# Sub-interval analysis and possibilities of its applications

Harin, Alexander

Modern Humanitarian Academy

30 December 2012

# Суб-интервальный анализ и возможности его применения

А.А. Харин
Современная гуманитарная академия

Статья представляет собой краткий обзор и развитие пленарного доклада в МФТИ. Рассматриваются три существующих инструмента суб-интервального анализа и представлены элементы двух его новых инструментов. Суб-интервальный анализ может использоваться, например, в микро- и макроэкономике, бухучете, эконометрике, теории полезности, интернете, военном деле статистике, Фурье-анализе.

Содержание

## Введение

Эта статья представляет собой краткий обзор и развитие пленарного доклада [1] «Суб-интервальный анализ и возможности его применения», представленного 23 ноября 2012 г. в Московском физико-техническом институте на факультете инноваций и высоких технологий.

Суб-интервальный анализ был основан в 2011 в [2] как новое направление интервального анализа (см., например, [3]-[6]). В настоящее время суб-интервальному анализу посвящено около 20 докладов и статей и книга [7].

## 1. Что это такое?

### 1.1. Что такое суб-интервал?

Упрощенно, суб-интервал – это часть интервала.

Более точно: Интервал представляет собой замкнутый отрезок. Суб-интервал – это интервал, который является частью другого интервала. Пример интервала $\boldsymbol{X}$ и его суб-интервала $\boldsymbol{X_1}$ приведен на рис. 1.

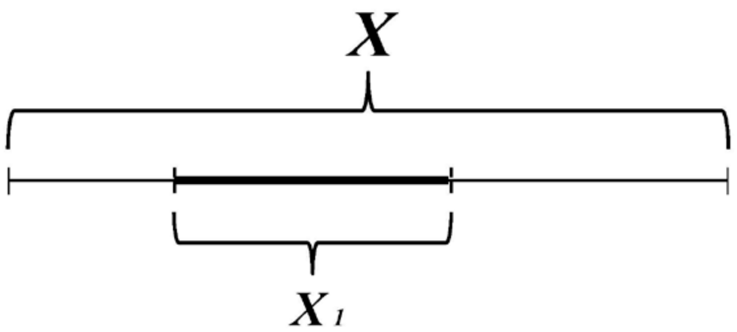


Рисунок 1. Пример интервала $\boldsymbol{X}$
и его суб-интервала $\boldsymbol{X_1}$

## 1.2. Что такое суб-интервальный анализ?

Суб-интервальный анализ – это средство:

1) **для расчета характеристик всего интервала и**
2) **для представления всего интервала**

**с помощью его суб-интервалов.**

Суб-интервальный анализ имеет достаточно большое количество возможных областей использования. Но он не является универсальным средством. Он может эффективно использоваться только тогда, когда имеет место хотя бы одна из двух главных предпосылок его применения.

## 1.3. В чем основные предпосылки применения суб-интервального анализа?

1) Первая главная предпосылка:

Суб-интервал определяют
всего **3 (три) значения** (три числа):

2 координаты границ + 1 вес.

Для хранения информации (и для ряда других операций) это может дать выигрыш на порядки по сравнению с непрерывным (поточечным) описанием, особенно для больших баз данных и сложных систем.

2) Вторая главная предпосылка:

Суб-интервальный анализ позволяет выполнять
расчеты **при неполных данных**.

А) Суб-интервальный анализ позволяет выполнять расчеты тогда, когда затруднительно или даже невозможно получение полных данных.*
Б) Суб-интервальный анализ позволяет управлять продолжительными процессами, для которых возможны и целесообразны корректировки.*
* Следует подчеркнуть, что и в А) и в Б) речь идет о расчетах **полных (конечных)** результатов **по неполным (промежуточным)** данным.

3) Третья основная предпосылка:

Суб-интервальный анализ позволяет выполнять
**точные** (**достоверные**) расчеты и оценки.

Суб-интервальный анализ позволяет выполнять расчеты и оценки не менее точные, чем исходные данные (ср., напр., с теорией вероятностей).

## 1.4. Пример. Расчет положения центра тяжести (среднего значения). Ящик с шарами

Можно рассмотреть в качестве примера ящик, который будет выполнять роль интервала. Роли суб-интервалов будут выполнять две коробки (см. Рис. 1). В ящике находятся шары, каждый из которых может свободно перекатываться в пределах своей коробки.

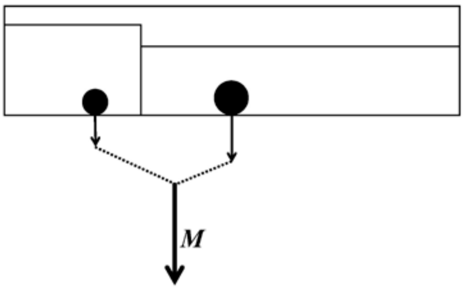


Рисунок 1. Ящик с двумя коробками

Но для начала рассмотрим интервал без суб-интервалов, то есть ящик без коробок.

### 1.4.1. Ящик без коробок

Представим себе ящик без коробок. Такой ящик изображен на Рис. 2. В ящике находятся два шара (яблоки, теннисные мячи, пушечные ядра или что-то еще). Шары могут быть расположены в любой точке ящика. Веса шаров изображены стрелками. Центр тяжести шаров обозначен $M$.

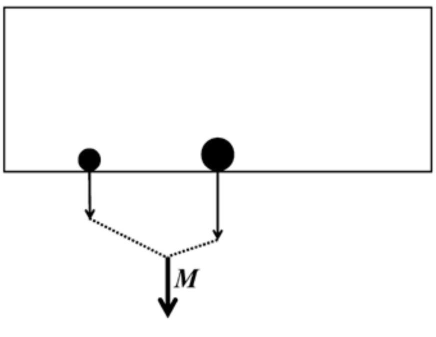


Рисунок 2. Ящик без коробок

Предположим, что шары находятся в крайней левой точке ящика (см. Рис. 3). При этом центр тяжести $M$ также находится в этой точке.

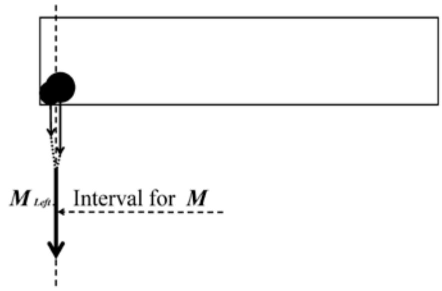


Рисунок 3. Шары в крайней левой точке

Предположим, что шары находятся в крайней правой точке ящика (см. Рис. 4). При этом центр тяжести $M$ также находится в этой точке.

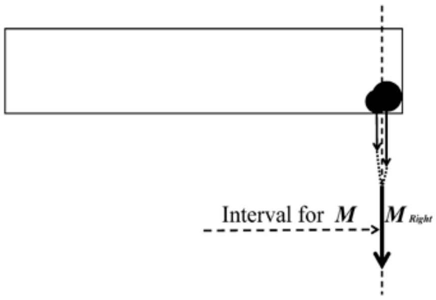


Рисунок 4. Шары в крайней правой точке

На рис. 5 представлен интервал возможных положений центра тяжести $M$. Видно, что если оба шара могут находиться в любой точке ящика, то центр тяжести $M$ тоже может быть расположен в любой точке ящика.

