12. Concept of Constrained Optimum Design (Kuhn-Tucker necessary conditions)

Hae-Jin Choi
School of Mechanical Engineering,
Chung-Ang University
Constrained Optimization

• Optimization with equality constraints
  • Lagrange Multiplier
  • Necessary condition

• Optimization with inequality constraints
  • Kuhn-Tucker (K-T) necessary condition

• Global optimality condition

• Sufficient condition for general optimization problem
Optimization with Equality Constraints

- Necessary condition for constrained minimization problem with \textit{only equality constraints}.
- Let assume a optimum problem of
  - Minimization of an objective function, $f(x_1, x_2)$.
  - With an equality constraint, $h(x_1, x_2) = 0$.
- The equality constraint, $h(x_1, x_2) = 0$, may be reformatted as
  \[ x_2 = \phi(x_1) \]
- The constrained problem may be converted to an \textit{unconstrained problem}
  - Minimization of an converted objective function,
    \[ f(x_1, \phi(x_1)) \]
However, in many cases, an explicit form, $x_2 = \phi(x_1)$, of an equality constraint is not available; therefore,

We introduce **Lagrange Multipliers**

Using the chain rule,

$$\frac{df}{dx_1} = \frac{\partial f(x_1, x_2)}{\partial x_1} + \frac{\partial f(x_1, x_2)}{\partial x_2} \frac{dx_2}{dx_1}$$

or

$$\frac{df}{dx_1} = \frac{\partial f(x_1, x_2)}{\partial x_1} + \frac{\partial f(x_1, x_2)}{\partial x_2} \frac{d\phi}{dx_1}$$

At an optimum point $(x_1^*, x_2^*)$, the necessary condition is

$$\frac{df(x_1^*, x_2^*)}{dx_1} = \frac{\partial f(x_1^*, x_2^*)}{\partial x_1} + \frac{\partial f(x_1^*, x_2^*)}{\partial x_2} \frac{d\phi}{dx_1} = 0 \quad (1)$$
Lagrange Multipliers

- Since $\phi(x_1)$ is unknown, we need to eliminate it from the equation, to eliminate it, we differentiate the constraint equation, $h(x_1, x_2)=0$, at the point $(x_1^*, x_2^*)$

$$\frac{dh(x_1^*, x_2^*)}{dx_1} = \frac{\partial h(x_1^*, x_2^*)}{\partial x_1} + \frac{\partial h(x_1^*, x_2^*)}{\partial x_2} \frac{d\phi}{dx_1} = 0$$

or

$$\frac{d\phi}{dx_1} = -\frac{\partial h(x_1^*, x_2^*) / \partial x_1}{\partial h(x_1^*, x_2^*) / \partial x_2} \tag{2}$$

- Replacing $\frac{d\phi}{dx_1}$ in Eq. (1) with Eq. (2)

$$\frac{\partial f(x_1^*, x_2^*)}{\partial x_1} \frac{\partial f(x_1^*, x_2^*)}{\partial x_2} \left( \frac{\partial h(x_1^*, x_2^*) / \partial x_1}{\partial h(x_1^*, x_2^*) / \partial x_2} \right) = 0 \tag{3}$$
Lagrange Multipliers

- Defining Lagrange Multiplier as

\[
v = -\frac{\partial f(x_1^*, x_2^*)}{\partial x_2} \frac{\partial h(x_1^*, x_2^*)}{\partial x_2}
\]  \hspace{1cm} (4)

- Eq.(3) becomes

\[
\frac{\partial f(x_1^*, x_2^*)}{\partial x_1} + v \frac{\partial h(x_1^*, x_2^*)}{\partial x_1} = 0
\]  \hspace{1cm} (5)

- Rearranging Eq. (4)

\[
\frac{\partial f(x_1^*, x_2^*)}{\partial x_2} + v \frac{\partial h(x_1^*, x_2^*)}{\partial x_2} = 0
\]  \hspace{1cm} (6)

