MTM4502-Optimization Techniques

Hale Gonce Köçken and Gökhan Göksu

KKT Week - 28/04/2025

Historical Background

- The conditions known as the KKT Conditions were first published in 1951 by Princeton University professors, American mathematician Harold William Kuhn and Canadian mathematician Albert William Tucker.

- H. W. Kuhn and A. W. Tucker.
Nonlinear Programming.
Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 481–492, University of California Press, Berkeley, California, 1951.

Historical Background

- Then, over time, it was realized that the necessary conditions of the nonlinear optimization problem were stated in the 1939 master's thesis of William Karush, who was then a graduate student at the University of Chicago.

- W. Karush.
Minima of Functions of Several Variables with Inequalities as Side Constraints.
MSc Thesis, Chicago University, Dept. of Mathematics, Chicago, Illinois.

Problem Statement

- The mathematical formulation of the problem of finding the minimum (maximum) of a given function under equality and inequality constraints is as follows:

$$\begin{aligned} \underset{\mathbf{x}\in\mathbb{R}^n}{\text{minimize}}\ & f(\mathbf{x}) \\ \text{subject to}\ & g_i(\mathbf{x}) \leq 0,\ i = 1, ..., m \\ & h_j(\mathbf{x}) = 0,\ j = 1, ..., \ell \end{aligned}$$

KKT Conditions

- The necessary conditions for the optimal solution of the nonlinear optimization problem are examined in four main conditions:
  - Stationarity Condition:
    - To minimize $f(\mathbf{x})$:
      $\nabla f(\mathbf{x}^*) - \sum_{i=1}^{m} \mu_i \nabla g_i(\mathbf{x}^*) - \sum_{j=1}^{\ell} \lambda_j \nabla h_j(\mathbf{x}^*) = 0,$
    - To maximize $f(\mathbf{x})$:
      $-\nabla f(\mathbf{x}^*) - \sum_{i=1}^{m} \mu_i \nabla g_i(\mathbf{x}^*) - \sum_{j=1}^{\ell} \lambda_j \nabla h_j(\mathbf{x}^*) = 0,$
  - Primary Feasibility Condition:
    $g_i(\mathbf{x}^*) \leq 0,\ i = 1, ..., m$
    $h_j(\mathbf{x}^*) = 0,\ j = 1, ..., \ell$
  - Dual Feasibility Condition:
    $\mu_i \geq 0,\ i = 1, ..., m$
  - Complementary Slackness Condition:
    $\mu_i g_i(\mathbf{x}^*) = 0,\ i = 1, ..., m.$

KKT Conditions

- Moreover,
  - if $f(\nu \mathbf{x}_1 + (1-\nu)\mathbf{x}_2) \leq \nu f(\mathbf{x}_1) + (1-\nu) f(\mathbf{x}_2)$ holds for any $\mathbf{x}_1 \neq \mathbf{x}_2$ from the domain of the function f with some $\nu \in [0,1]$, i.e. $f(\mathbf{x})$ is a convex function,
  - if $g_i(\nu_i \mathbf{x}_1 + (1-\nu_i)\mathbf{x}_2) \leq \nu_i g_i(\mathbf{x}_1) + (1-\nu_i) g_i(\mathbf{x}_2)$ holds for any $\mathbf{x}_1 \neq \mathbf{x}_2$ from the domain of each g_i function with some $\nu_i \in [0,1]$, i.e. $g_i(\mathbf{x})$'s are convex functions,
  - if h_i's are linear functions,

  then, these conditions are also sufficient conditions.

Example 1

Examine whether the minimization problem given below satisfies the KKT conditions under the given constraints.

$$\begin{aligned} \min\ f(x_1, x_2) &= 4x_1^2 + 2x_2^2 \\ \text{subject to } 3x_1 + x_2 &= 8 \\ 2x_1 + 4x_2 &\leq 15 \end{aligned}$$

Example 1

Examine whether the minimization problem given below satisfies the KKT conditions under the given constraints.

$$\begin{aligned} \min\ f(x_1, x_2) &= 4x_1^2 + 2x_2^2 \\ \text{subject to } 3x_1 + x_2 = 8 \quad &\rightarrow h(x_1, x_2) = 3x_1 + x_2 - 8 = 0, \\ 2x_1 + 4x_2 \leq 15 &\rightarrow g(x_1, x_2) = 2x_1 + 4x_2 - 15 \leq 0. \end{aligned}$$

