Multivariable problem with equality and inequality constraints

Rajib Bhattacharjya
Department of Civil Engineering
IIT Guwahati
Email: rkbc@iitg.ernet.in
General formulation

Min/Max \( f(X) \) \]

Subject to \( g_j(X) = 0 \) \( j = 1, 2, 3, \ldots, m \)

Where \( X = [x_1, x_2, x_3, \ldots, x_n]^T \)

This is the minimum point of the function

Now this is not the minimum point of the constrained function

This is the new minimum point

Rajib Bhattacharjya, IITG

CE 602: Optimization Method
Consider a two variable problem

Min/Max \quad f(x_1, x_2)

Subject to \quad g(x_1, x_2) = 0

Take total derivative of the function at \((x_1, x_2)\)

\[ df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 = 0 \]

If \((x_1, x_2)\) is the solution of the constrained problem, then

\[ g(x_1, x_2) = 0 \]

Now any variation \(dx_1\) and \(dx_2\) is admissible only when

\[ g(x_1 + dx_1, x_2 + dx_2) = 0 \]
Consider a two variable problem

\[ \text{Min/Max } f(x_1, x_2) \]

Subject to \[ g(x_1, x_2) = 0 \]

This can be expanded as

\[ g(x_1 + dx_1, x_2 + dx_2) = g(x_1, x_2) + \frac{\partial g(x_1, x_2)}{\partial x_1} dx_1 + \frac{\partial g(x_1, x_2)}{\partial x_2} dx_2 = 0 \]

\[ dg = \frac{\partial g}{\partial x_1} dx_1 + \frac{\partial g}{\partial x_2} dx_2 = 0 \]

\[ dx_2 = -\frac{\partial g}{\partial x_2} dx_1 \]
Consider a two variable problem

Min/Max \( f(x_1, x_2) \)

Subject to \( g(x_1, x_2) = 0 \)

\[ dx_2 = -\frac{\partial g}{\partial x_1} dx_1 \]

Putting in \( df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 = 0 \)

\[ df = \frac{\partial f}{\partial x_1} dx_1 - \frac{\partial f}{\partial x_2} \frac{\partial g}{\partial x_2} dx_1 = 0 \]

\[ \left( \frac{\partial f}{\partial x_1} \frac{\partial g}{\partial x_2} - \frac{\partial f}{\partial x_2} \frac{\partial g}{\partial x_1} \right) dx_1 = 0 \]

\[ \left( \frac{\partial f}{\partial x_1} \frac{\partial g}{\partial x_2} - \frac{\partial f}{\partial x_2} \frac{\partial g}{\partial x_1} \right) = 0 \]

This is the necessary condition for optimality for optimization problem with equality constraints
Lagrange Multipliers

Min/Max \( f(x_1, x_2) \)

Subject to \( g(x_1, x_2) = 0 \)

We have already obtained the condition that

\[
\frac{\partial f}{\partial x_1} - \frac{\partial f}{\partial x_2} \frac{\partial g}{\partial x_1} = 0
\]

By defining \( \lambda = -\frac{\partial f}{\partial x_2} \frac{\partial g}{\partial x_2} \)

We have

\[
\frac{\partial f}{\partial x_1} + \lambda \frac{\partial g}{\partial x_1} = 0
\]

We can also write

\[
\frac{\partial f}{\partial x_2} + \lambda \frac{\partial g}{\partial x_2} = 0
\]

Also put

\[
g(x_1, x_2) = 0
\]

Necessary conditions for optimality
Lagrange Multipliers

Let us define

\[ L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda g(x_1, x_2) \]

By applying necessary condition of optimality, we can obtain

\[ \frac{\partial L}{\partial x_1} = \frac{\partial f}{\partial x_1} + \lambda \frac{\partial g}{\partial x_1} = 0 \]

\[ \frac{\partial L}{\partial x_2} = \frac{\partial f}{\partial x_2} + \lambda \frac{\partial g}{\partial x_2} = 0 \]

\[ \frac{\partial L}{\partial \lambda} = g(x_1, x_2) = 0 \]

Necessary conditions for optimality
Lagrange Multipliers

Sufficient condition for optimality of the Lagrange function can be written as

\[ L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda g(x_1, x_2) \]

\[ H = \begin{bmatrix}
\frac{\partial^2 L}{\partial x_1 \partial x_1} & \frac{\partial^2 L}{\partial x_1 \partial x_2} & \frac{\partial^2 L}{\partial x_1 \partial \lambda} \\
\frac{\partial^2 L}{\partial x_2 \partial x_1} & \frac{\partial^2 L}{\partial x_2 \partial x_2} & \frac{\partial^2 L}{\partial x_2 \partial \lambda} \\
\frac{\partial^2 L}{\partial \lambda \partial x_1} & \frac{\partial^2 L}{\partial \lambda \partial x_2} & \frac{\partial^2 L}{\partial \lambda \partial \lambda}
\end{bmatrix} \]

If \( H \) is positive definite, the optimal solution is a minimum point.