- Solving Eq.(5), Eq. (6), and \( h(x_1, x_2) = 0 \), simultaneously for \( x_1, x_2 \) and \( v \)

- Here, \( v \) is an **Lagrange Multiplier**
Lagrange Multipliers

**Constrained Problem**

*Given*

- Problem definition

*Find*

- $x$

*Satisfy*

- $h_i(x) = 0 \quad ; \quad i = 1 \text{ to } p$

*Minimize*

- $f(x)$

**Unconstrained Problem**

*Given*

- Problem definition

*Find*

- $x, v$

*Satisfy*

- ...

*Minimize*

- $L(x, v) = f(x) + \sum_{j=1}^{p} v_j h_j(x) = f(x) + v^T h(x)$
Necessary Condition of Lagrange Function

- Objective function with Lagrange multiplier

\[ L(x, v) = f(x) + \sum_{j=1}^{p} v_j h_j(x) = f(x) + v^T h(x) \]

- Necessary condition in terms of a Lagrange function defined above is

\[ \nabla L(x^*, v^*) = 0 \quad \text{or} \quad \frac{\partial L(x^*, v^*)}{\partial x_i} = 0; \quad i = 1 \text{ to } n \]

and

\[ \frac{\partial L(x^*, v^*)}{\partial v_j} = h_j(x^*) = 0; \quad j = 1 \text{ to } p \]
Example 1: Lagrange Multiplier

**Given**

Problem statement

**Find**

$x_1, x_2$

Satisfy

$h(x_1, x_2) \equiv x_1 + x_2 - 2 = 0$

Minimize

$f(x_1, x_2) \equiv (x_1 - 1.5)^2 + (x_2 - 1.5)^2$

The Lagrange function with Lagrange multiplier

$L(x_1, x_2, \nu) = f(x_1 + x_2) + \nu h(x_1, x_2)$

$= (x_1 - 1.5)^2 + (x_2 - 1.5)^2 + \nu(x_1 + x_2 - 2)$

$\frac{\partial L}{\partial x_1} = 2(x_1 - 1.5) + \nu = 0 \quad (1)$

$\frac{\partial L}{\partial x_2} = 2(x_2 - 1.5) + \nu = 0 \quad (2)$

$\frac{\partial L}{\partial \nu} = x_1 + x_2 - 2 = 0 \quad (3)$

Solving Eqs. (1)-(3) simultaneously

$x_1 = 1, x_2 = 1 \text{ and } \nu = 1.$
Geometrical meaning of Lagrange Multipliers

\[ \nabla L(x^*, v^*) = \nabla f + v \nabla h = 0 \]

Rearranging
\[ \nabla f = -v \nabla h \]

At candidate minimum point, gradient of the cost and constraint functions are along the same line and proportional to each other.

The Lagrange multiplier \( v \) is the proportionality constant.
Example 2: Cylinder Tank Design

- Cylindrical tank design. The problem is to find radius $R$ and length $l$ of the cylinder to minimize surface area while achieving the storage volume $V$.

**Given**
- Assumption of thin sheet metal
- $V =$ given volume to achieve

**Find**
- $R =$ Diameter of cylindrical storage (cm)
- $l =$ length of cylindrical storage (cm)

**Satisfy**
- $g_1(R,l) = \pi R^2 l - V = 0$

**Minimize**
- $f(D,H) = R^2 + Rl$

**Solving**

\[
\begin{align*}
\frac{\partial L}{\partial R} &\equiv 2R + l + 2\pi v R l = 0 \\
\frac{\partial L}{\partial l} &\equiv R + \pi R^2 = 0 \\
\frac{\partial L}{\partial v} &\equiv h \equiv \pi R^2 l - V = 0
\end{align*}
\]

