Kuhn-Tucker theorem foundations and its application in mathematical economics

Josheski, Dushko and Gelova, Elena

University Goce Delcev-Stip

12 October 2013

KUHN-TUCKER THEOREM FOUNDATIONS AND ITS APPLICATION IN THE MATHEMATICAL ECONOMICS

Dushko Josheski   Elena Gelova

dusko.josevski@ugd.edu.mk   ; elena.gelova@ugd.edu.mk

University Goce Delcev-Stip,R.Macedonia

Abstract

In this paper the issue of mathematical programming and optimization has being revisited. The theory of optimization deals with the development of models and methods that determine optimal solutions to mathematical problems defined. Mathematical model must be some function of any solution that accompanies a value which is a measure of quality. In mathematics Kuhn-Tucker conditions are first order necessary conditions for a solution in non-linear programming. Under, certain specific circumstances, Kuhn-Tucker conditions are necessary and sufficient conditions as well. In this paper it is also introduced the use of these mathematical methods of optimization in economics.

Keywords: Kuhn-Tucker conditions, nonlinear optimization, mathematical economics

1. Introduction

The theory of optimization deals with the development of models and methods that determine optimal solutions to mathematical problems defined. Optimal solution of a mathematical problem defined is denoted by x^*.

To conclude that a solution is optimal, there must be a measure that determines its quality and allows its comparison with other possible solutions. The mathematical model must be some function of any solution that accompanies a value which is a measure of quality. This function is usually called objective or cost function and usually is marked with $F(X)$. Mathematical optimization task is to determine a solution that provides optimal (minimum or maximum) value of $F(X)$.

The value of the function $F^* = F(x^*)$ corresponding to the optimal solution is called optimal value.

Nonlinear programming (NP) belongs to a group of dynamic methods for solving a large class static control tasks. Each control task in which the objective function $F(X)$ and a set of constraints defined by nonlinear dependencies (objective function with nonlinear function, and set limits with nonlinear algebraic equations or inequalities), down to the task of NP, whose optimal solution is found by any of the convenient methods which is most suitable for finding the particular solution.

2. Setting the task

The general formulation of the NP task can be expressed as follows: Find the value of $n-$ dimensional vector

$$X = (x_1, x_2, ..., x_n),$$

for which the objective function

$$F(X), \tag{1}$$

gets the maximum (minimum) value, and thereby be satisfied constraints

$$G(X) \leq 0, \qquad (2)$$

$$X \geq 0,$$

where $G(X)$ is $m-$dimensional vector function whose components are $g_1(X), g_2(X), ..., g_m(X)$.

The mathematical model of the general task of NP (1) - (2), which is written in the most general form, involves determining those values $x_1, x_2, ..., x_n$ for which the objective function

$$F(x_1, x_2, ..., x_n),$$

gets the maximum (minimum) value when the constraints

$$g_i(x_1, x_2, ..., x_n) \leq 0, \ (i = 1, 2, ..., m),$$

$$x_j \geq 0, \ (j = 1, 2, ..., n).$$

Often, it is common expressions in equations (2) are called conditions and restrictions inequality. Indexes m and n mutually independent, i.e. m can be less, equal or greater than n. Functions $F(X)$ and $g_i(X), (i = 1, 2, ..., m)$, in general are nonlinear functions, hence the name nonlinear programming.

For convex functions the task of NP formulated in the form (1) - (2), which requires a minimum of objective function $F(X)$. However, concave objective function and a set of constraints, the task of NP is formulated through the maximization of the objective function $F(X)$ set limits $g_i(X) \geq 0, (i = 1, 2, ..., m)$. Reducing task of (1) - (2) the task of maximizing the objective function

$$F(X), \qquad (1')$$

the set of constraints

$$X \geq 0, \qquad (2')$$

and bring the task (1 ') - (2') to the task of (1) - (2), it is not difficult, given that in the first and in the second case it is necessary to alter the sign function, i.e. instead $F(X)$ you need to put $-$ $F(X)$, or instead $G(X)$ put $-$ $G(X)$ and change the sign in the inequality of set constraints. In the opposite case requests are handled in the same way.