Точность определения центра тяжести $M$ это – весь ящик.

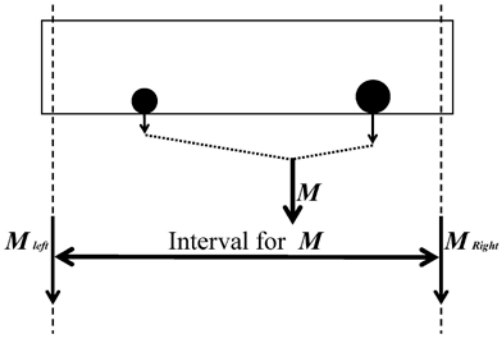


Рисунок 5. Интервал возможных положений центра тяжести $M$

Таким образом, в случае отсутствия коробок точность – наихудшая.

1.4.2. Ящик с двумя коробками

Предположим, что ящик – с двумя коробками. В каждой коробке находится один шар. Шары могут быть расположены в любой точке коробки. Положение центра тяжести (среднее значение) также обозначается через $M$.

Если шары находятся в крайних левых точках коробок, то центр тяжести $M$ также находится в крайнем левом положении (см. Рис. 6).

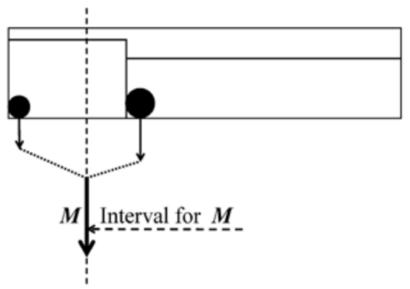


Рисунок 6. Центр тяжести $M$ при крайних левых позициях шаров

Если шары находятся в крайней правой позиции, то центр тяжести $M$ находится также в крайнем правом положении (см. Рис. 7)

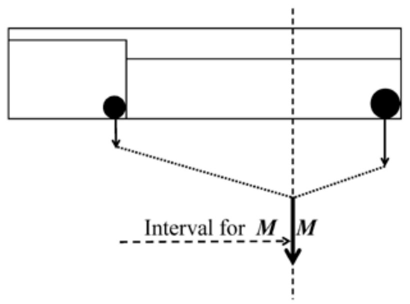


Рисунок 7. Центр тяжести $M$ при крайних правых позициях шаров

На Рис. 8 представлен интервал возможных положений $M$.

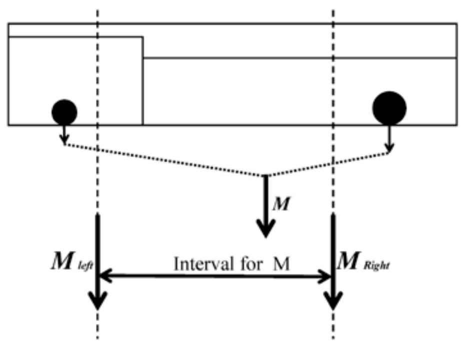


Рисунок 8. Интервал для возможных положений центра тяжести $M$.

Мы видим, что в то время, как первый шар может быть расположен в любой точке первой коробки и второй шар может быть расположен в любой точке второй коробки, интервал возможных положений центра тяжести $M$ может быть существенно меньше, чем размер всего ящика.

Мы видим, что точность определения $M$, стала значительно лучше, чем размер всего ящика. Таким образом, в случае наличия коробок (суб-интервалов), точность может быть лучше, чем в случае их отсутствия.

### Преимущества коробок (суб-интервалов)

Сравним интервал для центра тяжести для ящика без коробок и для ящика с коробками. Очевидно, что интервал для центра тяжести с коробками меньше, чем интервал для центра тяжести без коробок (см. Рис. 9)

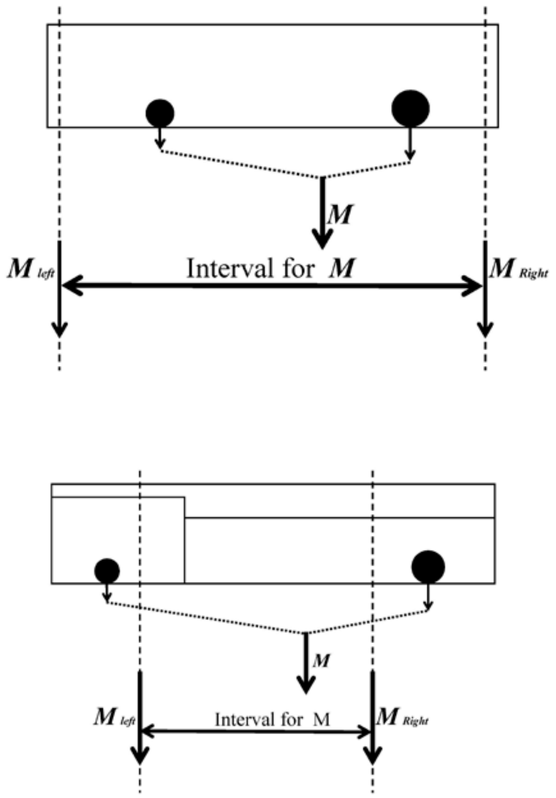


Рисунок 9. Интервалы для возможных положений центра тяжести $M$

Мы видим, что коробки (как частный случай суб-интервалов) повышают точность примерно в два раза (для двух коробок). Обратите внимание, что для описания ящика без коробок нужны 3 значения: 2 координаты границ и 1 вес. Для ящика с двумя коробками нужно 5 значений: 3 координаты границ и 2 веса. Таким образом, 2 (Два) дополнительных значения могут повысить точность примерно в два раза. При этом, интервал для центра тяжести (среднего значения) определяется точно, независимо от положения шаров.

## 2. Инструменты суб-интервального анализа

### 2.1. Суб-интервальная арифметика

**Суб-интервальная арифметика** – это инструмент для расчета аналогов моментов и других количественных характеристик всего интервала с помощью характеристик его суб-интервалов.

Рассмотрим общий случай системы из $S$ : $2 \le S < \infty$ смежных суб-интервалов. Рассмотрим интервал $\boldsymbol{X}=[A, B]$ : $0<(B-A)<\infty$, и величину или функцию $w(x_k)$, : $k=1, 2, \ldots K$ : $1 \le K \le \infty$ и $w(x_k) \ge 0$ for $A \le x_k \le B$ и

$$\sum_{k=1}^{K} w(x_k) = W < \infty \ .$$

Для простоты положим по определению $W=1$.

Определим момент (аналог момента) $n$-го : $1 \le n < \infty$, порядка для величины $w(x_k)$ относительно точки $x_0$ как

$$E(X - X_0)^n \equiv \sum_{k=1}^{K} (x_k - x_0)^n w(x_k)$$

и среднее значение величины $w(x_k)$ как

$$M \equiv \sum_{k=1}^{K} x_k w(x_k) \ .$$

Пусть величина $w(x_k)$ известна с точностью до системы смежных суб-интервалов $\{\boldsymbol{X}_s\}$ : $s=1, 2, \ldots S$ : $2 \le S \le \infty$, и обозначим $\boldsymbol{X}_{1..S} \equiv \boldsymbol{X}$ так, что

$$\begin{gathered} A \equiv \underline{X} \equiv \underline{X_{1..S}} = \underline{X_1} < \overline{X_1} = \underline{X_2} < \ldots \\ \ldots < \overline{X_S} = \overline{X_{1..S}} \equiv \overline{X} \equiv B \end{gathered} \ .$$

Очевидно, что при этом многие характеристики $\boldsymbol{X}$ будут интервалами.