- $R^* = \left( \frac{V}{2\pi} \right)^{1/3}$
- $l^* = \left( \frac{4V}{\pi} \right)^{1/3}$
- $v^* = - \frac{1}{\pi R} = \left( \frac{2}{\pi^2 V} \right)^{1/3}$

or Newton-Raphson method
The design problems formulated in Lecture 9 module often included inequality constraints of the form

\[ g_i(x) \leq 0; \quad i = 1 \text{ to } m \]

We can transform an inequality constraint to an equality by adding a new variable to it, called the \textit{slack variable}, which must always be nonnegative (i.e., positive or zero)

\[ g_i(x) + s_i = 0 \quad \text{where } s_i \geq 0 \]

To avoid additional constraint \( s_i \geq 0 \),

\[ g_i(x) + s_i^2 = 0 \]
Necessary Condition for Inequality Constraints

- The design problems formulated in Lecture 9 module often included inequality constraints of the form
  \[ g_i(x) \leq 0; \quad i = 1 \text{ to } m \]

- We can transform an inequality constraint to an equality by adding a new variable to it, called the \textit{slack variable}, which must always be nonnegative (i.e., positive or zero)
  \[ g_i(x) + s_i = 0 \quad \text{where } s_i \geq 0 \]

- To avoid additional constraint \( s_i \geq 0 \), use
  \[ g_i(x) + s_i^2 = 0 \]

- Now, this form can be used for Lagrange multiplier to find necessary condition
Necessary Condition for Inequality Constraints

• New equations with the necessary condition,

\[
\frac{\partial L}{\partial s_i} = 0
\]

for the Lagrangian \( L \) to be stationary with respect to the slack variables

• Lagrange multipliers, \( u \), of inequality conditions must be **non-negative.**

\[
u_j^* \geq 0; \quad j = 1 \ to \ m
\]

• If the constraint is inactive at the optimum, its associated Lagrange multiplier is zero.

• If it is active then associated multiplier must be non-negative.
Example 3: Resolving Example 1

- Resolving Example 1 with an inequality constraint

**Given**

Problem statement

**Find**

\( x_1, x_2 \)

**Satisfy**

\( h(x_1, x_2) \equiv x_1 + x_2 - 2 \leq 0 \)

**Minimize**

\( f(x_1, x_2) \equiv (x_1 - 1.5)^2 + (x_2 - 1.5)^2 \)

\[
L = (x_1 - 1.5)^2 + (x_2 - 1.5)^2 + u(x_1 + x_2 - 2 + s^2)
\]

\[
\frac{\partial L}{\partial x_1} \equiv 2(x_1 - 1.5) + u = 0
\]

\[
\frac{\partial L}{\partial x_2} \equiv 2(x_2 - 1.5) + u = 0
\]

\[
\frac{\partial L}{\partial u} \equiv x_1 + x_2 - 2 + s^2 = 0
\]

\[
\frac{\partial L}{\partial s} \equiv 2us = 0
\]
Example 3: Resolving Example 1

\[ L = (x_1 - 1.5)^2 + (x_2 - 1.5)^2 + u(x_1 + x_2 - 2 + s^2) \]

\[
\frac{\partial L}{\partial x_1} \equiv 2(x_1 - 1.5) + u = 0
\]

\[
\frac{\partial L}{\partial x_2} \equiv 2(x_2 - 1.5) + u = 0
\]

\[
\frac{\partial L}{\partial u} \equiv x_1 + x_2 - 2 + s^2 = 0
\]

\[
\frac{\partial L}{\partial s} \equiv 2us - 0
\]

The equations are nonlinear and they can have many roots

(Sol 1) Setting \( s=0 \) to satisfy \( 2us=0 \), \( x_1=x_2=1, u=1 \) and \( s=0 \)

(Sol 2) Setting \( u=0 \) to satisfy \( 2us=0 \), \( x_1=x_2=1.5, u=0 \), and \( s^2=-1 \)

This is not a valid solution since \( g=- s^2 > 0 \)