Therefore, no matter whether we are talking about the task (1) - (2) or task (1 ') - (2'). This feature will often use, and when it will have in mind that when it comes to the task (1) - (2), while talking about the task (1 ') - (2'), and vice versa. Because of these two tasks, which are formally different, is referred to as a single task.

3. Theorem of Kuhn and Tucker

Theorem of Kuhn and Tucker[1] in the literature often referred to as $-$ theorem saddle point. Occupies a central place in the theory of convex programming and is a generalization of the classical method of Lagrangian multipliers.

As is known, the method of Lagrangian multiples (multipliers) provides finding extreme values of functions which depend on several variables, and constraints that are set by default draws. However, the theorem of Kuhn and Tucker generalizes method of Lagrangian multipliers, expanding it by finding the extreme values of a function that depends on several variables, but when set limits are not given but only draws and inequality. Theorem of Kuhn and Tucker gives a necessary and sufficient condition must meet the vector $X = X^{*}$, which is the solution of the task (1) - (2) . The criteria for meeting the necessary and sufficient condition is established and verified based on

---

[1] Kuhn H.Tucker A., Non-Linear Programming, Proceedings of the Second Berkeley

Symposium on Mathematical Statistics and Probality, University of California Press,

Berkeley, California, 1950

generalized Lagrangian function $\Phi(X, \Lambda)$. The establishment of the post m - introduce new variables called Lagrange multiples or multipliers , which will contain the $\lambda_1, \lambda_2, ..., \lambda_m$. In other words, the Lagrangian multipliers are composed of components m - dimensional vector Λ of which depends generalizes Lagrangian function Φ which is dependent on the function of $n + m$ the variables (X, Λ) that are set as follows

$$\Phi(X, \Lambda) = F(X) + \sum_{i=1}^{m} \lambda_i g_i(X). \tag{3}$$

Even now you can give a precise definition of the theorem of Kuhn and Tucker as follows:

Vector $X = X^*$ is a solution to the task of NP, defined for finding the minimum of the function (1) with constraints (2), and then only when there is a vector $\Lambda = \Lambda^*$ such that

$$X^* \geq 0, \qquad \Lambda^* \geq 0, \tag{4}$$

$$\Phi(X^*, \Lambda) \leq \Phi(X^*, \Lambda^*) \leq \Phi(X, \Lambda^*) \tag{5}$$

for all values of $X \geq 0$, $\Lambda \geq 0$.

Then the function Φ at the point (X^*, Λ^*) must have a global minimum in the area $X^* \geq 0$ in terms of X and global maximum in the area $\Lambda^* \geq 0$, the terms of Λ, or in other words: (X^*, Λ^*) is a not negative saddle point for function Φ.

Therefore this theorem is often called saddle point theorem, given that the task of minimization $F(X)$ corresponds to the task of determining saddle point function Φ in which all the constraints are preserved only limitations to sign. The solution X^* of minimax task is both a solution of the minimization function $F(X)$ and vice versa. We will show that the conditions (4) and (5) are sufficient. Let's (X^*, Λ^*) sedlesta point function Φ in terms of the definition (5). Introducing replacement value of Φ over expression (3) and (5) we get

$$F\left(X^*\right)+\sum_{i=1}^{m}\lambda_i g_i\left(X^*\right)\le F\left(X^*\right)+\sum_{i=1}^{m}\lambda_i^* g_i\left(X^*\right)\le F(X)+\sum_{i=1}^{m}\lambda_i^* g_i(X),$$

for all values of

$$X\ge 0,\quad \Lambda\ge 0.$$

As the left inequality in the previous expression must be met for each Λ, then

$$g_i\left(X^*\right)\le 0,\quad (i=1,2,...,m),$$

i.e. X^* an admissible plan, because the area belongs to (D) which is defined by a set of constraints and conditions

$$\sum_{i=1}^{m}\lambda_i^* g_i\left(X^*\right)=0.$$

The right inequality thus taking shape

$$F\left(X^*\right)\le F(X)+\sum_{i=1}^{m}\lambda_i^* g_i(X),$$

for all values $X\ge 0$, where given the condition $\Lambda^*\ge 0$, it follows that

$$F\left(X^*\right)\le F(X),$$

for all values $X\ge 0$ that satisfy the conditions

$$g_i(X)\le 0,\qquad (i=1,2,...,m),$$

because the value of X^* is the solution of the task (1) - (2).

