краткосрочные затраты не могут быть меньше долгосрочных.

На практике получила большое распространение квадратичная регрессия, определяющая зависимость суммарных затрат от объема выпуска:

$$C(f) = \gamma_0 + \gamma_1 f + \gamma_2 f^2 + \varepsilon,$$

где γ_0, γ_1, γ_2 – параметры, устанавливаемые корреляционно-регрессионным анализом.

Глава 22. Производственные функции в экономическом анализе хозяйственной деятельности

22.1. Использование производственной функции в комплексном анализе хозяйственной деятельности

Одной нз основных задач экономнческого аналнза является качественное н количественное исследование влияния факторов на обобщающие экономические показатели. Эта задача решается в три этапа.

1. Формирование факторной системы, т. е. множества показателей, оказывающих наиболее существенное влияние на обобщающий показатель в анализируемом периоде.

2. Построение математической модели зависимости уровня обобщающего показателя от уровней показателей – факторов.

3. Количественная оценка влияния каждого из факторов или их группы на изменение обобщающего показателя.

Наиболее простым при этом является случай, когда обобщающий показатель выражается через факторы с помощью детерминированной *функциональной* зависимости. Например, если y – обобщающий показатель производительности труда, а факторную систему составляют показатели фондовооруженности (x_1) и фондоотдачи (x_2), то зависимость между ними выражается соотношением $y = x_1x_2$. Однако функциональное факторное разложение при заданных $y, x_1,...,x_n$ существует далеко не всегда. В большинстве случаев приходится строить статистические модели, приближенно выражающие факторную зависимость на определенном промежутке времени. Производственная функция $y = f(x_1,..., x_n)$, по существу, и представляет собой факторное разложение одного из обобщающих показателей – объема выпуска продукции. Построение этой функции дает возможность проводить анализ и оценку влияния факторов на динамику объема выпуска за определенный период [17].

Для оценки этого влияния сначала уточняется показатель, характеризующий итоговую динамику изменения y. Обычно в качестве такого показателя берется лнбо прирост за период $\Delta y = y(t_{кон}) - y(t_{нач})$, либо индекс роста $\text{ind}y = y(t_{кон})/y(t_{нач})$, либо темп роста $\tau(y) = \{y(t_{кон}) - y(t_{нач})\}/y(t_{нач})$.

В зависимости от этого выбора меняется и понятие «вклада» i-го фактора в итог изменения y.

При естественных предпосылках «вклад» A_i каждого i-го фактора в прирост Δy следует искать в виде слагаемых Δy, в сумме дающих велнчину Δy. Соответственно, вклады B_i i-го фактора в индекс $\text{ind}y$ следует считать сомножителями $\text{ind}y$, произведение которых равно $\text{ind}y$. Вклады C_i i-го фактора в темп роста $\tau(y)$ связаны с $\tau(y)$ формулой $\tau(y) = (C_1 + 1) \dots (C_n + 1) - 1$.

Если $y = f(x_1,...,x_n)$ и заданы траектории изменения каждого фактора $x_i(t)$, а сами факторы независимы, то

$$A_i = \int_{t_{нач}}^{t_{кон}} \frac{\partial f}{\partial x_i} \frac{dx_i}{dt} dt;$$

$$B_i = \exp \int_{t_{нач}}^{t_{кон}} \frac{\partial \ln f}{\partial x_i} \frac{dx_i}{dt} dt;$$

$$C_i = \exp \int_{t_{нач}}^{t_{кон}} \frac{\partial \ln f}{\partial x_i} \frac{dx_i}{dt} dt - 1.$$

В частности, если $y = a_0 x_1^{a_1} x_2^{a_2}$, а траектории изменения факторов в промежутке $t_{нач} \le t \le t_{кон}$ могут быть представлены в виде степенных функций $x_1(t) = b_1 t^{\beta_1}$, $x_2(t) = b_2 t^{\beta_2}$, то

$$A_1 = \frac{a_0 a_1 \beta_1 b_1^{a_1} b_2^{a_2}}{\beta_1 a_1 + \beta_2 a_2} (t_{кон}^{\beta_1 a_1 + \beta_2 a_2} - t_{нач}^{\beta_1 a_1 + \beta_2 a_2});$$

$$A_2 = \frac{a_0 a_2 \beta_2 b_1^{a_1} b_2^{a_2}}{\beta_1 a_1 + \beta_2 a_2} (t_{кон}^{\beta_1 a l_1 + \beta_2 a_2} - t_{нач}^{\beta_1 a_1 + \beta_2 a_2});$$

$$B_1 = \left(\frac{t_{кон}}{t_{нач}}\right)^{\beta_1 a_1}; \qquad B_2 = \left(\frac{t_{кон}}{t_{нач}}\right)^{\beta_2 a_2};$$

$$C_1 = \frac{t_{кон}^{\beta_1 a_1} - t_{нач}^{\beta_1 a_1}}{t_{нач}^{\beta_1 a_1}}; \qquad C_2 = \frac{t_{кон}^{\beta_2 a_2} - t_{нач}^{\beta_2 a_2}}{t_{нач}^{\beta_2 a_2}}.$$

Выражение

$$\Delta y = A_1 + ... + A_n = \int_{t_{нач}}^{t_{кон}} \frac{\partial f}{\partial x_1} \frac{dx_1}{dt} + ... + \int_{t_{нач}}^{t_{кон}} \frac{\partial f}{\partial x_n} \frac{dx_n}{dt}$$

для вклада i-го фактора в прирост результирующего показателя называется интегральным разложением прироста.

Для вычисления компонентов A_i, B_i, C_i интегрального разложения прироста, индекса и темпа роста показателя необходимо знать частные производные функции f и производные функции $x_i(t)$ на отрезке $[t_0, t_1]$. Кроме того, необходим метод определения величины определенного интеграла.

Рассмотрим конкретные задачи комплексного экономического анализа хозяйственной деятельности, в которых целесообразно применение производственных функций.

1. Анализ использования трудовых ресурсов, средств и предметов труда. Обычно для исследования эффективности использования этих факторов производства привлекаются показатели производительности труда, фондоотдачи, оборачиваемости оборотных средств. Однако движение каждого из этих показателей само по себе не полностью характеризует интенсификацию использования ресурсов.

Рассмотрим следующий условный пример. Предположим, что за некоторый период времени производительность труда предприятия (выработка на одного работающего) выросла в два раза. Можно ли утверждать, что трудовые ресурсы стали использоваться в два раза эффективнее? Если бы остальные ресурсы предприятия не изменились за это время, то да. Если же увеличение объема производства (в неизменных ценах и при неизменной в целом номенклатуре) произошло, например, за счет значительного увеличения основных фондов, то положительная динамика производительности труда будет сопровождаться падением фондоотдачи и не будет характеризовать интенсификацию использования трудовых ресурсов.

Для адекватного описания динамики производительности труда следовало бы в общем приросте объема выпуска выделить ту часть, которая непосредственно связана с приростом трудовых ресурсов. Отношение этой части прироста выпуска к приросту трудовых ресурсов и характеризует производительность дополнительно вовлеченных в производство трудовых ресурсов. Этот показатель допускает сравнение как с аналогичными показателями, относящимися к другим периодам, так и со средней выработкой одного работающего.

Аналогичным образом показатели, характеризующие динамику эффективности использования предметов труда, определяются как отношение доли прироста выпуска, обусловленной приростом стоимости вовлеченных в производство материалов, к величине этого прироста. Допустимо также использовать в качестве показателя объема предметов труда размер оборотных средств предприятия. В свою очередь, соответствующий динамический показатель фондоотдачи строится как отношение части прироста выпуска, обусловленной приростом основных фондов, к величине этого прироста.

Таким образом, в указанных случаях для построения динамической характеристики влияния факторов в течение периода времени необходимо получить пофакторное разложение прироста выпуска. Средством решения этой задачи является построение производственной функции и определение на ее основе вклада каждого из факторов в прирост объема выпуска продукции.

2. Анализ выпуска продукции. Одной из главных задач анализа выпуска продукции является оценка влияния производственных факторов на объем выпуска. Без применения производственной функции такой анализ для каждой группы факторов можно выполнить только в предположении неизменности наличных ресурсов других групп факторов, что на практике выполняется очень редко. Получение пофакторного разложения прироста объема выпуска дает возможность решать эту задачу в общем случае.

3. Анализ динамики технологического уровня производства. Для отдельных, не слишком крупных, предприятий технологические сдвиги являются следствием вполне определенных организационно-технических мероприятий и прослеживаются непосредственно. Существует целый ряд детализированных показателей, отражающих как уровень технологии, так и динамику. Для анализа технологии на уровне крупных производственных предприятий и отраслей, где состав, очередность и сроки выполнения мероприятий по совершенствованию технологии практически невыделимы, о динамике технологии можно судить в основном лишь по соотношению агрегированных показателей производства. На этом уровне технология – способ переработки сырья и материалов в продукцию – от ражается в коэффициентах в виде производственной функции, связывающей агрегированные ресурсные и выходные показатели производства.

Производственная функция в этом случае строится в динамическом варианте и явно зависит от времени. Относительно характера этой зависимости возможны два предположения. В первом случае предполагается, что смена одной технологии другой происходит дискретно. Это выражается в том, что зависимость производственной функции от времени является кусочно-постоянной. Для определения точек разрыва и высоты перепада уровней используются тонкие вероятностные методы, основанные на возможности выяснить, могут ли коэффициенты производственных функций, построенных по данным различных периодов, относиться к одной и той же генеральной совокупности. Таким образом, производится периодизация всего анализируемого промежутка времени, отражающая существенные сдвиги технологии.

Альтернативная гипотеза основывается на непрерывной зависимости производственной функции от времени, выражающейся в росте эффективности ее факторов. Здесь возникают различные типы технического прогресса, в частности различные типы нейтрального прогресса. В рамках этой гипотезы проблема разграничения эволюционных и революционных сдвигов в техническом уровне одного и того же объекта практически неразрешима, однако появляется возможность классификации объектов по характеру технологического дрейфа.

4. Анализ себестоимости и прибыли. Здесь могут быть применены функции удельных затрат и прибыли. Кроме того, методы получения интегрального разложения прироста, индекса и темпа роста удельных затрат позволяют оценить вклад каждого из факторов в динамику этих показателей.

22.2. Использование производственных функций в сравнительном экономическом анализе

Производственная функция представляет собой наиболее обобщенную модель функционирования экономического объекта. Поэтому ее можно использовать для сравнения эффективности работы различных объектов. Обычно сравнение работы двух предприятий за некий период базируется на попарном сопоставлении показателей производительности труда, фондоотдачи, оборачиваемости оборотных средств, материалоемкости. Показатель сравнительной экономии ресурса x на предприятии Б по сравнению с предприятием А за один и тот же период (или на одном и том же предприятии за разные периоды) вычисляется по формуле:

$$Э_x = x_А y_Б / y_А - x_Б,$$

где $x_А$, $x_Б$ – количества используемого ресурса на предприятиях А и Б; $y_А$, $y_Б$ – соответствующие объемы выпуска продукции. Уменьшаемое в этой формуле показывает количество ресурса x, необходимое для производства продукции $y_Б$ на предприятии Б при условии, что ресурсоемкость производства такая же, как и на предприятии А. Вычитаемое выражает количество ресурса, необходимое для выпуска продукции $y_Б$ в условиях производства на предприятии Б. Если $Э > 0$, то эффективность использования ресурса x на предприятии Б выше, чем на предприятии А.

Данная формула построена на сравнении реальной ситуации на предприятии Б с условной ситуацией, когда «технология» предприятия А переносится на предприятие Б. Понятие технологии при этом фактически сводится к ресурсоемкости. Эта формула может быть получена из факторного разложения величины $x_А - x_Б$ в сумму двух слагаемых:

$$x_А - x_Б = (x_А y_Б/y_А - x_Б) + (x_А - x_А y_Б/y_А),$$

первое из которых выражает вклад в $x_А - x_Б$ изменения ресурсоемкости продукции, второе – вклад изменения объема выпуска. Здесь используется тождество $x = y \cdot x/y$ и предполагается следующая условная последовательность перехода от ситуации предприятия А к ситуации предприятия Б: сначала ресурсоемкость при постоянном выпуске изменяется от $x_Б/y_Б$ до $x_А/y_А$, затем объем выпуска при постоянной ресурсоемкости изменяется от $y_Б$ до $y_А$. Поскольку данные для обоснования последнего предположения, также как и для обратной последовательности, обычно отсутствуют, интегральный метод рекомендует более сбалансированную формулу для экономии ресурса x:

$$Э_x^{инт} = 1/2(x_А \cdot y_Б/y_А - x_Б + x_А - x_Б \cdot y_А/y_Б).$$

Эта формула, однако, также не может использоваться в большинстве случаев для сравнения эффективности работы двух предприятий, поскольку в ней фигурирует лишь один вид ресурсов. Часто бывает, что производительность труда выше на одном предприятии, а фондоотдача – на другом. В таких случаях необходим показатель совокупной эффективности использования основных производственных ресурсов.

Для получения такого показателя рассмотрим вместо факторного разложения прироста (экономии) ресурса $x_А - x_Б$ аналогичное разложение прироста объема производства $y_Б - y_А$. Согласно интегральному методу это разложение имеет вид:

$$y_Б - y_А = 1/2(x_А \cdot y_Б/y_А - y_Б + y_А - x_Б \cdot y_А/y_Б) + 1/2(x_Б \cdot y_А/y_Б - y_А + y_Б - x_А \cdot y_Б/y_А),$$

где первое слагаемое выражает вклад в $y_Б - y_А$ изменения ресурсоотдачи. По сути дела, здесь суммируются результаты сравнения использования ресурсов $x_А$ в технологии предприятия Б и ресурсов $x_Б$ в технологии предприятия А. Технология здесь, как и в предыдущих формулах, характеризуется показателем ресурсоотдачи. Именно выражение

$$Э = 1/2(x_А \cdot y_Б/y_А - y_А + y_Б - x_Б \cdot y_А/y_Б)$$

допускает обобщение на случай нескольких ресурсов производства. Пусть $x_{А,Б} = (x_1^{А,Б}, ..., x_n^{А,Б})$ – вектор ресурсов предприятий А и Б, $f_{А,Б}(x) = f_{А,Б}(x_1, ..., x_n)$ их производственные функции с областями определения $G_А$ и $G_Б$. Сравнительная эффективность совокупного использования ресурсов в этих условиях в рамках интегрального метода факторного анализа оценивается по формуле:

$$Эфф = 1/2(f_Б(x_А) + y_А - y_Б - f_А(x_Б)).$$

Технология производства отражается здесь с помощью производственной функции каждого из предприятий.

Рассмотрим пример расчетов сравнительной эффективности двух предприятий. Предположим, что по статистическим данным за 6 лет работы предприятий А (рис. 22.1) и Б (рис. 22.2) для них построены производственные функции:

$$Y_А = 1,033K_А + 2,137L_А;$$

$$Y_Б = 1,408K_Б + 1,1105L_Б,$$

где Y – объем товарной продукции, K – основные фонды, L – численность промышленно-производственного персонала

		Предприятие А			
Год	Y	K	L	Y/K	Y/L
1	9,8	4,6	2,4	2,13	4,08
2	10,8	5	2,6	2,16	4,15
3	12,2	6	2,8	2,03	4,36
4	12,4	6,2	2,8	2,00	4,43
5	13,8	6,8	3,2	2,03	4,31
6	16,4	8	3,8	2,05	4,32
	2,136604	1,032769	0		
	0,190705	0,091275	#Н/Д		
	0,999362	0,06581	#Н/Д		
	3132,803	4	#Н/Д		
	27,13601	0,017324	#Н/Д		

Лист2

Рис. 22.1

		Предприятие Б			
Год	Y	K	L	Y/K	Y/L
1	19	8,6	6,4	2,21	2,97
2	20,6	9,2	6,8	2,24	3,03
3	21,4	9,8	6,8	2,10	3,15
4	22,4	10,2	7,4	2,20	3,03
5	24,2	10,8	8	2,24	3,03
6	26	11,6	8,6	2,24	3,02
	1,110553	1,408321	0		
	0,401832	0,293858	#Н/Д		
	0,996364	0,170256	#Н/Д		
	548,1293	4	#Н/Д		
	31,77738	0,115949	#Н/Д		

Лист2

Рис. 22.2

Из рис. 22.1 и 22.2 видно, что фондоотдача на предприятии А ниже, чем на предприятии Б, поэтому сравнительная экономия основных фондов (столбец N на рис. 22.3 и 22.4) за все 6 лет работы > 0, т. е. предприятие Б лучше использует основные фонды.

С другой стороны, производительность труда на предприятии А выше, чем на предприятии Б, так что сравнительная экономия трудовых ресурсов отрицательна (столбец О на рис. 22.3 и 22.4).

Сделать вывод на основании этих данных о том, какое предприятие работало более эффективно, затруднительно.

В столбце Р на рис. 22.3 и 22.4 произведен расчет сравнительной эффективности совокупного использования ресурсов на основе производственных функций. Как видно из рис. 22.4, величина Эфф за все годы работы отрицательна. Следовательно, во всех случаях более эффективно работало предприятие А.

Эфонды	Этруд	Эфф
=C3*H3/B3-I3	=D3*H3/B3-J3	=0,5*($I$11*C3+$H$11*D3-B3+H3-$C$10*I3-$B$10*J3)
=C4*H4/B4-I4	=D4*H4/B4-J4	=0,5*($I$11*C4+$H$11*D4-B4+H4-$C$10*I4-$B$10*J4)
=C5*H5/B5-I5	=D5*H5/B5-J5	=0,5*($I$11*C5+$H$11*D5-B5+H5-$C$10*I5-$B$10*J5)
=C6*H6/B6-I6	=D6*H6/B6-J6	=0,5*($I$11*C6+$H$11*D6-B6+H6-$C$10*I6-$B$10*J6)
=C7*H7/B7-I7	=D7*H7/B7-J7	=0,5*($I$11*C7+$H$11*D7-B7+H7-$C$10*I7-$B$10*J7)
=C8*H8/B8-I8	=D8*H8/B8-J8	=0,5*($I$11*C8+$H$11*D8-B8+H8-$C$10*I8-$B$10*J8)

Лист2

Рис. 22.3

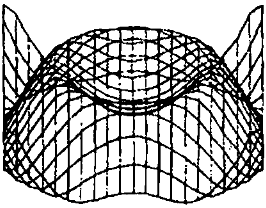
Эфонды	Этруд	Эфф
0,32	-1,75	-2,11
0,34	-1,84	-2,15
0,72	-1,89	-1,95
1,00	-2,34	-2,25
1,12	-2,39	-2,36
1,08	-2,58	-2,63

Лист2

Рис. 22.4

Глава 23. Моделирование спроса

23.1. Модель потребительского спроса



При моделировании спроса используются два основных подхода, аналогичные структурному и функциональному подходам по построению производственных функций. При первом подходе используется построение функций полезности и карт безразличия. При втором подходе устанавливается статистическая связь между спросом, доходом и ценами [11, 14, 21, 23].

Прежде всего рассмотрим модели, выражающие механизм изменения спроса. Наиболее простая модель такого типа имеет следующий вид. Рассматриваются n товаров, на которые имеется спрос со стороны населения. Модель предназначена для выражения зависимости вектора платежеспособного спроса на эти товары $y = (y_1, y_2, .., y_n)$ со стороны потребителя от денежных доходов d и вектора цен на эти товары: $p = (p_1, ..., p_n)$.

На вектор спроса y можно заранее наложить некоторые ограничения. Во-первых, компоненты этого вектора могут принимать лишь неотрицательные значения:

$$y_i \geq 0 \text{ при } i = 1, ..., n. \qquad (23.1)$$

Кроме того, если не учитывать возможность использования сбережения, то на вектор спроса y можно наложить бюджетное ограничение:

$$(p, y) = \sum_{i=1}^{n} p_i y_i \leq d. \qquad (23.2)$$

Смысл данного ограничения состоит в том, что затраты на потребление не могут превосходить доход. Для того чтобы выразить выбор одного из множества вектора y, удовлетворяющих ограничениям (23.1) и (23.2), используется функция полезности (или индикатора предпочтения $u(y)$, с помощью которого поведение потребителя можно представить себе как стремление получить такой набор товаров, который бы соответствовал возможно большему значению этой функции).

Чем больше величина $u(y)$, тем предпочтительнее для изучаемого индивидуума вектор товаров (y). Функцию полезности можно интерпретировать следующим образом: она выражает уровень удовлетворения потребностей потребителя, которой при своем выборе товаров старается максимизировать этот уровень.

Математически простейшая модель спроса, основанная на использовании функции полезности, выглядит так:

$$u(y) \to \max \quad \text{при } y \geq 0, (p, y) \leq d. \qquad (23.3)$$

23.2. Свойства функции полезности

Относительно функций полезности, используемых для выражения спроса на товары, обычно делаются некоторые предположения, которые выполняются для большинства функций. Эти предположения позволяют исследовать характерные свойства функции полезности.

Прежде всего предполагают, что количество товаров может изменяться непрерывно. Далее предполагают, что функция $u(y)$ меняется непрерывно и имеет все необходимые частные производные. Эти предположения по своей природе аналогичны соответствующим предположениям о производственных функциях и имеют математическую природу.

Предположение, имеющее экономическую основу, формулируется следующим образом. **Во-первых**, считается, что функция полезности возрастает или, по крайней мере, не убывает при увеличении количества товаров, точнее говоря:

$$u(y^2_i) \geq u(y^1_i), \tag{23.4}$$

как только $y^2_i \geq y^1_i$. Это свойство можно записать следующим образом:

$$u'(y_i) \equiv \partial u / \partial y_i \geq 0,\ i = 1,\ldots, n. \tag{23.5}$$

Это означает, что возрастание потребления одного продукта при постоянном потреблении других продуктов ведет к росту потребительской оценки и заключается в положительности первых частных производных.

Величины $u'(y_i)$ – первых частных производных функции полезности – характеризуют предельную полезность i-го продукта.

Во-вторых, предельная полезность каждого продукта уменьшается, если объем его потребления растет (это свойство предельной полезности называется законом убывания предельной полезности). Второе свойство выражается в отрицательности вторых частных производных по каждому i-му продукту:

$$\frac{\partial^2 u}{\partial y_i^2} < 0. \tag{23.6}$$

Третье свойство заключается в том, что при росте потребления одного продукта предельная полезность каждого продукта увеличивается. В этом случае продукт, количество которого фиксировано, оказывается относительно дефицитным. Поэтому дополнительная его единица приобретает большую ценность и может быть потреблена более эффективно. Данное свойство выражается в положительности смешанных производных. Так, для двух продуктов y_1 и y_2 данное свойство описывается как:

$$\frac{\partial^2 u}{\partial y_1 \partial y_2} > 0. \tag{23.7}$$

Это свойство справедливо не для всех продуктов; если продукты могут полностью замещать друг друга в потреблении, свойство не выполняется. Третье предположение вводится не всегда, но оно гарантирует выпуклость вниз кривой безразличия.

Линия, соединяющая потребительские наборы (y_1,y_2), имеющие один и тот же уровень удовлетворения потребностей, называется кривой безразличия. Кривая безразличия есть не что иное, как изолиния уровня функции полезности. Множество кривых безразличия называется картой безразличия.

Поверхность безразличия определяется как множество всех таких векторов y, для которых выполняется условие:

$$u(y) = c. \tag{23.8}$$

Поверхность безразличия – это множество наборов продуктов, которые имеют одинаковую ценность для потребителя. Выбирая различные значения постоянной c в выражении (23.8), можно построить семейство поверхности безразличия. Это семейство характеризует свойства функции полезности.