Note: the necessary condition \( u \geq 0 \) insures that the gradients of the cost and the constraint functions point in opposite directions. This way \( f \) cannot be reduced any further by stepping in the negative gradient direction without violating the constraints.
Kuhn-Tucker (K-T) Necessary Conditions

- The equality and inequality constraints can be summed up in what are commonly known as the **Kuhn-Tucker (K-T) necessary conditions**

\[
L(x, v, u, s) = f(x) + \sum_{i=1}^{p} v_i h_i(x) + \sum_{j=1}^{m} u_j (g_j(x) + s_j^2)
= f(x) + v^T h(x) + u^T (g(x) + s^2)
\]

\[
\frac{\partial L}{\partial x_k} = \frac{\partial f}{\partial x_k} + \sum_{i=1}^{p} v_i^{*} \frac{\partial h_i}{\partial x_k} + \sum_{j=1}^{m} u_j^{*} \frac{\partial g_j}{\partial x_k} = 0; \quad k = 1 \text{ to } n
\]

\[
h_i(x^*) = 0; \quad i = 1 \text{ to } p
\]

\[
g_j(x^*) + s_j^2 = 0; \quad j = 1 \text{ to } m
\]

\[
u_j^{*} s_j = 0; \quad j = 1 \text{ to } m
\]

\[
u_j^{*} \geq 0; \quad j = 1 \text{ to } m
\]

all derivatives are evaluated at point \( x^* \)
Kuhn-Tucker (K-T) Necessary Conditions

- Negative gradient direction (steepest descent direction) for the objective function is a linear combination of the gradients of constraints with Lagrange multipliers.
  \[- \frac{\partial f}{\partial x_k} = \sum_{i=1}^{p} v_i^* \frac{\partial h_i}{\partial x_k} + \sum_{j=1}^{m} u_j^* \frac{\partial g_j}{\partial x_k}; \quad k = 1 \text{ to } n\]

- With \(m\) inequality constraints, the switching condition \((u_j s_j = 0)\) lead to \(2^m\) distinct solution cases

- Evolution of \(s_i\) essentially implies evaluation of the constraints function \(g_i(x)\) since \(s_i^2 = -g_i(x)\).
  - If an inequality constraint \(g_i(x) \leq 0\) is inactive at the candidate minimum point \(x^*\), \((i.e., \ g_i(x^*) < 0 \ or \ s_i^2 > 0)\), then the corresponding Lagrange multiplier, \(u_i^* = 0\).
  - If it is active \((i.e., \ g_i(x^*) = 0 \ or \ s_i^2 = 0)\), then \(u_i^* \geq 0\)
Example 4: K-T necessary conditions

Minimize \( f(x) = \frac{1}{3} x^3 - \frac{1}{2} (b + c)x^2 + bcx + f_0 \)

Subject to \( a \leq x \leq d \)

where

\( 0 < a < b < c < d, \text{ and } f_0 \text{ are constants} \)

\( g_1 \equiv a - x \leq 0 \)

\( g_2 \equiv x - d \leq 0 \)

\[ L = \left\{ \frac{1}{3} x^3 - \frac{1}{2} (b + c)x^2 + bcx + f_0 \right\} + u_1(a - x + s_1^2) + u_2(x - d + s_2^2) \]

K-T necessary condition

\[ \frac{\partial L}{\partial x} = x^2 - (b + c)x + bc - u_1 + u_2 = 0 \]

\( (a - x) + s_1^2 = 0 \)

\( (x - d) + s_2^2 = 0 \)

\( u_1 s_1 = 0, \quad u_2 s_2 = 0 \)

\( u_1 \geq 0, \quad u_2 \geq 0 \)

Four cases
Example 4: K-T necessary conditions

Case 1: \( u_1 = 0, u_2 = 0 \), which give
\[(x - b)(x - c) = 0\]

At \( x = b \)
\[s_1^2 = b - a > 0, \quad s_2^2 = d - b > 0 \quad \text{(Valid)}\]