Система формул суб-интервальной арифметики базируется в основном на так называемом "Кольце формул"

$$\begin{aligned} & wid M_{1..S} = \\ & = \sum_{s=1}^{S} w_s wid X_s = \\ & = wid X_{1..S} - \sum_{s=1}^{S} w_s \sum_{m=1,..,N|m \ne s} wid X_m = \\ & = wid X_{1..S} - \sum_{s=1}^{S} wid X_s \sum_{m=1,..,N|m \ne s} w_m = \\ & = wid M_{1..S} \end{aligned}$$

где $wid$ - ширина (*width*);

$M_{1..S}$ - интервал для среднего значения;

$w_s$ - относительный вес (*weight*) $s$-го суб-интервала.

Первая из формул называется "формула Новоселова". Аналогичные (но более сложные) формулы позволяют рассчитывать интервалы и границы для аналогов дисперсии (момента инерции), эксцесса, и т.д.

## 2.2. Суб-интервальный анализ неполных данных

Суб-интервальный анализ позволяет сделать точные расчеты и точные оценки при неполной информации.

**Суб-интервальный анализ неполных данных** (или **анализ неполных данных**) - это инструмент для расчета аналогов моментов и других количественных характеристик всего интервала с помощью характеристик его суб-интервалов в условиях, когда информация о характеристиках суб-интервалов является неполной.

В случае, когда информация является неполной, например, когда известны только $wid\_X_{First}$ и $w_{First}$ суб-интервала $\boldsymbol{X_{First}}$, то кольцо формул можно переписать как

$$\begin{aligned}
& w_{First} wid\, X_{First} \leq \\
& \leq wid\, M_{1..S} \leq \\
& \leq wid\, X_{1..S} - w_{First}(wid\, X_{1..S} - wid\, X_{First}) = \\
& = w_{First} wid\, X_{First} + wid\, X_{1..S}\ (1 - w_{First})
\end{aligned}$$

Для примера рассмотрим интервал $\boldsymbol{X}=[A, B] =[0, 10]\equiv\boldsymbol{X_{1..S}}$. Пусть единственное или первое измерение дает вес $w_{First}=0.7$ суб-интервала $\boldsymbol{X_{First}}=[2, 4]$ (см. Рис. 10)

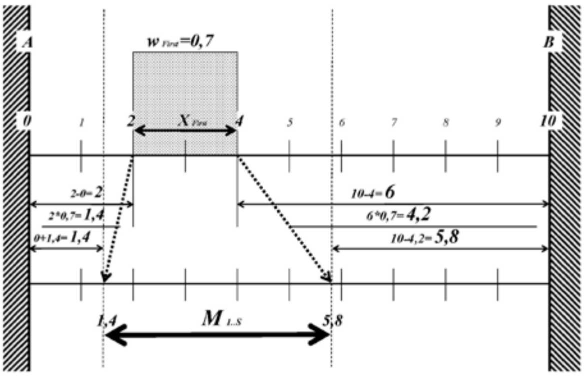


Рисунок 10. Иллюстративный пример вычисления интервала $\boldsymbol{M_{1..S}}$ для среднего значения с помощью единственного (или первого) измерения.

Интервал $\boldsymbol{M_{1..S}}$ для среднего значения удовлетворяет соотношениям

$$\underline{M_{1..S}} \geq \underline{X_{1..S}} + wid(\underline{X_{First}} - \underline{X_{1..S}})\ w_{First} = 0 + 2 \times 0.7 = 1.4$$

и

$$\overline{M_{1..S}} \leq \overline{X_{1..S}} - wid(\overline{X_{1..S}} - \overline{X_{First}})\, w_{First} = 10 - 6 \times 0.7 = 5.8$$

и

$$\begin{aligned}
& 1.4 \leq wid\, M_{1..S} \leq wid\, X_{1..S} - w_{First}(wid\, X_{1..S} - wid\, X_{First}) = \\
& = 10 - 0.7 \times 8 = 4.4
\end{aligned}$$

## 2.3. Суб-интервальные образы

**Суб-интервальный образ** изображения (*N*-мерного объекта) – это суб-интервальное представление изображения, требующее объема памяти на порядки меньшего, чем исходное изображение (*N*-мерный объект).

Примеры суб-интервальных образов (для простоты взяты цифры "4" и "5") показаны на Рис. 11.

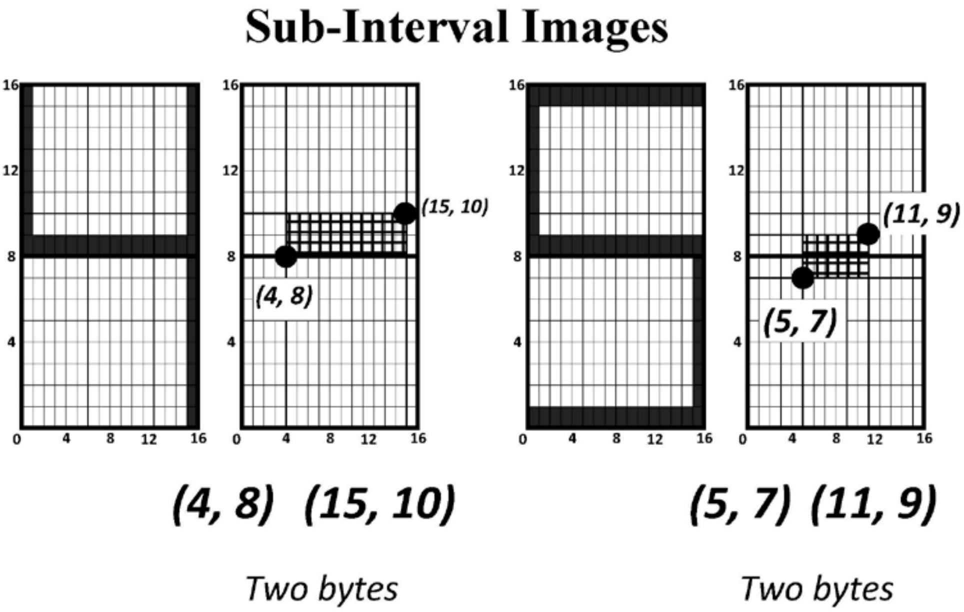


Рисунок 11. Цифры "4" и "5" и их суб-интервальные образы

Для наглядности рассмотрим пример формирования суб-интервального образа. Для простоты вначале возьмем равномерно закрашенную картину, чтобы была видна пропорциональность суб-интервалов.

Необходимо отметить, что во всей данной главе рассматриваются примеры цифр для почтовых конвертов. Поэтому решетка суб-интервалов везде удлинена в два раза в вертикальном направлении.

### Пример формирования суб-интервального образа

Возьмем любые двумерные изображения, например, изображение цифры "4" или "5" (см. Рис. 11) для почтового конверта.