Согласно сделанным предположениям, потребитель выбирает такой вектор y, на котором достигается максимум функции полезности при выполнении ограничений (23.1) и (23.2), поэтому с точки зрения описания зависимости спроса от цен и дохода семейство, или карта, поверхностей безразличия содержит ту же информацию, что и функция полезности. Поэтому вместо построения функций полезности часто сразу строят семейство поверхностей безразличия.

Свойства функции полезности означают, что кривая безразличия убывает и выпукла к началу координат. Рассмотрим дифференциал (главную линейную часть приращения) функции полезности для двух продуктов $u(y_1,y_2)$. Если двигаться вдоль кривой безразличия, величина функции полезности не изменится и приращение функции $u(y_1,y_2)$ будет равно нулю. Будет равна нулю и его главная линейная часть. Дифференциал функции полезности выражается следующим образом:

$$du(y_1, y_2) = \frac{\partial u}{\partial y_1}dy_1 + \frac{\partial u}{\partial y_2}dy_2 = u_1' dy_1 + u_2' dy_2 = 0,$$

откуда:

$$\frac{dy_2}{dy_1} = -\frac{u_1'}{u_2'} < 0.$$

Зависимость y_2 от y_1 вдоль кривой безразличия является убывающей, поскольку производная ее отрицательна. Вторая производная функции $y_2(y_1)$ выглядит следующим образом:

$$d\left(\frac{dy_2}{dy_1}\right)\Big/dy_1 = -\frac{u_{11}'' \cdot u_2' - u_1' \cdot u_{21}''}{(u_2')^2} > 0.$$

Ее положительность вытекает из свойств функций полезности, и, следовательно, кривые безразличия выпуклы вниз.

Величину $-u'_1/u'_2$ принято называть предельной нормой замены первого продукта вторым.

Если в модели рассматриваются только два продукта, то поверхность безразличия, которая в этом случае является кривой безразличия, может быть изображена на плоскости (рис. 23.1).

Кривые безразличия, соответствующие разным уровням удовлетворения потребностей, не касаются и не пересекаются. Чем выше и правее расположена кривая безразличия, тем большему уровню удовлетворения потребности она соответствует.

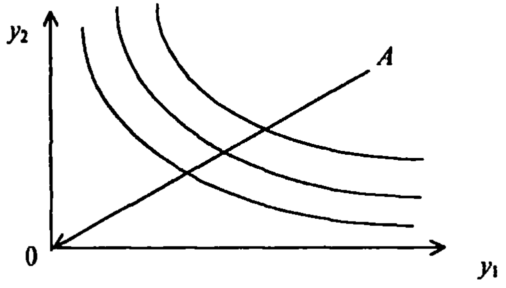


Рис. 23.1

23.3. Анализ модели спроса

Проведем анализ исходной модели спроса (23.3), заключающийся в выборе вектора y, составляющего максимум функции полезности при выполнении ограничений (23.1) и (23.2). При выполнении (23.5) неравенство (23.2) можно заменить равенством. Модель приобретает вид:

$$u(y) \Rightarrow \max \text{ при } y \geq 0 \ (p, y) = d. \tag{23.9}$$

Если $u(y) > 0$ только при $y > 0$, то для решения задачи модель можно проанализировать при помощи метода неопределенных множителей Лагранжа. Для задачи (23.9) функция Лагранжа:

$$L(y,\lambda) = u\,(y) + \lambda\,(d - \Sigma\, p_i\, y_i).$$

Необходимым условием максимума функции $u(y)$ в точке y^* при выполнении ограничений задачи (23.9) является существование такого множителя Лагранжа λ^*, что при $y = y^*$ и $\lambda = \lambda^*$ выполняются условия:

$$\begin{gathered} \frac{\partial L}{\partial y_i} \equiv \frac{\partial u}{\partial y_i} - \lambda \cdot p_i = 0,,\ i = 1,\ldots, n; \\ \frac{\partial L}{\partial \lambda} \equiv d - \sum_{i=1}^{n} p_i y_i = 0. \end{gathered} \tag{23.10}$$

Предельные полезности $u'(y^*)$ должны быть пропорциональны ценам:

$$u'_1(y^*)\,/\,p_1 = u'_2(y^*)\,/\,p_2 = \ldots = u'_n(y^*)\,/\,p_n.$$

Соотношение (23.10) позволяет оценить воздействие на спрос y^* дохода d и цен p.

Если необходимо с помощью выбора дохода и цен так воздействовать на поведение потребителя, чтобы он приобрел товары, описанные вектором y^*, цены и доход должны быть назначены в соответствии с (23.10).

Для случая двух товаров задача формулируется следующим образом:

$$u(y_1,y_2) \Rightarrow \max \text{ при } y_{1,2} \geq 0 \ p_1 y_1 + p_2 y_2 = d. \tag{23.11}$$

Функция Лагранжа:

$$L(y_1,y_2, \lambda) = u\,(y_1,y_2) + \lambda\,(d - p_1 y_1 + p_2 y_2).$$

Находим первые частные производные функции Лагранжа по переменным y_1, y_2, λ и приравниваем эти частные производные к нулю:

$$\partial L/\partial y_1 = \partial u/\partial y_1 - \lambda p_1 = 0;$$
$$\partial L/\partial y_2 = \partial u/\partial y_2 - \lambda p_2 = 0;$$
$$\partial L/\partial \lambda = d - p_1 y_1 - p_2 y_2 = 0.$$

Исключив из полученной системы трех уравнений с тремя неизвестными неизвестную λ, получим систему двух уравнений с двумя неизвестными y_1 и y_2:

$$\partial u/\partial y_1 \,/\, \partial u/\partial y_2 = p_1/p_2,$$
$$p_1 y_1 + p_2 y_2 = d.$$

Решением этой системы будет точка локального оптимума (y^*_1, y^*_2). Подставив решение (y^*_1, y^*_2) в левую часть равенства $\partial u/\partial y_1 \,/\, \partial u/\partial y_2 = p_1/p_2$, получим, что в точке локального рыночного равновесия отношение предельных полезностей товаров равно отношению рыночных цен p_1 и p_2:

$$\partial u(y^*_1, y^*_2)/\partial y_1 \,/\, \partial u(y^*_1, y^*_2)/\partial y_2 = p_1/p_2.$$

Отношение $\partial u(y^*_1, y^*_2)/\partial y_1 \,/\, \partial u(y^*_1, y^*_2)/\partial y_2$ представляет собой предельную норму замены первого продукта вторым в точке локального рыночного равновесия (y^*_1, y^*_2). Отсюда следует, что предельная норма замены равна отношению рыночных цен на эти товары в точке (y^*_1, y^*_2).

Полученные результаты можно представить графически для случая двух товаров (рис. 23.2).

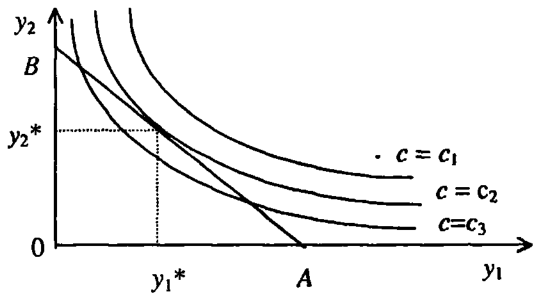


Рис. 23.2

Прямая AB отражает бюджетное ограничение, т. е. описывается соотношением $p_1y_1 + p_2y_2 = d$. Поэтому треугольник $0AB$ представляет собой множество всех допустимых векторов y. Точка A соответствует максимально возможным закупкам первого товара, точка B – второго, т. е.

$$A = d / p_1;\ B = d / p_2.$$

Кривые безразличия c_1, c_2 и c_3 позволяют найти точку y^*, соответствующую спросу с максимальной функцией полезности.

Поскольку функция полезности растет с удалением от начала координат, то точка y^* лежит на наиболее удаленной от нуля кривой безразличия, пересекающейся с треугольником $0AB$.

Из (23.10) следует:

$$\frac{1}{p_1}u_1(y^*) = \frac{1}{p_2}u_2(y^*), \tag{23.12}$$

т. е. угол наклона касательной к кривой безразличия, задаваемой как отношение u_1/u_2, совпадает с наклоном прямой, задаваемой бюджетным ограничением. Поэтому отрезок AB касается кривой безразличия $c = c_2$, на которой лежит точка y^*.

23.4. Построение функций полезности

Построение функций полезности, или карт поверхностей безразличия, обычно основывается на изучении расходов потребителей на покупку товаров. Оценка коэффициентов функции полезности осуществляется на основе методов регрессионного анализа. Однако поскольку непосредственно наблюдается поведение потребителей, т. е. вектор y, а не функция $u(y)$, здесь требуется использовать специальную методику. Основная идея этой методики состоит в применении регрессионных методов для определения величин $u_i(y)$ на основе соотношения (23.10) Зная величины $u_i(y)$, можно построить карту поверхностей безразличия.

Наиболее часто используется функция, имеющая вид

$$u(y) = \sum_{i=1}^{n} a_i \cdot \log(y_i - y_i^0), \qquad (23.13)$$

заданная при $y_i > y_i^0$, где y_i^0 – неотрицательные минимальные уровни потребления. Величины a_i – некоторые положительные коэффициенты. Значения a_i, y_i^0 определяются на основе анализа реальной структуры потребления.

Построим кривые безразличия для функции полезности (23.13). Предельная полезность $u_i(y)$ в этом случае имеет вид:

$$u'(y) \equiv \frac{\partial u(y)}{\partial y_i} \equiv \frac{a_i}{y_i - y_i^0}, \; i = 1,\dots, n.$$

При движении вдоль поверхности безразличия предельные полезности удовлетворяют условию

$$\sum_{i=1}^{n} \frac{a_i}{y_i - y_i^0} \cdot dy_i = 0.$$

Отсюда получаем, что поверхность безразличия выражается соотношением

$$\prod_{i=1}^{n} (y_i - y_i^0)^{a_i} = c.$$

Проанализируем модель спроса (23.9) для функции полезности (23.13). Согласно (23.10), спрос y^* должен удовлетворять условиям, которые для функции (23.13) приобретают вид:

$$\frac{a_1}{p_1} \cdot \frac{1}{y_1 - y_1^0} = \frac{a_2}{p_2} \cdot \frac{1}{y_2 - y_2^0} = \dots = \frac{a_n}{p_n} \cdot \frac{1}{y_n - y_n^0},$$

$$\sum_{i=1}^{n} p_i \cdot y_i = d.$$

На основе этих соотношений можно построить функцию, выражающую зависимость спроса от цен и дохода. Из первой системы равенств получаем:

$$y^*_i - y^0_i = a_i / (\lambda^* p_i), \; i = 1,\dots, n\,. \qquad (23.14)$$

Подставляем это соотношение в бюджетное равенство и получаем:

$$\lambda^* = \sum_{i=1}^{n} a_i \Big/ \left(d - \sum_{i=1}^{n} p_i \cdot y_i^0\right).$$

Поэтому из (23.14) следует, что

$$y_i^* = y_i^0 + \frac{1}{p_i} \cdot \frac{a_i}{\sum\limits_{i=1}^{n} a_i} (d - \sum_{i=1}^{n} p_i \cdot y_i^0), \ i = 1,\ldots,n.$$

Умножение функции полезности на некоторое число не меняет картины. Поэтому можно считать, что функция (23.13) удовлетворяет условию $\sum\limits_{i=1}^{n} a_i = 1$. Тогда зависимость спроса от цен и дохода приобретает вид:

$$y_i^* = y_i^0 + \frac{a_i}{p_i} (d - \sum_{i=1}^{n} p_i \cdot y_i^0), \ i = 1,\ldots,n. \qquad (23.15)$$

Функции (23.15) принято называть функциями Стоуна, по имени английского экономиста, предложившего их.

23.5. Функции спроса

Функции спроса базируются на непосредственной обработке статистики доходов, цен и потребления. То есть необходимо построить непосредственную зависимость между вектором спроса y и экономическими факторами спроса доходом d, а также вектором цен p в виде функций:

$$y_i = f_i(d,p), i = 1,\ldots,n. \qquad (23.16)$$

Функция $f_i(d,p)$ содержит параметры, которые определяются на основе методов регрессионного анализа, а сами функции выбираются таким образом, чтобы они удовлетворяли некоторым априорным предположениям о свойствах функции спроса. В большинстве случаев функции (23.15) удовлетворяют условию

$$\partial f_i(d,p) / \partial p_i < 0, \ i = 1,\ldots,n,$$

которое означает, что при увеличении цены на данный вид продукции платежеспособный спрос на него падает. Однако надо отметить, что данное предположение выполняется не всегда. Существуют такие товары, называемые товарами Гиффена, для которых имеет место соотношение:

$$\partial f_i(d,p) / \partial p_i > 0,$$

Эта ситуация, когда происходит рост спроса на товар при повышении цены, возникает в том случае, когда потребители с низким уровнем дохода из-за повышения цен на относительно дешевый товар, теряют возможность покупать более дорогие товары.

При росте дохода потребление одних товаров может увеличиваться, потребление других – уменьшаться. Товары, для которых имеет место соотношение

$$\partial f_i(d,p) / \partial d < 0,$$

принято называть малоценными.

В том случае, когда выполняется обратное условие:

$$\partial f_i(d,p) / \partial d > 0,$$

товары принято называть ценными.

Ценные товары делятся на 3 группы, в зависимости от значения эластичности спроса по отношению к доходу:

$$\varepsilon_i^d = \frac{d}{y_i}\frac{\partial f_i(d,p)}{\partial d},$$

которое показывает, на сколько процентов увеличивается спрос на данный товар при росте дохода на 1%.

К первой группе относят товары с малой эластичностью, лежащей между 0 и 1. Ко второй группе – товары с эластичностью, близкой к 1. К третьей группе товары с высокой эластичностью, превышающей 1.

Эластичность товаров по отношению к ценам будет выражаться соотношением:

$$\varepsilon_i^p = \frac{p_j}{y_i}\frac{\partial f_i(d,p)}{\partial p_j},\ i = 1,\dots,n, j = 1,\dots,m.$$

Эластичности товаров по доходам и по ценам являются локальными характеристиками функции спроса, поскольку эти эластичности обычно сами зависят от величин доходов и цен. Исключением являются степенные функции спроса:

$$y_i = a_i p_i^{\beta_i} \cdot d^{g_i},\ i = 1,\dots,n,$$

для которых:

$$\varepsilon_j^d = g_i;\ \varepsilon_{ji}^p = \beta_i,$$

где a_i, β_i, g_i – параметры, оцениваемые на основе экономической статистики.

Эмпирические исследования свидетельствуют, что колебания спроса на потребительские товары можно достаточно точно выразить функцией трех переменных: цены продукции, общего дохода потребителей на рынке и индекса цен всех остальных потребительских товаров на рынке:

$$y_i = ad^g p^{-\beta} p_0^{\beta - f},$$

где d – доход; p – цена продукции; p_0 – индекс цен всех остальных потребительских товаров; a, g, β, f – положительные параметры.

Недостатком данной функции является постоянство эластичности по доходам и ценам.

Глава 24. Системы одновременных уравнений

24.1. Основные понятия

При моделировании достаточно сложных экономических объектов часто приходится вводить не одно, а несколько связанных между собой уравнений. А значит, при проведении регрессионного анализа модели может возникнуть необходимость оценивать систему уравнений [11, 22]. Оценка систем уравнений требует введения новых понятий и методов.

Рассмотрим классический пример системы одновременных уравнений, который демонстрирует основные проблемы, возникающие при попытке оценить неизвестные параметры. Таким примером является исследование зависимости спроса и предложения некоторого товара от его цены и дохода:

$$S_t = a_1 + a_2 P_t + \varepsilon_t \text{ (предложение)};$$
$$D_t = b_1 + b_2 P_t + b_3 Y_t + u_t \text{ (спрос)};$$
$$S_t = D_t \text{ (равновесие)},$$

где P_t – цена товара; Y_t – доход в момент времени t.

Записывая каждое уравнение для упрощения в отклонениях от средних значений[1], получаем следующую систему:

$$s_t = a_2 p_t + \varepsilon_t;$$
$$d_t = b_2 p_t + b_3 y_t + u_t.$$

Данная система уравнений называется структурной формой модели, соответственно, коэффициенты этих уравнений называются структурными коэффициентами. В соответствии с этой моделью цена и величина спроса-предложения определяются одновременно и поэтому эти переменные должны считаться эндогенными, а доход y_t, в отличие от них, – экзогенной переменной. Деление переменных на экзогенные и эндогенные определяется содержательной стороной модели. Предполагается, что в каждом уравнении экзогенные переменные не коррелированы с ошибкой. В то же время эндогенные переменные, стоящие в правых частях уравнений, как правило, имеют ненулевую корреляцию с ошибкой в соответствующем уравнении. В модели со стохастическими регрессорами наличие корреляции между регрессорами и ошибками приводит к смещенности и несостоятельности МНК-оценок.

---

[1] Уравнение в отклонениях. Обозначим через $x_t = X_t - \overline{X}$, $y_t = Y_t - \overline{Y}$ отклонения от средних по выборке зачений X_t и Y_t, $\overline{X} = \frac{1}{n}\sum X_t$, $\overline{Y} = \frac{1}{n}\sum Y_t$. Решением задачи оценки параметров линейной функции $y = a + bx$ будет та же прямая на плоскости (x,y), что и для исходных данных X_t, Y_t. Переход от X, Y к отклонениям x,y означает лишь перенос начала координат в точку $(\overline{X}, \overline{Y})$. Учитывая, что $\bar{x} = \bar{y} = \sum x_t = \sum y_t = 0$, получим:

$$\hat{a} = 0, \quad \hat{b} = \frac{\sum x_t y_t}{\sum x_t^2}.$$

Обозначая $d_t = s_t = g_t$, разрешим систему относительно g_t и p_t:

$$g_t = \frac{a_2 b_3}{a_2 - b_2} y_t + \frac{a_2 u_2 - b_2 \varepsilon_t}{a_2 - b_2};$$

$$p_t = \frac{b_3}{a_2 - b_2} + \frac{u_t - \varepsilon_t}{a_2 - b_2}.$$

Данная система называется приведенной формой модели. Обозначая

$$h_1 = \frac{a_2 b_3}{a_2 - b_2}; \qquad v_{1t} = \frac{a_2 u_2 - b_2 \varepsilon_t}{a_2 - b_2};$$

$$h_2 = \frac{b_3}{a_2 - b_2}; \qquad v_{2t} = \frac{u_t - \varepsilon_t}{a_2 - b_2},$$

перепишем ее в виде:

$$g_t = h_1 y_t + v_{1t};$$
$$p_1 = h_2 y_t + v_{2t}.$$

Здесь уже в каждом уравнении экзогенная переменная y_t не коррелирована с ошибкой, поэтому метод наименьших квадратов даст состоятельные оценки $\hat{h}_1$ и $\hat{h}_2$ коэффициентов h_1 и h_2. Величина $a_2 = h_1/h_2$ и будет состоятельной оценкой структурного параметра a_2. Такой способ оценивания структурных коэффициентов с помощью оценок коэффициентов приведенной формы называется косвенным методом наименьших квадратов. Следовательно, для структурного коэффициента первого уравнения можно построить состоятельную оценку, используя косвенный метод наименьших квадратов.

Усложним исходную модель, включив в уравнение спроса процентную ставку r_t:

$$d_t = b_2 p_t + b_3 y_t + b_4 r_t + u_t,$$

считая эту переменную экзогенной. Тогда получим следующую приведенную форму:

$$g_t = h_{11} y_t + h_{12} r_t + v_{1t};$$
$$p_1 = h_{21} y_t + h_{22} r_t + v_{2t},$$

где

$$h_{11} = \frac{a_2 b_3}{a_2 - b_2}; \qquad h_{12} = \frac{a_2 b_4}{a_2 - b_2};$$

$$h_{21} = \frac{b_3}{a_2 - b_2}; \qquad h_{22} = \frac{b_4}{a_2 - b_2}.$$

Отсюда $a_2 = h_{11}/h_{21} = h_{12}/h_{22}$. Поэтому при использовании косвенного метода наименьших квадратов можно в качестве оценки структурного параметра a_2 брать либо $\hat{h}_{11}/\hat{h}_{21}$, либо $\hat{h}_{12}/\hat{h}_{22}$, причем в общем случае это будут разные оценки. Даже знание точных значений коэффициентов приведенной формы и для исходной, и для усложненной моделей не позволяет сделать никаких выводов относительно структурных параметров второго уравнения.

Можно сделать несколько выводов из изложенного.

- Переменные в системах одновременных уравнений делятся на экзогенные и эндогенные. Первые отличаются от вторых тем, что в каждом уравнении они не коррелированы с соответствующей ошибкой.
- Из-за наличия корреляции между эндогенными переменными и ошибками непосредственное применение метода наименьших квадратов к структурной форме модели приводит к смещенным и несостоятельным оценкам структурных коэффициентов.
- Коэффициенты приведенной формы модели могут быть состоятельно оценены методом наименьших квадратов. Эти оценки могут быть использованы для оценивания структурных параметров (косвенный метод наименьших квадратов). При этом возможны три ситуации: структурный коэффициент однозначно выражается через коэффициенты приведенной системы, структурный коэффициент допускает несколько разных оценок косвенного метода наименьших квадратов, структурный коэффициент не может быть выражен через коэффициенты приведенной системы. В последнем случае соответствующее структурное уравнение является неидентифицируемым. Неидентифицируемость уравнения не связана с числом наблюдений.

24.2. Системы одновременных уравнений в матричной форме

Предположим теперь, что имеется следующая система уравнений, называемая структурной формой модели:

$$\beta_{11}Y_{1t} + \beta_{12}Y_{2t} + \ldots + \beta_{1m}Y_{mt} + \gamma_{11}X_{1t} + \gamma_{12}X_{2t} + \ldots + \gamma_{1k}X_{kt} = \varepsilon_{1t};$$
$$\beta_{21}Y_{1t} + \beta_{22}Y_{2t} + \ldots + \beta_{2m}Y_{mt} + \gamma_{21}X_{1t} + \gamma_{22}X_{2t} + \ldots + \gamma_{2k}X_{kt} = \varepsilon_{2t};$$
$$\ldots$$
$$\beta_{m1}Y_{1t} + \beta_{m2}Y_{2t} + \ldots + \beta_{mm}Y_{mt} + \gamma_{m1}X_{1t} + \gamma_{m2}X_{2t} + \ldots + \gamma_{mk}X_{kt} = \varepsilon_{mt}.$$

Переменные $Y_t, \ldots, Y_m$, определяемые внутри системы, называются эндогенными, а переменные $X_t, \ldots, X_k$ могут быть включены как внешние по отношению к системе (экзогенные) переменные, так и лагированные значения эндогенных переменных, которые называются предопределенными переменными.

