Ch.21 Problems with Inequality Constraints

21.1 Karush-Kuhn-Tucker Condition

- In this chapter, we discuss extremum problems with inequality constraints.
- Problems with inequality constraints can also be treated using Lagrange multipliers.
- We consider the following problem:

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & h(x) = 0, \\ & g(x) \leq 0 \end{aligned}$$

  where $f : \Re^n \to \Re$, $h : \Re^n \to \Re^m$, $m \leq n$, and $g : \Re^n \to \Re^p$.

- **Def. 21.1** *An inequality constraint $g_j(x) \leq 0$ is said to be active at x^* if $g_j(x^*) = 0$. It is inactive at x^* if $g_j(x^*) < 0$.*
- **Def. 21.2** *Let x^* satisfy $h(x^*) = 0$, $g(x^*) \leq 0$, and let $J(x^*)$ be the index set of active inequality constraints, that is,*

$$J(x^*) \equiv \{j : g_j(x^*) = 0\}$$

Then, we say that x^* *is a regular point if the vectors*

$$\nabla h_i(x^*),\ \nabla g_j(x^*),\ 1 \le i \le m,\ j \in J(x^*)$$

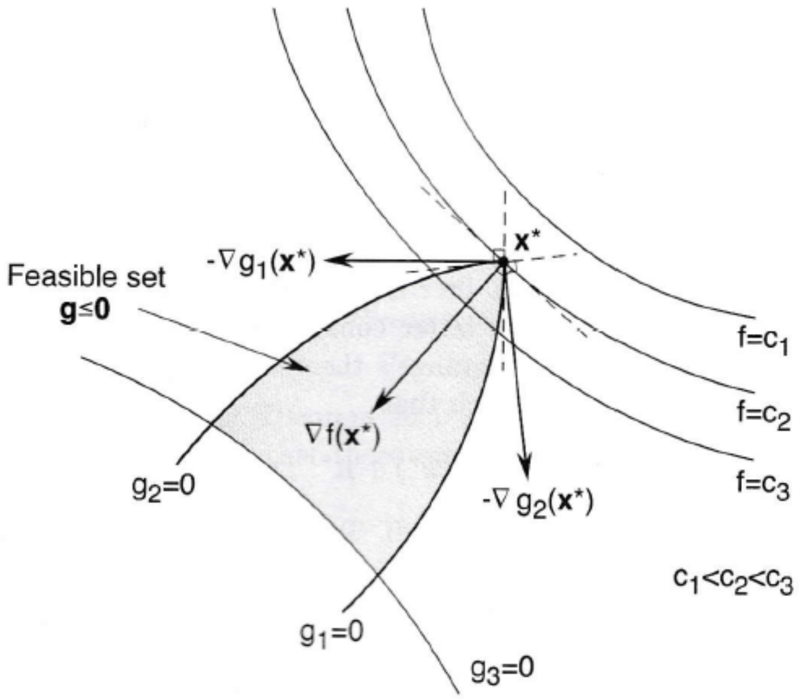
are linearly independent.

- **Theorem 21.1 Karush-Kuhn-Tucker (KKT) Theorem** *Let* $f, h, g \in C^1$. *Let* x^* *be a regular point and a local minimizer for the problem of minimizing* f *subject to* $h(x) = 0$, $g(x) \le 0$. *Then, there exist* $\lambda^* \in \Re^m$ *and* $\mu^* \in \Re^p$ *such that*
  1. $\mu^* \ge 0$
  2. $Df(x^*) + \lambda^{*T} Dh(x^*) + \mu^{*T} Dg(x^*) = 0^T$
  3. $\mu^{*T} g(x^*) = 0$
  4. $h(x^*) = 0$
  5. $g(x^*) \le 0$
- We refer to λ^* as the Lagrange multiplier vector, and μ^* as the Karush-Kuhn-Tucker (KKT) multiplier.
- The condition

$$\mu^{*T} g(x^*) = \mu_1^* g_1(x^*) + \cdots + \mu_p^* g_p(x^*) = 0$$

  implies that if $g_j(x^*) < 0$, then $\mu_j^* = 0$.
- KKT multipliers μ_j^* corresponding to inactive constraints are zero.
- **Ex. 21.1** A graphical illustration of KKT theorem

♣ Fig. 21.1 on page 399.