Эта или любая другая двумерная картинка делится сеткой 16x16, на 3x3 суб-интервала. Затем формируются суб-интервальные образы (см. рис. 12-17).

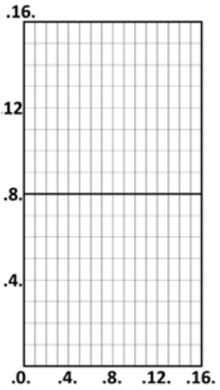


Рисунок 12. Сетка 16х16

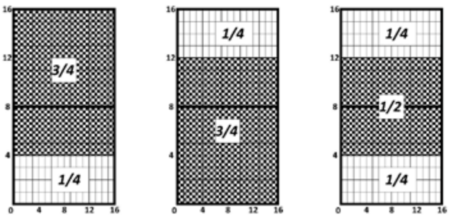


Рисунок 13. Формирование вертикальных суб-интервалов
1/4, 1/2, 1/4 на равномерном изображении

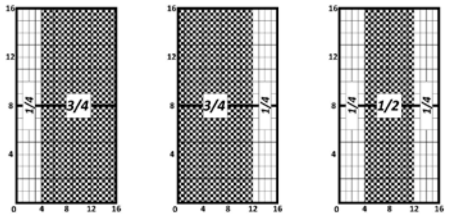


Рисунок 14. Формирование горизонтальных суб-интервалов
1/4, 1/2, 1/4 на равномерном изображении

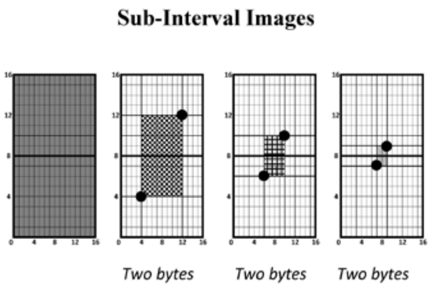


Рисунок 15. Суб-интервальные образы с весом центрального суб-интервала 1/2, 1/4, 1/8 для равномерного изображения

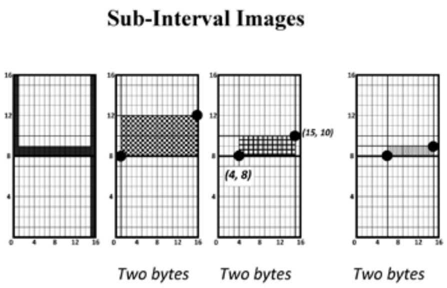


Рисунок 16. Суб-интервальные образы с весом центрального суб-интервала 1/2, 1/4, 1/8 для изображения цифры "4".

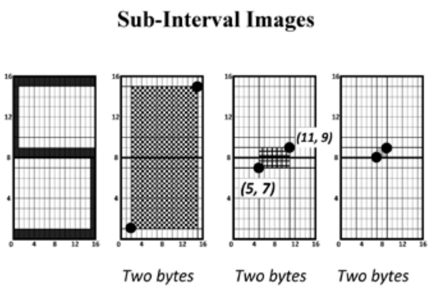


Рисунок 17. Суб-интервальные образы с весом центрального суб-интервала 1/2, 1/4, 1/8 для изображения цифры "5".

Суб-интервальные образы позволяют отображать $N$-мерное изображение с размером изображения, например, $1000^N$ пикселей в объем $N$ байт для предварительного анализа и распознавания.

Суб-интервальные образы могут быть использованы, напр., для предварительного поиска, анализа и распознавания в больших базах данных, в Интернете, в номенклатурах крупных корпораций, в базах данных, содержащих информацию о жителях города, области, республики и т.д.

## 2.4. Суб-интервальное сглаживание

**Суб-интервальное сглаживание** (суб-интервальное усреднение, суб-интервальное исчисление) – это инструмент для суб-интервального представления информации для применения аналитических методов (напр., дифференциального и интегрального исчисления). Оно может быть использовано для других областей суб-интервального анализа.

### Суб-интервальная функция (суб-интервальное сглаживание, усреднение)

Пусть на интервале $\boldsymbol{X}$ имеется некоторая функция, величина $f(x)$, определенная на некоторых точках и/или суб-интервалах интервала $\boldsymbol{X}$, и некоторое множество суб-интервалов $\{X_s\} : s=1, 2, \dots, S : S<\infty$.

Возьмем дополнительно суб-интервал $\boldsymbol{X_\delta}$ шириной $wid\_X_\delta=\delta\geq\delta_0>0$ и $\delta\ll Min(wid\_X_s)$ и будем непрерывно перемещать его от левой границы интервала $\boldsymbol{X}$ до правой. При этом, нормированный вес $w_\delta=weight\_X_\delta/weight\_X$ суб-интервала $\boldsymbol{X_\delta}$ будет пробегать некоторую непрерывную последовательность значений, формируя, определяя таким образом некоторую функцию.

Назовем такую функцию $\boldsymbol{SI(f(x))}$ или $\boldsymbol{SI(f(x), \delta)}$ **суб-интервальной функцией** (суб-интервальным сглаживанием, усреднением) с параметром $\delta$

$$SI(f(x)) \equiv SI(f(x),\delta) \equiv \frac{1}{\delta}\int_{x-\frac{\delta}{2}}^{x+\frac{\delta}{2}} f(x_1)dx_1 \ .$$

Определим суб-интервальную функцию $\boldsymbol{SI(f(x), N)}$ или $\boldsymbol{SI(f(x), N, \delta)}$ (суб-интервальное сглаживание, усреднение) ***N*–го** порядка с параметром $\delta$

$$SI(f(x),N) \equiv (\frac{N}{\delta})^N \int_{x-\frac{\delta}{2N}}^{x+\frac{\delta}{2N}} .. \int_{x_{N-1}-\frac{\delta}{2N}}^{x_{N-1}+\frac{\delta}{2N}} f(x_N)dx_1..dx_N \ .$$

Определим суб-интервальную функцию $SI(f(x), 0)$ как $SI(f(x), 0)\equiv f(x)$.

### Некоторые свойства суб-интервальных функций

Интервал определения суб-интервальной функции $SI(f(x), N)$ меньше, чем интервал $\boldsymbol{X}$, на $\delta$.

Если суб-интервал $X_s$ больше, чем $\boldsymbol{X_\delta}$ и первая производная функции $f(x)$ (или функция $f(x)$) постоянна на $X_s$, тогда на суб-интервале $X_s$-$X_\delta$ суб-интервальная функция точно равна функции $f(x)$. То есть

$$SI(f(x),N) = f(x) \,|\, x \in [\underline{X_s} + \frac{wid\ X_\delta}{2}, \overline{X_s} - \frac{wid\ X_\delta}{2}] \ .$$

Дискретная величина $f(x_a)$ преобразуется сглаживанием (функцией) $SI(f(x))$ в прямоугольный столбик шириной $\delta$ и высотой $f(x_a)/\delta$.

Ступенька преобразуется $SI(f(x))$ в склон шириной $\delta$.