Индекс t означает номер наблюдения, $t = 1, \ldots, n$, а $\varepsilon_{1t}, \ldots, \varepsilon_{mt}$ – случайные ошибки. Будем считать, что в каждом уравнении один из коэффициентов β при какой-либо эндогенной переменной равен 1 – это естественное условие нормировки. Оно позволяет представить каждое уравнение в привычном виде, когда в левой части стоит одна эндогенная переменная, а в правой части – остальные переменные с неизвестными коэффициентами плюс случайная ошибка. Обозначим:

$$Y_t = \begin{pmatrix} Y_{1t} \\ Y_{2t} \\ \cdot \\ Y_{mt} \end{pmatrix}, \quad X_t = \begin{pmatrix} X_{1t} \\ X_{2t} \\ \cdot \\ X_{kt} \end{pmatrix}, \quad \varepsilon_t = \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \cdot \\ \varepsilon_{mt} \end{pmatrix},$$

$$B=\begin{pmatrix}\beta_{11} & \beta_{12} & \cdots & \beta_{1m}\\ \beta_{21} & \beta_{22} & \cdots & \beta_{2m}\\ \cdot & \cdot & \cdots & \cdot\\ \beta_{m1} & \beta_{m2} & \cdots & \beta_{mm}\end{pmatrix},\qquad \Gamma=\begin{pmatrix}\gamma_{11} & \gamma_{12} & \cdots & \gamma_{1k}\\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2k}\\ \cdot & \cdot & \cdots & \cdot\\ \gamma_{m1} & \gamma_{m2} & \cdots & \gamma_{mk}\end{pmatrix}.$$

Тогда систему можно представить в следующем виде:

$$\mathrm{B}Y_t+\Gamma X_t=\varepsilon_t.$$

Деление переменных на эндогенные и экзогенные должно быть проведено вне модели. Одно из основных требований к экзогенным переменным – некоррелированность векторов X_t и ε_t в каждом наблюдении t. Полагая, что матрица B не вырождена, умножим обе части равенства на B^{-1}:

$$Y_t=-\mathrm{B}^{-1}\Gamma X_t+\mathrm{B}^{-1}\varepsilon_t=\Pi X_t+\nu_t,$$

где $\Pi=-\mathrm{B}^{-1}\Gamma X_t$, $\nu=\mathrm{B}^{-1}\varepsilon_t$.

Данная система является приведенной формой модели, элементы матриц B и Г – структурными коэффициентами, а элементы матрицы П – коэффициентами приведенной формы.

В общем случае эндогенные переменные и ошибки в структурной системе коррелированы, поэтому применение к какому-либо из уравнений метода наименьших квадратов даст смещенные и несостоятельные оценки структурных коэффициентов. В то же время коэффициенты приведенной формы могут быть состоятельно оценены, поскольку переменные X_t не коррелированы со структурными ошибками ε_t и, следовательно, с ошибками приведенной формы модели ν_t.

Проблема идентификации структурной модели связана с числом структурных коэффициентов и числом коэффициентов приведенной формы, и, соответственно, число уравнений – с числом неизвестных. Структурный коэффициент идентифицируем, если он может быть вычислен на основе коэффициентов приведенной формы. Соответственно, какое – либо уравнение в структурной форме модели будет называться идентифицируемым, если идентифицируемы все его коэффициенты. Проблема идентифицируемости логически предшествует задаче оценивания, отсутствие идентифицируемости означает, что существует бесконечно много моделей, совместимых с имеющимися наблюдениями, и это никак не связано с количеством наблюдений.

Система неидентифицируема в случае превышения числа структурных коэффициентов над числом коэффициентов приведенной формы. В общем случае система неидентифицируема. Основная причина возможности идентификации – наличие априорных ограничений на структурные коэффициенты. Ограничения имеют наиболее простой вид, когда часть структурных коэффициентов равна 0.

При идентифицируемости уравнения оценки структурных коэффициентов можно найти, оценив методом наименьших квадратов приведенную форму модели, а затем решить систему уравнений, заменяя элементы матрицы П их оценками. Этот способ носит название косвенного метода наименьших квадратов. Однако у этого метода есть серьезный недостаток: если уравнение сверхидентифицируемо, то число уравнений превышает число неизвестных, а это значит, что один и тот же структурный коэффициент допускает разные выражения через коэффициенты приведенной формы. Это сужает область его применения как с теоретической (не ясно, какую же оценку следует предпочесть), так и с практической точки зрения (трудность алгоритмизации). Поэтому, как правило, используются другие методы.

24.3. Двухшаговый метод наименьших квадратов

Методы оценивания систем одновременных уравнений можно разделить на методы, позволяющие оценивать каждое из уравнений поочередно, и методы, предназначенные для оценивания всех уравнений сразу, т. е. всей модели в целом. Примерами первой группы методов служат двухшаговый МНК и метод ограниченной информации для одного уравнения, а примерами методов второй группы – трехшаговый метод наименьших квадратов и метод максимального правдоподобия полной информации.

Наиболее важным и широко применяемым методом оценивания отдельного уравнения системы является двухшаговый МНК.

Представим первое уравнение в следующем виде:

$$y_{1t} = -\beta_{12}Y_{2t} - \ldots - \beta_{1q}Y_{qt} - \gamma_{11}X_{1t} - \gamma_{12}X_{2t} - \ldots - \gamma_{1p}X_{pt} + \varepsilon_{1t},\ t = 1,\ldots,n.$$

Обозначим:

$$y_1 = \begin{pmatrix} y_{11} \\ y_{12} \\ . \\ y_{1n} \end{pmatrix}, \quad Y_1 = \begin{pmatrix} Y_{21} & \ldots & Y_{q1} \\ Y_{22} & \ldots & Y_{q2} \\ . & \ldots & . \\ Y_{2n} & \ldots & Y_{qn} \end{pmatrix}, \quad X_1 = \begin{pmatrix} X_{11} & \ldots & X_{p1} \\ X_{12} & \ldots & X_{p2} \\ \vdots & \ldots & . \\ X_{1n} & \ldots & X_{pn} \end{pmatrix},$$

$$\beta_1 = \begin{pmatrix} -\beta_{12} \\ . \\ -\beta_{1q} \end{pmatrix}, \quad \gamma_1 = \begin{pmatrix} -\gamma_{11} \\ . \\ -\gamma_{1p} \end{pmatrix}, \quad \varepsilon_1 = \begin{pmatrix} \varepsilon_{11} \\ . \\ \varepsilon_{1t} \end{pmatrix}, \quad X = \begin{pmatrix} X_{11} & \ldots & X_{p1} \\ X_{12} & \ldots & X_{p2} \\ . & \ldots & . \\ X_{1n} & \ldots & X_{kn} \end{pmatrix},$$

и перепоишем его в виде:

$$y_1 = Y_1\beta_1 + X_1\gamma_1 + \varepsilon_1.$$

Итак, в данном уравнении модели: y_1 – вектор размерности $n\times 1$ наблюдений над зависимой переменной; Y_1 – матрица порядка $n\times q$ наблюдений над другими текущими значениями эндогенных переменных, входящих в уравнение; β – вектор размерности $q\times 1$ структурных коэффициентов, относящихся к переменным из матрицы Y_1; X_1 – матрица порядка $n\times p$ наблюдений над предопределенными переменными, присутствующими в уравнении (включая столбец единиц, если необходим свободный член); γ – вектор размерности $p\times 1$ коэффициентов, относящихся к переменным из матрицы X_1; ε – вектор возмущающих воздействий, испытываемых рассматриваемым уравнением.

Осложнения, связанные с применением к уравнению обыкновенного МНК возникают из-за наличия корреляций между Y_1 и ε. Существо двухшагового метода состоит в замене матрицы Y_1 расчетной матрицей $\hat{Y}_1$, которая будет очищена от влияния стохастического элемента, после чего и вычисляется обыкновенная регрессия y на $\hat{Y}_1$ и X_1.

Матрица $\hat{Y}_1$ вычисляется на первом шаге метода с помощью регрессии каждой из переменных, входящих в Y_1, на все предопределенные переменные полной модели, после чего наблюдения, входящие в матрицу Y_1, заменяются соответствующими значениями, рассчитанными с помощью этих регрессий.

Таким образом,

$$\hat{Y}_1 = X(X^T X)^{-1} X^T Y_1,$$

где $X = [X_1 X_2]$ – есть матрица порядка $n \times k$ наблюдений над всеми предопределенными переменными, входящими в полную модель, а X_2 – матрица наблюдений над теми предопределенными переменными модели, которые не включены в рассматриваемое нами уравнение.

На втором шаге вычисляется регрессия y на $\hat{Y}_1$ и X_1, что приводит к процедуре оценивания:

$$\begin{bmatrix} \hat{Y}_1^T \hat{Y}_1 & \hat{Y}_1^T X_1 \\ X_1^T \hat{Y}_1 & X_1^T X_1 \end{bmatrix} \cdot \begin{bmatrix} \bar{\beta} \\ \bar{\gamma} \end{bmatrix} = \begin{bmatrix} \hat{Y}_1^T & y \\ X_1^T & y \end{bmatrix}.$$

Для вычисления оценок $\begin{bmatrix} \bar{\beta} \\ \bar{\gamma} \end{bmatrix}$ на самом деле нет необходимости в вычислении регрессионных значений для матрицы $\hat{Y}_1$. Можно вывести альтернативное выражение, которое будет зависеть только от матриц реальных наблюдений. Выразим в виде:

$$Y_1 = \hat{Y}_1 + v_1,$$

где v_1 – матрица размерности $n \times q$ остатков регрессий наименьших квадратов Y_1 на X.

Обычные свойства остатков наименьших квадратов дают нам:

$$\hat{Y}_1^T v_1 = 0 = X' v_1.$$

Поэтому

$$\hat{Y}_1^T \hat{Y}_1 = \hat{Y}_1^T (Y_1 - v_1) = \hat{Y}_1^T Y_1 = Y_1^T X (X^T X)^{-1} X^T Y_1$$

и

$$\hat{Y}_1^T X_1 = (Y_1 - v_1)^T X_1 = Y_1^T X_1.$$

Таким образом, уравнение для вычисления оценок двухшагового метода наименьших квадратов можно записать в виде:

$$\begin{bmatrix} Y_1^T X (X^T X)^{-1} X^T Y_1 & Y_1^T X_1 \\ X_1^T Y_1 & X_1^T X_1 \end{bmatrix} \cdot \begin{bmatrix} \bar{\beta} \\ \hat{\gamma} \end{bmatrix} = \begin{bmatrix} Y_1^T X (X^T X)^{-1} X^T y \\ X_1^T y \end{bmatrix}.$$

Можно использовать следующую последовательность расчетов:

1. Проводится регрессия каждого столбца матрицы Y_1 на *все* экзогенные переменные, т. е. рассматривается регрессия:

$$Y_1 = X\Pi_1 + v_1,$$

где Π_1 – $k \times (q - 1)$ матрица коэффициентов приведенной формы.

2. Строится прогнозное значение

$$\hat{Y}_1 = X\hat{\Pi}_1,$$

где

$$\hat{\Pi}_1 = (X^T X)^{-1} X Y_1.$$

3. В исходном уравнении в правой части заменяется Y1 на $\hat{Y}_1$ и определяются МНК-оценки структурных параметров β1 и γ1 в регрессии

$$Y_1 = \hat{Y}_1\beta_1 + X_1\gamma_1 + \varepsilon_1.$$

24.4. Двухшаговый метод наименьших квадратов в Excel

В Excel *нет встроенного двухшагового метода наименьших квадратов.* Поэтому основные возможности следующие: последовательные вычисления с использованием функции ЛИНЕЙН, учитывая, что она выводит вектор-строку коэффициентов регрессии в обратном порядке, поэтому вектор-столбец коэффициентов при транспонировании также, к сожалению, выходит в обратном порядке. Следующая возможность – использовать матричные функции или комбинацию матричных функций и функции ЛИНЕЙН.

Последовательность действий рассмотрим на конкретном примере. Возьмем систему уравнений «спрос–предложение». Обозначим y_1 – спрос, y_2 – предложение, Y_1, Y_2 – цена, X_1, X_2 – доход, a_1, a_2, β_1, β_2, γ_1, γ_2 – искомые коэффициенты регрессий:

$$y_1 = a_1 + Y_1\beta_1 + X_1\gamma_1 + \varepsilon_1,$$

$$y_2 = a_2 + Y_2\beta_2 + X_2\gamma_2 + \varepsilon_2.$$

Введем исходные данные: y_1 – в A2:A11, Y_1 – в B2: B11, X_1 – в C2:C11, y_2 – в D2:D11, Y_2 – в E2:E11, X_2 – в F2:F11 (табл. 24.1).

Таблица 24.1

y1	Y1	X1	y2	Y2	X2
79	1	5	15	2	10
74	3	7	26	4	15
77	2	9	36	6	17
74	5	13	46	8	20
70	6	8	60	11	23
56	4	15	70	15	18
48	8	20	67	13	28
49	6	18	86	17	30
44	9	22	96	19	34
34	11	24	125	25	40

Определим для сравнения коэффициенты уравнений обычным МНК, т. е. применим функцию ЛИНЕЙН к обоим уравнениям (табл. 24.2).

Таблица 24.2

– 1,890	– 0,954	92,398	0,661	3,856	0,885
0,570	1,211	4,028	0,132	0,170	1,474
0,919	5,220	#Н/Д	0,999	1,353	#Н/Д
39,667	7,000	#Н/Д	2791,519	7,000	#Н/Д
2161,757	190,743	#Н/Д	10213,295	12,805	#Н/Д

Таким образом, уравнения выглядят:

$$y_1 = 92{,}398 - 0{,}954\ Y_1 - 1{,}89\ X_1,$$
$$y_2 = 0{,}885 + 3{,}856\ Y_2 + 0{,}661\ X_2.$$

Теперь перейдем к двухшаговому методу наименьших квадратов. Для удобства образуем матрицу X: запишем в столбцы G и H векторы X_1 и X_2. Теперь определим коэффициенты приведенной формы для первого уравнения. Для этого применим функцию ЛИНЕЙН к Y_1 и матрице $X = [X_1X_2]$ и выведем результат в ячейки A13:B13 (рис. 24.1).

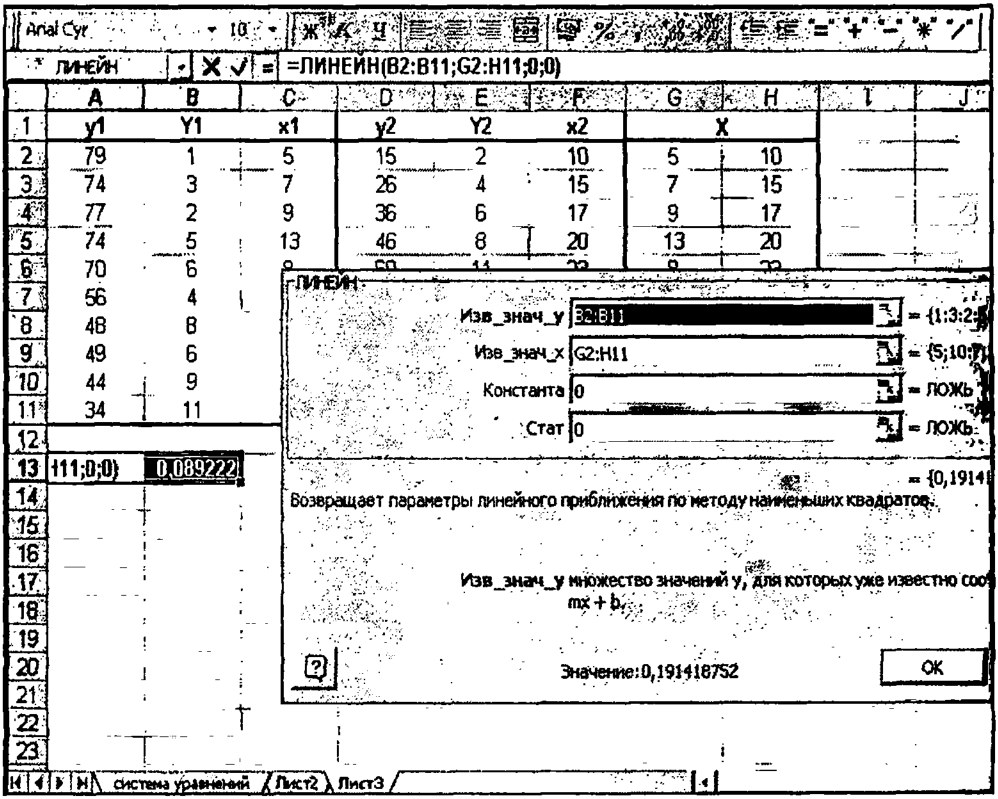


Рис. 24.1. Первое применение МНК в двухшаговой процедуре

Сформируем вектор коэффициентов приведенной формы, скопировав значение ячейки A13 в ячейку B14. Рассчитаем вектор прогнозных значений $\hat{Y}_{11}$ в столбец J2:J11, перемножив матрицу X на вектор коэффициентов приведенной формы:

=МУМНОЖ(G2:H11;B13:B14).

Теперь определим коэффициенты регрессии для первого уравнения между y_1 и матрицей $[\hat{Y}_{11}\ X_1]$, сформированной в ячейках J2:K11, функцией ЛИНЕЙН, введя ответ в ячейки F13:H17:

=ЛИНЕЙН(A2:A11;J2:K11,1,1).

Теперь проделаем аналогичные вычисления для второго уравнения. В ячейки D13:E13 введем коэффициенты приведенной формы:

=ЛИНЕЙН(E2:E11);G2:H11;0;0).

Сформируем вектор коэффициентов приведенной формы, вставив значение ячейки D13 в ячейку E14. Рассчитаем в L2:L11 вектор прогнозных значений $\hat{Y}_{12}$:

=МУМНОЖ(G2:H11;E13:E14).

Сформируем матрицу $[\hat{Y}_{12}\ X_2]$ в ячейках L2:M11 и определим коэффициенты регрессии для второго уравнения между y_2 и матрицей $[\hat{Y}_{12}\ X_2]$ функцией ЛИНЕЙН, введя ответ в ячейки I13:K17:

=ЛИНЕЙН(D2:D11;L2:M11,1,1).

Решение приведено на рис. 24.2.

	F	G	H	I	J	K	L	M
1	x2	X			Y1^	x1	Y2^	x2
2	10	5	10		2,3603	5	4,8032425	10
3	15	7	15		3,495838	7	6,9420441	15
4	17	9	17		4,057121	9	8,4283318	17
5	20	13	20		4,988267	13	11,183403	20
6	23	8	23		5,116411	8	9,2077202	23
7	18	15	18		4,783874	15	11,799672	18
8	28	20	28		7,144174	20	16,602915	28
9	30	18	30		7,348566	18	15,986645	30
10	34	22	34		8,471131	22	18,959221	34
11	40	24	40		9,798089	24	21,315527	40
12								
13	-1,53666	-2,27676	95,27273	2,000804	2,551019	-16,265		
14	0,808824	2,321323	4,656124	1,457119	2,492113	9,909965		
15	0,9224	5,106788	#Н/Д	0,919132	10,86915	#Н/Д		
16	41,60285	7	#Н/Д	39,7802	7	#Н/Д		
17	2169,945	182,555	#Н/Д	9399,132	826,9682	#Н/Д		
18								
19	-1,89987	-0,9808	20,46181	1,373123	1,023637	-1,64127		
20								
21	-3,31677	-0,78753	22,93902	5,018305	22,64345	0,600193		

Рис. 24.2. Окончательный результат применения двухшагового метода

Таким образом, система уравнений выглядит так:

$$y_1 = 95{,}272 - 2{,}277\ Y_1 - 1{,}537\ X_1,$$
$$y_2 = -\ 16{,}265 + 2{,}551\ Y_2 + 2{,}001\ X_2.$$

Теперь рассмотрим пример использования матричных функций. В общем случае использование встроенных функций для определения коэффициентов β уравнения

$$\beta = (X^TX)^{-1}X^TY$$

имеет вид:

=МУМНОЖ(МУМНОЖ(МОБР(МУМНОЖ(ТРАНСП(X);(X)); ТРАНСП(X)); Y).

Напомню, что заканчивать ввод нужно нажатием клавиш <Shift>+<Ctrl>+ +<Enter>. Если необходим свободный член уравнения, то матрица X должна иметь единичный вектор.

Коэффициенты приведенной формы для первого уравнения устанавливаются по выражению:

=МУМНОЖ(МУМНОЖ(МОБР(МУМНОЖ(ТРАНСП(X);(X)); ТРАНСП(X)); Y₁),

где $X = [X_1 X_2]$. Использование вместо Y_1 матрицы $Y = [Y_1 Y_2]$ даст сразу матрицу коэффициентов приведенной формы для обоих уравнений. Тогда матрицу прогнозных значений $\hat{Y}_1 = [\hat{Y}_{11}\ \hat{Y}_{12}]$ получим вводом функции

=МУМНОЖ(X;Y).

Учитывая, что Excel позволяет использовать 7 встроенных функций, матрицу прогнозных значений $\hat{Y}_1$ можно получить сразу, без промежуточных вычислений:

=МУМНОЖ(X;МУМНОЖ(МУМНОЖ(МОБР(МУМНОЖ(ТРАНСП(X);(X)); ТРАНСП(X)); (Y))).

Обратимся к нашему примеру. Для удобства работы сформируем матрицы. Скопируем значения в ячейки A24:A33 – y_1; в B24:B33 – y_2; в C24:C33 – Y_1; в D24:D33 – Y_2; в E24:E33 – X_1; в F24:F33 – X_2. В ячейках C35:D35 рассчитаем матрицу коэффициентов приведенной формы:

=МУМНОЖ(МУМНОЖ(МОБР(МУМНОЖ(ТРАНСП(E24:F33);E24:F33)); ТРАНСП(E24:F33));C24:D33).

В ячейках H24:I33 рассчитаем матрицу прогнозных значений $\hat{Y}_1$:

=МУМНОЖ(E24:F33;МУМНОЖ(МУМНОЖ(МОБР(МУМНОЖ (ТРАНСП(E24:F33);E24:F33));ТРАНСП(E24:F33));(C24:D33))).

Теперь образуем матрицы для расчета коэффициентов регрессий. Сделаем две матрицы. В первую должны войти единичный вектор, вектор прогнозных значений $\hat{Y}_{11}$ и X_1. Во вторую, соответственно, – единичный вектор, $\hat{Y}_{12}$ и X_2. Первую матрицу введем в диапазон J24:L33, вторую – в M24:O33. Выделим ячейки I36:I38 и определим параметры первого уравнения:

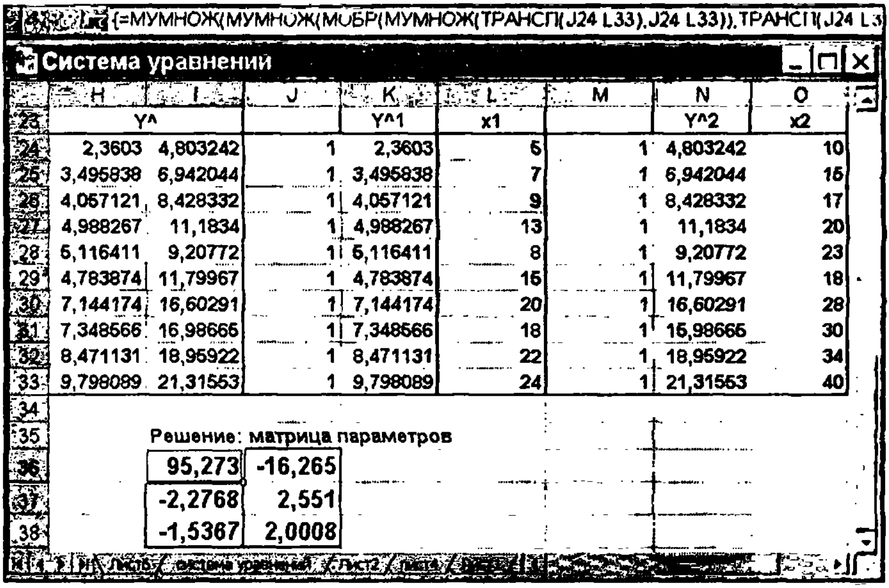
=МУМНОЖ(МУМНОЖ(МОБР(МУМНОЖ(ТРАНСП(J24:L33);J24:L33)); ТРАНСП(J24:L33));A24:A33).

