LAGRANGE MULTIPLIERS AND KUHN-TUCKER CONDITIONS

INTRODUCTION

In the previous lecture the optimization of functions of multiple variables subjected to equality constraints using the method of constrained variation was dealt. Optimization of functions of multiple variables subjected to equality constraints using Lagrange multiplier and inequality constraints using Kuhn-Tucker conditions will be discussed in the present lecture with examples.

CONSTRAINTED OPTIMIZATION PROBLEM WITH EQUALITY CONSTRAINTS

Solution by method of Lagrange multipliers

As discussed in the previous lecture, a function of multiple variables, $f(x)$, is to be optimized subject to one or more equality constraints of many variables. The problem statement is as follows:

Maximize (or minimize) $f(\mathbf{X})$, subject to $g_j(\mathbf{X}) = 0$, j = 1, 2, … , m

where

$$\mathbf{X} = \begin{Bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{Bmatrix} \tag{1}$$

with the condition that $m \le n$; or else if $m > n$ then the problem becomes an over defined one and there will be no solution. Let us consider a specific case with $n = 2$ and m=1. Consider a quantity λ, called the *Lagrange multiplier* as

$$\lambda = -\left.\frac{\partial f / \partial x_2}{\partial g / \partial x_2}\right|_{(x_1^*, x_2^*)} \tag{2}$$

Using this in the constrained variation of f [given in the previous lecture in eqn. 15 as

$$df = \left.\left(\frac{\partial f}{\partial x_1} - \frac{\partial g / \partial x_1}{\partial g / \partial x_2}\frac{\partial f}{\partial x_2}\right)\right|_{(x_1^*, x_2^*)} dx_1 = 0$$

$$\left.\left(\frac{\partial f}{\partial x_1} + \lambda \frac{\partial g}{\partial x_1}\right)\right|_{(x_1^*, x_2^*)} = 0 \tag{3}$$

And (2) written as

$$\left(\frac{\partial f}{\partial x_2} + \lambda \frac{\partial g}{\partial x_2} \right) \Bigg|_{(x_1^*,\, x_2^*)} = 0 \tag{4}$$

Also, the constraint equation has to be satisfied at the extreme point

$$g(x_1, x_2)\big|_{(x_1^*, x_2^*)} = 0 \tag{5}$$

Hence equations (2) to (5) represent the necessary conditions for the point [x_1*, *x2**] to be an extreme point.

Note that λ could be expressed in terms of $\partial g / \partial x_1$ as well $\partial g / \partial x_2$ has to be non-zero. Thus, these necessary conditions require that at least one of the partial derivatives of $g(x_1, x_2)$ be non-zero at an extreme point.

The conditions given by equations (2) to (5) can also be generated by constructing a function ***L***, known as the Lagrangian function, as

$$L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda g(x_1, x_2) \tag{6}$$

Alternatively, treating ***L*** as a function of x_1,x_2 and λ, the necessary conditions for its extremum are given by

$$\begin{aligned}
\frac{\partial L}{\partial x_1}(x_1, x_2, \lambda) &= \frac{\partial f}{\partial x_1}(x_1, x_2) + \lambda \frac{\partial g}{\partial x_1}(x_1, x_2) = 0 \\
\frac{\partial L}{\partial x_2}(x_1, x_2, \lambda) &= \frac{\partial f}{\partial x_2}(x_1, x_2) + \lambda \frac{\partial g}{\partial x_2}(x_1, x_2) = 0 \\
\frac{\partial L}{\partial \lambda}(x_1, x_2, \lambda) &= g(x_1, x_2) = 0
\end{aligned} \tag{7}$$

The necessary and sufficient conditions for a general problem are discussed next.

<u>Necessary conditions for a general problem</u>

For a general problem with *n* variables and *m* equality constraints the problem is defined as shown earlier

Maximize (or minimize) $f(\mathbf{X})$, subject to $g_j(\mathbf{X}) = 0$, j = 1, 2, … , *m*

where
$$\mathbf{X} = \begin{Bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{Bmatrix}$$

In this case the Lagrange function, ***L***, will have one Lagrange multiplier λ_j for each constraint $g_j(\mathrm{X})$ as

$$L(x_1, x_2, \ldots, x_n, \lambda_1, \lambda_2, \ldots, \lambda_m) = f(\mathbf{X}) + \lambda_1 g_1(\mathbf{X}) + \lambda_2 g_2(\mathbf{X}) + \ldots + \lambda_m g_m(\mathbf{X}) \tag{8}$$