Examples - Example 1

Examples - Example 1

- Stationarity Condition:

$$\nabla f - \mu \nabla g - \lambda \nabla h = \begin{bmatrix} 8x_1 - 2\mu - 3\lambda \\ 4x_2 - 4\mu - \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Examples - Example 1

- Stationarity Condition:
  $$\nabla f - \mu \nabla g - \lambda \nabla h = \begin{bmatrix} 8x_1 - 2\mu - 3\lambda \\ 4x_2 - 4\mu - \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
- Primal Feasibility Condition:
  $3x_1 + x_2 = 8$
  $2x_1 + 4x_2 \leq 15$

Examples - Example 1

- Stationarity Condition:
  $$\nabla f - \mu \nabla g - \lambda \nabla h = \begin{bmatrix} 8x_1 - 2\mu - 3\lambda \\ 4x_2 - 4\mu - \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
- Primal Feasibility Condition:
  $3x_1 + x_2 = 8$
  $2x_1 + 4x_2 \leq 15$
- Dual Feasibility Condition:
  $\mu \geq 0$

Examples - Example 1

- Stationarity Condition:
$$\nabla f - \mu \nabla g - \lambda \nabla h = \begin{bmatrix} 8x_1 - 2\mu - 3\lambda \\ 4x_2 - 4\mu - \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
- Primal Feasibility Condition:
$3x_1 + x_2 = 8$
$2x_1 + 4x_2 \leq 15$
- Dual Feasibility Condition:
$\mu \geq 0$
- Complementary Slackness Condition:
$\mu(2x_1 + 4x_2 - 15) = 0.$

There will be two cases, based on the dual feasibility condition and the complementary slackness condition:

There will be two cases, based on the dual feasibility condition and the complementary slackness condition:

- $\mu > 0$

$$\left.\begin{aligned} 8x_1 - 2\mu - 3\lambda &= 0 \\ 4x_2 - 3\mu - \lambda &= 0 \\ 3x_1 + x_2 &= 8 \\ 2x_1 + 4x_2 &= 15 \end{aligned}\right\}$$

There will be two cases, based on the dual feasibility condition and the complementary slackness condition:

- $\mu > 0$

$$\left.\begin{aligned} 8x_1 - 2\mu - 3\lambda = 0 \\ 4x_2 - 3\mu - \lambda = 0 \\ 3x_1 + x_2 = 8 \\ 2x_1 + 4x_2 = 15 \end{aligned}\right\} \rightarrow \begin{bmatrix} 8 & 0 & -2 & -3 \\ 0 & 4 & -4 & -1 \\ 3 & 1 & 0 & 0 \\ 2 & 4 & 0 & 0 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \mu \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 8 \\ 15 \end{bmatrix}$$

Examples - Example 1

There will be two cases, based on the dual feasibility condition and the complementary slackness condition:

- $\mu > 0$

$$\left.\begin{aligned} 8x_1 - 2\mu - 3\lambda &= 0 \\ 4x_2 - 3\mu - \lambda &= 0 \\ 3x_1 + x_2 &= 8 \\ 2x_1 + 4x_2 &= 15 \end{aligned}\right\} \rightarrow \begin{bmatrix} 8 & 0 & -2 & -3 \\ 0 & 4 & -4 & -1 \\ 3 & 1 & 0 & 0 \\ 2 & 4 & 0 & 0 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \mu \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 8 \\ 15 \end{bmatrix}$$

- The solution will be $x_1 = \frac{17}{10}$, $x_2 = \frac{29}{10}$, $\mu = \frac{53}{25}$ and $\lambda = \frac{78}{25}$.