Коэффициенты второго уравнения введем в ячейки J36:J38:

=МУМНОЖ(МУМНОЖ(МОБР(МУМНОЖ(ТРАНСП(M24:O33);M24:O33)); ТРАНСП(M24:O33));B24:B33).

Решение приведено на рис. 24.3.

Удобнее, конечно, окончательный результат выводить функцией ЛИНЕЙН для вывода статистик.



{=МУМНОЖ(МУМНОЖ(МОБР(МУМНОЖ(ТРАНСП(J24:L33),J24:L33)),ТРАНСП(J24:L3

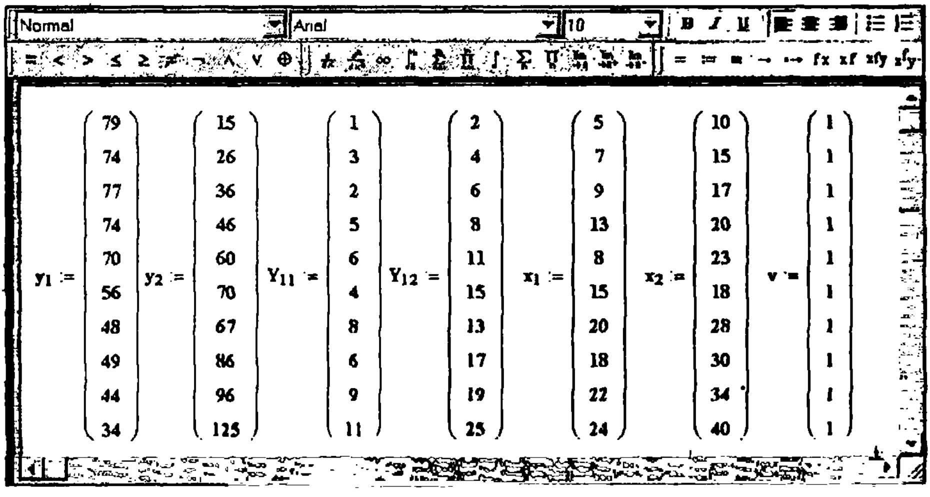
Система уравнений

	H	I	J	K	L	M	N	O
23	Y^			Y^1	x1		Y^2	x2
24	2,3603	4,803242	1	2,3603	5	1	4,803242	10
25	3,495838	6,942044	1	3,495838	7	1	6,942044	15
26	4,057121	8,428332	1	4,057121	9	1	8,428332	17
27	4,988267	11,1834	1	4,988267	13	1	11,1834	20
28	5,116411	9,20772	1	5,116411	8	1	9,20772	23
29	4,783874	11,79967	1	4,783874	15	1	11,79967	18
30	7,144174	16,60291	1	7,144174	20	1	16,60291	28
31	7,348566	16,98665	1	7,348566	18	1	16,98665	30
32	8,471131	18,95922	1	8,471131	22	1	18,95922	34
33	9,798089	21,31553	1	9,798089	24	1	21,31553	40
34								
35		Решение: матрица параметров						
36		95,273	-16,265					
37		-2,2768	2,551					
38		-1,5367	2,0008					

Рис. 24.3. Применение матричных функций для двухшагового МНК

24.5. Двухшаговый метод наименьших квадратов в Mathcad

Введем значения векторов y_1, y_2, Y_1, Y_2, x_1, x_2 и единичный вектор v, как показано на рис. 24.4. Для создания требующихся нам матриц будем пользоваться функцией augment, объединяющей две матрицы. Сначала создадим матрицу $[x_1\ x_2]$, обозначив ее X (рис. 24.5).



$$y_1 := \begin{pmatrix} 79 \\ 74 \\ 77 \\ 74 \\ 70 \\ 56 \\ 48 \\ 49 \\ 44 \\ 34 \end{pmatrix} \quad y_2 := \begin{pmatrix} 15 \\ 26 \\ 36 \\ 46 \\ 60 \\ 70 \\ 67 \\ 86 \\ 96 \\ 125 \end{pmatrix} \quad Y_{11} := \begin{pmatrix} 1 \\ 3 \\ 2 \\ 5 \\ 6 \\ 4 \\ 8 \\ 6 \\ 9 \\ 11 \end{pmatrix} \quad Y_{12} := \begin{pmatrix} 2 \\ 4 \\ 6 \\ 8 \\ 11 \\ 15 \\ 13 \\ 17 \\ 19 \\ 25 \end{pmatrix} \quad x_1 := \begin{pmatrix} 5 \\ 7 \\ 9 \\ 13 \\ 8 \\ 15 \\ 20 \\ 18 \\ 22 \\ 24 \end{pmatrix} \quad x_2 := \begin{pmatrix} 10 \\ 15 \\ 17 \\ 20 \\ 23 \\ 18 \\ 28 \\ 30 \\ 34 \\ 40 \end{pmatrix} \quad v := \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

Рис. 24.4. Задание векторов для двухшагового метода наименьших квадратов

Затем запишем формулы векторов прогнозных значений Y_{11} и Y_{12}, обозначив их YP_1 и YP_2. С помощью функции **augment** создадим матрицу $Z_1 = [YP_1\ x_1]$ и присоединим ее к единичному вектору $Z_{11} = [v\ Z_1]$. Для второго уравнения аналогично получим матрицу Z_{12}. Обозначим через β_1 и β_2 параметры первого и второго уравнений и введем формулы их определения. Результаты расчетов приведены на рис. 24.5.

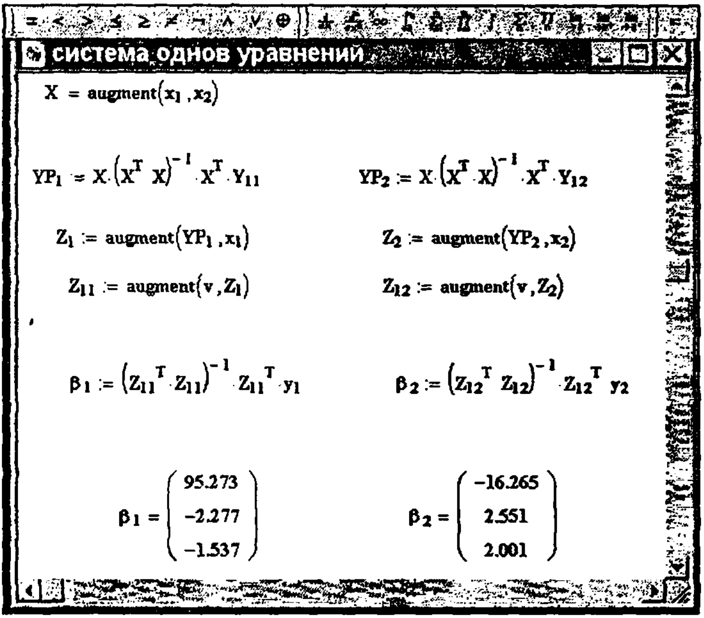


Рис. 24.5. Алгоритм двухшагового метода наименьших квадратов в Mathcad

24.6. Трехшаговый метод наименьших квадратов

Трехшаговый МНК предложили Зельнер и Тейл в качестве метода оценивания всех уравнений модели, который к тому же в определенных обстоятельствах может оказаться асимптотически эффективнее двухшагового метода. Рассмотрим общую линейную модель, содержащую Q взаимозависимых и K предопределенных переменных. Можно записать i-ое уравнение этой модели в виде:

$$y_i = Y_i\beta_i + X_i\gamma_i + u_i,\ i = 1,\ldots, Q,$$

где y_i – вектор, образованный из $n\times 1$ наблюдений над зависимой переменной i-го уравнения; Y_i – матрица порядка $n\times q$ наблюдений над остальными эндогенными переменными этого уравнения; X_i – матрица порядка $n\times k$ наблюдений над предопределенными переменными, входящими в i-ое уравнение; β_i и γ_i – векторы параметров, u_i – вектор возмущений.

Перепишем в виде:

$$y_i = Z_i\delta_i + u_i,\ i = 1,\dots,Q,$$

где $Z_i = [Y_iX_i]$ и $\delta_i = \begin{bmatrix}\beta_i \\ \gamma_i\end{bmatrix}$. (24.1)

Если последнее уравнение умножить слева на X^T, где X – матрица порядка $n\times k$, образованная всеми предопределенными переменными модели, то

$$X^T y_i = X^T Z_i\delta_i + X^T u_i,\ i = 1,\dots,Q. \quad (24.2)$$

Рассмотрим данное уравнение как соотношение между зависимой переменной X^Ty_i, объясняющими переменными X^TZ_i и вектором возмущений X^Tu_i. Тогда ковариационная матрица возмущений имеет вид:

$$E(X^T u_i u_i^T X) = \sigma_{ii} X^T X,$$

где σ_{ii} – постоянная дисперсия возмущения, испытываемого i-тым уравнением. В силу данного уравнения вектор δ_i, входящий в (24.2), может быть оценен с помощью обобщенного метода наименьших квадратов (обозначим оценку δ_i через d_i):

$$d_i = [Z_i^TX(X^TX)^{-1}X^TZ_i]^{-1}Z_i^TX(X^TX)^{-1}X^Ty_i. \quad (24.3)$$

Данное уравнение представляет собой другой способ записи двухшагового оператора оценивания для (24.1), в чем можно убедиться, если взять вместо матриц Y_i и X_i матрицу Z_i, а вместо β_i и γ_i вектор δ_i.

Теперь перепишем всю систему уравнений в виде (24.2) и воспользуемся такой матричной формой:

$$\begin{bmatrix} X^Ty_1 \\ X^Ty_2 \\ . \\ X^Ty_q \end{bmatrix} = \begin{bmatrix} X^TZ_1 & 0 & \dots & 0 \\ 0 & X^TZ_2 & \dots & 0 \\ . & 0 & . & . \\ 0 & 0 & \dots & X^TZ_q \end{bmatrix} \cdot \begin{bmatrix} \delta_1 \\ \delta_2 \\ . \\ \delta_q \end{bmatrix} + \begin{bmatrix} X^Tu_1 \\ X^Tu_2 \\ . \\ X^Tu_q \end{bmatrix}. \quad (24.4)$$

Матрица ковариаций для входящих в данное уравнение вектора возмущений будет тогда иметь вид:

$$V = \begin{bmatrix} \sigma_{11}X^TX & \sigma_{12}X^TX & \dots & \sigma_{1q}X^TX \\ \sigma_{21}X^TX & \sigma_{22}X^TX & \dots & \sigma_{2q}X^TX \\ . & . & \dots & . \\ \sigma_{q1}X^TX & \sigma_{q2}X^TX & \dots & \sigma_{qq}X^TX \end{bmatrix},$$

где через σ_{ij} обозначена ковариация одновременных структурных возмущений i-го и j-го уравнений. Образуем из элементов σ_{ij} матрицу Σ, тогда:

$$V = \Sigma\otimes(X^TX)$$

$$\text{и } V^{-1} = \Sigma^{-1}\otimes(X^TX)^{-1}.$$

Обобщенный МНК дает приближение оценки параметров (24.4). Однако для получения этих оценок необходимо знать матрицу V, которая, в свою очередь, зависит от неизвестной матрицы Σ. Предложение Зельнера–Тейла состоит в оценивании величин σ_{ij} на основе рассчитанных с помощью двухшаговой

...едуры значений возмущающих воздействий. Другими словами, двухшаго... ...роцедура применяется при оценивании по формуле (24.3) вектора d_i для ка... ...о структурного уравнения, после чего полученные оценки d_i подставляются ...2). Это позволяет рассчитать каждый из векторов $\hat{u}_i$, а с их помощью полу... оценки S_{ij} для σ_{ij}. Таким образом, трехшаговый оператор, позволяющий ...чить оценку $\bar{\bar{\delta}}$, имеет вид:

$$\left\{\begin{bmatrix} Z_1^T X & 0 & \dots & 0 \\ 0 & Z_2^T X & \dots & 0 \\ . & . & . & . \\ 0 & 0 & \dots & Z_q^T X \end{bmatrix} \times \begin{bmatrix} S^{11}(X^TX)^{-1} & S^{12}(X^TX)^{-1} & \dots & S^{1q}(X^TX)^{-1} \\ . & . & . & . \\ S^{q1}(X^TX)^{-1} & S^{q2}(X^TX)^{-1} & \dots & S^{qq}(X^TX)^{-1} \end{bmatrix} \times \right.$$

$$\left. \times \begin{bmatrix} X^T Z_1 & 0 & \dots & 0 \\ 0 & X^T Z_1 & \dots & 0 \\ . & . & . & . \\ 0 & 0 & \dots & X^T Z_1 \end{bmatrix} \right\}^{-1} \times$$

$$\begin{bmatrix} Z_1^T X & 0 & \dots & 0 \\ 0 & Z_2^T X & \dots & 0 \\ . & . & . & . \\ 0 & 0 & \dots & Z_q^T X \end{bmatrix} \times \begin{bmatrix} S^{11}(X^TX)^{-1} & S^{12}(X^TX)^{-1} & \dots & S^{1q}(X^TX)^{-1} \\ . & . & . & . \\ S^{q1}(X^TX)^{-1} & S^{q2}(X^TX)^{-1} & \dots & S^{qq}(X^TX)^{-1} \end{bmatrix} \times \begin{bmatrix} X^T y_1 \\ X^T y_q \end{bmatrix}.$$

...ерез S^{ij} обозначены элементы матрицы $[S_{ij}]^{-1}$. После упрощения получим:

$$\begin{bmatrix} S^{11} Z_1^T X (X^TX)^{-1} X^T Z_1 & \dots & S^{1q} Z_1^T X (X^TX)^{-1} X^T Zq \\ . & . & . \\ S^{q1} Z_q^T X (X^TX)^{-1} X^T Z_1 & \dots & S^{qq} Z_q^T X (X^TX)^{-1} X^T Z_q \end{bmatrix} \times \begin{bmatrix} \sum_{j=1}^{q} S^{1j} Z_1^T X (X^TX)^{-1} X^T y_j \\ . \\ \sum_{j=1}^{q} S^{qj} Z_q^T X (X^TX)^{-1} X^T y_j \end{bmatrix}.$$

Оценку асимптотической матрицы ковариаций даст обратная матрица, со... ...кащаяся в правой части.

Трехшаговая процедура обеспечивает лучшую по сравнению с двухшаговым ...одом асимптотическую эффективность оценок лишь в случае, когда матрица ... является диагональной, т. е. когда одновременные возмущения, входящие в ...ичные структурные уравнения, коррелируют друг с другом.

Применение этого подхода на практике требует нескольких дополнительных ...аний.

- Каждое уравнение, являющееся определением (т. е. все тождества), нужно исключить из системы прежде, чем приступить к вычислениям.
- Аналогично следует исключить каждое неидентифицируемое уравнение.

- Остаются только идентифицируемые и сверхидентифицируемые уравнения, причем с вычислительной точки зрения целесообразно применять трехшаговую процедуру к каждой из этих групп уравнений отдельно.
- Для группы сверхидентифицируемых уравнений вычисляют оценки трехшаговым методом на основе последней формулы, взяв в качестве q число сверхидентифицируемых уравнений. Если имеется только одно сверхидентифицируемое уравнение, то трехшаговая процедура вырождается в двухшаговую.

Если матрица ковариаций Σ для структурных возмущений блочно-диагональная, то вся процедура трехшагового оценивания может быть применена отдельно к каждой группе уравнений, соответствующих каждому блоку.

Глава 25. Оптимизационные модели производства

25.1. Формализация задачи оптимизации производства в условиях конкуренции

Будем считать, что предприятие стремится максимизировать прибыль, которую оно извлекает из производства в условиях технологических ограничений. Будем также считать, что предприятие находится в условиях совершенной конкуренции, если цена каждого блага определена экзогенно для каждого предприятия и, следовательно, не зависит от производственных решений производителя и если по данной системе цен предприятие может приобрести все необходимое количество ресурсов и сбыть всю производимую им продукцию.

Предполагается, что предприятие достаточно мало́ по сравнению с рынком, так что его действия не сказываются на ценах. Более того, предполагается, что спрос и предложения других участников экономического процесса абсолютно гибки, т. е. что они немедленно реагируют на любое предложение или любой спрос рассматриваемого предприятия.

Основная цель фирмы заключается в максимизации прибыли путем рационального распределения затрачиваемых ресурсов.

Рассмотрим вопрос в случае, если предприятие выпускает один товар, используя два ресурса (фактора производства) x_1 и x_2. Выпуск продукции определяется производственной функцией $y = f(x_1,x_2)$. Рыночная цена продукции – p, цены ресурсов – c_1 и c_2. Функция затрат $C = c_1x_1 + c_2x_2$. Тогда задача максимизации прибыли (PR) в случае долговременного промежутка будет иметь следующий вид:

$$PR(x_1,x_2) = p\,f(x_1,x_2) - (c_1x_1 + c_2x_2) \to \max$$

при $x_1 > 0, x_2 > 0$.

В случае короткого промежутка времени предприятие должно учитывать неизбежные ограничения на объемы используемых ресурсов:

$$h(x_1,x_2) \le b.$$

Следовательно, задача максимизации прибыли для короткого промежутка времени имеет вид задачи на условный максимум:

$$PR(x_1,x_2) = p\,f(x_1,x_2) - (c_1x_1 + c_2x_2) \to \max$$

при $h(x_1,x_2) \le b$,

$$x_1 > 0, x_2 > 0.$$

25.2. Предельные свойства равновесия

В случае долговременного промежутка задача максимизации прибыли представляет собой классическую задачу безусловной оптимизации. Решение задачи (x_1^0, x_2^0) находится из необходимых условий первого порядка:

$$PR'_{x_1} = 0,\ PR'_{x_2} = 0,$$

или

$$p\,f'_{x_1} = c_1,\ p\,f'_{x_2} = c_2.$$

Условие второго порядка:

$$PR''_{x1} < 0,$$

$$PR''_{x1}\,PR''_{x2} - (PR''_{x1x2})^2 > 0.$$

В условиях первого порядка $f'_{x_1} = \dfrac{\partial f(x_1, x_2)}{\partial x_1}$ и $f'_{x_2} = \dfrac{\partial f(x_1, x_2)}{\partial x_2}$ предельные производительности ресурсов – x_1 и, соответственно, x_2. В точке равновесия предельные производительности ресурсов равны:

$$f'_{x1} = c_1/p,\ f'_{x2} = c_2/p,$$

т. е. предельная производительность ресурса x_i должна равняться отношению цены этого ресурса к цене выпускаемой продукции.

Если тождества из необходимых условий первого порядка почленно поделить, то получим

$$f'_{x1}/f'_{x2} = c_1/c_2.$$

Это означает, что в точке равновесия отношение предельных производительностей ресурсов равно отношению их цен. Нам уже известно, что отношение предельных производительностей ресурсов является предельной нормой замещения. Таким образом, в точке равновесия предельная норма замещения между x_1 и x_2 равна отношению их цен.

Из условий второго порядка (функция имеет максимум, когда в стационарной точке второй дифференциал есть отрицательно определенная или полуопределенная квадратичная форма) приходим к предположению невозрастания предельной выработки. Это предположение выполняется в окрестности равновесия для предприятия.

Это условие второго порядка вскрывает важное свойство: предприятие не может находиться в конкурентном равновесии в точке, в окрестности которой удельный выпуск возрастает.

Таким образом, равновесие при совершенной конкуренции несовместимо с возрастанием удельного выпуска, которое обычно характеризует недифференцируемый сектор. Поддержание равновесия для этого сектора требует установления других, нежели совершенная конкуренция, форм институциональной организации.

Изокоста и изокванта

При рассмотрении производственных функций были введены понятия изокванты и изоклинали. Напомним, что изокванта – это функция одного ресурса от другого при постоянном выпуске. Изоклиналь – тоже функция одного ресурса от другого при постоянной норме замещения.

Введем понятие изокосты. Это изолиния функции затрат. То есть это функция одного ресурса от остальных при постоянном уровне затрат. Она показывает сочетание количеств ресурсов, обеспечивающих данный постоянный уровень затрат. Если функция затрат будет иметь вид:

$$C = c_1x_1 + c_2x_2,$$

то для определенного уровня затрат C_n функция изокосты определяется из выражения:

$$C_n = c_1x_1 + c_2x_2,$$

т. е. она имеет вид:

$$x_2(x_1) = (C_n - c_1x_1) / c_2.$$

Построим в Excel изокосты для функции затрат вида $C = 2x_1 + 4x_2$. Прежде всего построим график функции затрат (рис. 25.1). Изолинии на этом графике и есть изокосты, построенные по цене деления.

Этот же график, представленный в виде контурной диаграммы, наглядно показывает изокосты в осях ресурсов (рис. 25.2). Видно, что для данной функции изокосты представляют собой прямые линии.

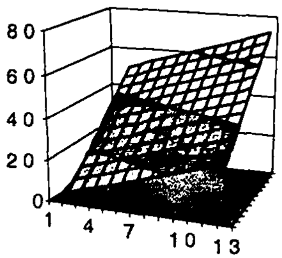


Рис. 25.1

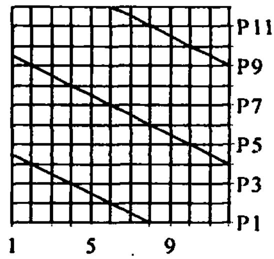


Рис. 25.2

Построим изокосты, например, для уровня затрат 20, 40, 60. Рассчитаем значения функций изокост (рис. 25.3).

Arial Cyr 10

R2 =($O$3-2*P2)/4

	O	P	Q	R	S	T
1	уровень	x1	x2	изокоста-20	изокоста-40	изокоста-60
2	затрат	0	0	5	10	15
3	20	1	1	4,5	9,5	14,5
4	40	2	2	4	9	14
5	60	3	3	3,5	8,5	13,5
6		4	4	3	8	13
7		5	5	2,5	7,5	12,5
8		6	6	2	7	12
9		7	7	1,5	6,5	11,5
10		8	8	1	6	11
11		9	9	0,5	5,5	10,5
12		10	10	0	5	10
13		11	11	-0,5	4,5	9,5
14						

Рис. 25.3

Графики изокост показаны на рис. 25.4.

Как видно из рисунков, изокосты представляют собой прямые линии с определенным углом наклона. Этот угол наклона для изокост вида

$$x_2(x_1) = (C_n - c_1x_1)/c_2$$

равен x_2/x_1. По отношению отрезков на осях координат это отношение равно

$$(C_n/c_2) / (C_n/c_1) = c_1/c_2.$$

То есть угол наклона изокост равен отношению цен ресурсов c_1/c_2.

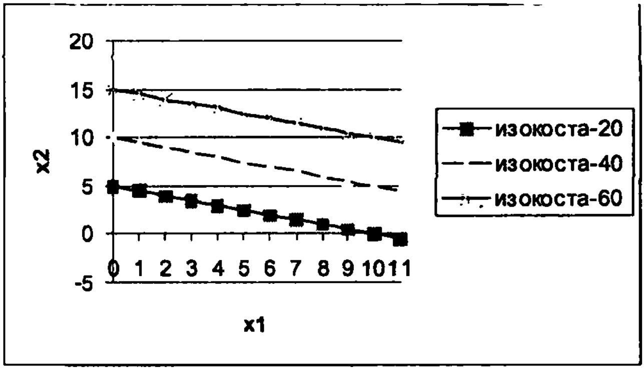


Рис. 25.4

Решив задачу максимизации прибыли и найдя рыночное равновесие предприятия (x_1^0, x_2^0), построим для данного оптимального соотношения ресурсов изокванту и изокосту. Уравнение изокванты имеет вид: $f(x_1,x_2) = y_0$, где $y_0 = f(x_1^0,x_2^0)$. Уравнение изокосты соответствует $c_1x_1 + c_2x_2 = C_0$, где $C_0 = c_1x_1^0 + c_2x_2^0$.