That the conditions (4) and (5) are required to obtain the regularity assumption upon which, in a pinch, there is at least one point $X=\overline{X}$ (admissible plan) such that

$$g_i\left(\overline{x}\right)<0, (i=1,2,...,m),$$

it is necessary to emphasize that the introduction of the presumption of regularity is unnecessary for all functions $g_i(X)$ which appear as linear features (details of the proof

here will not say). In any case, when the set limits is linear theorem of Kuhn and Tucker does not have any restrictions.

When functions $F(X)$ and $g_i(X)$ are differential, conditions (4) and (5) are equivalent to the following 'local' conditions of Kuhn - Tucker:

$$
\left.\begin{array}{l}
\left(\dfrac{\partial \Phi}{\partial x_j}\right)_{X^*,\Lambda^*} \geq 0, \\[1.2em]
x_j^*\left(\dfrac{\partial \Phi}{\partial x_j}\right)_{X^*,\Lambda^*} = 0, \\[1.2em]
x_j^* \geq 0,
\end{array}\right\} \qquad (j = 1,2,...,n),
\tag{6}
$$

$$
\left.\begin{array}{l}
\left(\dfrac{\partial \Phi}{\partial \lambda_i}\right)_{X^*,\Lambda^*} \leq 0, \\[1.2em]
\lambda_i\left(\dfrac{\partial \Phi}{\partial \lambda_i}\right)_{X^*,\Lambda^*} = 0, \\[1.2em]
\lambda_i^* \geq 0.
\end{array}\right\} \qquad (i = 1,2,...,m),
\tag{7}
$$

The terms of the Kuhn - Tucker remain valid and in some changes to the setup of the task (1) - (2). Thus, for example, you happen to setting task constraints not present $x_j \geq 0,\ (j = 1,2,...,n)$. In this case the three conditions (6) are replaced with only one condition

$$
\left(\frac{\partial \Phi}{\partial x_j}\right)_{X^*,\Lambda^*} = 0.
\tag{8}
$$

For the case when the functions $g_i(X)$ are linear, the conditions (7) is replaced by the condition

$$
\left(\frac{\partial \Phi}{\partial \lambda_i}\right)_{X^*,\Lambda^*} = 0,
\tag{9}
$$

which is a second way of writing requirement
$$
g_i(X) = 0, \qquad (i = 1,2,...,m).
$$

Multipliers λ_i , $(i=1,2,...,m)$ here are not restricted in terms of sign.

Finally, when the constraints $g_i(X)$ are linear and defined by the equation

$$g_i(X)=0,$$

conditions (6) and (7) are reduced to conditions (8) and (9), which represent a classic case of the method of Lagrange multipliers.

4.Constrained and unconstrained optimization

In the classical optimization problem, the first condition for optimization is the first partial derivative of the Lagrangian function. Now, in nonlinear programming[2]there also exists a similar type of first-order condition, Chiang, Wainwright (2005)[3]. First will take one variable case and then two variables case. So let's consider this problem:

$$max \ \pi = f(x_i) \tag{10}$$

$$\text{subject to} \quad x_i > 0 \tag{11}$$

In the previous expression it is assumed that f is differentiable. Chiang ,Wainwright(2005),here pose three solutions, first $\frac{d\pi}{dx_i}=f'(x)=0$,and this is an interior solution of the problem[4]. Second, also there may be solution where $x_1=0$,and $f'(x_1)=0$ .An the third solution is $f'(x_1)<0$. And these three conditions can be consolidated into one single condition:

$$f'(x_1)=0 \qquad x_i \geq 0 \quad \wedge \quad x_1 f'(x_1)=0$$
(12)

---

[2]Kuhn,H.,W.,,Tucker,W.,A.,(1951), *Nonlinear programming*, Second Berkeley symposium on Mathematical statistics and probability

[3] Chiang, A., Wainwright, K.,(2005), *Fundamental methods of mathematical economics*, 4[th]ed.,Mcgraw Hill

[4] This solution is interior because it lies is the region below the curve, where the feasible region is.