Examples - Example 1

There will be two cases, based on the dual feasibility condition and the complementary slackness condition:

- $\mu > 0$

$$\left.\begin{aligned} 8x_1 - 2\mu - 3\lambda = 0 \\ 4x_2 - 3\mu - \lambda = 0 \\ 3x_1 + x_2 = 8 \\ 2x_1 + 4x_2 = 15 \end{aligned}\right\} \rightarrow \begin{bmatrix} 8 & 0 & -2 & -3 \\ 0 & 4 & -4 & -1 \\ 3 & 1 & 0 & 0 \\ 2 & 4 & 0 & 0 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \mu \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 8 \\ 15 \end{bmatrix}$$

- The solution will be $x_1 = \frac{17}{10}$, $x_2 = \frac{29}{10}$, $\mu = \frac{53}{25}$ and $\lambda = \frac{78}{25}$.
- $f(17/10, 29/10) = \frac{1419}{50}$

Examples - Example 1

There will be two cases, based on the dual feasibility condition and the complementary slackness condition:

Examples - Example 1

There will be two cases, based on the dual feasibility condition and the complementary slackness condition:

- $\mu = 0$

$$\left.\begin{aligned} 8x_1 - 3\lambda &= 0 \\ 4x_2 - \lambda &= 0 \\ 3x_1 + x_2 &= 8 \end{aligned}\right\}$$

There will be two cases, based on the dual feasibility condition and the complementary slackness condition:

- $\mu = 0$

$$\left.\begin{aligned} 8x_1 - 3\lambda &= 0 \\ 4x_2 - \lambda &= 0 \\ 3x_1 + x_2 &= 8 \end{aligned}\right\} \rightarrow \begin{bmatrix} 8 & 0 & -3 \\ 0 & 4 & -1 \\ 3 & 1 & 0 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 8 \end{bmatrix}$$

There will be two cases, based on the dual feasibility condition and the complementary slackness condition:

- $\mu = 0$

$$\left.\begin{aligned} 8x_1 - 3\lambda &= 0 \\ 4x_2 - \lambda &= 0 \\ 3x_1 + x_2 &= 8 \end{aligned}\right\} \rightarrow \begin{bmatrix} 8 & 0 & -3 \\ 0 & 4 & -1 \\ 3 & 1 & 0 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 8 \end{bmatrix}$$

- The solution will be $x_1 = \frac{24}{11}$, $x_2 = \frac{16}{11}$, $\mu = 0$ and $\lambda = \frac{64}{11}$.

There will be two cases, based on the dual feasibility condition and the complementary slackness condition:

- $\mu = 0$

$$\left.\begin{aligned} 8x_1 - 3\lambda &= 0 \\ 4x_2 - \lambda &= 0 \\ 3x_1 + x_2 &= 8 \end{aligned}\right\} \rightarrow \begin{bmatrix} 8 & 0 & -3 \\ 0 & 4 & -1 \\ 3 & 1 & 0 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 8 \end{bmatrix}$$

- The solution will be $x_1 = \frac{24}{11}$, $x_2 = \frac{16}{11}$, $\mu = 0$ and $\lambda = \frac{64}{11}$.
- $f(24/11, 16/11) = \frac{256}{11} \rightarrow$ Global Minimum ✓

Examples - Example 1

Here,

Here,

- $f(x_1, x_2) = 4x_1^2 + 2x_2^2$ is a convex function,

Here,

- $f(x_1, x_2) = 4x_1^2 + 2x_2^2$ is a convex function,
- $g(x_1, x_2) = 2x_1 + 4x_2 - 15$ is a convex function,

Here,

- $f(x_1, x_2) = 4x_1^2 + 2x_2^2$ is a convex function,
- $g(x_1, x_2) = 2x_1 + 4x_2 - 15$ is a convex function,
- $h(x_1, x_2) = 3x_1 + x_2 - 8$ is a linear function.

Example 2

Examine whether the following KKT conditions are met for the revenue optimization problem of a company trying to maximize its revenues ($R(Q)$) under a certain minimum profit constraint ($G_{\min} \leq R(Q) - C(Q)$).

$$\begin{aligned} \min\ & f(Q) = -R(Q) \\ \text{subject to}\ & G_{\min} \leq R(Q) - C(Q) \\ & Q \geq 0 \end{aligned}$$

Example 2

Examine whether the following KKT conditions are met for the revenue optimization problem of a company trying to maximize its revenues ($R(Q)$) under a certain minimum profit constraint ($G_{\min} \leq R(Q) - C(Q)$).