Касательная к изокванте в точке с оптимальными ресурсами (x_1^0, x_2^0) будет равна

$$\operatorname{tg}\varphi = -\frac{dh(x_1^0)}{dx_1} = \frac{\partial f(x_1^0,x_2^0)}{\partial x_1} \Big/ \frac{\partial f(x_1^0,x_2^0)}{\partial x_2}.$$

Изокоста, построенная для данного оптимального соотношения ресурсов (x_1^0, x_2^0), будет иметь тот же наклон, равный предельной норме замещения. Таким образом, изокоста будет являться касательной к изокванте в точке (x_1^0, x_2^0).

Рассмотрим пример в Excel. Примем производственную функцию степенного вида $y = a_0\, x_1^a\, x_2^b$, функцию затрат – $C = c_1x_1 + c_2x_2$. Функция прибыли будет равна $PR = p\ \ a_0\ x_1^a\ x_2^b - (c_1x_1 + c_2x_2)$. Значения коэффициентов приведены на рис. 25.5. Установим максимум функции прибыли, используя опцию **Поиск решения** (рис. 25.5).

Для найденных количеств ресурсов $x_1^0 = 9{,}152$ и $x_2^0 = 12{,}203$ установим соответствующие значения выпуска y_0 = B4*B1^B5*B2^B6, затрат C_0 = B7*B1 + B8*B2 и предельную норму замещения γ = $B$5*$B$2/$B$6/$B$1 (рис. 25.6). Используя установленные значения, рассчитаем функции:

- изокванты = ($B$9/($B$4*C2^$B$5))^(1/$B$6);
- изокосты = ($B$10 – $B$7*C2)/$B$8;
- изоклинали = $B$11*$B$6/$B$5*C2.

Построенные по рассчитанным данным графики изокванты, изокосты и изоклинали приведены на рис. 25.7.

Как видно из рисунка, в точке (x_1^0, x_2^0) изокванта и изоклиналь пересекаются, а изокоста является к ней касательной.

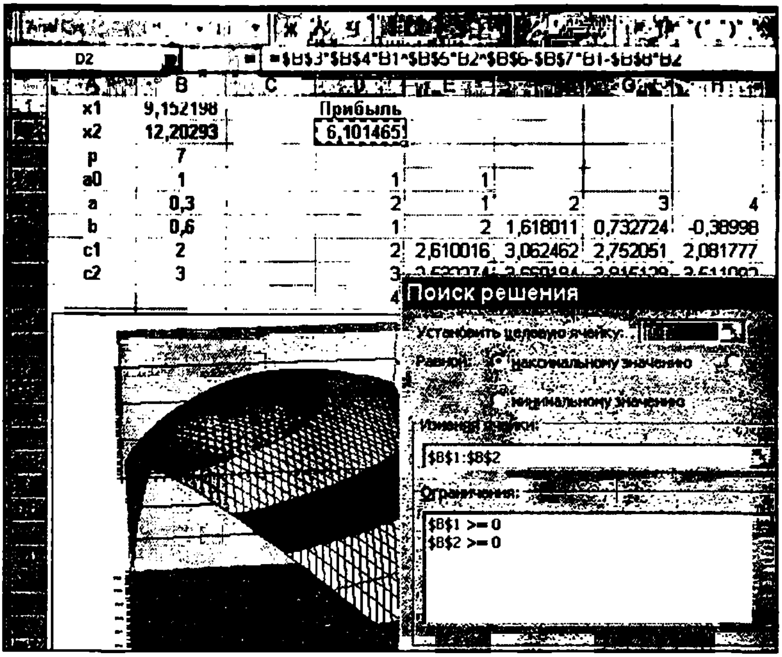


Рис. 25.5

Arial Cyr 10

E11 =($B$10-$B$7*C11)/$B$8

x1	9,152198	x1	изокванта	изокоста	изоклиналь
x2	12,20293	0,1	116,7	18,23773	0,133333
p	7	0,5	52,2	17,97106	0,666666
a0	1	1	36,9	17,63773	1,333333
a	0,3	1,5	30,1	17,30439	1,999999
b	0,6	2	26,1	16,97106	2,666666
c1	2	3	21,3	16,30439	3,999999
c2	3	4	18,5	15,63773	5,333331
y_0	8,716378	5	16,5	14,97106	6,666664
C_0	54,91318	6	15,1	14,30439	7,999997
y	0,666668	9,152198	12,20293	12,20293	12,20293
		10	11,7	11,63773	13,33333
		11	11,1	10,97106	14,66666
		12	10,7	10,30439	15,99999
		13	10,2	9,637726	17,33333
		14	9,9	8,971059	18,66666
		15	9,5	8,304393	19,99999
		16	9,2	7,637726	21,33333
		17	9,0	6,971059	22,66666
		18	8,7	6,304393	23,99999
		19	8,5	5,637726	25,33332
		20	8,3	4,971059	26,66666

Рис. 25.6

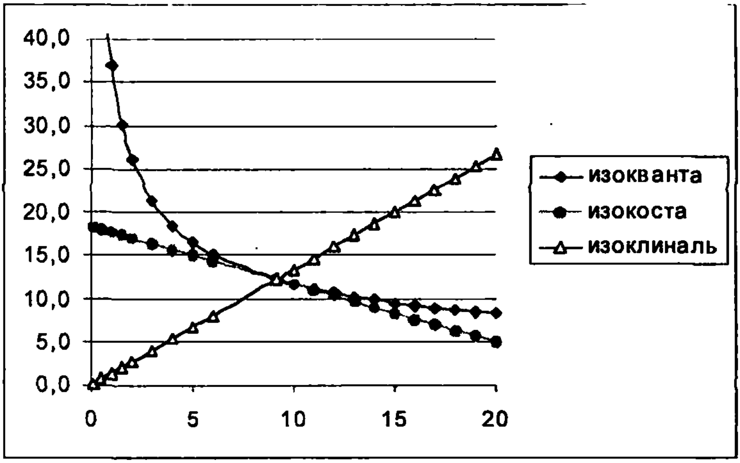


Рис. 25.7

25.3. Разновидности задач оптимизации предприятия

Различают следующие случаи оптимизации предприятия: в зависимости от времени решения задачи (долговременные и краткосрочные задачи); при наличии ограничений на ресурсы и без них; задача максимизации прибыли и задача минимизации затрат при постоянном выпуске; модели с учетом фактора времени и неопределенности и без них, и, наконец, институциональные условия – совершенная конкуренция и различные случаи несовершенной конкуренции [14, 21, 23, 37].

Задача максимизации прибыли в случае долговременного промежутка времени рассмотрена выше.

В случае **короткого промежутка времени** предприятие не может изменить второй ресурс (например, основные фонды). В этом случае добавляется ограничение по данному ресурсу. Так, в рассматриваемом ранее примере примем ограничение по первому ресурсу $x_1 = 7$. Решение задачи приведено на рис. 25.8.

Построим изокосту для этого случая (рис. 25.9).

Как видно из рисунка, изокоста не является касательной к изокванте, а пересекает ее в точке оптимума.

Ограничения на затраты

При долговременном промежутке задача максимизации прибыли в случае ограничения затрат на приобретение ресурсов будет выглядеть так:

$$PR(x_1,x_2) = p\,f(x_1,x_2) - (c_1x_1 + c_2x_2) \to \max$$

при $$x_1 > 0,\ x_2 > 0.$$

$$c_1x_1 + c_2x_2 \le C_c.$$

Это задача на условный экстремум. Решим ее с помощью метода Лагранжа в случае $c_1x_1 + c_2x_2 = C_c$. Функция Лагранжа будет равна:

$$L(x_1,x_2,\lambda) = p\,f(x_1,x_2) + \lambda(C_c - c_1x_1 - c_2x_2).$$

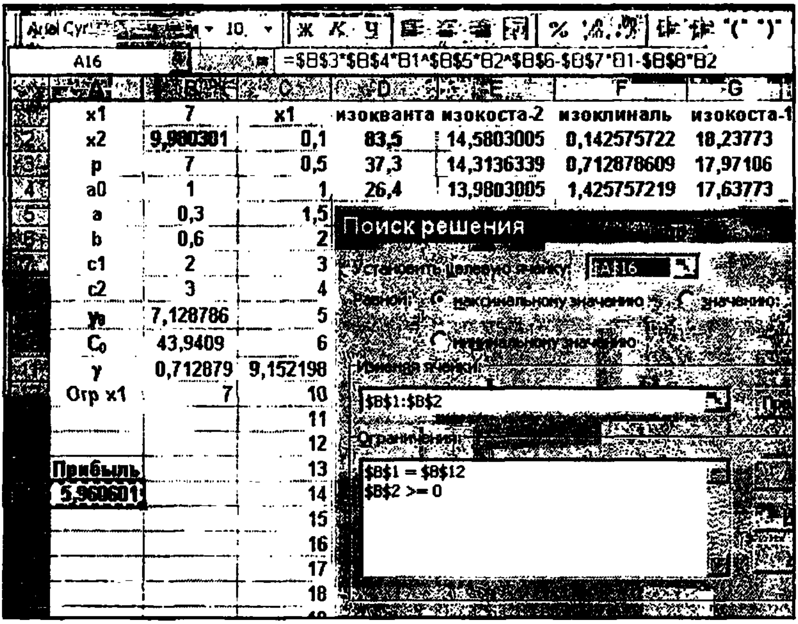


Рис. 25.8

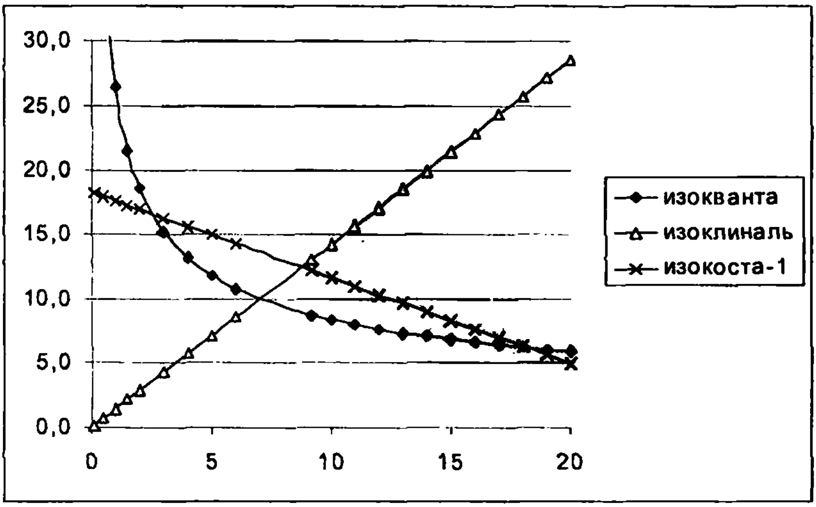


Рис .25.9

Решение задачи находится из уравнений:

$$L'_{x1} = 0,\ L'_{x2} = 0,\ L'_{\lambda} = 0.$$

Или в развернутом виде:

$$L'_{x1} = p f'_{x1} - \lambda c_1 = 0;$$
$$L'_{x2} = p f'_{x2} - \lambda c_2 = 0;$$
$$L'_{\lambda} = C_c - c_1 x_1 - c_2 x_2 = 0.$$

Из первых двух уравнений получаем:

$$p f'_{x1} = \lambda c_1;$$
$$p f'_{x2} = \lambda c_2.$$

Откуда в состоянии равновесия соблюдается:

$$f'_{x1}/f'_{x2} = c_1/c_2.$$

То есть получен аналитический вывод, что в состоянии равновесия предельная норма замещения ресурсов равна отношению цен на эти ресурсы.

Рассмотрим решение задачи в Excel. Для приведенного выше примера введем ограничение

$$2x_1 + 3x_2 \le 40.$$

Решение задачи показано на рис. 25.10.

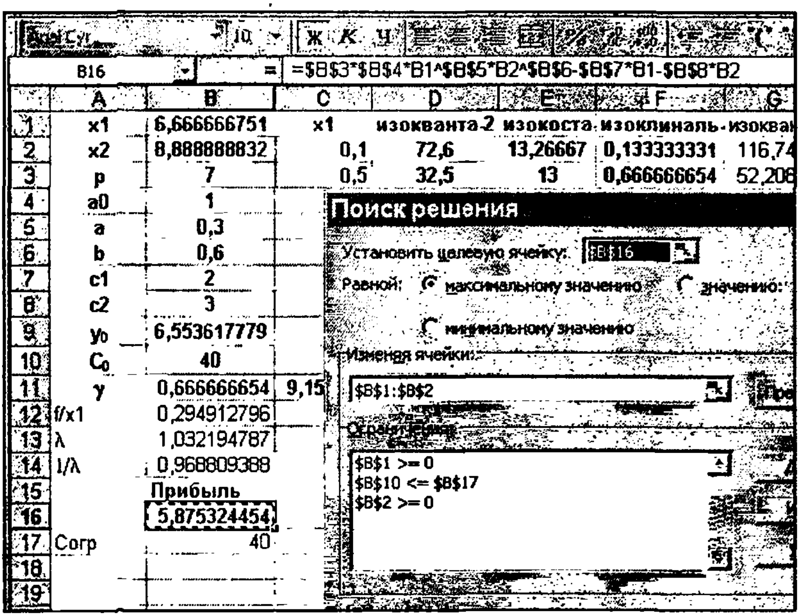


Рис. 25.10

Построим изокванту и изокосту для нового оптимума (рис. 25.11).

При ограничениях на ресурсы оптимальное состояние характеризуется меньшим количеством задействованных ресурсов, меньшим оптимальным выпуском. Поэтому изокванта и изокоста находятся ниже соответствующих изокванты и изокосты, установленных из условия безусловного максимума.

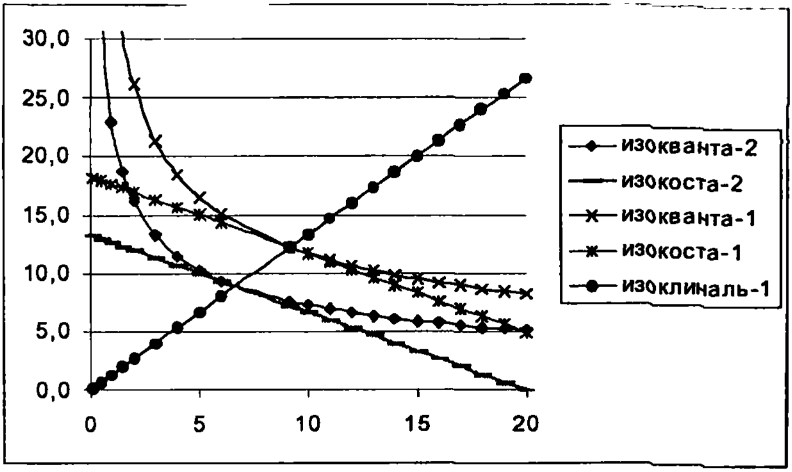


Рис. 25.11

Минимизация затрат

В случае постоянно заданного выпуска рациональное сочетание используемых ресурсов можно установить и из задачи минимизации затрат.

Задача *максимизации прибыли* при постоянном выпуске (y_0) будет формализована следующим образом:

$$PR(x_1,x_2) = p\,y_0 - (c_1x_1 + c_2x_2) \to \max$$

при $$x_1 > 0,\ x_2 > 0.$$

$$f(x_1,x_2) = y_0.$$

Решение ее методом Лагранжа будет выглядеть так:

$$L(x_1,x_2,\lambda) = p\lambda\,(y_0 - f(x_1,x_2) - (c_1x_1 - c_2x_2);$$

$$L'_{x1} = \lambda\,p\,f'_{x1} - c_1 = 0;$$

$$L'_{x2} = \lambda\,p\,f'_{x2} - c_2 = 0;$$

$$L'_\lambda = p\,(y_0 - f(x_1,x_2) = 0.$$

Задачу *минимизации затрат* при постоянном выпуске (y_0) можно формализовать следующим образом:

$$C(x_1,x_2) = (c_1x_1 + c_2x_2) \to \min$$

при $$x_1 > 0,\ x_2 > 0.$$

$$f(x_1,x_2) = y_0.$$

Решение задачи минимизации затрат методом Лагранжа будет выглядеть так:

$$L(x_1,x_2,\lambda) = (c_1x_1 - c_2x_2) + p\lambda\,(y_0 - f(x_1,x_2);$$

$$L'_{x1} = c_1 - \lambda\,p\,f'_{x1} = 0;$$

$$L'_{x2} = c_2 - \lambda\,p\,f'_{x2} = 0;$$

$$L'_\lambda = p\,(y_0 - f(x_1,x_2) = 0.$$

Очевидно, что задачи идентичны.

Найдем в Excel решение задачи минимизации затрат при постоянном выпуске $y_0 = 5$. Решение приведено на рис. 25.12.

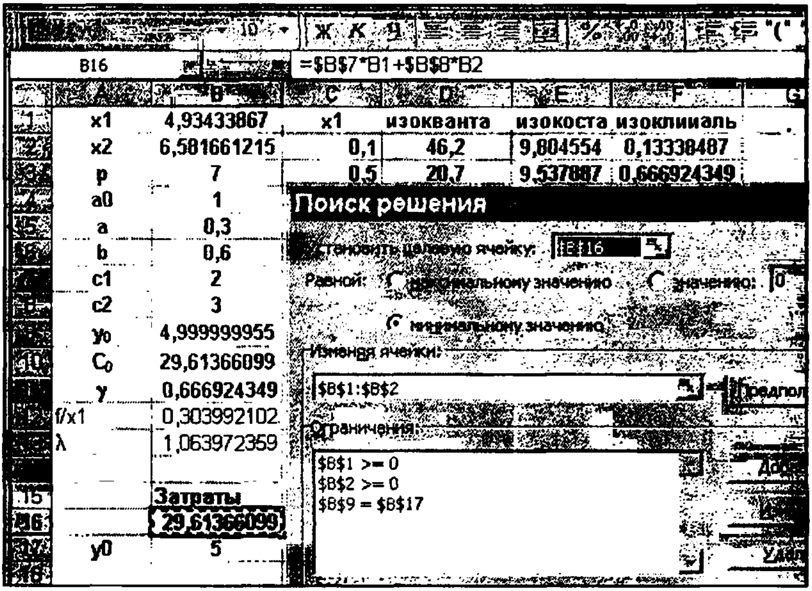


Рис. 25.12

Графически поиск решения в данной задаче представляет собой поиск такой изокосты, которая является касательной к фиксированной изокванте для y_0. Графическое представление изокосты и изокванты показано на рис. 25.13.

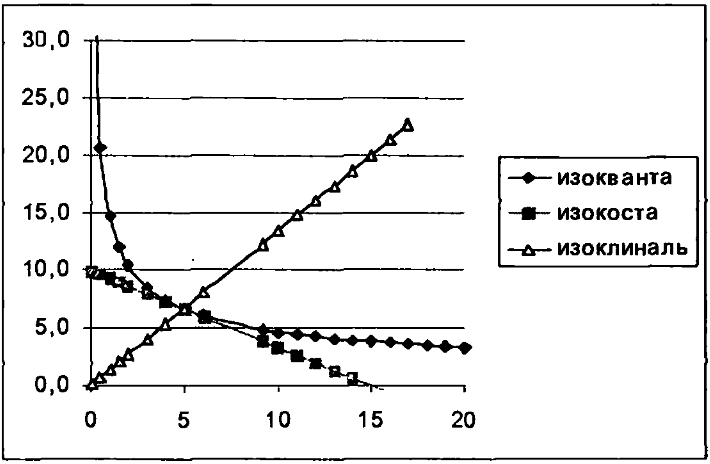


Рис. 25.13

Долговременная линия развития предприятия

На графиках изолиний можно представить линию долговременного развития предприятия как переход из одного оптимального состояния в другое при прогнозировании цены, затрат и технологических ограничений в виде производственной функции.

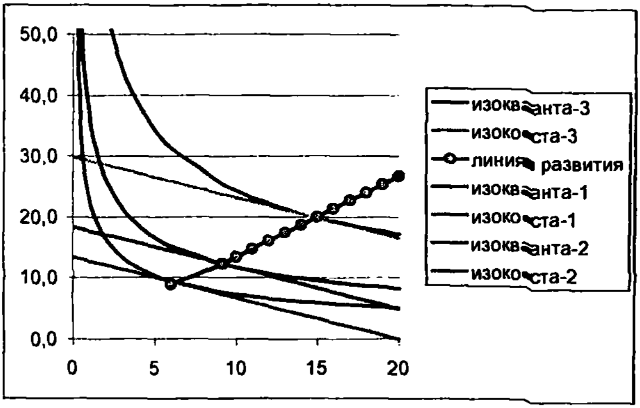


Рис. 25.14

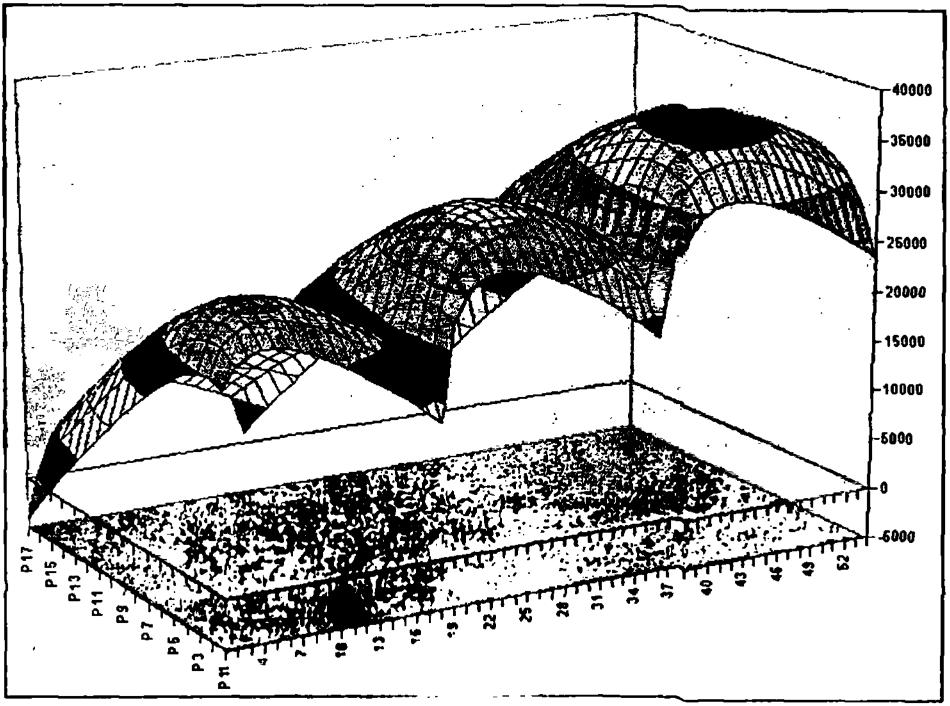


Рис. 25.15 Графики прибыли

Рассмотрим следующий пример для трех состояний, которые будут соответ-зовать определенному времени, например трем годам. В первом году примем вновесное состояние, взятое из задачи с ограничением на затраты. В следую-м году будем считать, что таких ограничений нет. И в третьем году будем итать, что цена вырастет на 5%. Рассчитаем соответствующие оптимумы для ех состояний и построим для них графики изоквант и изокост (рис. 25.14).

Данная линия долговременного развития будет находиться на изоклинали, сколько все оптимальные состояния рассчитывались при одной и той же про-зодственной функции и, следовательно, предельная норма замещения одина-ва. Как только предприятие перейдет на другую производственную функцию, ния долговременного развития будет соединять точки оптимального состоя-я, переходя на другую изоклиналь.