At \( x = c \)
\[s_1^2 = c - a > 0, \quad s_2^2 = d - c > 0 \quad \text{(Valid)}\]

Checking sufficient condition
\[
\frac{d^2f}{dx^2} = 2x - (b + c)
\]
\[
\frac{d^2f}{dx^2} > 0 \quad \text{at} \quad x = c
\]

The cost function is a local minimum at \( x = c \)

Satisfy sufficient condition
\( c \) is the local minimum
Example 4: K-T necessary conditions

Case 2: \( u_1 = 0, \ s_2 = 0 \)

\[ x = d \quad \text{(Point D)} \]
\[ u_2 = -(d - c)(d - b) < 0 \quad \text{(Not Valid)} \]

Case 3: \( s_1 = 0, \ u_2 = 0 \)

\[ x = a \quad \text{(Point A)} \]
\[ u_1 = (a - b)(a - c) > 0 \quad \text{(Valid)} \]

Case 4: \( s_1 = 0, s_2 = 0 \)

\[ x = a \text{ and } x = d, \text{ which is not possible} \quad \text{(Not Valid)} \]
Example 5: K-T necessary conditions

Minimize
\[ f(x) = x_1^2 + x_2^2 - 3x_1x_2 \]
Subject to
\[ g(x) = x_1^2 + x_2^2 - 6 \leq 0 \]

(Sol) Lagrange function is
\[ L = x_1^2 + x_2^2 - 3x_1x_2 + u(x_1^2 + x_2^2 - 6 + s^2) \]
\[ \frac{\partial L}{\partial x_1} = 2x_1 - 3x_2 + 2ux_1 = 0 \]
\[ \frac{\partial L}{\partial x_2} = 2x_2 - 3x_1 + 2ux_2 = 0 \]
\[ x_1^2 + x_2^2 - 6 + s^2 = 0 \]
\[ us = 0 \]
\[ u \geq 0 \]

Case 1: \( u=0 \)
\[ 2x_1 - 3x_2 = 0 \]
\[ 2x_2 - 3x_1 = 0 \]
\[ x_1 = x_2 = 0, u = 0, s^2 = 6 \] (valid point)

Case 2: \( s=0 \)
Solving three equations for \( x_1, x_2, \) and \( u \) simultaneously
\[ u = -1 + 3x_2 / 2x_1 \]
Substituting for \( u \)
\[ x_1^2 = x_2^2 \]
\[ x_1 = x_2 = \sqrt{3}, u = \frac{1}{2} \]
\[ x_1 = x_2 = -\sqrt{3}, u = \frac{1}{2} \]
\[ x_1 = -x_2 = \sqrt{3}, u = -\frac{5}{2} \] (Not valid)
\[ x_1 = -x_2 = -\sqrt{3}, u = -\frac{5}{2} \] (Not valid)
Example 5: K-T necessary conditions

Case 3: \( u = s = 0 \)

All K-T conditions cannot be satisfied (not valid)

<table>
<thead>
<tr>
<th>( x_1 )</th>
<th>( x_2 )</th>
<th>( u )</th>
<th>( f )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( \sqrt{3} )</td>
<td>( \sqrt{3} )</td>
<td>( 1/2 )</td>
<td>-3</td>
</tr>
<tr>
<td>( -\sqrt{3} )</td>
<td>( -\sqrt{3} )</td>
<td>( 1/2 )</td>
<td>-3</td>
</tr>
</tbody>
</table>

Point A and B satisfy the sufficient condition for local minima.

Any feasible move from the points results in an increase in the cost, and any further reduction in the cost results in violation of the constraint.