Прямоугольный столбик шириной $\varepsilon : \varepsilon<\delta$ и высотой $H$ преобразуется сглаживанием суб-интервальной функцией $SI(f(x), \delta)$ с суб-интервалом $\boldsymbol{X_\delta}$ в усеченную пирамиду с шириной её плоской вершины $\delta\text{-}\varepsilon$, высотой ее плоской вершины $H\varepsilon/\delta$ и шириной её склона $\varepsilon$ (см. Рис. 18)

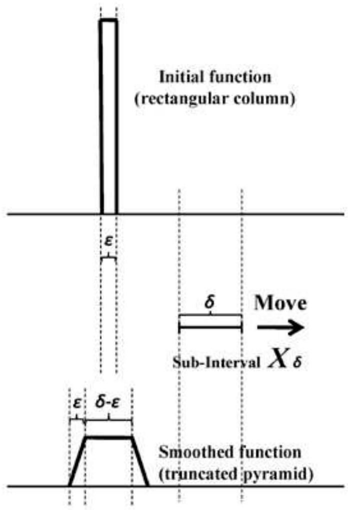


Рисунок 18. Суб-интервальное сглаживание. Прямоугольный столбик преобразуется суб-интервальной функцией $SI(f(x), \delta)$ в усеченную пирамиду

Суб-интервальная функция второго порядка $SI(f(x), 2)$ непрерывна (даже для дискретной функции $f(x_k)$).

$n$-я производная суб-интервальной функции $(n+2)$-го порядка является непрерывной (даже для дискретной функции $f(x_k)$).

Основополагающим свойством суб-интервальных функций является то, что они могут моделировать и представлять процессы протяженных (напр., линейных и объемных) измерений.

---

Примечания

Примечание 1. Идея суб-интервального поуровневого анализа родилась благодаря совету [8] ознакомиться с Астанинским Экономическим Форумом и идее G-Global [9] о том, что G-8 или G-20 могут быть недостаточными для того, чтобы адекватно решить некоторые глобальные проблемы. Суб-интервальное исчисление было создано в первую очередь как вспомогательное средство для суб-интервального поуровневого анализа.

Примечание 2. В [10] было предложено устремить порядок суб-интервальной функции к бесконечности и использовать обобщенные функции.

Примечание 3. В [11] было предложено устремить количество суб-интервалов к бесконечности, чтобы прийти к непрерывному исчислению и для проверки принципов суб-интервального анализа.

Примечание 4. Суб-интервальное исчисление (сглаживание) было стимулировано также работами [12] по интервальному дифференцированию.

## 2.5. Поуровневый суб-интервальный анализ

Поуровневый суб-интервальный анализ– это инструмент для определения и анализа уровней описания описываемого явления.

### Пример уровней описания

Глобальные экономические и политические явления могут быть описаны, например, в рамках G-2, G-3 G-8, G-20, ...

### Разделы, части суб-интервального анализа

1) Определение и анализ необходимой точности
2) Определение минимально необходимого количества суб-интервалов
3) Определение минимального размера суб-интервала
4) Определение и анализ особенностей для минимального представления
5) Определение и анализ обобщенных характеристик

Рассмотрим некоторые из этих частей, а именно : части 1 и 4.

### 1) Определение и анализ необходимой точности
### Равномерная суб-интервальная сетка

Для равномерной суб-интервальной сетки точность равна

$$wid\,M_{1..S} = \sum_{s=1}^{S} wid\,X_s\, w_s = wid\,X_k \sum_{s=1}^{S} w_s = = wid\,X_k \times 1 = wid\,X_k = \frac{wid\,X_{1..S}}{S}.$$

### Аппроксимация и интерполяция

А) Аппроксимируем описываемую функцию *f(x)* средними значениями суб-интервалов.

Б) Выполним интерполяцию *Interp(x)* путем соединения средних значений (медиан) суб-интервалов прямыми линиями.

Если функция *f(x)* подчиняется условию Липшица

$$| f(x_2) - f(x_1) | \le L\,| x_2 - x_1 | .$$

тогда, для случаев А) и Б), точность, $\Delta_{Precis}$, т.е. разница между средними значениями и функцией *f(x)* (точность среднего), и разница между интерполяцией *Interp(x)* и функцией *f(x)* для *s*-го суб-интервала, не хуже

$$\Delta_{\Pr ecis}(x \,|\, x \in X_s) \le L\frac{wid\,X_s}{2} .$$

### 4) Определение и анализ особенностей для минимального представления

Можно выделить некоторые характерные черты, которые должны быть отображены, в том числе экстремумы, области постоянства и области изменений. Рассмотрим некоторые из этих черт.

#### Области постоянства

Если существует область **X** такая, что ее ширина *wid_X* больше, чем некоторая максимальная ширина $wid_{Max}$

$$wid\ X > wid_{Max}$$

и изменения суб-интервальной функции *SI(f(x), 2)* второго порядка (суб-интервальные изменения функции *f(x)*) меньше минимальной величины $\Delta_{Min}$

$$|\Delta(SI(f(x \mid x \in X),2))| < \Delta_{Min} \ ,$$

тогда эта область должна быть выделена с помощью не менее, чем одного суб-интервала.

#### Области изменений. Области резких изменений

Если существует область **X** такая, что ее ширина *wid_X* меньше некоторой минимальной ширины $wid_{Min}$

$$wid\ X < wid_{Min}$$

и изменения суб-интервальной функции *SI(f(x), 2)* второго порядка (суб-интервальные изменения *f(x)*) больше, чем максимальное значение $\Delta_{Max}$

$$|\Delta(SI(f(x \mid x \in X),2))| > \Delta_{Max} \ ,$$

тогда граница между суб-интервалами должна пройти через эту область.

#### Области изменений. Области монотонных изменений

Если существует область **X** такая, что ее ширина *wid_X* больше, чем некоторая максимальная ширина $wid_{Max}$

$$wid\ X > wid_{Max} \ ,$$

суб-интервальная функция *SI(f(x), 2)* второго порядка монотонна, например, если $x_1 \leq x_2$ , то

$$SI(f(x_1),2) \leq SI(f(x_2),2) \ ,$$

и изменения суб-интервальной функции *SI(f(x), 2)* второго порядка (суб-интервальные изменения *f(x)*) больше, чем минимальное значение $\Delta_{Max}$

$$|\Delta(SI(f(x \mid x \in X),2))| > \Delta_{Max} ,$$

тогда такая область должна быть представлена с помощью не менее, чем одного дополнительного суб-интервала.

# 3. Утверждения. Теоремы. Гипотезы

## 3.1. Утверждения

### 3.1.1. Утверждения об объеме памяти

Неравенства

**1 суб-интервал.** 1 суб-интервал требует меньшего объема памяти для хранения информации (и для работы с ней), чем 2 точки.

Действительно: 1 суб-интервал требует хранить 2 координаты (концов) и 1 вес, а 2 точки требуют хранить 2 координаты и 2 веса.

Итого 3 значения для суб-интервалов и 4 значения для точек.

Для смежных суб-интервалов это тем более справедливо.

**3 смежных стандартных суб-интервала**. 3 смежных стандартных суб-интервала (1 интервал с тремя смежными суб-интервалами с заранее определенными соотношениями весов суб-интервалов для всей базы данных) требуют на порядок меньшего объема памяти для хранения информации (и для работы с ней), чем 3 точки.

Действительно: 3 смежных стандартных суб-интервала требуют хранить 4 координаты и 1 общий вес (вес всего интервала), а 3 точки требуют хранить 3 координаты и 3 веса.