Графики прибыли, характеризующие 3 состояния предприятия, приведены рис. 25.15. Соответственно, линия долговременного развития предприятия дет представлять собой линию, соединяющую максимумы прибыли в каждом тоянии.

25.4. Монополия

Рассмотрим кратко классическую теорию монополии. В прикладных иссле-заниях структуры рынка говорят, что предприятие имеет монополию на рын-продукта h, если только оно одно снабжает рынок этим продуктом и если рос на этот продукт исходит от многочисленных небольших и действующих ависимо друг от друга участников экономического процесса. Классическая рия монополии представляет эту ситуацию исходя из предположения, что ка-дя единица продукта h будет продаваться по одной и той же цене p_h, завися-й от общего количества y_h продукта h, которое монополист согласится выбро-ь на рынок. Таким образом, монополист сталкивается со спросом, величина орого изменяется вместе с ценой на продукцию, но сам характер изменения нкция спроса) не зависит от его решений.

Представим, например, что предприятие производит продукт 1 и продает его на нке, где спрос многочисленных потребителей определяется только ценой p_1 данно-родукта. Мы можем представить спрос в виде соотношения между p_1 и y_1:

$$p_1 = \pi_1(y_1),$$

π_1 – функция, определяющая цену p_1, по которой монополист предполагает одать количество y_1 своей продукции.

Точно так же может оказаться, что только одно предприятие использует не-орый фактор h (например, если только оно использует рабочую силу в неко-ом городе). В этом случае говорят, что предприятие пользуется положением онсонии. Оно считает в этом случае, что цена p_h зависит от количества $a_h = y_h$, орое и используется в качестве ресурсов производства. Если предприятие не нимает во внимание существующую зависимость между p_h и ценами других г, то оно будет принимать решения на основании функции предложения:

$$p_h = \pi_h(y_h).$$

Данная функция определяет поведение участников, предоставляющих благо цену p_h, которую предприниматель считает нужным заплатить за предостав-ие ему блага h в количестве y_h.

Заметим, что случай совершенной конкуренции соответствует особой ситуации, при которой функции π_l и π_h являются константами. Мы можем одновременно иметь дело как с монополией, так и с монопсонией на один или несколько факторов, рассматривая ситуацию, когда предприятие стремится максимизировать прибыль, принимая во внимание функции π_h, связывающие цену каждого блага с чистой продукцией этого блага h ($h = 1,2,\ldots, l$).

1. Если цены факторов производства фиксированы, производственные затраты монополиста могут быть выражены как функция объема продукции. Пусть $C(y)$ является общими затратами на производство объема продукции y.

Если цена продукции равна p_n, а $p_n = \pi_n(y)$ – кривая спроса, то прибыль монополиста от продажи выпущенной продукции y составит:

$$PR = \pi_n(y)\, y - C(y).$$

Прибыль достигает максимума, когда

$$\pi_n'(y)\, y + \pi_n(y) - C'(y) = 0,$$

или

$$\pi_n'(y)\, y + \pi_n(y) = C'(y).$$

Выражение в левой части называется предельным доходом (выручкой):

$$R'(y) = \pi_n'(y)\, y + \pi_n(y),$$

а в правой части – предельными затратами. Таким образом, первое условие монополистического равновесия состоит в равенстве предельной выручки предельным затратам $R'(y) = C'(y)$.

Предельную выручку можно представить в виде:

$$R'(y) = \pi_n(y) \cdot \left(\frac{\pi_n'(y)\, y}{\pi_n(y)} + 1 \right).$$

Учитывая, что обратная ценовая эластичность спроса равна:

$$\frac{1}{\varepsilon_{d/p}} = \frac{y}{\pi_n(y)} \cdot \pi_n'(y),$$

$$R'(y) = \pi_n(y) \cdot \left(1 - \frac{1}{\varepsilon_{d/p}} \right),$$

или

$$R'(y) = \pi_n(y) \cdot \left(\frac{\varepsilon - 1}{\varepsilon} \right),$$

тогда можно представить условие равновесия в виде:

$$\pi_n(y)\, (\varepsilon - 1)/\varepsilon = C'(y).$$

Цена в условиях равновесия:

$$\pi_n(y) = C'(y)\, \varepsilon/(\varepsilon - 1).$$

Из второго условия максимизации следует, что вторая производная функции прибыли должна быть отрицательна, что можно представить как:

$$R''(y) < C''(y).$$

Пока кривая предельной выручки понижается круче, чем кривая предельных затрат, монополистическое равновесие стабильно. Все случаи, где кривая предельной выручки понижается, а кривая предельных затрат растет, стабильны.

15 Зак. 1010

При понижающейся кривой предельных затрат стабильное равновесие возможно до тех пор, пока понижение кривой затрат меньше понижения кривой предельной выручки, а также пока общая выручка превышает общие затраты на величину, достаточную, чтобы монополист удержался в бизнесе.

2. Монополия и факторы производства. Для анализа данной проблемы удобно рассматривать монополиста как того, кто владеет некоторыми факторами производства и нанимает другие. Если он не может изменить предложение этих частных факторов, то весьма справедливо предположение, что он будет стремиться довести до максимума свои прибыли, иначе говоря, довести до максимума чистые доходы от своих факторов [37].

Если даны количества нанятых монополистом факторов – a, b, c, ..., их цены π_a, π_b, π_c, ... и их кривые предложения, то прибыль монополии:

$$PR = py - a\pi_a - b\pi_b - c\pi_c - \ldots$$

Выражение максимизируется, когда:

$$\left(p + y\frac{dp}{dy}\right)dy - \left(\pi_a + a\frac{d\pi_a}{da}\right)da - \left(\pi_a + b\frac{d\pi_b}{db}\right)db - \ldots = 0,$$

что можно записать как:

$$R'_y\,dy - C'_a\,da - C'_b\,db - \ldots = 0,$$

где $C'_a = \pi_a + a\dfrac{d\pi_a}{da} = MC_a$ и т. д.

Возьмем $y = f(a, b, c, \ldots)$ как технически данную производственную функцию, тогда:

$$dy = \frac{\partial y}{\partial a}da + \frac{\partial y}{\partial b}db + \frac{\partial y}{\partial c}dc + \ldots$$

Подставив это выражение в предыдущее уравнение, получим:

$$\left(R'_y\frac{\partial y}{\partial a} - C'_a\right)da + \left(R'_y\frac{\partial y}{\partial b} - C'_b\right)db + \ldots = 0.$$

Поскольку полученное выражение должно оставаться справедливым для всех значений da, db, ..., следует, что

$$R'_y\frac{\partial y}{\partial a} = C'_a,\quad R'_y\frac{\partial y}{\partial b} = C'_b, \ldots$$

для всех факторов.

C'_a, C'_b, ... – предельные затраты монополиста на найм дополнительной единицы факторов a, b, ... Если кривые предложения факторов направлены вверх, эти предельные затраты превысят цены факторов на $a(d\pi_a/da)$ и т. д. соответственно, т. е. на дополнительные суммы, которые должны быть выплачены за предшествующие единицы с целью поддержания их цен на уровне цены предельной единицы фактора.

$R'_y(\partial y/\partial a)$ определяется как "предельная ценность продукта" фактора a; это прирост общей ценности продукта, создаваемый применением дополнительной единицы фактора a. Условие равновесия на рынке фактора заключается в том, что предельная ценность продукта какого-либо фактора должна равняться его предельным затратам.

Представим это условие в виде:

$$(p_n + yp'_y)\, y'_a = \pi_a + y\pi'_a,$$

или, используя ценовую эластичность спроса, как:

$$(p_n(\varepsilon_p - 1)/\varepsilon_p)\, y'_a = \pi_a\, (\varepsilon_a - 1)/\varepsilon_a.$$

Отсюда предельная эффективность фактора a будет равна:

$$y'_a = \frac{p_n}{\pi_a} \cdot \frac{(\varepsilon_p - 1)/\varepsilon_p}{(\varepsilon_a - 1)/\varepsilon_a}.$$

Как видно, предельная эффективность фактора a не равна больше отношению цен продукции и фактора, как в условиях равновесия совершенной конкуренции, а равна этому отношению, умноженному на выражение, зависящее от относительных эластичностей фактора a и выпускаемой продукции.

Предельная норма замещения факторов a и b будет равна

$$\frac{y'_a}{y'_b} = \frac{\pi_a}{\pi_b} \cdot \frac{(\varepsilon_a - 1)/\varepsilon_a}{(\varepsilon_b - 1)/\varepsilon_b}.$$

Аналогично, предельная норма замещения в условиях монопольного равновесия равна не только отношению цен факторов, но зависит и от эластичностей факторов.

Рассмотрение других институциональных условий – различных случаев несовершенной конкуренции, двусторонней монополии, дуополии, монополистической конкуренции, олигополии – выходит за пределы нашей книги. Ознакомиться с формализацией равновесия в этих случаях можно в книгах Дж. Хикса, П. Самуэльсона, Э. Маленво.

25.5. Оптимизационная модель предприятия в условиях монополии в Excel

Краткосрочный период

Разработаем модель предприятия, выпускающего один продукт, для короткого периода при фиксированных ценах факторов производства. За функцию выпуска примем степенную производственную функцию:

$$f(x_1, x_2) = a_0 x_1^{a1} x_2^{a2},$$

где x_1 – основные производственные фонды (K), x_2 – трудовые ресурсы (L).

Функцию спроса примем в виде: $D = B\,p^{-g}$, где p – цена продукции.

Функцию затрат от выпуска определяем из условия минимизации затрат (см. главу 21).

Тогда функция затрат для короткого периода из степенной производственной функции будет равна:

$$C_k(f) = p_1 K + \left(\frac{p_2 K^{-a1}}{a_0}\right)^{\frac{1}{a2}} f_k^{\frac{1}{a2}}.$$

где f_k – соответствующая функция выпуска:

$$f_k = a_0 x_1^{a1} x_2^{a2} \quad \text{при } x_1 < K;$$

$$f_k = a_0 K^{a1} x_2^{a2} \quad \text{при } x_1 > K;$$

p_1 и p_2 – цены ресурсов x_1 (основных фондов – K) и, соответственно, x_2 (труд – L).

Предельные затраты для короткого периода будут равны:

$$\frac{dC}{df} = \frac{1}{a_2}\left(\frac{p_2 K^{-a1}}{a_0}\right)^{\frac{1}{a2}} f_k^{\left(\frac{1}{a2}-1\right)}.$$

Функцию цены можно выразить:

$$\pi(f) = b_0\, p^{-1/g}.$$

Прибыль монополиста от продажи выпущенной продукции f составит:

$$PR = \pi_n(f)\, f - C(f).$$

Прибыль достигает максимума, когда

$$\pi_n'(f)\, f + \pi_n(f) - C'(f) = 0,$$

или

$$\pi_n'(f)\, f + \pi_n(f) = C'(f), \text{ или } R'(f) = C'(f).$$

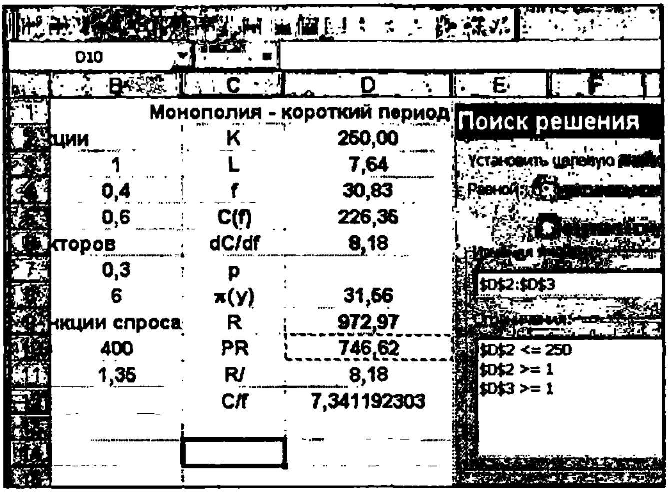
Параметры производственной функции, цены факторов, параметры функции спроса и формализация модели в Excel приведены на рис. 25.16.

B14

A	B	C	D
Параметры			Монополия - короткий период
производ.функции		K	250
a0	1	L	7,63955725193229
a1	0,4	f	=$B$3*D2^$B$4*D3^$B$6
a2	0,6	C(f)	=$B$7*D2+($B$8*D2^(-$B$4)/$B$3)^(1/$B$5)*D4^(1/$B$5)
Цены факторов		dC/df	=1/$B$5*($B$8*D2^(-$B$4)/$B$3)^(1/$B$5)*D4^(1/$B$5-1)
p1	0,3	p	
p2	6	π(y)	=$B$10*D4^(-1/$B$11)
функции спроса		R	=D8*D4
b0	400	PR	=D8*D4-D5
g	1,35	R/	=-$B$10/$B$11*D4^(-1/$B$11-1)*D4+D8
		C/f	=D5/D4

Рис. 25.16

Оптимизация данной модели производится максимизацией функции прибыли при положительности переменных – капитала и труда – и ограничении по капиталу с помощью опции **Поиск решения**. Примем ограничение по капиталу для краткосрочного периода ≤ 250. Диалоговое окно **Поиск решения** и оптимизация модели показаны на рис. 25.17.

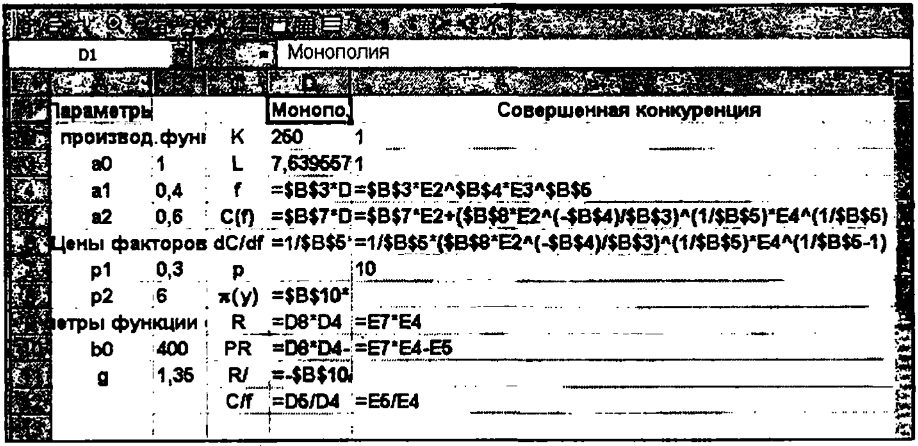


	B	C	D
1	Монополия - короткий период		
2	ции	K	250,00
3	1	L	7,64
4	0,4	f	30,83
5	0,6	C(f)	226,36
6	кторов	dC/df	8,18
7	0,3	p	
8	6	π(y)	31,56
9	нкции спроса	R	972,97
10	400	PR	746,62
11	1,35	R/	8,18
12		C/f	7,341192303

Рис. 25.17

Сравнение монополии с совершенной конкуренцией

Разместим модель фирмы в условиях совершенной конкуренции в столбце E. Она будет иметь вид, показанный на рис. 25.18.



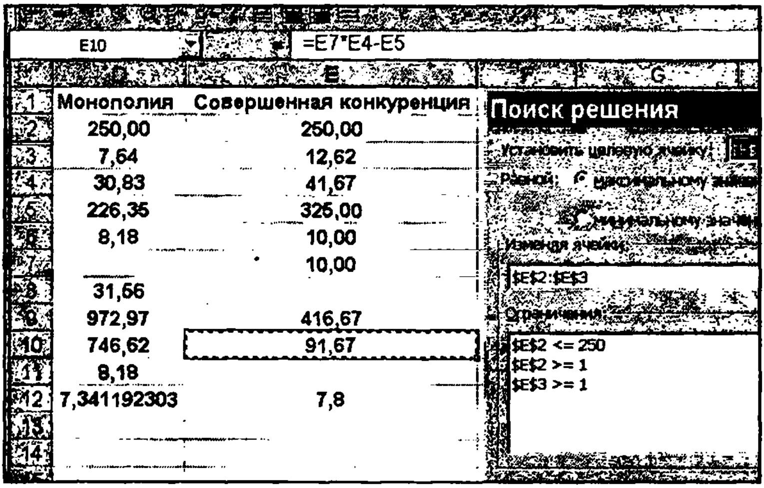
	A	B	C	D	E
1	Параметры			Монопо	Совершенная конкуренция
2	производ.функ		K	250	1
3	a0	1	L	7,639557	1
4	a1	0,4	f	=$B$3*D	=$B$3*E2^$B$4*E3^$B$5
5	a2	0,6	C(f)	=$B$7*D	=$B$7*E2+($B$8*E2^(-$B$4)/$B$3)^(1/$B$5)*E4^(1/$B$5)
6	Цены факторов		dC/df	=1/$B$5	=1/$B$5*($B$8*E2^(-$B$4)/$B$3)^(1/$B$5)*E4^(1/$B$5-1)
7	p1	0,3	p		10
8	p2	6	π(y)	=$B$10*	
9	етры функции		R	=D8*D4	=E7*E4
10	b0	400	PR	=D8*D4-	=E7*E4-E5
11	g	1,35	R/	=-$B$10	
12			C/f	=D5/D4	=E5/E4

Рис. 25.18

Примем цену равной $p = 10$ и оптимизируем модель (рис. 25.19).

Теперь проведем классическое сравнение фирмы в условиях монопольного равновесия и совершенной конкуренции. Для этого надо найти цену для совершенной конкуренции, оптимальный объем производства и рациональное распределение факторов производства, а затем сравнить. Цена для совершенной конкуренции находится из равенства предельных затрат функции спроса

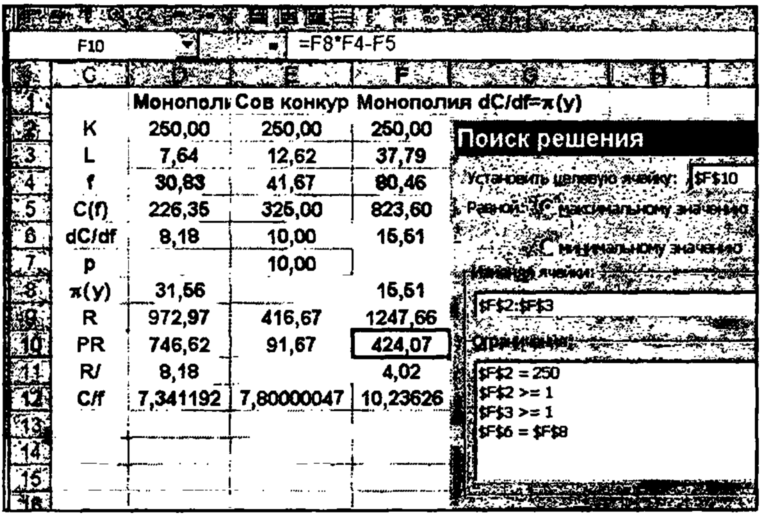
$C'(f) = \pi(f)$. Разместим в столбце D еще одну модель фирмы в условиях монополии и оптимизируем ее при данном ограничении (рис. 25.20). Кроме того, в данном случае существует еще одно ограничение. Поскольку данная область целевой функции находится в пределах уже действующего ограничения по капиталу, то необходимо принять постоянное значение капитала $K = 250$. Решение без данного ограничения будет лежать не на границе эффективного множества, что можно будет увидеть, если построить график.



E10 =E7*E4-E5

	D	E
1	Монополия	Совершенная конкуренция
2	250,00	250,00
3	7,64	12,62
4	30,83	41,67
5	226,35	325,00
6	8,18	10,00
7		10,00
8	31,56	
9	972,97	416,67
10	746,62	91,67
11	8,18	
12	7,341192303	7,8
13		
14		

Рис. 25.19



F10 =F8*F4-F5

	C	D	E	F	G
1		Монопол	Сов конкур	Монополия	dC/df=π(y)
2	K	250,00	250,00	250,00	
3	L	7,64	12,62	37,79	
4	f	30,83	41,67	80,46	
5	C(f)	226,35	325,00	823,60	
6	dC/df	8,18	10,00	15,61	
7	p		10,00		
8	π(y)	31,56		15,61	
9	R	972,97	416,67	1247,66	
10	PR	746,62	91,67	424,07	
11	R/	8,18		4,02	
12	C/f	7,341192	7,80000047	10,23626	
13					
14					
15					
16					

Рис. 25.20

Полученную цену, равную 15,51, поставим теперь в модель фирмы в условиях совершенной конкуренции и оптимизируем (рис. 25.21).

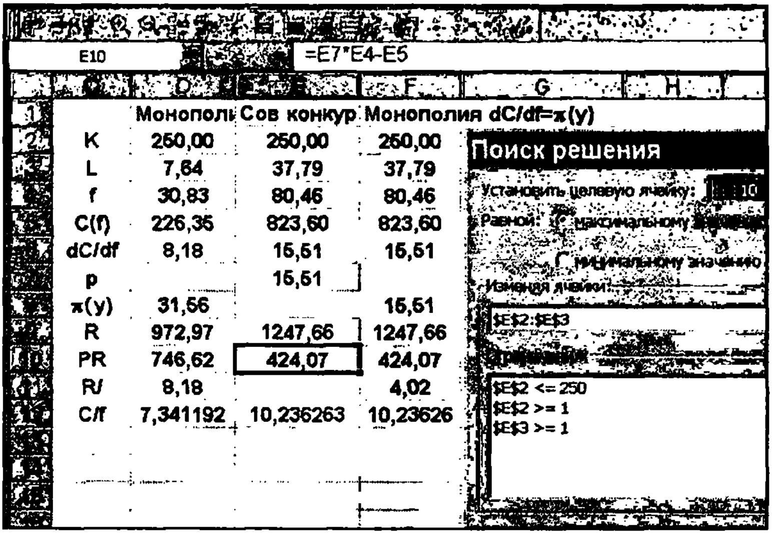


Рис. 25.21

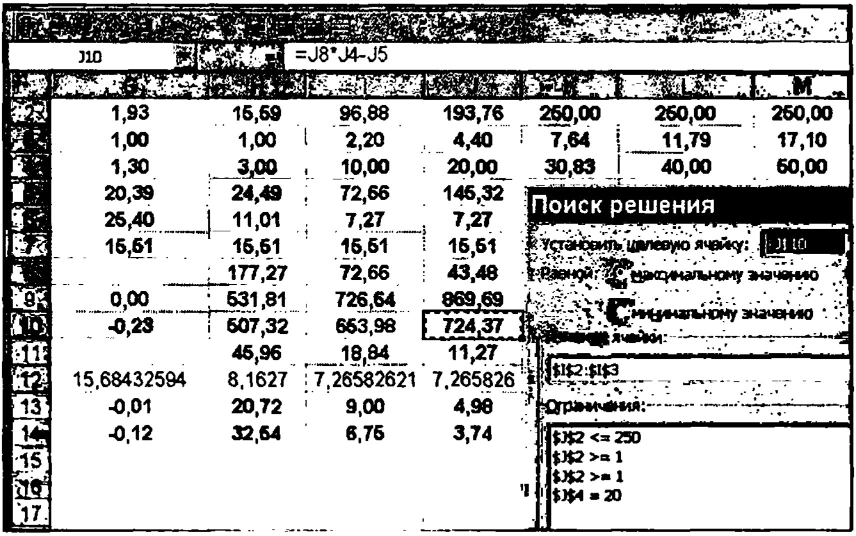


Рис. 25.22

Получили тот же результат. Сравнение показывает, что монополист использует меньше факторов производства, производит меньшее количество продукции и имеет бóльшую прибыль. Представим этот результат графически. Столбец F2:F12 скопируем в ряд других столбцов, например G:P. Затем необходимо установить значения минимума функции затрат для соответствующих значений выпуска. Для решения этой задачи в каждом столбце необходимо произвести оптимизацию при ограничении по значению выпуска, планируемого по оси абсцисс, и ограничении положительности переменных. Оптимизацию можно выполнять как максимизацию прибыли и как минимизацию затрат (рис. 25.22).