Now, for the two variables case,Wainwright(2007)[5], let's consider a utility function, and for the maximization problem,(x, y) are some goods that constitute the bundle of goods, consumer maximizes his/hers utility by consumption of this gods but,every consumer faces budget constraint ,which is given with the following expression :

$$maxU = U(x, y)$$

(13)

$$\text{subject to} B = P_x x + P_y y \qquad (14)$$

$$\text{and} \qquad \bar{x} \geq x$$

(15)

With a ration $\bar{x}$[6]has been imposed to be grater or equal to x.We now have two constraints but with the Lagrange method, this can be solved by adding second constraint in a same manner as first :

$$\max_{x,y} U = \max_{x,y} U(x, y) + \lambda_1 (B - P_x x - P_y y) + \lambda_2 (\bar{x} - x)$$

(16)

In the previous expression we know that budget constraint $B = P_x x + P_y y$is binding, but $\bar{x} \geq x$may be not binding. The Kuhn Tucker conditions for previous problem are given as;

$$L_x = U_x - P_x \lambda_1 - \lambda_2 = 0 \qquad x \geq 0 \ (17)$$

$$L_y = U_y - P_y \lambda_1 = 0 \qquad y \geq 0 (18)$$

$$L_{\lambda_1} = B - P_x x - P_y y \geq 0 \qquad \lambda_1 \geq 0 \ (19)$$

$$L_{\lambda_2} = \bar{x} - x \geq 0 \qquad \lambda_2 \geq 0 \ (20)$$

---

[5] Wainwright K.,(2007), *Econ 400 lecture notes*, Simon Fraser University
[6] In the expression $\bar{x}$ is the ration or the fixed portion of food allotted to some person.

Now from $\max_{x,y} U(x,y) + \lambda_1(B - P_x x - P_y y) + \lambda_2(\bar{x} - x)$, we require

that $\lambda_1(B - P_x x - P_y y) = 0$, and either $\lambda_1 = 0$,or $B - P_x x - P_y y = 0$. In the previous expression λ_1 can be interpreted as the marginal utility of the budget. If the budget constraint is not met that, $\lambda_1 = 0$.

the ration constraint also is $\bar{x} - x \geq 0$,or $\lambda_2 = 0$. Here λ_2 is interpreted as the marginal utility of relaxing ratio constraint. If we assume that $\lambda_2 = 0$, $\lambda_1 > 0$, the first order condition for maximization will become:

$$L_x = U_x - P_x \lambda_1 - \lambda_2 = 0$$
(21)

$$L_y = U_y - P_y \lambda_1 = 0 \tag{22}$$

$$L_{\lambda_1} = B - P_x x - P_y y = 0$$
(23)

First, we find a solution x^* and y^*,and these two might be first solution if we did not ignored the constraints. Now, if the solution is not correct by iteration we assume that $\lambda_2 > 0$ and $\lambda_1 > 0$,so here we use both constraints and we assume that they are binding

$$L_x = U_x - P_x \lambda_1 - \lambda_2 = 0 \tag{24}$$

$$L_y = U_y - P_y \lambda_1 = 0 \tag{25}$$

$$L_{\lambda_1} = B - P_x x - P_y y = 0 \tag{26}$$

$$L_{\lambda_2} = \bar{x} - x = 0 \tag{27}$$

With such constraints the solution will be where the two constraints intercept. Now about setting up the Lagrange conditions, this conditions can be set up for a minimum or a maximum,now let the function $f: R^n \to R$ and other function $g_i: R^n \to R$ for every $i = 1, \ldots n$,and we have to consider following optimization problem:

10

$$maxf(x) \quad\quad\quad (28)$$

$$\text{Subject to} \quad g_i(x) \leq 0 \quad\quad i = 1, \dots . n \quad\quad (29)$$