Итого 5 значений для суб-интервалов и 6 значений для точек.

Превосходство на порядок.

Одномерные случаи

**1 суб-интервал.** 1 суб-интервал требует на порядок меньшего объема памяти для хранения информации (и для работы с ней), чем 20 точек.

Действительно: 1 суб-интервал требует хранить 2 координаты (концов) и 1 вес, а 20 точек требуют хранить 20 координат и 20 весов.

Итого 3 значения для суб-интервалов и 40 значений для точек.

Для смежных суб-интервалов это тем более справедливо.

**3 смежных стандартных суб-интервала (минимальный стандартный одномерный суб-интервальный образ)**. 3 смежных стандартных суб-интервала (суб-интервалы с заранее определенными соотношениями весов для всей базы данных) требуют на порядок меньшего объема памяти для хранения информации (и для работы с ней), чем 25 точек.

Действительно: 3 смежных стандартных суб-интервала требуют хранить 4 координаты и 1 общий вес (вес всего интервала), а 25 точек требуют хранить 25 координат и 25 весов.

Итого 5 значений для суб-интервалов и 50 значений для точек.

Превосходство на порядок.
Многомерные случаи

**Двумерный случай. $3^2$ смежных стандартных суб-интервала (минимальный стандартный двумерный суб-интервальный образ)**.

$3^2$ смежных стандартных суб-интервала (двумерный интервал с тремя смежными стандартными суб-интервалами по каждой из двух координат) требуют на порядок меньшего объема памяти для хранения информации (и для работы с ней), чем $7^2$ точек.

Действительно: $3^2$ смежных стандартных суб-интервала требуют хранить 4х2 координаты и 1 общий вес (вес всего интервала), а $7^2$ точек требуют хранить $7^2$=49 координат и 49 весов.

Итого 9 значений для суб-интервалов и 98 значений для точек.

**N-мерный случай. $3^N$ смежных стандартных суб-интервалов (минимальный стандартный N-мерный суб-интервальный образ)**.

$3^N$ смежных суб-интервалов (N-мерный интервал (брус) с тремя смежными стандартными суб-интервалами по каждой из N координат) требуют на порядок меньшего объема памяти для хранения информации (и для работы с ней), чем $5^3$ точек для 3-мерного и более случаев.

Действительно: для 3-мерного случая, $3^3$ смежных суб-интервала требуют хранить 4х3 координат и 1 общий вес (вес всего интервала), а $5^3$=125 точек требуют хранить 125 координат и 125 весов.

Итого 13 значений для суб-интервалов и 250 значений для точек.

Почему так увеличивается выигрыш с увеличением числа координат?

Минимальный N-мерный суб-интервальный образ формируется по каждой координате независимо и сразу для всего образа. Поэтому, с увеличением числа координат, количество суб-интервалов растет как произведение числа «4» на число координат, в то время, как количество точек растет как число точек по одной координате в степени числа координат.

Таким образом:

Использование минимального стандартного суб-интервального одномерного образа дает выигрыш на порядок и более для объема памяти для хранения информации и для работы с ней, по сравнению с одномерным случаем 25 и более точек (или пикселей).

Использование минимального стандартного суб-интервального двумерного образа дает выигрыш на порядок и более для объема памяти для хранения информации и для работы с ней, по сравнению с двумерным случаем $7^2$ и более точек (или пикселей).

Использование минимального стандартного суб-интервального N-мерного (N>2) образа дает выигрыш на порядок и более для объема памяти для хранения информации и для работы с ней, по сравнению со случаем $5^2$ и более точек (или пикселей).

## 3.2. Теоремы

### 3.2.1. Теоремы об интервальном характере неполных знаний

Суть теорем: 1) ненулевой вес удаленный на ненулевое расстояние от центра тяжести системы, изменит все моменты системы. 2) Из двух точек разделенных ненулевым расстоянием, одна из точек будет удалена от центра тяжести системы на ненулевое расстояние.

Если на конечном отрезке дискретная величина известна во всех точках кроме двух точек, в этих двух точках величина может принимать значения не менее, чем в ненулевом интервале и эти две точки находятся на ненулевом расстоянии друг от друга, то среднее значение и аналог любого конечного момента этой величины известны с точностью до ненулевого интервала.

**Теорема.** Если на отрезке $[A, B]$ дана дискретная величина $\{w(x_k)\}$, соответствующая условиям с раздела 1.1. и ее значения точно известны во всех точках множества $\{x_k\}$ (обозначим эти точки как $\{x_{kExact}\}$), кроме двух точек $x_{Inexact1}$ и $x_{Inexact2}$, т.е. $\{w(x_k)\}=\{w(x_{kExact})\}+w(x_{Inexact1})+w(x_{Inexact2})$, значения $w(x_{Inexact1})$ и $w(x_{Inexact2})$ могут принимать значения не менее, чем в ненулевом интервале $Max(w(x_{Inexact1}))-Min(w(x_{Inexact1}))\geq\Delta>0$ и $Max(w(x_{Inexact2}))-Min(w(x_{Inexact2}))\geq\Delta>0$, и эти две точки находятся на ненулевом расстоянии $x_{Inexact2}-x_{Inexact1}\geq 2L>0$ друг от друга, то любой конечный $(n<\infty)$ момент (аналог любого конечного момента) известен с точностью до некоторого ненулевого интервала.

**Доказательство**. Обозначим точно известные части моментов $E(X-X_0)^n$

$$E(X-X_0)^n{}_{Exact} \equiv \sum_{kExact=1}^{K-2}(x_{kExact}-x_0)^n w(x_{kExact}),$$

в частности, точно известную часть $M_{Exact}\equiv E(X)_{Exact}$ от $M\equiv E(X)$

$$\begin{aligned}E(X-X_0)^n &= \sum_{k=1}^{K}(x_k-x_0)^n w(x_k) = \sum_{kExact=1}^{K-2}(x_{kExact}-x_0)^n w(x_{kExact}) + \\ &+ (x_{Inexact1}-x_0)^n w(x_{Inexact1}) + (x_{Inexact2}-x_0)^n w(x_{Inexact2}) = \\ &= E(X-X_0)^n{}_{Exact} + (x_{Inexact1}-x_0)^n w(x_{Inexact1}) + (x_{Inexact2}-x_0)^n w(x_{Inexact2})\end{aligned}.$$

Если точка $x_{Inexact1}$ удалена от точки $x_0$ на расстояние не менее, чем $|x_{Inexact1}-x_0|\geq L$, то для четного $n$-го момента получаем

$$\begin{aligned}&Max(E(X-X_0)^n) = E(X-X_0)^n{}_{Exact} + \\ &+ Max[(x_{Inexact1}-x_0)^n w(x_{Inexact1}) + (x_{Inexact2}-x_0)^n w(x_{Inexact2})] \geq \\ &\geq E(X-X_0)^n{}_{Exact} + Max((x_{Inexact1}-x_0)^n w(x_{Inexact1})) + 0 \geq \\ &\geq E(X-X_0)^n{}_{Exact} + L^n Max(w(x_{Inexact1}))\end{aligned}$$

и

$$\begin{aligned}&Min(E(X-X_0)^n) = E(X-X_0)^n{}_{Exact} + \\ &+ Min[(x_{Inexact1}-x_0)^n w(x_{Inexact1}) + (x_{Inexact2}-x_0)^n w(x_{Inexact2})] = \\ &= E(X-X_0)^n{}_{Exact} + Min((x_{Inexact1}-x_0)^n w(x_{Inexact1})) + 0 = \\ &= E(X-X_0)^n{}_{Exact} + L^n Min(w(x_{Inexact1}))\end{aligned}.$$

Интервал для $n$-го момента формируется разностью этих выражений

$$\begin{aligned}&Max(E(X-X_0)^n)-Min(E(X-X_0)^n)=\\&\geq L^n Max(w(x_{Inexact1}))-L^n Min(w(x_{Inexact1}))=\\&=L^n(Max(w(x_{Inexact1}))-Min(w(x_{Inexact1})))\geq\\&\geq L^n\Delta\end{aligned}.$$

Видно, что при $L>0$ и $\Delta>0$ интервал для четного $n$-го момента будет больше нуля, что и требовалось доказать.