По рассчитанным данным уже нетрудно построить график – рис. 25.23, на котором точка пересечения предельных затрат и предельного дохода есть точка монопольного равновесия, а точка пересечения предельных затрат и цены 15,51 (или кривой спроса) – это точка равновесия в условиях совершенной конкуренции.

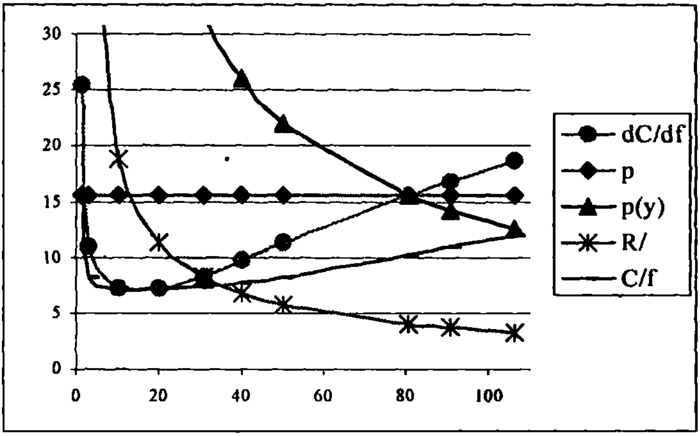


Рис. 25.23

Влияние налогов

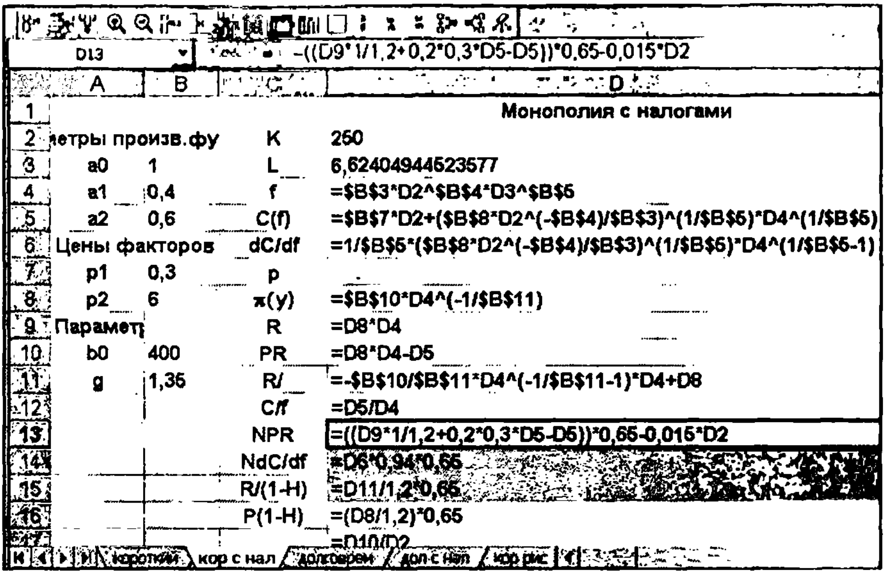
Налоги по-разному влияют на оптимум фирмы. Так, налог на прибыль не влияет на размер фирмы, в то время как налоги на выручку (акцизы, НДС) действуют угнетающе. Рассмотрим три главных налога: налог на добавленную стоимость, налог на прибыль и налог на имущество. НДС примем в размере 20% от величины выручки, а также примем, что доля материалов в затратах составляет 30%. Налог на прибыль составляет 35%. База налогообложения – валовая прибыль после уплаты НДС. Налог на имущество – 1,5% от основных фондов из прибыли. Таким образом, чистая прибыль (NPR) будет равна:

$$NPR = (R/1{,}2 + 0{,}2 \cdot 0{,}3 \cdot C) \cdot 0{,}65 - 0{,}015 \cdot K.$$

Модель монополии с учетом указанных трех налогов приведена на рис. 25.24.

Сравнение с равновесием в условиях совершенной конкуренции достигается следующим образом. Сначала, максимизируя чистую прибыль, устанавливается цена, при которой она, за минусом налогов, равна предельным затратам, за минусом налогов при ограничении – равенстве по капиталу:

$$NPR \to \max \quad NdC/df = \pi(y)(1 - H), \quad K = 250.$$



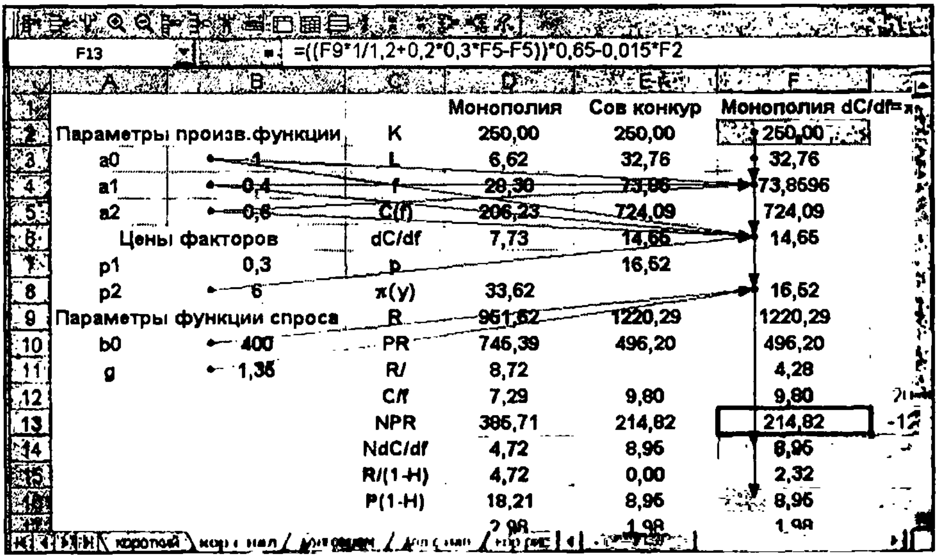
D13 =((D9*1/1,2+0,2*0,3*D5-D5))*0,65-0,015*D2

	A	B	C	D
1				Монополия с налогами
2	іетры произв.фу		K	250
3	a0	1	L	6,62404944523577
4	a1	0,4	f	=$B$3*D2^$B$4*D3^$B$5
5	a2	0,6	C(f)	=$B$7*D2+($B$8*D2^(-$B$4)/$B$3)^(1/$B$5)*D4^(1/$B$5)
6	Цены факторов		dC/df	=1/$B$5*($B$8*D2^(-$B$4)/$B$3)^(1/$B$5)*D4^(1/$B$5-1)
7	p1	0,3	p	
8	p2	6	π(y)	=$B$10*D4^(-1/$B$11)
9	Парамет		R	=D8*D4
10	b0	400	PR	=D8*D4-D5
11	g	1,35	R/	=-$B$10/$B$11*D4^(-1/$B$11-1)*D4+D8
12			C/f	=D5/D4
13			NPR	=((D9*1/1,2+0,2*0,3*D5-D5))*0,65-0,015*D2
14			NdC/df	=D6*0,94*0,65
15			R/(1-H)	=D11/1,2*0,65
16			P(1-H)	=(D8/1,2)*0,65
17				=D10/D2

короткий / кор с нал / долговрем / дол с нал / кор рис

Рис. 25.24

Подставив полученную цену в столбце F = 16,52 в модель совершенной конкуренции в столбец E (p = 16,52) и оптимизируя чистую прибыль при ограничении по капиталу, получим равновесие при совершенной конкуренции (рис. 25.25).



F13 =((F9*1/1,2+0,2*0,3*F5-F5))*0,65-0,015*F2

	A	B	C	D	E	F
1				Монополия	Сов конкур	Монополия dC/df=π
2	Параметры произв.функции		K	250,00	250,00	250,00
3	a0	1	L	6,62	32,76	32,76
4	a1	0,4	f	28,30	73,86	73,8696
5	a2	0,6	C(f)	206,23	724,09	724,09
6	Цены факторов		dC/df	7,73	14,66	14,65
7	p1	0,3	p		16,52	
8	p2	6	π(y)	33,62		16,52
9	Параметры функции спроса		R	951,62	1220,29	1220,29
10	b0	400	PR	745,39	496,20	496,20
11	g	1,35	R/	8,72		4,28
12			C/f	7,29	9,80	9,80
13			NPR	385,71	214,82	214,82
14			NdC/df	4,72	8,96	8,96
15			R/(1-H)	4,72	0,00	2,32
16			P(1-H)	18,21	8,96	8,96
17				2,98	1,98	1,98

короткий / кор с нал / долговрем / дол с нал / кор рис

Рис. 25.25

Этот результат можно представить графически (рис. 25.26).

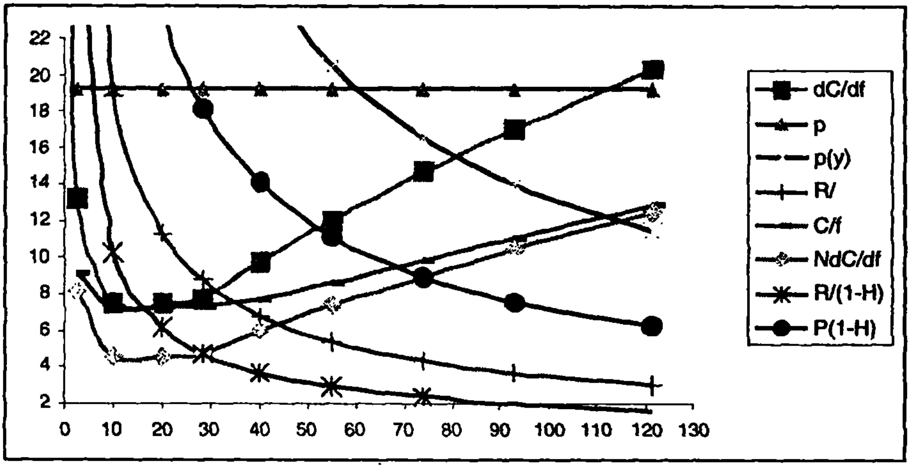


Рис. 25.26

Размеры фирмы в обоих случаях уменьшаются, что является следствием ияния в основном НДС. Монопольное равновесие можно интерпретировать к равенство предельных затрат и предельного дохода за вычетом налогов. вновесие в условиях совершенной конкуренции достигается при равенстве едельных затрат и цены за вычетом налогов.

Долгосрочное равновесие

Для долгосрочного периода функция затрат от выпуска будет равна:

$$C(f)=p_2\left(\frac{a_1+a_2}{a_2}\right)\left(\frac{a_2p_1}{a_1p_2}\right)^{\frac{a_1}{a_1+a_2}}\left(\frac{f}{a_0}\right)^{\frac{1}{a_1+a_2}}.$$

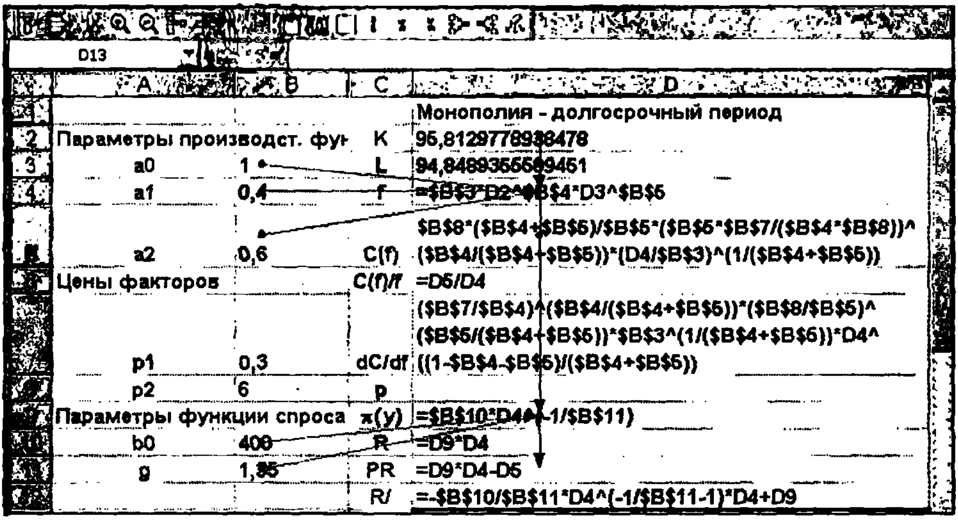
Предельные затраты будут равны:

$$\frac{dC}{df}=\frac{p_2}{a_2}\frac{\left[\left(\frac{a_2p_1}{a_1p_2}\right)^{a_1}\frac{f}{a_0}\right]^{\frac{1}{a_1+a_2}}}{f},$$

і после преобразований:

$$\frac{dC}{df}=\left(\frac{p_1}{a_1}\right)^{\frac{a_1}{a_1+a_2}}\left(\frac{p_2}{a_2}\right)^{\frac{a_2}{a_1+a_2}}a_0^{\left(-\frac{1}{a1+a2}\right)}f^{\frac{1-a_1-a_2}{a_1+a_2}}.$$

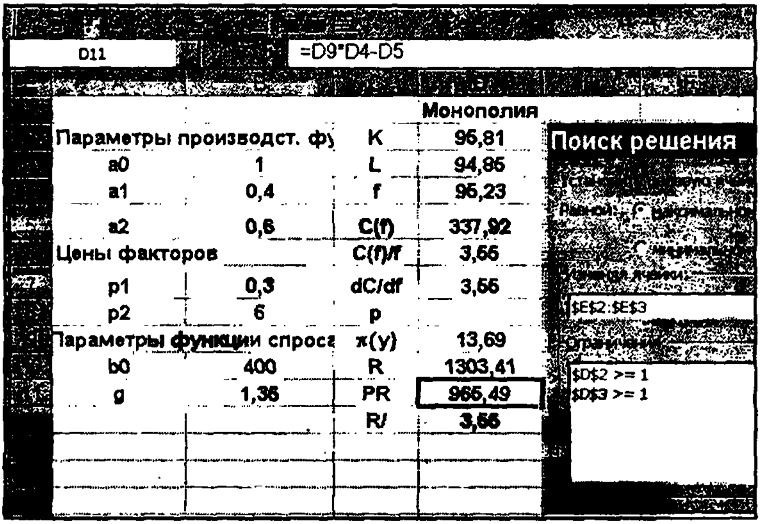
Модель фирмы при монопольном равновесии в долгосрочном периоде при- ена на рис. 25.27.



	A	B	C	D
1				Монополия - долгосрочный период
2	Параметры производст. фун		K	95,8129778938478
3	a0	1	L	94,8489355509451
4	a1	0,4	f	=$B$3*D2^$B$4*D3^$B$5
5	a2	0,6	C(f)	$B$8*($B$4+$B$5)/$B$5*($B$5*$B$7/($B$4*$B$8))^($B$4/($B$4+$B$5))*(D4/$B$3)^(1/($B$4+$B$5))
6	Цены факторов		C(f)/f	=D5/D4
7	p1	0,3	dC/df	($B$7/$B$4)^($B$4/($B$4+$B$5))*($B$8/$B$5)^($B$5/($B$4+$B$5))*$B$3^(1/($B$4+$B$5))*D4^((1-$B$4-$B$5)/($B$4+$B$5))
8	p2	6	p	
9	Параметры функции спроса		π(y)	=$B$10*D4^(-1/$B$11)
10	b0	400	R	=D9*D4
11	g	1,35	PR	=D9*D4-D5
12			R/	=-$B$10/$B$11*D4^(-1/$B$11-1)*D4+D9

Рис. 25.27

Модель оптимизации и ее решение показаны на рис. 25.28.



			Монополия
Параметры производст. фу		K	95,81
a0	1	L	94,85
a1	0,4	f	95,23
a2	0,6	C(f)	337,92
Цены факторов		C(f)/f	3,55
p1	0,3	dC/df	3,55
p2	6	p	
Параметры функции спроса		π(y)	13,69
b0	400	R	1303,41
g	1,35	PR	965,49
		R/	3,55

Рис. 25.28

При эластичности производства $a_1 + a_2 = 1$ предельные и средние затраты будут совпадать и представлять собой прямую линию – рис. 25.29.

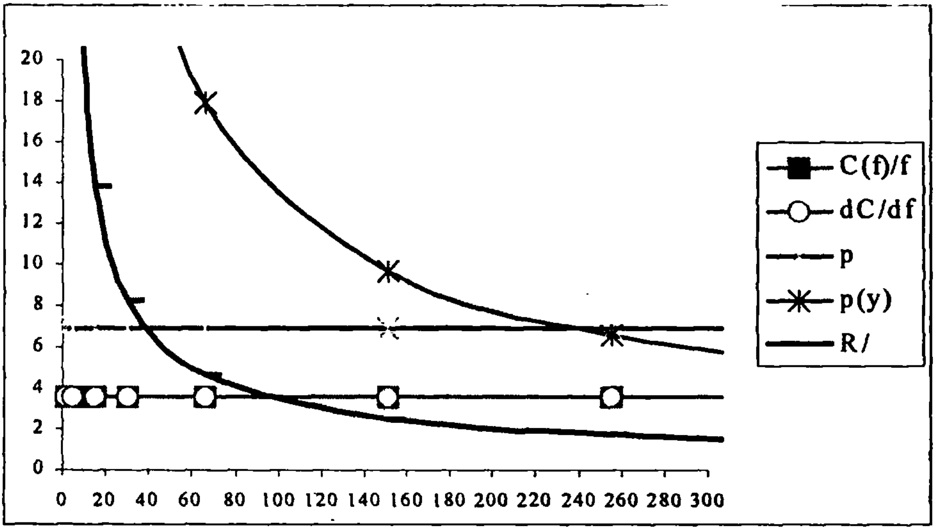


Рис. 25.29

При убывающем масштабе производства $a_1 + a_2 < 1$ графики предельных и средних затрат будут вогнуты вверх. Примем значение $a_1 = 0{,}25$. Тогда график примет вид, показанный на рис. 25.30.

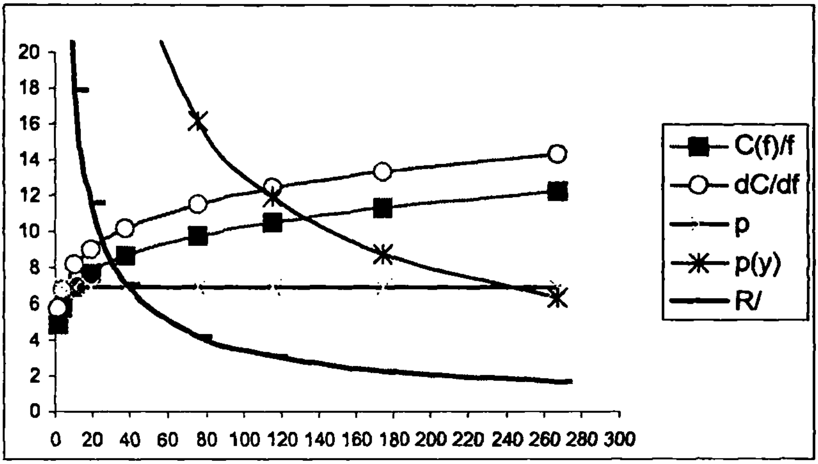


Рис. 25.30

При возрастающем эффекте от масштаба производства кривые затрат будут вогнуты вниз. Примем $a_1 = 0{,}3$, $a_2 = 0{,}8$, тогда график будет иметь вид, показанный на (рис. 25.31).

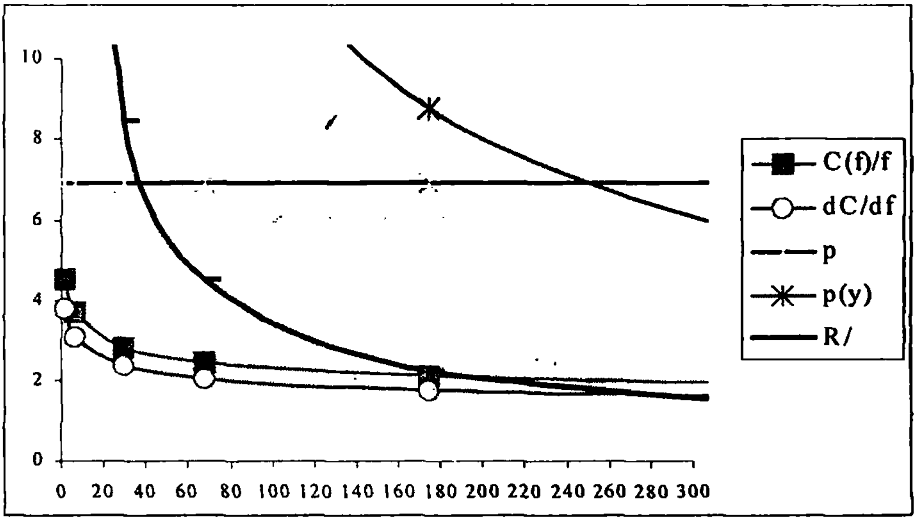


Рис. 25.31

Для убывающего эффекта от масштаба производства сравнение монопольного равновесия с равновесием в условиях совершенной конкуренции приведено на рис. 25.32.

	A	B	C	D	E	F
1				Монополия	Сов конкур	Монополия - dc/df
2	Параметры производст. функции		K	71,84	126,14	179,41
3	a0	1	L	36,68	337,19	291,16
4	a1	0,25	f	25,28	110,14	110,14
5	a2	0,6	C(f)	203,65	1150,46	1150,46
6	Цены факторов		C(f)/f	8,06	10,45	10,45
7	p1	0,3	dC/df	9,48	12,29	12,29
8	p2	6	p		12,29	12,29
9	Параметры функции спроса		π(y)	36,56		12,29
10	b0	400	R	924,14	1353,48	1353,48
11	g	1,35	PR	720,49	203,02	203,02
12			R/	9,48		3,19

Рис. 25.32

Графически этот результат представлен на рис. 25.33.

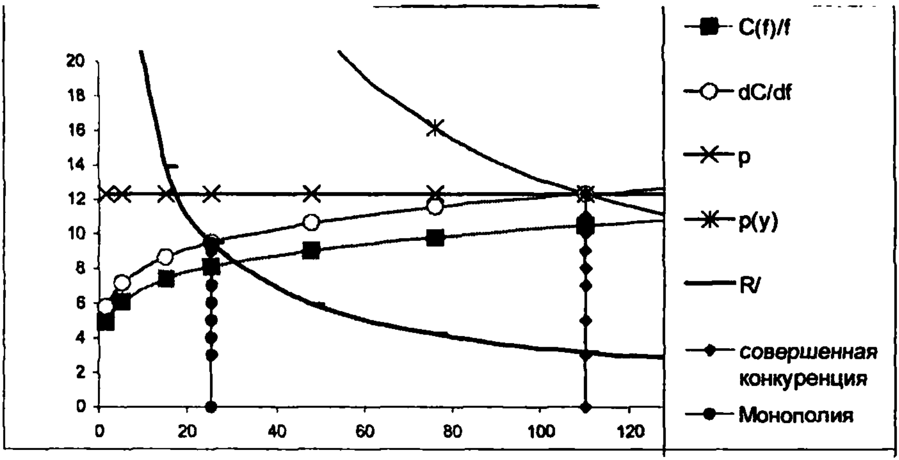


Рис. 25.33

25.6. Оптимизационная модель предприятия в Mathcad

Рассмотрим модель предприятия в условиях монополии в долгосрочный период с учетом налогов в Mathcad. Учтем два налога – налог на добавленную стоимость и налог на прибыль. Оптимизацию будем осуществлять методом нелинейного программирования, используя функцию **Maximaze**. Введем значения параметров функций, формализуем модель с помощью функций выпуска, затрат, спроса и чистой прибыли. Введем инициализацию переменных – капитала и труда. В решающем блоке введем слово Given, строгую положительность переменных и применим функцию **Maximaze** (рис. 25.34).