The set of points $\{x : g_i(x *) \leq 0\}, i = 1, \dots . n$, is a feasible set. It means that is every point x lies in that area the solution is optimal. Now if there exist such a solution that $x *$ where $g_i(x *) = 0$, now we say that i^{th} constraint is a binding constraint, Varian (1992)[7]. Otherwise if $g_i(x *) \neq 0$, we say that i^{th} constraint is a slack constraint or that is not binding. Now, the Kuhn-Tucker theorem states as in Varian (1992), that, if there such a point $x *$ that solves the optimization problem $maxf(x)$ subject to $g_i(x) \leq 0 ; i = 1, \dots . n$, and the constraint qualification holds at $x *$, then there exist a set of Kuhn-Tucker multipliers $A > 0, \forall i = 1, \dots . n$, such that this equality holds i.e. $Df(x *) = \sum_{i-1}^{n} \lambda_i Dg_i(x *)$, furthermore this conditions for slackness hold such that $A \geq 0; \forall z$, [8] and the second condition $A = 0; if g_i(x*) < 0$. Comparing the Kuhn-Tucker theorem to the Lagrange multiplier theorem the major difference is that the signs of the Kuhn-Tucker multipliers can be positive, and the Lagrange multipliers signs can be positive or negative. The Lagrangian for this problem $maxf(x)$ subject to $g_i(x) \leq 0 ; i = 1, \dots . n$, is given as, $\mathcal{L} = f(x) - \sum_{i-1}^{n} \lambda_i g_i(x)$, so when the problem is set to be like this Kuhn-Tucker conditions will always be non-negative[9].Now, the *envelope theorem*[10], exist in its regular version (unconstrained), and irregular version constrained, this is basic theorem for solving problems in microeconomics. Now lets consider some arbitrary problem[11]:

$$M(a) = \max_x f(x, a) \quad\quad\quad (30)$$

Where the function, where $M(a)$ gives the maximized value of the objective function f as a function of parameter a. Now let $x(a)$ be the argument of the maximum value of x,

---

[7] Varian, R.,H.,(1992),*Microeconomic analysis*, third edition
[8] z is some whole real number
[9] That is because $g_i(x) \leq 0$, and the sum of negative numbers is negative number
[10] Kimball, W. S., *Calculus of Variations by Parallel Displacement*.London: Butterworth, p. 292, 1952.
[11] E3m-lab,Lecture notes,(2011),*Basics for mathematical economics*,National technical university of Athens, Institute of communications and computer systems

that solves the maximization problem in terms of the parameter (a). Now, that

$$M(a) = f(x(a), a)$$

$$\frac{dM(a)}{da} = \frac{\partial f(x*, a)}{\partial a}, x* = x(a) \tag{31}$$

That is derivative of M with respect to a is given by the partial derivative of $f(x, a)$ with respect to a.

5. Conclusion

In mathematics Kuhn-Tucker conditions are first order necessary conditions for a solution in non-linear programming. Under, certain specific circumstances, Kuhn-Tucker conditions are necessary and sufficient conditions as well. Mathematical programming is capable of handling inequality constraints, and apart from its obvious application to industrial problems and in business management, it also enables economists to see the theory of consumption, production, and resource allocation in a new light, Chiang (1984).But some of the limitations of mathematical programming are: variables are assumed to be continuous (in practice more of the variables may admit integer values),and the static nature of the solution.

References

1. Kuhn H.Tucker(1950). **A, *Non-Linear Programming***, Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probality, University of California Press, Berkeley, California, 1950

2. Chiang, A., Wainwright, K.,(2005), ***Fundamental methods of mathematical economics***, 4^{th}ed.,Mcgraw Hill

3. Wainwright K.,(2007), ***Econ 400 lecture notes***, Simon Fraser University

4. Varian, R.,H.,(1992),***Microeconomic analysis***, third edition

5. Kimball, W. S., ***Calculus of Variations by Parallel Displacement***.London: Butterworth, p. 292, 1952.

6. E3m-lab,Lecture notes,(2011),***Basics for mathematical economics***,National technical university of Athens, Institute of communications and computer systems