If \( H \) is negative definite, the optimal solution is a maximum point.

Else it is neither minima nor maxima.
Lagrange Multipliers

Necessary conditions for general problem

Min/Max $f(X)$ Where $X = [x_1, x_2, x_3, ..., x_n]^T$

Subject to $g_j(X) = 0$ $j = 1, 2, 3, ..., m$

$L(x_1, x_2, ..., x_n, \lambda_1, \lambda_2, \lambda_3, ..., \lambda_m) = f(X) + \lambda_1 g_1(X) + \lambda_2 g_2(X), ..., \lambda_m g_m(X)$

Necessary conditions

$$\frac{\partial L}{\partial x_i} = \frac{\partial f}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i} = 0$$

$$\frac{\partial L}{\partial \lambda_j} = g_j(X) = 0$$
Lagrange Multipliers

Sufficient condition for general problem

The hessian matrix is

\[
H = \begin{bmatrix}
L_{11} & L_{12} & L_{13} & \ldots & L_{1n} & g_{11} & g_{21} & \cdots & g_{m1} \\
L_{21} & L_{22} & L_{23} & \ldots & L_{2n} & g_{12} & g_{22} & \cdots & g_{m2} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
L_{n1} & L_{n2} & L_{n3} & \ldots & L_{nn} & g_{1n} & g_{1n} & \cdots & g_{mn} \\
g_{11} & g_{12} & g_{13} & \ldots & g_{1n} & 0 & 0 & \cdots & 0 \\
g_{21} & g_{22} & g_{23} & \vdots & g_{2n} & \vdots & \vdots & \ddots & \vdots \\
g_{31} & g_{32} & g_{33} & \ldots & g_{3n} & 0 & \cdots & \cdots & \cdots \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
g_{m1} & g_{m2} & g_{m3} & \ldots & g_{mn} & 0 & \cdots & \cdots & 0
\end{bmatrix}
\]

Where,

\[
L_{ij} = \frac{\partial^2 L}{\partial x_i \partial x_j} \\
g_{ij} = \frac{\partial g_i}{\partial x_j}
\]
Lagrange Multipliers

Min/Max \( f(X) \) Where \( X = [x_1, x_2, x_3, \ldots, x_n]^T \)

Subject to \( g(X) = b \) Or, \( b - g(X) = 0 \)

Applying necessary conditions

\[
\frac{\partial f}{\partial x_i} - \lambda \frac{\partial g}{\partial x_i} = 0 \quad \text{Where, } i = 1, 2, 3, \ldots, n
\]

\( g = 0 \)

\[
\frac{\partial g}{\partial x_i} = \frac{\partial f}{\partial x_i} / \lambda
\]

There may be three conditions

\[ \lambda^* > 0 \]
\[ \lambda^* < 0 \]
\[ \lambda^* = 0 \]

Further

\[
db - dg = 0
\]

\[
db = dg = \sum_{i=1}^{n} \frac{\partial g}{\partial x_i} dx_i
\]

\[
db = \sum_{i=1}^{n} \frac{1}{\lambda} \frac{\partial f}{\partial x_i} dx_i
\]

\[
l = \frac{df}{db}
\]

\[
df = \lambda db
\]
Multivariable problem with inequality constraints

Minimize \[ f(X) \] Where \( X = [x_1, x_2, x_3, \ldots, x_n]^T \)

Subject to \[ g_j(X) \leq 0 \quad j = 1, 2, 3, \ldots, m \]

We can write \[ g_j(X) + y_j^2 = 0 \]

Thus the problem can be written as

Minimize \[ f(X) \]

Subject to \[ G_j(X, Y) = g_j(X) + y_j^2 = 0 \quad j = 1, 2, 3, \ldots, m \]

Where \( Y = [y_1, y_2, y_3, \ldots, y_m]^T \)
Multivariable problem with inequality constraints

Minimize \( f(X) \) \quad \text{where} \quad X = [x_1, x_2, x_3, \ldots, x_n]^T

Subject to \( G_j(X,Y) = g_j(X) + y_j^2 = 0 \quad j = 1,2,3,\ldots,m \)

The Lagrange function can be written as

\[
L(X,Y,\lambda) = f(X) + \sum_{j=1}^{m} \lambda_j G_j(X,Y)
\]

The necessary conditions of optimality can be written as

\[
\frac{\partial L(X,Y,\lambda)}{\partial x_i} = \frac{\partial f(X)}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j(X)}{\partial x_i} = 0 \quad i = 1,2,3,\ldots,n
\]

\[
\frac{\partial L(X,Y,\lambda)}{\partial \lambda_j} = G_j(X,Y) = g_j(X) + y_j^2 = 0 \quad j = 1,2,3,\ldots,m
\]

\[
\frac{\partial L(X,Y,\lambda)}{\partial y_j} = 2\lambda_j y_j = 0 \quad j = 1,2,3,\ldots,m
\]
Multivariable problem with inequality constraints

