Optimization

Fourth Class

2020 - 2021

By

Dr. Jawad Mahmoud Jassim

Dept. of Math.

Education College for Pure Sciences

University of Basrah

Iraq

Chapter Six

Constrained Optimization

Lecture 4

الدكتور
جواد
محمود
جاسم

الدكتور
جواد
محمود
جاسم

3: Multivariate Optimization with Inequality Constraints

This section is concerned with the solution of the following problem:

Minimize $f(X) \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots . \quad (20)$

Subject to

$$g_j(X) \le 0, j = 1, 2, \dots, m \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \quad (21)$$

The inequality constraints can be transformed to equality constraints by adding nonnegative slack variables y_j^2 **as**

$$g_j(X) + y_j^2 = 0\ , j = 1, 2, \dots, m \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \quad (22)$$

Where the values of the slack variables are yet unknowns.

The problem is now in a form suitable for the application of the methods discussed in the preceding section.

Therefore our problem becomes

Minimize $f(X) \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots$ (23)

Subject to

$$G_j(X,Y) = g_j(X) + y_j^2 = 0\ , j = 1,2,\ldots,m \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \quad (24)$$

Where $Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$ *is the vector of the slack variables.*

This problem can be conveniently solved by the method of Lagrange multipliers.

For this, we construct the Lagrange function L as

$$L(X,Y,\lambda) = f(X) + \sum_{j=1}^{m} \lambda_j G_j(X,Y) \ \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots . \quad (25)$$

Where $\lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{bmatrix}$ *is the vector of Lagrange multipliers.*

The necessary conditions are

$$\frac{\partial L}{\partial x_i} = \frac{\partial f}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i}(X) = 0, i = 1, 2, \dots, n \quad \dots \dots \dots \dots \dots \dots \dots \quad (26)$$

$$\frac{\partial L}{\partial \lambda_j} = G_j(X, Y) = g_j(X) + y_j^2 = 0 \ , j = 1, 2, \dots, m \quad \dots \dots \dots \dots \dots \dots \dots \quad (27)$$

$$\frac{\partial L}{\partial y_j} = 2\lambda_j y_j = 0, , j = 1, 2, \dots, m \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \quad (28)$$

<u>Note (5):</u>

Equation (27) ensure that the constraints $g_j(X) \le 0, \ j = 1, 2, \dots, m$ are satisfied, while Equation (28) imply that either $\lambda_j = 0$ or $y_j = 0$.

$If \ \lambda_j = 0$, it means that the constraint is inactive and hence it can be ignored. If $y_j = 0$, it means that the constraint is active $(g_j = 0)$ at the optimum point.

Consider the division of the constraints into two subsets J_1 and J_2 where $J_1 + J_2$ $represent$ the total set of constraints.

Let the set J_1 indicate the indices of those constraints which are active at the optimum point and J_2 include the indices of all inactive constraints.

Thus for $j \epsilon J_1$, $y_j = 0$ (constraints are active) and for $j \epsilon J_2$, $\lambda_j = 0$ (constraints are inactive) and Equation (26) can be simplified as

$$\frac{\partial f}{\partial x_i} + \sum_{j \epsilon J_1} \lambda_j \frac{\partial g_j}{\partial x_i} = 0, i = 1, 2, \ldots, n \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (29)$$

Similarly Equation (27) can be written as

$$g_j(X) = 0 \; for \;\; j \epsilon J_1 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (30)$$

And

$$g_j(X) + y_j^2 = 0 \; for \; j \epsilon J_2 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (31)$$

Note (6):

Equations (29), (30) and (31) represent $\boldsymbol{n+p+(m-p)=n+m}$ **equations in the** $\boldsymbol{n+m}$ **unknowns** $\boldsymbol{x_i}$, $\boldsymbol{(i=1,2,\ldots,n), \lambda_j\ (j \epsilon J_1)}$ **and** $\boldsymbol{y_j\ (j \epsilon J_2)}$ **where** $\boldsymbol{p}$ **denotes the number of active constraints.**

Assuming that the first $\boldsymbol{p}$ **constraints are active. Equation (29) can be expressed as**

$$-\frac{\partial f}{\partial x_i}=\lambda_1\frac{\partial g_1}{\partial x_i}+\lambda_2\frac{\partial g_2}{\partial x_i}+\cdots+\lambda_p\frac{\partial g_p}{\partial x_i}\ ,i=1,2,\ldots,n\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(32)$$

Equation (32) can be written as

$$-\nabla f=\lambda_1\nabla g_1+\lambda_2\nabla g_2+\cdots+\lambda_p\nabla g_p\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(33)$$

Where $\boldsymbol{\nabla f\ and\ \nabla g_j}$ **are the gradients of the objective function and** $\boldsymbol{jth\ constraint\ given, respectively, by}$**:**

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix} \ and \ \nabla g_j = \begin{bmatrix} \frac{\partial g_j}{\partial x_1} \\ \vdots \\ \frac{\partial g_j}{\partial x_n} \end{bmatrix}.$$

<u>Note (7):</u>

Equation (33), means that the negative of the gradient of the objective function can be expressed as a linear combination of the gradients of the active constraints at the optimum point.

Kuhn Tucker Conditions

The conditions to be satisfied at a constrained minimizer point X^* of the problem stated in Equations (20) and (21) can be expressed as:

$$\frac{\partial f}{\partial x_i} + \sum_{j \,\epsilon\, J_1} \lambda_j \frac{\partial g_j}{\partial x_i} = 0, i = 1, 2, \dots, n \dots \dots \dots \dots \dots \dots \dots \dots \quad (34)$$

And

$$\lambda_j > 0 \,, j \,\epsilon\, J_1 \quad \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \quad (35)$$

These are called Kuhn – Tucker conditions and are the necessary conditions to be satisfied at a local minimizer of $f(X)$.

<u>Note (8):</u>

If the set of active constraints is not known, the Kuhn – Tucker conditions can be stated as follows:

$$\frac{\partial L}{\partial x_i} = \frac{\partial f}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i} = 0, i = 1, 2, \dots, n \quad \dots\dots\dots\dots\dots\dots\dots \quad (36)$$

$$\lambda_j g_j = 0 \,, j = 1, 2, \dots, m \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \quad (37)$$

$$g_j \le 0 \,, \; j = 1, 2, \dots, m \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \quad (38)$$

And

$$\lambda_j \ge 0 \,, \;\; j = 1, 2, \dots, m \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \quad (39)$$

Note (9):

If the problem is one of maximization or if the constraints are of the type $g_j \geq 0$, then λ_j have to be $non\ positive$ in Equations (36) – (39).