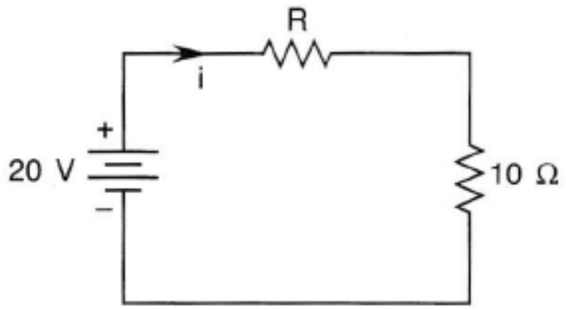


We have only inequality constraints $g_j(x) \leq 0$, $j = 1, 2, 3$. The constraint $g_3(x) \leq 0$ is inactive, that is, $g_3(x^*) < 0$; hence $\mu_3^* = 0$. By the KKT theorem,

$$\nabla f(x^*) + \mu_1^* \nabla g_1(x^*) + \mu_2^* \nabla g_2(x^*) = 0$$
$$\nabla f(x^*) = -\mu_1^* \nabla g_1(x^*) - \mu_2^* \nabla g_2(x^*)$$

where $\mu_1^* > 0$ and $\mu_2^* > 0$.

- $\nabla f(x^*)$ must be a linear combination of the vectors $-\nabla g_1(x^*)$ and $-\nabla g_2(x^*)$ with positive coefficients.
- **Ex. 21.2** Consider the circuit of Fig. 21.2.
  ♣ Fig. 21.2 on page 401.



1. Find the value of the resistor $R \geq 0$ such that the power absorbed by this resistor is maximized.
2. Find the value of the resistor $R \geq 0$ such that the power delivered to the 10Ω resistor is maximized.

Sol.

1. The power absorbed by the resistor R is $p = i^2 R$, where $i = \frac{20}{10+R}$. The optimization problem can be represented as

$$\begin{aligned} \text{minimize} \quad & -\frac{400R}{(10+R)^2} \\ \text{subject to} \quad & -R \leq 0 \end{aligned}$$

The KKT condition is

$$\begin{aligned} -\frac{400(10-R)}{(10+R)^3} - \mu &= 0 \\ \mu &\geq 0 \\ \mu R &= 0 \\ -R &\leq 0 \end{aligned}$$

From the above conditions

(a) $\mu > 0 \rightarrow R = 0 \rightarrow -\frac{400(10-R)}{(10+R)^3} - \mu \neq 0.$

(b) $\mu = 0 \rightarrow R \geq 0 \rightarrow -\frac{400(10-R)}{(10+R)^3} = 0 \rightarrow R = 10.$

2. The power absorbed by the 10Ω resistor is $p = i^2 10$, where $i = 20/(10+R)$. The optimization problem can be represented as

$$\begin{array}{rl} \text{minimize} & -\dfrac{4000}{(10+R)^2} \\ \text{subject to} & -R \leq 0 \end{array}$$

The KKT condition is

$$\begin{aligned} \frac{8000}{(10+R)^3} - \mu &= 0 \\ \mu &\geq 0 \\ \mu R &= 0 \\ -R &\leq 0 \end{aligned}$$

(a) $\mu > 0 \rightarrow R = 0 \rightarrow \frac{8000}{10^3} = 8 = \mu$

(b) $\mu = 0 \rightarrow R \geq 0 \rightarrow \frac{8000}{(10+R)^3} \neq 0$

- When the objective function is to be maximized,

$$\begin{array}{rl} \text{maximize} & f(x) \\ \text{subject to} & h(x) = 0 \\ & g(x) \leq 0 \end{array}$$

The KKT condition is;

1. $\mu^* \leq 0$

2. $Df(x^*) + \lambda^{*T} Dh(x^*) + \mu^{*T} Dg(x^*) = 0^T$
3. $\mu^{*T} g(x^*) = 0$
4. $h(x^*) = 0$
5. $g(x^*) \leq 0$