Рассмотрим нечетный $n$-й момент. Если $x_{Inexact2}$-$x_0>0$, то справедливы расчеты для четного момента. Если $x_{Inexact2}$-$x_0<0$, то для нечетного $n$-го момента получаем

$$\begin{aligned}&Max(E(X-X_0)^n)=E(X-X_0)^n{}_{Exact}+\\&+Max[(x_{Inexact1}-x_0)^n w(x_{Inexact1})+(x_{Inexact2}-x_0)^n w(x_{Inexact2})]=\\&=E(X-X_0)^n{}_{Exact}+\\&+Max((x_{Inexact1}-x_0)^n w(x_{Inexact1}))-Min(|(x_{Inexact2}-x_0)^n w(x_{Inexact2})|)\geq\\&\geq E(X-X_0)^n{}_{Exact}+L^n Max(w(x_{Inexact1}))-L^n Min(w(x_{Inexact2}))=\\&=E(X-X_0)^n{}_{Exact}+L^n[Max(w(x_{Inexact1}))-Min(w(x_{Inexact2}))]\end{aligned}.$$

и

$$\begin{aligned}&Min(E(X-X_0)^n)=E(X-X_0)^n{}_{Exact}+\\&+Min[(x_{Inexact1}-x_0)^n w(x_{Inexact1})+(x_{Inexact2}-x_0)^n w(x_{Inexact2})]=\\&=E(X-X_0)^n{}_{Exact}+\\&+Min((x_{Inexact1}-x_0)^n w(x_{Inexact1}))-Max(|(x_{Inexact2}-x_0)^n w(x_{Inexact2})|)=\\&=E(X-X_0)^n{}_{Exact}+L^n Min(w(x_{Inexact1}))-L^n Max(w(x_{Inexact2}))=\\&=E(X-X_0)^n{}_{Exact}+L^n[Min(w(x_{Inexact1}))-Max(w(x_{Inexact2}))]\end{aligned}.$$

Интервал для $n$-го момента формируется разностью этих выражений

$$\begin{aligned}&Max(E(X-X_0)^n)-Min(E(X-X_0)^n)\geq\\&\geq L^n[Max(w(x_{Inexact1}))-Min(w(x_{Inexact2}))-\\&-Min(w(x_{Inexact1}))+Max(w(x_{Inexact2}))]=\\&=L^n[Max(w(x_{Inexact1}))-Min(w(x_{Inexact1}))+\\&+Max(w(x_{Inexact2}))-Min(w(x_{Inexact2}))]\geq\\&\geq L^n[\Delta+\Delta]=2L^n\Delta\end{aligned}.$$

Видно, что при $L>0$ и $\Delta>0$ интервал для четного $n$-го момента будет больше нуля, что и требовалось доказать.

Теорема доказана.

### 3.2.2. Теоремы о существовании ограничений (разрывов)

**Теорема.** Если ширина любого суб-интервала не меньше некоторой ненулевой величины $\boldsymbol{wid_{Min}}$ и вес не больше общего веса интервала за вычетом некоторой ненулевой величины $\boldsymbol{weight_{Min}}$, то для общей средней величины существуют ненулевые ограничения (разрывы) $\boldsymbol{R_{Restriction}}$ у границ общего интервала.

**Доказательство:** Ограничение (разрыв) $\boldsymbol{R_{Restriction}}$ (для левой границы)

$$\begin{aligned}
&R_{\mathrm{Re}striction} = \underline{M_{1..S}} - \underline{X_{1..S}} = \\
&= \sum_{s=2}^{S} w_s \sum_{m=1}^{s-1} wid\, X_m \geq \\
&\geq weight_{Min}(s-1) wid_{Min} \geq \\
&\geq weight_{Min} wid_{Min} > 0
\end{aligned}$$

Теорема доказана.

Данная теорема решает известные проблемы теории полезности, не решенные в течение полувека (см. [13] 1953 и [14] 2006).

## 3.3. Гипотезы

### 3.3.1. Гипотеза о неполноте интерпретации измерений

Исходное измерение или исходную серию измерений можно рассматривать как часть более полной серии измерений.

При этом, количество $\boldsymbol{N}$ и границы $\boldsymbol{X}$ измерений более полной серии должны быть больше и шире исходных $\boldsymbol{n}$ и $\boldsymbol{x}$. Получаем

$$\begin{aligned}
&\sum_{s=1}^{n} w_s wid\ x_s \leq \\
&\leq wid\ M_{1..N} \leq \\
&\leq wid\ X_{1..N} - \sum_{s=1}^{n} w_s \sum_{m=1,..,N|m\neq s} wid\ X_m
\end{aligned},$$

где

$\boldsymbol{n < N}$ - исходное и полное количество измерений

$\boldsymbol{x < X}$ - исходные и полные границы измерений

Или, упрощенно и для точечных измерений (при $wid\_x_s \sim 0$)

$$\begin{aligned}
&wid\ M_{1..N} \leq \\
&\leq wid\ X_{1..N} - \sum_{s=1}^{n} w_s (wid\ X_{1..N} - wid\ x_s)
\end{aligned}.$$

## 4. **Возможные применения суб-интервального анализа**

Суб-интервальный анализ может применяться в широком спектре областей и направлений. Такое разнообразие приложений определяется и обусловлено целым рядом предпосылок.

### 4.1. Общие предпосылки применений

Суб-интервальный анализ может быть эффективно применим, когда:

1. Важен объем памяти
2. Важна скорость
3. Важна надежность
4. Трудно получить достаточно данных

Некоторые другие предпосылки применения суб-интервального анализа:

#### Временн'ые суб-интервалы

Суб-интервальный анализ может использоваться для расчета или точных оценок интервалов средних значений, дисперсии (или других моментов), широко распространенных временн'ых суб-интервалов, таких как год, квартал, месяц, день.

#### Пространственные суб-интервалы

Суб-интервальный анализ может использоваться для решения широко распространенных пространственных суб-интервалов, таких как континент, остров, страна, область, город, район, …

#### Анализ неполных данных

Суб-интервальный анализ неполных данных позволяет проводить мониторинг, планирование и корректировку долговременных процессов, как научных, так и практических. Для первичного анализа достаточно даже единственного или первого измерения.