L is now a function of $n + m$ unknowns, $x_1, x_2, \ldots, x_{n,} \lambda_1, \lambda_2, \ldots, \lambda_m$, and the necessary conditions for the problem defined above are given by

$$\begin{aligned} &\frac{\partial L}{\partial x_i} = \frac{\partial f}{\partial x_i}(\mathbf{X}) + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i}(\mathbf{X}) = 0, \qquad i = 1, 2, \ldots, n; \quad j = 1, 2, \ldots, m \\ &\frac{\partial L}{\partial \lambda_j} = g_j(\mathbf{X}) = 0, \qquad j = 1, 2, \ldots, m \end{aligned} \tag{9}$$

which represent $n + m$ equations in terms of the $n + m$ unknowns, x_i and λ_j. The solution to this set of equations gives us

$$\mathbf{X} = \begin{Bmatrix} x_1^* \\ x_2^* \\ \vdots \\ x_n^* \end{Bmatrix} \quad \text{and} \quad \lambda^* = \begin{Bmatrix} \lambda_1^* \\ \lambda_2^* \\ \vdots \\ \lambda_m^* \end{Bmatrix} \tag{10}$$

The vector **X** corresponds to the relative constrained minimum of $f(\mathbf{X})$ (subject to the verification of sufficient conditions).

Sufficient conditions for a general problem

A sufficient condition for $f(\mathbf{X})$ to have a relative minimum at $\mathbf{X}^*$ is that each root of the polynomial in $\in$, defined by the following determinant equation be positive.

$$\begin{vmatrix} L_{11} - \in & L_{12} & \cdots & L_{1n} & g_{11} & g_{21} & \cdots & g_{m1} \\ L_{21} & L_{22} - \in & & L_{2n} & g_{12} & g_{22} & & g_{m2} \\ \vdots & & \ddots & \vdots & \vdots & & \ddots & \vdots \\ L_{n1} & L_{n2} & \cdots & L_{nn} - \in & g_{1n} & g_{2n} & \cdots & g_{mn} \\ & & & & & & & \\ g_{11} & g_{12} & \cdots & g_{1n} & 0 & \cdots & \cdots & 0 \\ g_{21} & g_{22} & & g_{2n} & \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots & \vdots & & & \vdots \\ g_{m1} & g_{m2} & \cdots & g_{mn} & 0 & \cdots & \cdots & 0 \end{vmatrix} = 0 \tag{11}$$

where

$$\begin{aligned} &L_{ij} = \frac{\partial^2 L}{\partial x_i \partial x_j}(\mathbf{X}^*, \lambda^*), \qquad \text{for } i = 1, 2, \ldots, n; \quad j = 1, 2, \ldots, m \\ &g_{pq} = \frac{\partial g_p}{\partial x_q}(\mathbf{X}^*), \qquad \text{where } p = 1, 2, \ldots, m \text{ and } q = 1, 2, \ldots, n \end{aligned} \tag{12}$$

Similarly, a sufficient condition for $f(\mathbf{X})$ to have a relative maximum at **X*** is that each root of the polynomial in $\in$, defined by equation (11) be negative. If equation (11), on solving yields roots, some of which are positive and others negative, then the point **X*** is neither a maximum nor a minimum.

Example

Minimize $f(\mathbf{X}) = -3x_1^2 - 6x_1x_2 - 5x_2^2 + 7x_1 + 5x_2$

Subject to $x_1 + x_2 = 5$

Solution

$g_1(\mathbf{X}) = x_1 + x_2 - 5 = 0$

$\mathbf{L}(x_1, x_2, ..., x_n, \lambda_1, \lambda_2, ..., \lambda_m) = f(\mathbf{X}) + \lambda_1 g_1(\mathbf{X}) + \lambda_2 g_2(\mathbf{X}) + ... + \lambda_m g_m(\mathbf{X})$ with $n = 2$ and $m = 1$

$$\mathbf{L} = -3x_1^2 - 6x_1x_2 - 5x_2^2 + 7x_1 + 5x_2 + \lambda_1(x_1 + x_2 - 5)$$

$$\frac{\partial \mathbf{L}}{\partial x_1} = -6x_1 - 6x_2 + 7 + \lambda_1 = 0$$

$$=> x_1 + x_2 = \frac{1}{6}(7 + \lambda_1)$$

$$=> 5 = \frac{1}{6}(7 + \lambda_1)$$

or $$\lambda_1 = 23$$

$$\frac{\partial \mathbf{L}}{\partial x_2} = -6x_1 - 10x_2 + 5 + \lambda_1 = 0$$

$$=> 3x_1 + 5x_2 = \frac{1}{2}(5 + \lambda_1)$$

$$=> 3(x_1 + x_2) + 2x_2 = \frac{1}{2}(5 + \lambda_1)$$

$$x_2 = \frac{-1}{2}$$

and, $$x_1 = \frac{11}{2}$$

Hence $\mathbf{X}^* = \left[\frac{11}{2}, \frac{-1}{2}\right]; \boldsymbol{\lambda}^* = 23$