- When the inequality constraint is of the form $g(x) \geq 0$;

$$\begin{array}{rl} \text{minimize} & f(x) \\ \text{subject to} & h(x) = 0 \\ & g(x) \geq 0 \end{array}$$

The KKT condition is;

1. $\mu^* \leq 0$
2. $Df(x^*) + \lambda^{*T} Dh(x^*) + \mu^{*T} Dg(x^*) = 0^T$
3. $\mu^{*T} g(x^*) = 0$
4. $h(x^*) = 0$
5. $g(x^*) \geq 0$

- For the problem;

$$\begin{array}{rl} \text{maximize} & f(x) \\ \text{subject to} & h(x) = 0 \\ & g(x) \geq 0 \end{array}$$

KKT condition is exactly equal to conditions in Theorem 21.1.

- **Ex. 21.4** Consider the problem

$$\begin{aligned} \text{minimize} \quad & f(x_1, x_2) \\ \text{subject to} \quad & x_1, x_2 \geq 0 \\ f(x_1, x_2) \quad = \quad & x_1^2 + x_2^2 + x_1 x_2 - 3x_1 \end{aligned}$$

The KKT condition for this problem is

1. $\mu_1 \leq 0$, $\mu_2 \leq 0$.
2. $Df(x) + \mu^T = 0^T$

$$\begin{aligned} 2x_1 + x_2 - 3 + \mu_1 &= 0 \\ 2x_2 + x_1 + \mu_2 &= 0 \end{aligned}$$

3. $\mu_1 x_1 + \mu_2 x_2 = 0$.
4. $x_1 \geq 0$, $x_2 \geq 0$.

To find a solution (x^*, μ^*), we first try as $\mu_1^* = 0$, $x_2^* = 0$. Then, $x_1^* = 3/2$, $\mu_2^* = -3/2$. The above satisfies all the KKT and feasibility conditions.
However, there is no guarantee that the point is a minimizer. KKT condition is only necessary.

21.2 Second-Order Conditions

- Second-order necessary and sufficient conditions for extremum problems involving inequality constraints.

$$L(x, \lambda, \mu) = F(x) + [\lambda H(x)] + [\mu G(x)]$$

where $F(x)$ is the Hessian matrix of f at x, and

$[\lambda H(x)]$ and $[\mu G(x)]$ represents

$$\begin{aligned}[\lambda H(x)] &= \lambda_1 H_1(x) + \cdots + \lambda_m H_m(x)\\ [\mu G(x)] &= \mu_1 G_1(x) + \cdots + \mu_p G_p(x)\end{aligned}$$

where $G_k(x)$ is the Hessian of g_k at x,

$$G_k(x) = \begin{bmatrix} \frac{\partial^2 g_k}{\partial x_1^2}(x) & \cdots & \frac{\partial^2 g_k}{\partial x_n \partial x_1}(x) \\ \vdots & \vdots & \vdots \\ \frac{\partial^2 g_k}{\partial x_1 \partial x_n}(x) & \cdots & \frac{\partial^2 g_k}{\partial^2 x_n}(x) \end{bmatrix}$$

- The tangent space to the surface defined by active constraints.

  $T(x^*) = \{y \in \Re^n : Dh(x^*)y = 0,\ Dg_j(x^*)y = 0,\ j \in J(x^*)\}$

- **Theorem 21.2 Second-Order Necessary Conditions** *Let x^* be a local minimizer of $f : \Re^n \to \Re$ subject to $h(x) = 0$, $g(x) \leq 0$, $h : \Re^n \to \Re^m$, $m \leq n$, $g : \Re^n \to \Re^p$, and $f, h, g \in C^2$. Suppose x^* is regular. Then, there exist $\lambda^* \in \Re^m$ and $\mu^* \in \Re^p$ such that;*
  1. *$\mu^* \geq 0,\ Df(x^*) + \lambda^{*T} Dh(x^*) + \mu^{*T} Dg(x^*) = 0^T,\ \mu^{*T} g(x^*) = 0$*
  2. *For all $y \in T(x^*)$ we have $y^T L(x^*, \lambda^*, u^*) y \geq 0$.*

- Second-Order sufficient conditions for extremum problems involving inequality constraints.