#### Простота и точность

Суб-интервальный анализ прост для использования и понимания. Его выводы и результаты являются не менее точными, чем первоначальные данные и предположения.

## 4.2. Экономика: микроэкономика, макроэкономика, бухгалтерский учет и аудит, эконометрика, теория полезности

Экономика является одной из наиболее естественных и широких сфер для применения суб-интервального анализа.

Микроэкономика, макроэкономика, бухгалтерский учет и аудит, теория принятия решений, эконометрика – это неполный перечень возможных областей применения суб-интервального анализа в экономике.

### 4.2.1. Микроэкономика

Суб-интервалы в микроэкономике могут быть представлены, к примеру, в формах суб-интервалов времени, например месяцы, кварталы и годы и суб-интервалов, деталей номенклатуры корпораций, складов, розничных торговых фирм и т.д.

### 4.2.2. Макроэкономика

Макроэкономика является областью естественного применения пространственных суб-интервалов как город, город, провинция, штат, и т.д. Она является областью применения временных суб-интервалов, таких как годовой суб-интервал ВВП. Суб-интервалы в макроэкономике могут быть представлены также в виде, например, суб-интервалов в Леонтьевских межотраслевых балансах.

### 4.2.3. Бухгалтерский учет

Бухгалтерский учет является естественным применением временных суб-интервалов, таких как месяцы и кварталы для выручки, прибыли и т.д. Дополнительные типы суб-интервалов могут быть представлены в виде разделов и подразделов бухгалтерской отчетности. Так, суб-интервалы таких активов баланса, как пассивы и капитал распределяются между суб-интервалами более низких уровней, таких как текущие и долгосрочные задолженности и т.д.

### 4.2.4. Аудит

Аудиторская проверка нередко касается только части от общего объема информации. В этом случае, данные аудита могут быть обработаны методом анализа неполных данных. Это может позволить сделать достоверные выводы обо всей информации с использованием только ее части.

### 4.2.5. Теория принятия решений. Теория полезности

Теоремы существования ограничений могут помочь объяснить основные парадоксы теории принятия решений и полезности, таких, как недооценка высокой и переоценка низкой вероятности, неприятие риска, и т.д.

### 4.2.6. Эконометрика

Результат измерения и результаты измерений могут быть истолкованы как только часть серии измерений. Эта гипотеза может быть использована в эконометрике и улучшить ее результаты.

## 4.3. Интернет

Интернет – самый большой общедоступный массив информации. Суб-интервальный анализ может позволить на порядки ускорить работу в Интернете, в т.ч. с помощью суб-интервальных образов.

Интернет представляет собой перспективное поле для суб-интервального анализа. Легко видеть, что веб-страница, экранная страница, сайт, страница поиска – это, в определенном смысле, интервалы и суб-интервалы.

Интернет базы данных, Интернет-магазины, форумы, блоги, Интернет-поиск и т.д. могут быть областями приложений суб-интервального анализа.

## 4.4. Военное дело

Требованиями и условиями военных применений являются высокая скорость и неполнота информации. Им естественно удовлетворяют интервальный и суб-интервальный анализ.

Например, уже ведутся интервальные разработки по навигации военных кораблей и военным роботам.

Суб-интервальный анализ не только эффективен. Он доступен, прост и допускает массовое применение в армейских условиях с минимальной подготовкой.

## 4.5. Общественные науки

Общественные науки могут быть полем для применения, в т.ч. суб-интервальные образов. Суб-интервальные образы могут быть использованы, напр., для предварительного поиска, анализа и распознавания в больших данных, содержащих информацию о жителях города, области, республики и т.д.

### 4.6. Статистика

Доверительные интервалы, как суб-интервальные образы, уже широко используются на практике.

### 4.7. Фурье-анализ

Фурье-анализ – мощный и универсальный математический инструмент. Суб-интервалы могут быть включены в анализ Фурье и могут дать новые результаты в различных областях своих традиционных и новых приложений как Фурье суб-интервальный анализ.

## Заключение

В статье представлен краткий обзор и развитие пленарного доклада [1] "Суб-интервальный анализ и возможности его применения" в Московском Физико-техническом институте на факультете Инноваций и Высоких Технологий.

Рассмотрены существующие инструменты суб-интервального анализа: суб-интервальная арифметика, суб-интервальный анализ неполных данных и суб-интервальные образы. Описаны сглаживание (суб-интервальное исчисление) и суб-интервальный поуровневый анализ, как новые инструменты суб-интервального анализа.

Следует отметить, что 3 из 5 основных инструментов, суб-интервального анализа были впервые предложены на конференциях [1] и [15] Московского Физико-технического института на факультете Инноваций и Высоких Технологий.

## Благодарности

Автор заранее благодарен всем тем, кто выскажет свои замечания и предложения по настоящей статье и затронутым в ней вопросам по адресу aaharin@yandex.ru

## Литература

1. Харин, А. Суб-интервальный анализ и возможности его применения. *Труды 55-й научной конференции МФТИ*, Москва, 2012.
2. Харин, А. О возможных дополнениях к интервальной арифметике. *X Международная конференция по финансово-актуарной математике и эвентоконвергенции технологий*, 2011.
3. Moore, R. *Interval Analysis*. Prentice-Hall, Englewood Cliffs N. J., 1966.
4. Шокин, Ю. *Интервальный анализ.* Новосибирск: Сибирское отделение изд-ва "Наука", 1981.
5. Nguyen, H. Kreinovich, V. Wu, B. and Xiang, G. *Computing Statistics under Interval and Fuzzy Uncertainty.* Springer Verlag, Berlin, Heidelberg, 2012, ISBN 978-3-642-24904-4.
6. Шарый, С. *Конечномерный интервальный анализ*. Новосибирск, “XYZ”, 2010.
7. Harin, A. Introduction to Sub-Interval Analysis and its Applications (Selected Chapters). *ZBW - German National Library of Economics*, EconStor Books, № 62286, 2012. Selected chapters of the book are in open access http://ideas.repec.org/b/zbw/esmono/62286.html
8. Астафьев, Г. Частное сообщение 1. 2012.
9. Nazarbayev, N. "Address by the President of the Republic of Kazakhstan Nursultan Nazarbayev on Astana Economic Forum. 23.05.2012". http://www.akorda.kz/ru/page/vystuplenie-prezidenta-respubliki-kazakhstan-nursultana-nazarbaeva-na-v-astaninskom-ekonomicheskom-forume-23_1340714814
10. Кривцов, В. Комментарии к пленарному докладу, *55-я научная конференция МФТИ*, Москва, 2012.
11. Астафьев, Г. Частное сообщение 2. 2012.
12. Терехов, Л. К построению аналога производной интервальных величин. *Международная конференция «Современные проблемы прикладной математики и механики: теория, эксперимент и практика»*, Новосибирск, 2011.
13. Allais, M. Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école Américaine. *Econometrica* 21:503-46, 1953.
14. Kahneman, D. and Thaler, R. Anomalies: Utility Maximization and Experienced Utility. *Journal of Economic Perspectives*, 20 (1):221-234, 2006.
15. Харин А. Интервальный анализ распределений. Интервальные образы текста, речи, музыки, изображений и видеоинформации. *Труды 54-й научной конференции МФТИ*. Москва, 2011.