Point 0 is a saddle point (stationary point)
Example 6: K-T necessary condition

Minimize

\[ f(x_1, x_2) = x_1^2 + x_2^2 - 2x_1 - 2x_2 + 2 \]

Subject to

\[ g_1 = -2x_1 - x_2 + 4 \leq 0 \]
\[ g_2 = -x_1 - 2x_2 + 4 \leq 0 \]

K-T necessary conditions are

\[ \frac{\partial L}{\partial x_1} = 2x_1 - 2 - 2u_1 - u_2 = 0 \]
\[ \frac{\partial L}{\partial x_2} = 2x_2 - 2 - u_1 - 2u_2 = 0 \]
\[ g_1 = -2x_1 - x_2 + 4 + s_1^2 = 0 \]
\[ g_2 = -x_1 - 2x_2 + 4 + s_2^2 = 0 \]
\[ u_1 s_1 = 0, u_2 s_2 = 0, \]
\[ u_1 \geq 0, u_2 \geq 0 \]

\[ L = x_1^2 + x_2^2 - 2x_1 - 2x_2 + 2 \]
\[ + u_1(-2x_1 - x_2 + 4 + s_1^2) + u_2(-x_1 - 2x_2 + 4 + s_2^2) \]

Four cases

1. \( u_1 = 0, \quad u_2 = 0 \)
2. \( u_1 = 0, \quad s_2 = 0 \) (or \( g_2 = 0 \))
3. \( s_1 = 0 \) (or \( g_1 = 0 \)), \( u_2 = 0 \),
4. \( s_1 = 0 \) (or \( g_1 = 0 \)), \( s_2 = 0 \) (or \( g_2 = 0 \))
Example 6: K-T necessary condition

Case 1: \( u_1 = u_2 = 0 \)
This gives
\[ x_1 = x_2 = 1, \text{ and } s_1^2 = -1, s_2^2 = -1 \]
(Not Valid)

Case 2: \( u_1 = 0, s_2 = 0 \)
\[ 2x_1 - 2 - u_2 = 0 \]
\[ 2x_2 - 2 - 2u_2 = 0 \]
\[ -x_1 - 2x_2 + 4 = 0 \]
Solving for the unknowns
\[ x_1 = 1.2; x_2 = 1.4; u_1 = 0; u_2 = 0.4; f = 0.2 \]
This will give \( s_2^2 = -0.2 < 0 \) (Not Valid)

Case 3: \( s_1 = 0, u_2 = 0 \)
\[ 2x_1 - 2 - 2u_1 = 0 \]
\[ 2x_2 - 2 - u_1 = 0 \]
\[ -2x_1 - x_1 + 4 = 0 \]
Solving for the unknowns
\[ x_1 = 1.4; x_2 = 1.2; u_1 = 0.4; u_2 = 0; f = 0.2 \]
This will give \( s_2^2 = -0.2 < 0 \) (Not Valid)

Case 4: \( s_1 = 0, s_2 = 0 \)
Solving the following equations for the unknowns
\[ 2x_1 - 2 - 2u_1 - u_2 = 0 \]
\[ 2x_2 - 2 - u_1 - 2u_2 = 0 \]
\[ -2x_1 - x_2 + 4 = 0 \]
\[ -x_1 - 2x_2 + 4 = 0 \]
\[ x_1 = \frac{4}{3}, x_2 = \frac{4}{3}, u_1 = \frac{2}{9}, u_2 = \frac{2}{9} \] (Candidate of local minimum)
Example 6: K-T necessary condition

Minimum at Point A
\[ x^+ = (4/3, 4/3) \]
\[ f(x^+) = 2/9 \]
**HOMEWORK**

**Given**

Problem definition

**Find**

$x, y$

**Subject to**

$g_1 = x + y - 12 \leq 0,$

$g_2 = x - 8 \leq 0$

**Minimize**

$f(x, y) = (x - 10)^2 + (y - 8)^2$

- Find the candidates of local minima applying K-T necessary conditions
Constrained Optimization

- Optimization with equality constraints
  - Lagrange Multiplier
  - Necessary condition
- Optimization with inequality constraints
  - Kuhn-Tucker (K-T) necessary condition
- Global optimality condition
- Sufficient condition for general optimization problems