Ch.5 Constrained Optimality Criteria (2)

-Review of the book "Engineering Optimization"

Yi Zhang
Networks Lab
UC Davis
July 2, 2010

Outline

- 5.6 Saddlepoint Conditions
- 5.7 Second-Order Optimality Conditions
- 5.8 Generalized Lagrange Multiplier Method
- 5.9 Generalization of Convex Functions
- 5.10 Summary

5.6 SADDLEPOINT CONDITIONS

The discussion of Kuhn–Tucker optimality conditions of Sections 5.4 and 5.5 assume that the objective function and the constraints are differentiable. We now discuss constrained optimality criteria for nondifferentiable functions.

Definition

A function $f(x, y)$ is said to have a *saddlepoint* at (x^*, y^*) if $f(x^*, y) \leqslant f(x^*, y^*) \leqslant f(x, y^*)$ for all x and y.

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

saddlepoint at the point $x^* = 2,\ y^* = 4$.

$$f(2, y) \leqslant f(2, 4) \leqslant f(x, 4) \qquad \text{for all } y \geqslant 0 \text{ and all real } x$$

Consider the general NLP problem:

$$\begin{aligned} &\text{Minimize} \quad && f(x) \\ &\text{Subject to} \quad && g_j(x) \geq 0 \qquad \text{for } j = 1, \ldots, J \\ & && x \in S \end{aligned}$$

The *Kuhn–Tucker saddlepoint problem* (KTSP) is as follows: Find (x^*, u^*) such that

$$L(x^*, u) \leq L(x^*, u^*) \leq L(x, u^*)$$

$$\text{all } u \geq 0 \qquad \text{and} \qquad \text{all } x \in S$$

where

$$L(x, u) = f(x) - \sum_j u_j g_j(x)$$

Theorem 5.3 Sufficient Optimality Theorem

If (x^*, u^*) is a saddlepoint solution of a KTSP, the x^* is an optimal solution to the NLP problem.

A proof of this theorem is available in Mangasarian [2, Chap. 3].