$$\begin{pmatrix} L_{11} - \in & L_{12} & g_{11} \\ L_{21} & L_{22} - \in & g_{21} \\ g_{11} & g_{12} & 0 \end{pmatrix} = 0$$

$$L_{11} = \left.\frac{\partial^2 \mathbf{L}}{\partial x_1^2}\right|_{(\mathbf{X}^*, \boldsymbol{\lambda}^*)} = -6$$

$$L_{12} = L_{21} = \left.\frac{\partial^2 \mathbf{L}}{\partial x_1 \partial x_2}\right|_{(\mathbf{X}^*, \boldsymbol{\lambda}^*)} = -6$$

$$L_{22} = \left.\frac{\partial^2 \mathbf{L}}{\partial x_2^2}\right|_{(\mathbf{X}^*, \boldsymbol{\lambda}^*)} = -10$$

$$g_{11} = \left.\frac{\partial g_1}{\partial x_1}\right|_{(\mathbf{X}^*, \boldsymbol{\lambda}^*)} = 1$$

$$g_{12} = g_{21} = \left.\frac{\partial g_1}{\partial x_2}\right|_{(\mathbf{X}^*, \boldsymbol{\lambda}^*)} = 1$$

The determinant becomes

$$\begin{pmatrix} -6 - \in & -6 & 1 \\ -6 & -10 - \in & 1 \\ 1 & 1 & 0 \end{pmatrix} = 0$$

or

$$(-6 - \in)[-1] - (-6)[-1] + 1[-6 + 10 + \in] = 0$$
$$=> \in = -2$$

Since $\in$ is negative, $\mathbf{X}^*, \boldsymbol{\lambda}^*$ correspond to a maximum.

KUHN-TUCKER CONDITIONS

It was previously established that for both an unconstrained optimization problem and an optimization problem with an equality constraint the first-order conditions are sufficient for a global optimum when the objective and constraint functions satisfy appropriate concavity/convexity conditions. The same is true for an optimization problem with inequality constraints.

The Kuhn-Tucker conditions are both necessary and sufficient if the objective function is concave and each constraint is linear or each constraint function is concave, i.e., the problems belong to a class called the *convex programming problems.*

Consider the following optimization problem:

Minimize $f(\mathbf{X})$ subject to $g_j(\mathbf{X}) \leq 0$ for $j = 1,2,\ldots,p$; where $\mathbf{X} = [x_1, x_2, \ldots x_n]$

Then the Kuhn-Tucker conditions for $\mathbf{X}^* = [x_1^* \ x_2^* \ \ldots x_n^*]$ to be a local minimum are

$$
\begin{aligned}
\frac{\partial f}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial g}{\partial x_i} &= 0 \qquad i = 1,2,\ldots,n \\
\lambda_j g_j &= 0 \qquad j = 1,2,\ldots,m \\
g_j &\leq 0 \qquad j = 1,2,\ldots,m \\
\lambda_j &\geq 0 \qquad j = 1,2,\ldots,m
\end{aligned} \tag{13}
$$

In case of minimization problems, if the constraints are of the form $g_j(\mathbf{X}) \geq 0$, then λ_j have to be nonpositive in (13). On the other hand, if the problem is one of maximization with the constraints in the form $g_j(\mathbf{X}) \geq 0$, then λ_j have to be nonnegative.

It may be noted that sign convention has to be strictly followed for the Kuhn-Tucker conditions to be applicable.

Example 1

Minimize $f = x_1^2 + 2x_2^2 + 3x_3^2$ subject to the constraints

$$g_1 = x_1 - x_2 - 2x_3 \leq 12$$
$$g_2 = x_1 + 2x_2 - 3x_3 \leq 8$$

using Kuhn-Tucker conditions.

Solution:

The Kuhn-Tucker conditions are given by

a) $\frac{\partial f}{\partial x_i} + \lambda_1 \frac{\partial g_1}{\partial x_i} + \lambda_2 \frac{\partial g_2}{\partial x_i} = 0$

i.e.

$$2x_1 + \lambda_1 + \lambda_2 = 0 \tag{14}$$
$$4x_2 - \lambda_1 + 2\lambda_2 = 0 \tag{15}$$
$$6x_3 - 2\lambda_1 - 3\lambda_2 = 0 \tag{16}$$

b) $\lambda_j g_j = 0$

i.e.,

$$\lambda_1(x_1 - x_2 - 2x_3 - 12) = 0 \tag{17}$$
$$\lambda_2(x_1 + 2x_2 - 3x_3 - 8) = 0 \tag{18}$$

c) $g_j \leq 0$

i.e.,

$$x_1 - x_2 - 2x_3 - 12 \leq 0 \tag{19}$$
$$x_1 + 2x_2 - 3x_3 - 8 \leq 0 \tag{20}$$

d) $\lambda_j \geq 0$

i.e.,

$$\lambda_1 \geq 0 \tag{21}$$
$$\lambda_2 \geq 0 \tag{22}$$

From (17) either $\lambda_1 = 0$ or, $x_1 - x_2 - 2x_3 - 12 = 0$

Case 1: $\lambda_1 = 0$

From *(14), (15) and (16)* we have $x_1 = x_2 = -\lambda_2/2$ and $x_3 = \lambda_2/2$.