  $\tilde{T}(x^*, \mu^*) = \{y : Dh(x^*)y = 0,\ Dg_i(x^*)y = 0, i \in \tilde{J}(x^*, \mu^*)\}$

where $\tilde{J}(x^*, \mu^*) = \{i : g_i(x^*) = 0, \mu_i^* > 0\}$

- **Theorem 21.3 Second-Order Sufficient Conditions** *Suppose $f, g, h \in C^2$ and there exist a feasible point $x^* \in \Re^n$ and vectors $\lambda^* \in \Re^m$ and $\mu^* \in \Re^p$, such that:*
  1. $\mu^* \geq 0, Df(x^*) + \lambda^{*T} Dh(x^*) + \mu^{*T} Dg(x^*) = 0^T, \ \mu^{*T} g(x^*) = 0.$
  2. *For all $y \in \tilde{T}(x^*, \mu^*), y \neq 0$, we have* $y^T L(x^*, \lambda^*, \mu^*) y > 0.$

  Then, x^* is a strict local minimizer of f subject to $h(x) = 0, g(x) \leq 0$.

- **Ex. 21.6** We wish to minimize $f(x) = (x_1 - 1)^2 + x_2 - 2$ subject to

$$\begin{aligned} h(x) &= x_2 - x_1 - 1 = 0 \\ g(x) &= x_1 + x_2 - 2 \leq 0 \\ Dh(x) &= [-1, 1], \ Dg(x) = [1, 1], \ Df(x) = [2x_1 - 2, 1] \end{aligned}$$

The KKT condition is;

$$\begin{aligned} Df(x) + \lambda Dh(x) + \mu Dg(x) = & \\ [2x_1 - 2 - \lambda + \mu, 1 + \lambda + \mu] &= 0^T \\ \mu(x_1 + x_2 - 2) &= 0 \\ \mu &\geq 0 \\ x_2 - x_1 - 1 &= 0 \end{aligned}$$

$$x_1 + x_2 - 2 \le 0$$

We first try $\mu > 0$,

$$\begin{aligned} 2x_1 - 2 - \lambda + \mu &= 0 \\ 1 + \lambda + \mu &= 0 \\ x_2 - x_1 - 1 &= 0 \\ x_1 + x_2 - 2 &= 0 \end{aligned}$$

$$x_1 = \frac{1}{2},\ x_2 = \frac{3}{2},\ \lambda = -1,\ \mu = 0$$

This contradicts the assumption that $\mu > 0$. In the second try, we assume $\mu = 0$.

$$\begin{aligned} 2x_1 - 2 - \lambda &= 0 \\ 1 + \lambda &= 0 \\ x_2 - x_1 - 1 &= 0 \end{aligned}$$

$$g(x_1, x_2) = x_1 + x_2 - 2 \le 0$$

$$x_1 = \frac{1}{2},\ x_2 = \frac{3}{2},\ \lambda = -1.$$

The point x^* satisfying the KKT necessary condition is therefore the candidate for being a minimizer.
The second-order sufficient conditions,

$$\begin{aligned} L(x^*, \lambda^*, \mu^*) &= F(x^*) + \lambda^* H(x^*) + \mu^* G(x^*) \\ &= \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} + (-1) \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \end{aligned}$$

$$+(0)\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$$

Because $\mu^* = 0$, the active constraint $g(x^*) = 0$ does not enter the computation of $T(x^*, \mu^*)$.

$$\begin{aligned} \tilde{T}(x^*, \mu^*) &= \{y : Dh(x^*)y = 0\} \\ &= \{y : [-1, 1]y = 0\} = \{[a, a]^T : a \in \Re\} \end{aligned}$$

For positive definiteness of $L(x^*, \lambda^*, \mu^*)$ on $\tilde{T}(x^*, \mu^*)$.

$$y^T L(x^*, \lambda^*, \mu^*) y = [a,\ a] \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ a \end{bmatrix} = 2a^2$$

By the second-order sufficient conditions, we conclude that $x^* = [1/2, 3/2]^T$ is a strict local minimizer.