$$
\begin{aligned}
&\min \ f(Q) = -R(Q) \\
\text{subject to } &G_{\min} \leq R(Q) - C(Q) \rightarrow g_1(Q) = G_{\min} - R(Q) + C(Q) \leq 0, \\
&Q \geq 0 \qquad\qquad\qquad\qquad\quad \rightarrow g_2(Q) = -Q \leq 0.
\end{aligned}
$$

Examples - Example 2

Examples - Example 2

- Stationarity Condition:
$$-\frac{dR}{dQ} - \mu_1\left[-\frac{dR}{dQ} + \frac{dC}{dQ}\right] - \mu_2(-1) = 0$$

Examples - Example 2

- Stationarity Condition:
$$-\frac{dR}{dQ} - \mu_1\left[-\frac{dR}{dQ} + \frac{dC}{dQ}\right] - \mu_2(-1) = 0$$
$$\rightarrow (\mu_1 - 1)\frac{dR}{dQ} - \mu_1\frac{dC}{dQ} + \mu_2 = 0,$$

Examples - Example 2

- Stationarity Condition:
  $-\frac{dR}{dQ} - \mu_1 \left[-\frac{dR}{dQ} + \frac{dC}{dQ} \right] - \mu_2(-1) = 0$
  $\rightarrow (\mu_1 - 1)\frac{dR}{dQ} - \mu_1 \frac{dC}{dQ} + \mu_2 = 0,$
- Primal Feasibility Condition:
  $G_{\text{min}} - R(Q) + C(Q) \leq 0,$
  $-Q \leq 0.$

Examples - Example 2

- Stationarity Condition:
  $-\frac{dR}{dQ} - \mu_1\left[-\frac{dR}{dQ} + \frac{dC}{dQ}\right] - \mu_2(-1) = 0$
  $\rightarrow (\mu_1 - 1)\frac{dR}{dQ} - \mu_1\frac{dC}{dQ} + \mu_2 = 0,$
- Primal Feasibility Condition:
  $G_{\text{min}} - R(Q) + C(Q) \leq 0,$
  $-Q \leq 0.$
- Dual Feasibility Condition:
  $\mu_1, \mu_2 \geq 0,$

Examples - Example 2

- Stationarity Condition:
$$-\frac{dR}{dQ}-\mu_1\left[-\frac{dR}{dQ}+\frac{dC}{dQ}\right]-\mu_2(-1)=0$$
$$\rightarrow(\mu_1-1)\frac{dR}{dQ}-\mu_1\frac{dC}{dQ}+\mu_2=0,$$
- Primal Feasibility Condition:
$$G_{\text{min}}-R(Q)+C(Q)\le 0,$$
$$-Q\le 0.$$
- Dual Feasibility Condition:
$$\mu_1,\mu_2\ge 0,$$
- Complementary Slackness Condition:
$$\mu_1(G_{\text{min}}-R(Q)+C(Q))=0,$$
$$\mu_2 Q=0.$$

Examples - Example 2

- Stationarity Condition:
  $-\frac{dR}{dQ} - \mu_1 \Big[-\frac{dR}{dQ} + \frac{dC}{dQ} \Big] - \mu_2(-1) = 0$
  $\to (\mu_1 - 1)\frac{dR}{dQ} - \mu_1 \frac{dC}{dQ} + \mu_2 = 0,$
- Primal Feasibility Condition:
  $G_{\text{min}} - R(Q) + C(Q) \leq 0,$
  $-Q \leq 0.$
- Dual Feasibility Condition:
  $\mu_1, \mu_2 \geq 0,$
- Complementary Slackness Condition:
  $\mu_1(G_{\text{min}} - R(Q) + C(Q)) = 0,$
  $\mu_2 Q = 0. \ \to \mu_2 = 0.$

Examples - Example 2

Since $\mu_2 = 0$ is satisfied, we have

Examples - Example 2

Since $\mu_2 = 0$ is satisfied, we have

- $\frac{dC}{dQ} = \frac{\mu_1 - 1}{\mu_1}\frac{dR}{dQ}$,

Examples - Example 2

Since $\mu_2 = 0$ is satisfied, we have

- $\frac{dC}{dQ} = \frac{\mu_1 - 1}{\mu_1} \frac{dR}{dQ}$,
- which indicates that marginal revenue of the company is greater than its marginal costs.