Для построения графика выразим зависимости через объем производства (рис. 25.34, окончание).

Можно примерно установить объем производства и цену при совершенной конкуренции, открыв контекстно-зависимое меню **Trace** и щелкнув мышью на пересечении кривых цены без налогов и предельных затрат без налогов – рис. 25.35.

Решение данной задачи сравнения с совершенной конкуренцией приведено на рис. 25.36. Найднм объем производства, соответствующий равенству цены за минусом налогов и предельных затрат за минусом налогов (рис. 25.36).

Mathcad позволяет представить сравнение монополии и совершенной конкуренции в прямом виде – в виде графиков прибыли от двух переменных – капитала и труда. Для этого надо рассчитать две соответствующие матрицы и построить графики поверхности на одном графике. Пример приведен на рис. 25.37 и 25.38.

И в заключение рассмотрим возможности Mathcad в представлении состояния развития предприятия на одном графике – рис. 25.39–25.41.

Данные графики можно представить в виде, более соответствующем тому положению, что они существуют в разное время (рис. 25.41).

$a0 := 1 \qquad b0 := 900 \qquad p1 := 0.3$

$a1 := 0.25 \qquad g := 1.35 \qquad p2 := 6$

$a2 := 0.6$

$L := 1 \qquad K := 1$

$$f(K,L) := a0\ K^{a1}\ L^{a2}$$

$$C(K,L) := p2\left[\frac{(a1+a2)}{a2}\right]\left(a2\ \frac{p1}{a1\ p2}\right)^{\left[\frac{a1}{(a1+a2)}\right]}\left(\frac{f(K,L)}{a0}\right)^{\left[\frac{1}{(a1+a2)}\right]}$$

$$D(K,L) := b0\ f(K,L)^{\left(\frac{-1}{g}\right)}$$

$$RET(K,L) := D(K,L)\ f(K,L)$$

$$NPR(K,L) := \left[\frac{RET(K,L)}{1.2} - 0.94\,C(K,L)\right]0.65$$

Функция чистой прибыли

Given

$K \ge 1 \qquad L \ge 1$

Оптимизация

$$\begin{pmatrix} K \\ L \end{pmatrix} := \mathrm{Maximize}(NPR, K, L)$$

$$\begin{pmatrix} K \\ L \end{pmatrix} = \begin{pmatrix} 68.838 \\ 130.938 \end{pmatrix}$$

$f(K,L) = 53.666 \qquad D(K,L) = 47.096$

$C(K,L) = 493.772 \qquad RET(K,L) = 2527.448$

$NPR(K,L) = 1067.34$

Рис. 25.34. Модель предприятия в условиях монополии в долгосрочный период с учетом налогов в Mathcad (начало)

$$x := f(K, L)$$

$$C(x) := p2\left[\frac{(a1 + a2)}{a2}\right]\left(a2\ \frac{p1}{a1\ p2}\right)^{\left[\frac{a1}{(a1+a2)}\right]}\left(\frac{x}{a0}\right)^{\left[\frac{1}{(a1+a2)}\right]}$$

$$C(x) = 493.772$$

$$dC(x) := \frac{d}{dx}C(x) \rightarrow 5.36007608841760888412\ x^{.17647058823529411776}$$

$$dC(x) = 10.825 \qquad NMC(x) := 0.94\ 0.65\ dC(x)\ *$$

$$D(x) := b0\ x^{\left(\frac{-1}{g}\right)} \qquad AC(x) := \frac{C(x)}{x}$$

$$NMC(x) = 6.614$$

$$RET(x) := D(x)\ x \qquad NAR(x) := D(x)\ \frac{0.65}{1.2}\ *$$

$$dRET(x) := \left(\frac{d}{dx}RET(x)\right)$$

$$NMR(x) := dRET(x)\ \frac{0.65}{1.2}$$

$$dRET(x) = 12.21$$

$$NMR(x) = 6.614$$

$$x := 1..250$$

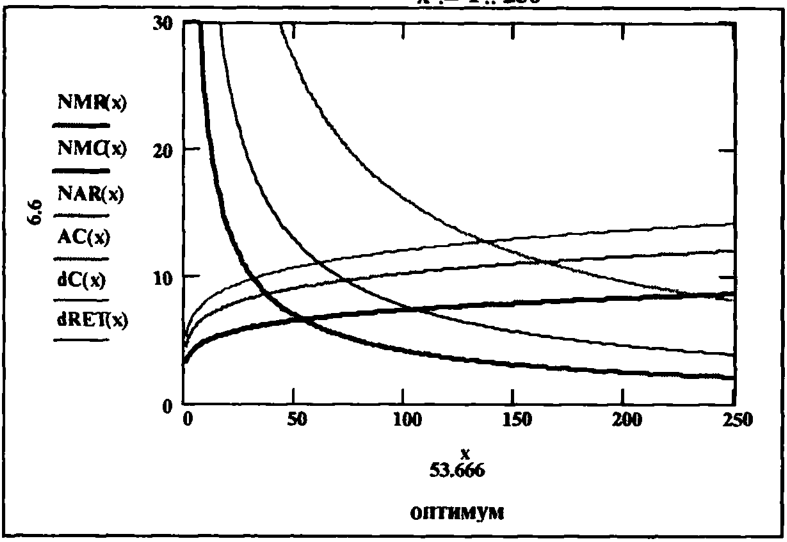


Рис. 25.34. Модель предприятия в условиях монополии в долгосрочный период с учетом налогов в Mathcad (окончание)

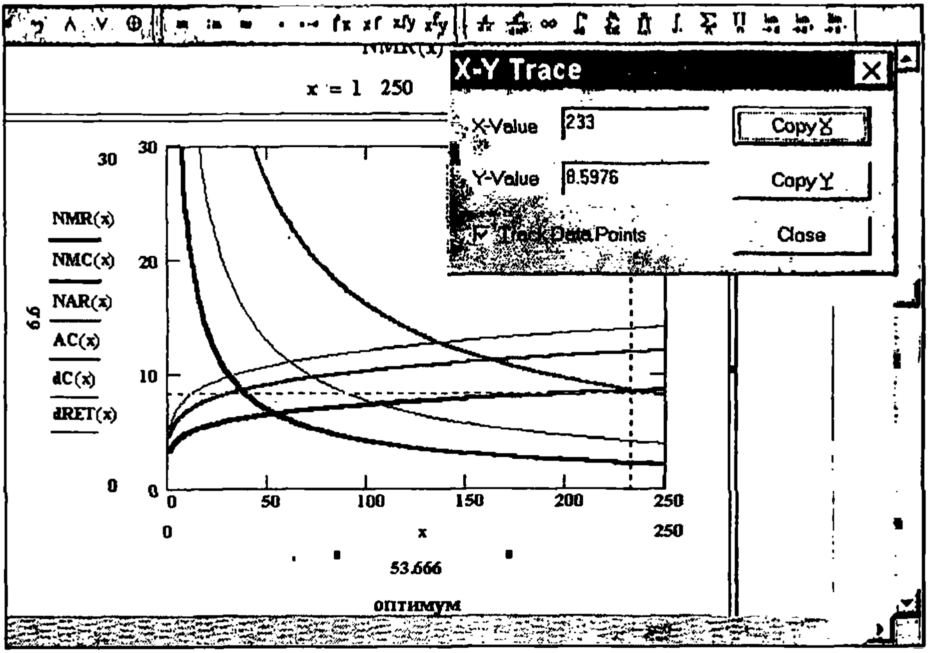


Рис. 25.35

$$x := 1$$

$$NPR(x) := \left[\left[\left(\frac{RET(x)}{1.2}\right) - 0.94\ C(x)\right] 0.65\right] *$$

Given

$$K \geq 1 \qquad\qquad L \geq 1$$

$$D(x)\ \frac{0.65}{1.2} = 0.94\ 0.65\ dC(x)$$

$$x := Maximize(NPR, x)$$

$$x = 233.827$$

Рис. 25.36. Установление оптимального объема производства в Mathcad (начало)

Введем ограничение – равенство по установленному объему производства, и установим цену в условиях совершенной конкуренции и объема капитала и труда.

$$f(K,L) := a0\ K^{a1}\ L^{a2} \qquad K := 1 \qquad L := 1$$

$$C(K,L) := p2\left[\frac{(a1+a2)}{a2}\right]\left(a2\ \frac{p1}{a1\ p2}\right)^{\left[\frac{a1}{(a1+a2)}\right]}\left(\frac{f(K,L)}{a0}\right)^{\left[\frac{1}{(a1+a2)}\right]}$$

$$D(K,L) := b0\ f(K,L)^{\left(\frac{-1}{g}\right)} \qquad dC(x) := \frac{d}{dx}C(x)$$

$$RET(K,L) := D(K,L)\ f(K,L)$$

$$NPR(K,L) := \left[\left[\left(\frac{RET(K,L)}{1.2}\right) - 0.94\ C(K,L)\right]0.65\right] *$$

Optimized Result

NPR(K,L) := .54166666666666666666 · RET(K,L) – .61100000000000000000 · C(K,L)

Given

$$K \ge 1 \qquad L \ge 1$$

$$f(K,L) = x$$

$$\begin{pmatrix} K \\ L \end{pmatrix} := \mathrm{Maximize}(NPR, K, L)$$

$$\begin{pmatrix} K \\ L \end{pmatrix} = \begin{pmatrix} 527.571 \\ 651.442 \end{pmatrix}$$

$f(K,L) = 233.828$ $\qquad D(K,L) = 15.831$ **Цена**

$C(K,L) = 2789.498$ $\qquad NPrise := D(K,L)\ \frac{0.65}{1.2}$

$RET(K,L) = 3701.695$

$NPR(K,L) = 300.702$ $\qquad NPrise = 8.575$ **Цена за минусом налогов**

$NPise(x) := NPrise$

$x := 1..250$

Рис. 25.36. Установление оптимального объема производства в Mathcad (продолжение)

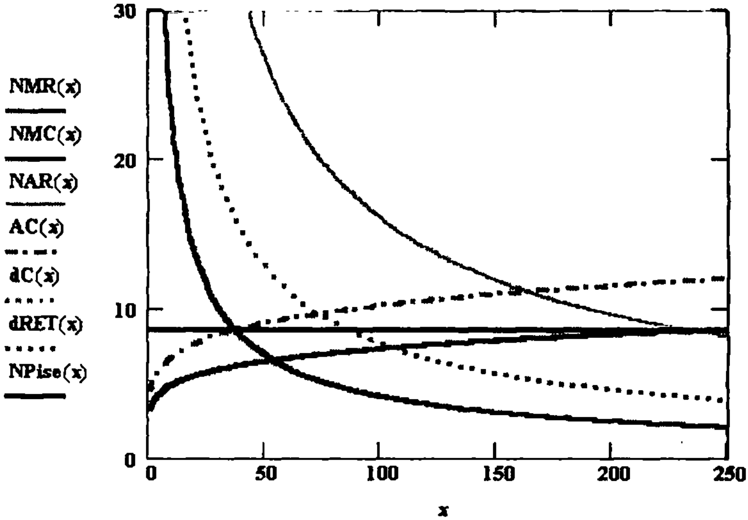


Рис. 25.36. Установление оптимального объема производства в Mathcad (окончание)

$$RET(K,L) := D(K,L)\ f(K,L)$$

$$NPR(K,L) := \left[\left(\frac{RET(K,L)}{1.2} - 0.94C(K,L)\right) 0.65\right]$$

$$i := 0..80 \qquad j := 0..80$$

$$K_i := 1 + \frac{i}{0.1} \qquad L_j := 1 + \frac{j}{0.1}$$

$$M_{i,j} := NPR(K_i, L_j)$$

$$RET2(K,L) := 15.831\ f(K,L)$$

$$NPR2(K,L) := \left[\left(\frac{RET2(K,L)}{1.2} - 0.94C(K,L)\right) 0.65\right]$$

$$t := 0..120 \qquad h := 0..120$$

$$K_t := 1 + \frac{t}{0.1} \qquad L_h := 1 + \frac{h}{0.1}$$

$$P_{t,h} := NPR2(K_t, L_h)$$

Рис. 25.37. Расчет матриц

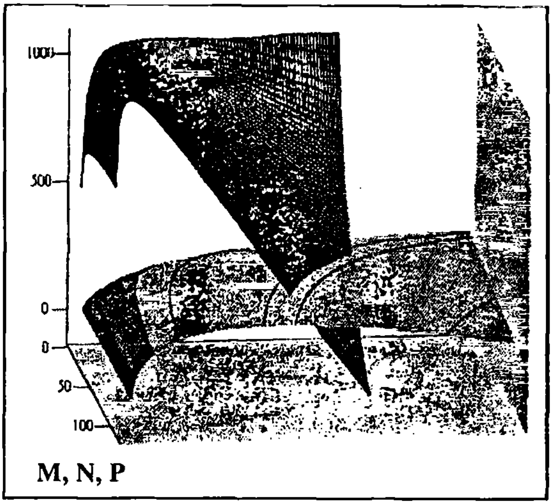


Рис. 25.38. Графики прибыли в условиях монополии и совершенной конкуренции

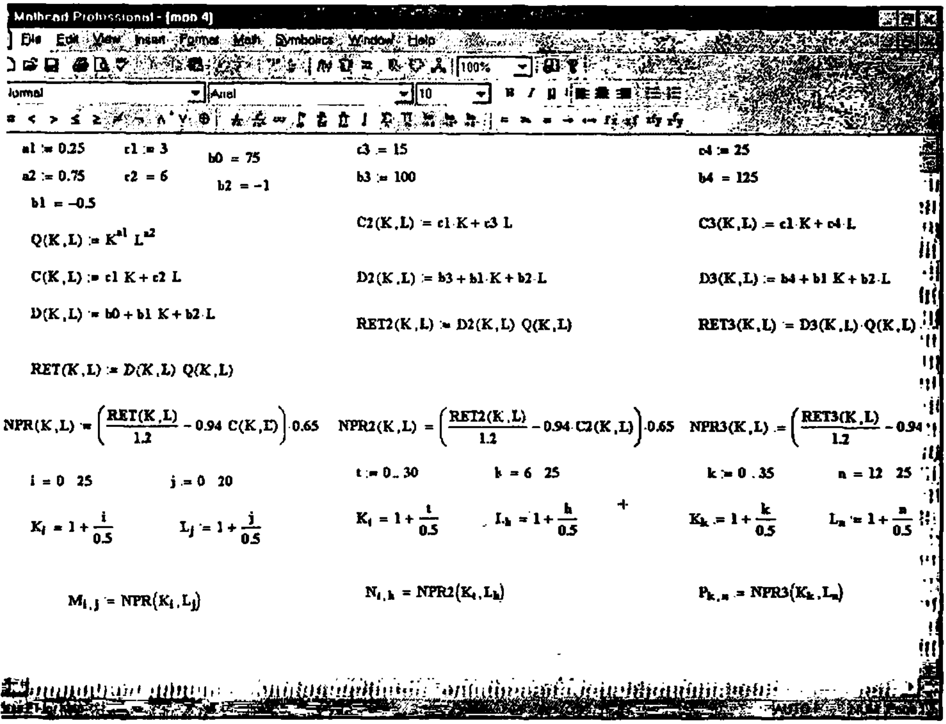


Рис. 25.39. Расчет трех состояний предприятий при увеличении спроса и затрат на один из ресурсов и соответствующих им матриц

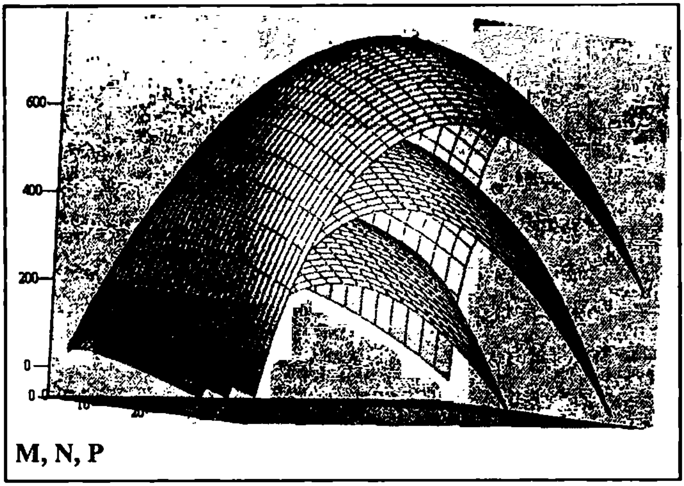


Рис. 25.40. Графики прибыли для трех состояний предприятия

$t := 0..30$ $h := 6..25$ $k := 0..35$ $n := 12..25$

$K_t := 1 + \frac{t}{0.5}$ $L_h := 1 + \frac{h}{0.5}$ $K_k := 1 + \frac{k}{0.5}$ $L_n := 1 + \frac{n}{0.5}$

$N_{t,h} := NPR2(K_t, L_h)$ $P_{k,n} := NPR3(K_k, L_n)$

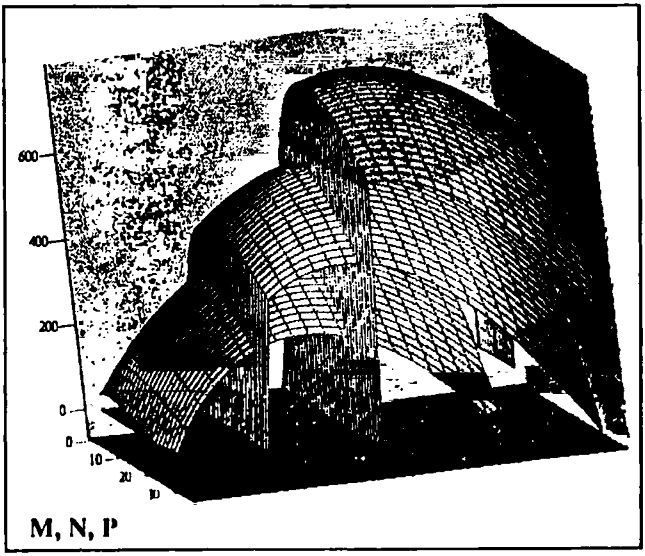


Рис. 25.41. Графики прибыли предприятия, соответствующие его развитию при росте спроса

Литература

1. Алчиан А. Затраты и выпуски. /Теория фирмы. СПб.: Экономическая школа, 1995.
2. Базара М., Шетти К. Нелинейное программирование. Теория и алгоритмы. М.: Мир, 1985.
3. Васильев Ф. П. Численные методы решения экстремальных задач. М.: Наука, 1988.
4. Вентцель Е. С. Исследование операций. Задачи, принципы, методология. М.: Наука, 1980.
5. Высшая математика для экономистов: Учебник для вузов/Н. Ш. Кремер, Б. А. Путко, И. М. Тришин, М. Н. Фридман/Под ред. проф. Н. Ш. Кремера. – М.: Банки и биржи. ЮНИТИ, 1998.
6. Галеев Э. М., Тихомиров В. М. Краткий курс теории экстремальных задач. – М.: Издательство Московского университета, 1989.
7. Демиденко Е. З. Оптимизация и регрессия. М.: Наука, 1989.
8. Джонсон Дж. Эконометрические методы. М.: Статистика, 1980.
9. Джошуа Носитер. EXCEL 7.0 без проблем для Windows 95. Руководство пользователя. М.: BINOM, 1996.
10. Додж М., Кината К., Стинсон К. The Cobb Group. Эффективная работа с EXCEL 7.0 для Windows 95. М.: Питер, 1997.
11. Доугерти К. Введение в эконометрику: Пер. с англ. М.: ИНФРА-М, 1999.
12. Дрейпер Н., Смит Г. Прикладной регрессионный анализ. В 2-х книгах, Кн. 1. М.: Финансы и статистика, 1986. Кн. 2. М.: Финансы и статистика, 1987.
13. Дьяконов В. П. Справочник по Mathcad 7 Pro. М.: СК Пресс, 1998.
14. Замков О. О., Толстопятенко А. В., Черемных Ю. Н. Математические методы в экономике. Учебник. М.: МГУ, Издательство «ДИС», 1997.
15. Исследование операций в экономике. Учебное пособие для вузов/Н. Ш. Кремер, Б. А. Путко, И. М. Тришин. М.: Банки и биржи, ЮНИТИ, 1997.
16. Карлберг К. Бизнес-анализ с помощью Excel: Пер. с англ. К.: Диалектика, 1997.
17. Клейнер Г. Б. Производственные функции: Теория, методы, применение. М.: Финансы и статистика, 1986.
18. Корн Г., Корн Т. Справочник по математике. Для научных работников и инженеров. М.: Наука, 1978.
19. Курицкий Б. Я. Поиск оптимальных решений средствами Excel 7.0. СПб.: BHV, 1997.
20. Леонтьев В. В. Межотраслевая экономика: Пер. с англ. М.: ОАО «Издательство Экономика», 1997.

21. Лотов А. В. Введение в экономико-математическое моделирование. М.: Наука, 1984.
22. Магнус Я. Р., Катышев П. К., Пересецкий А. А. Эконометрика. Начальный курс. М.: Дело, 1997.
23. Маленво Э. Лекции по микроэкономическому анализу: Пер. с фран. Под ред. К. А. Багриновского. М.: Наука. 1985.
24. Математическая экономика на персональном компьютере: Пер. с яп.: М. Кубонива, М. Табата, Ю. Хасбэ. М.: Финансы и статистика, 1991.
25. MathCAD 6.0 PLUS. Руководство пользователя: Пер. с англ. М.: Филинъ, 1996.
26. Мэйндоналд Дж. Вычислительные алгоритмы в прикладной статистике. М.: Финансы и статистика, 1988.
27. Marquardt D. W. An algoritm for least-squarres estimation of nonlinear parameters. – Journ. Soc. Indust. Appl. Math., 1963, 11, No 2.
28. Николь Н., Альбрехт Р. Электронные таблицы Excel для квалифицорованных пользователей: Практ. пособ.: Пер. с нем. М.: ЭКОМ, 1996.
29. Очков В. Ф. Mathcad PLUS 6.0 для студентов и инженеров. М.: ТОО фирма «КомпьютерПресс», 1996.
30. Поллард Дж. Справочник по вычислительным методам статистики. М.: Финансы и статистика, 1982.
31. Резниченко С. С., Подольский М. П., Ашихмин А. А. Экономико-математические методы и моделирование в планировании и управлении горным производством. Учеб. для вузов. М.: Недра, 1991.
32. Справочник по математике для экономистов/Под ред. В. И. Ермакова. М.: Высшая школа, 1987.
33. Сухарев А. Г., Тимохов А. В., Федоров В. В. Курс методов оптимизации. М.: Наука, 1986.
34. Уолтерс А. А. Производственные функции и функции затрат: эконометрический обзор / Теория фирмы. СПб.: Экономическая школа, 1995.
35. Уотшем Т. Дж., Паррамоу К. Количественные методы в финансах: Пер. с англ. М.: Финансы, ЮНИТИ, 1999.
36. Шелобаев С. И. Математические методы и модели в экономике, финансах. бизнесе: Учеб. пособоие для вузов. М.: ЮНИТИ-ДАНА, 2000.
37. Хикс Дж. Р. Стоимость и капитал. М.: Прогресс, 1988.