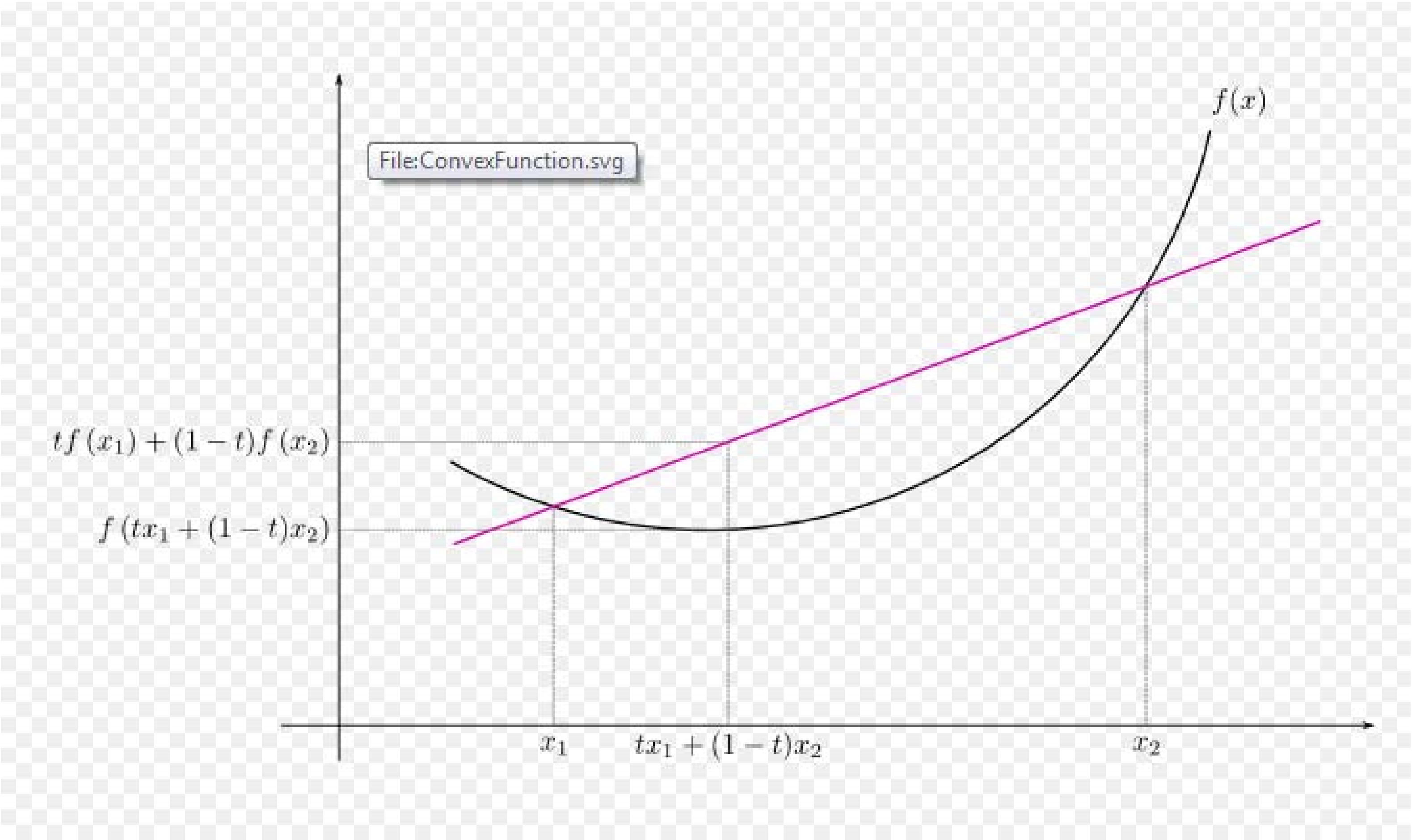
Remarks

1. No convexity assumptions of the functions have been made in Theorem 5.3.
2. No constraint qualification is invoked.
3. Nonlinear equality constraints of the form $h_k(x) = 0$ for $k = 1, \ldots, K$ can be handled easily by redefining the Lagrangian function as

$$L(x, u, v) = f(x) - \sum_j u_j g_j(x) - \sum_k v_k h_k(x)$$

   Here the variables v_k for $k = 1, \ldots, K$ will be unrestricted in sign.
4. Theorem 5.3 provides only a sufficient condition. There may exist some NLP problems for which a saddlepoint does not exist even though the NLP problem has an optimal solution.

Existence of Saddlepoints. There exist necessary optimality theorems that guarantee the existence of a saddlepoint solution without the assumption of differentiability. However, they assume that the constraint qualification is met and that the functions are convex.


File:ConvexFunction.svg
$f(x)$
$tf(x_1) + (1-t)f(x_2)$
$f(tx_1 + (1-t)x_2)$
x_1
$tx_1 + (1-t)x_2$
x_2

Theorem 5.4 Necessary Optimality Theorem

Let x^* minimize $f(x)$ subject to $g_j(x) \geqslant 0$, $j = 1, \ldots, J$ and $x \in S$. Assume S is a convex set, $f(x)$ is a convex function, and $g_j(x)$ are concave functions on S. Assume also that there exists a point $\bar{x} \in S$ such that $g_j(\bar{x}) > 0$ for all $j = 1, 2, \ldots, J$. then there exists a vector of multipliers $u^* \geqslant 0$ such that (x^*, u^*) is a saddlepoint of the Lagrangian function

$$L(x, u) = f(x) - \sum_j u_j g_j(x)$$

satisfying

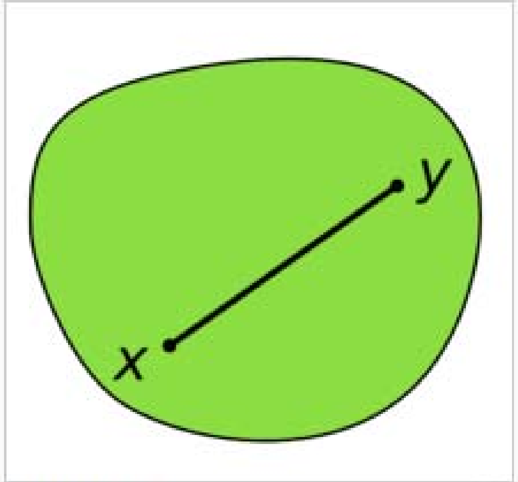
$$L(x^*, u) \leqslant L(x^*, u^*) \leqslant L(x, u^*)$$

for all $x \in S$ and $u \geqslant 0$.

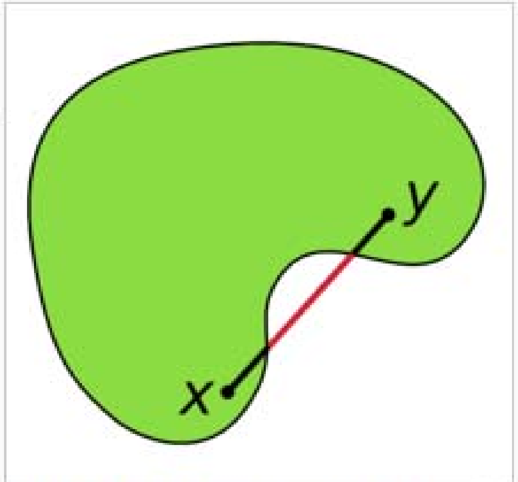
For a proof of this theorem, refer to the text by Lasdon [3, Chap. 1].