Using these in *(18)* we get $\lambda_2^2 + 8\lambda_2 = 0$

Therefore, $\lambda_2 = 0 \; or \; -8$

From *(22),* $\lambda_2 \geq 0$, therefore, λ_2=0, $\mathbf{X}^*$ = [0, 0, 0], this solution set satisfies all of *(18)* to *(21)*

Case 2: $x_1 - x_2 - 2x_3 - 12 = 0$

Using *(14), (15) and (16)*, we have $\frac{-\lambda_1 - \lambda_2}{2} - \frac{\lambda_1 - 2\lambda_2}{4} - \frac{2\lambda_1 + 3\lambda_2}{3} - 12 = 0$ or,

$17\lambda_1 + 12\lambda_2 = -144$. But conditions *(21)* and *(22)* give us $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$ simultaneously, which cannot be possible with $17\lambda_1 + 12\lambda_2 = -144$.

Hence the solution set for this optimization problem is $\mathbf{X}^*$ = [0 0 0]

Example 2

Minimize $f = x_1^2 + x_2^2 + 60x_1$ subject to the constraints

$g_1 = x_1 - 80 \geq 0$

$g_2 = x_1 + x_2 - 120 \geq 0$

using Kuhn-Tucker conditions.

Solution

The Kuhn-Tucker conditions are given by

a) $\frac{\partial f}{\partial x_i} + \lambda_1 \frac{\partial g_1}{\partial x_i} + \lambda_2 \frac{\partial g_2}{\partial x_i} + \lambda_3 \frac{\partial g_3}{\partial x_i} = 0$

i.e.

$$2x_1 + 60 + \lambda_1 + \lambda_2 = 0 \quad (23)$$

$$2x_2 + \lambda_2 = 0 \quad (24)$$

b) $\lambda_j g_j = 0$

i.e.

$$\lambda_1 (x_1 - 80) = 0 \quad (25)$$

$$\lambda_2 (x_1 + x_2 - 120) = 0 \quad (26)$$

c) $g_j \leq 0$

i.e.,

$$x_1 - 80 \geq 0 \quad (27)$$

$$x_1 + x_2 + 120 \geq 0 \quad (28)$$

d) $\lambda_j \leq 0$

i.e.,

$$\lambda_1 \leq 0 \quad (29)$$

$$\lambda_2 \leq 0 \quad (30)$$

From (25) either $\lambda_1 = 0$ or, $(x_1 - 80) = 0$

<u>*Case 1*</u>: $\lambda_1 = 0$

From *(23) and (24)* we have $x_1 = -\lambda_2/2 - 30$ and $x_2 = -\lambda_2/2$

Using these in *(26)* we get $\lambda_2 \;\; \lambda_2 - 150 \; = 0; \quad \therefore \lambda_2 = 0 \; or \; -150$

Considering $\lambda_2 = 0$, $\mathbf{X}^* = [\,30, 0]$.

But this solution set violates *(27)* and *(28)*

For $\lambda_2 = -150$, $\mathbf{X}^* = [\,45, 75]$.

But this solution set violates *(27)* .

Case 2: $(x_1 - 80) = 0$

Using $x_1 = 80$ in *(23)* and *(24)*, we have

$$\begin{aligned} \lambda_2 &= -2x_2 \\ \lambda_1 &= 2x_2 - 220 \end{aligned} \qquad (31)$$

Substitute (31) in (26), we have

$$-2x_2 \;\; x_2 - 40 \;\; = 0 \,.$$

For this to be true, either $x_2 = 0$ *or* $x_2 - 40 = 0$

For $x_2 = 0$, $\lambda_1 = -220$. This solution set violates *(27)* and *(28)*

For $x_2 - 40 = 0$, $\lambda_1 = -140$ *and* $\lambda_2 = -80$. This solution set is satisfying all equations from (27) to (31) and hence the desired. Therefore, the solution set for this optimization problem is $\mathbf{X}^* = [\ 80, 40\ ]$.

BIBLIOGRAPHY / FURTHER READING:

1. Rao S.S., *Engineering Optimization – Theory and Practice*, Fourth Edition, John Wiley and Sons, 2009.
2. Ravindran A., D.T. Phillips and J.J. Solberg, *Operations Research – Principles and Practice*, John Wiley & Sons, New York, 2001.
3. Taha H.A., *Operations Research – An Introduction*, 8th edition, Pearson Education India, 2008.