From equation \( \frac{\partial L(X,Y,\lambda)}{\partial y_j} = 2\lambda_j y_j = 0 \)

Either \( \lambda_j = 0 \) or \( y_j = 0 \)

If \( \lambda_j = 0 \), the constraint is not active, hence can be ignored

If \( y_j = 0 \), the constraint is active, hence have to consider

Now, consider all the active constraints,

Say set \( J_1 \) is the active constraints
And set \( J_2 \) is the active constraints

The optimality condition can be written as

\[
\frac{\partial f(X)}{\partial x_i} + \sum_{j \in J_1} \lambda_j \frac{\partial g_j(X)}{\partial x_i} = 0 \quad i = 1, 2, 3, \ldots, n
\]

\( g_j(X) = 0 \) \quad \text{for } j \in J_1

\( g_j(X) + y_j^2 = 0 \) \quad \text{for } j \in J_2
Multivariable problem with inequality constraints

\[-\frac{\partial f}{\partial x_i} = \lambda_1 \frac{\partial g_1}{\partial x_i} + \lambda_2 \frac{\partial g_2}{\partial x_i} + \lambda_3 \frac{\partial g_3}{\partial x_i} + \cdots + \lambda_p \frac{\partial g_p}{\partial x_i} \quad i = 1,2,3,\ldots,n\]

\[-\nabla f = \lambda_1 \nabla g_1 + \lambda_2 \nabla g_2 + \lambda_3 \nabla g_3 + \cdots + \lambda_m \nabla g_m\]

This indicates that negative of the gradient of the objective function can be expressed as a linear combination of the gradients of the active constraints at optimal point.

\[-\nabla f = \lambda_1 \nabla g_1 + \lambda_2 \nabla g_2\]

Let \( S \) be a feasible direction, then we can write

\[-S^T \nabla f = \lambda_1 S^T \nabla g_1 + \lambda_2 S^T \nabla g_2\]

Since \( S \) is a feasible direction, \( S^T \nabla g_1 < 0 \) and \( S^T \nabla g_2 < 0 \)

If \( \lambda_1, \lambda_2 > 0 \)

Then the term \( S^T \nabla f \) is +ve

This indicates that \( S \) is a direction of increasing function value

Thus we can conclude that if \( \lambda_1, \lambda_2 > 0 \), we will not get any better solution than the current solution.
Infeasible region

Feasible region

\( g_1(X) = 0 \)

\( g_1(X) < 0 \)

\( g_2(X) = 0 \)

\( g_2(X) < 0 \)

\( g_2(X) > 0 \)

\( g_1(X) > 0 \)
Multivariable problem with inequality constraints

The necessary conditions to be satisfied at constrained minimum points $X^*$ are

$$\frac{\partial f(X)}{\partial x_i} + \sum_{j \in J_1} \lambda_j \frac{\partial g_j(X)}{\partial x_i} = 0 \quad i = 1,2,3,\ldots,n$$

$$\lambda_j \geq 0 \quad j \in J_1$$

These conditions are called **Kuhn-Tucker conditions**, the necessary conditions to be satisfied at a relative minimum of $f(X)$.

These conditions are in general not sufficient to ensure a relative minimum, However, in case of a convex problem, these conditions are the necessary and sufficient conditions for global minimum.
Multivariable problem with inequality constraints

If the set of active constraints are not known, the Kuhn-Tucker conditions can be stated as

\[
\frac{\partial f(X)}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j(X)}{\partial x_i} = 0 \quad i = 1, 2, 3, \ldots, n
\]

\[
\begin{align*}
\lambda_j g_j &= 0 \\
g_j &\leq 0 \\
\lambda_j &\geq 0
\end{align*} \quad j = 1, 2, 3, \ldots, m
\]
Multivariable problem with equality and inequality constraints

For the problem

Minimize \( f(X) \)

Subject to

\( g_j(X) = 0 \) \quad j = 1, 2, 3, \ldots, m
\( k_k(X) = 0 \) \quad k = 1, 2, 3, \ldots, p

The Kuhn-Tucker conditions can be written as

\[
\frac{\partial f(X)}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j(X)}{\partial x_i} + \sum_{k=1}^{p} \beta_k \frac{\partial h_k(X)}{\partial x_i} = 0 \quad i = 1, 2, 3, \ldots, n
\]

\( \lambda_j g_j = 0 \) \quad j = 1, 2, 3, \ldots, m
\( g_j \leq 0 \) \quad j = 1, 2, 3, \ldots, m
\( h_k = 0 \) \quad k = 1, 2, 3, \ldots, p
\( \lambda_j \geq 0 \) \quad j = 1, 2, 3, \ldots, m
Thanks for your attention