Even though Theorem 5.3 and the KTSP provide sufficient conditions for optimality without invoking differentiability and convexity, determination of a saddlepoint to a KTSP is generally difficult. However, the following theorem makes it computationally more attractive.

How to find a saddlepoint?



A convex set.



A non-convex set, with a line-segment outside the set.

Theorem 5.5

A solution (x^*, u^*) with $u^* \geqslant 0$ and $x^* \in S$ is a saddlepoint of a KTSP if and only if the following conditions are satisfied:

(i) x^* minimizes $L(x, u^*)$ over all $x \in S$

(ii) $g_j(x^*) \geqslant 0$ for $j = 1, \ldots, J$

(iii) $u_j g_j(x^*) = 0$ for $j = 1, \ldots, J$

For a proof, see Lasdon [3, Chap. 1].

5.7 SECOND-ORDER OPTIMALITY CONDITIONS

In Sections 5.4–5.6, we discussed the first-order necessary and sufficient conditions, called the Kuhn–Tucker conditions, for constrained optimization problems using the gradients of the objective function and constraints. Second-order necessary and sufficient optimality conditions that apply to twice-differentiable functions have been developed by McCormick [5], whose main results are summarized in this section. Consider the following NLP problem.

Problem P1

$$
\begin{aligned}
&\text{Minimize} && f(x) && \\
&\text{Subject to} && g_j(x) \geqslant 0 && j = 1, 2, \ldots, J \\
& && h_k(x) = 0 && k = 1, 2, \ldots, K \\
& && x \in R^N &&
\end{aligned}
$$

The first-order KTCs are given by

$$\nabla f(x) - \sum_j u_j \nabla g_j(x) - \sum v_k \nabla h_k(x) = 0 \qquad (5.40)$$

$$g_j(x) \geqslant 0 \qquad j = 1, \ldots, J \qquad (5.41)$$

$$h_k(x) = 0 \qquad k = 1, \ldots, K \qquad (5.42)$$

$$u_j g_j(x) = 0 \qquad j = 1, \ldots, J \qquad (5.43)$$

$$u_j \geqslant 0 \qquad j = 1, \ldots, J \qquad (5.44)$$

Definitions

$\bar{x}$ is a *feasible solution* to an NLP problem when $g_j(\bar{x}) \geqslant 0$ for all j and $h_k(\bar{x}) = 0$ for all k.

x^* is a *local minimum* to an NLP problem when x^* is feasible and $f(x^*) \leqslant f(\bar{x})$ for all feasible $\bar{x}$ in some small neighborhood $\delta(x^*)$ of x^*.

x^* is a *strict* (*unique* or *isolated*) *local minimum* when x^* is feasible and $f(x^*) < f(\bar{x})$ for feasible $\bar{x} \neq x^*$ in some small neighborhood $\delta(x^*)$ of x^*.

Let us first consider the basic motivation for the second-order optimality conditions. For simplicity, consider an equality-constrained NLP problem as follows:

$$\begin{aligned} &\text{Minimize} \quad f(x) \\ &\text{Subject to} \quad h_k(x) = 0 \qquad k = 1, 2, \ldots, K \end{aligned}$$

The first-order KTCs are given by $h_k(x) = 0$, $k = 1, \ldots, k$,

$$\nabla f(x) - \sum_k v_k \nabla h_k(x) = 0 \tag{5.45}$$

Consider a point $\bar{x}$ that satisfied the first-order conditions. To check further whether it is a local minimum, we can write down the Taylor series expansion at the point $\bar{x}$ using higher order terms for each function f and h_k as follows:

$$\begin{aligned} \Delta \cancel{\nabla} f(\bar{x}) &= f(\bar{x} + \Delta x) - f(\bar{x}) \\ &= \nabla f(\bar{x})\, \Delta x + \tfrac{1}{2}\, \Delta x^{\mathrm{T}}\, \mathbf{H}_f\, \Delta x + O(\Delta x) \end{aligned} \tag{5.46}$$

where $O(\Delta x)$ are very small higher order terms involving Δx.

$$\begin{aligned} \Delta h_k(\bar{x}) &= h_k(\bar{x} + \Delta x) - h_k(\bar{x}) \\ &= \nabla h_k(\bar{x})\, \Delta x + \tfrac{1}{2}\, \Delta x^{\mathrm{T}}\, \mathbf{H}_k\, \Delta x + O(\Delta x) \end{aligned} \tag{5.47}$$

where $\mathbf{H}_k$ is the Hessian matrix of $h_k(x)$ evaluated at $\bar{x}$. Multiply Eq. (5.47) by the Kuhn–Tucker multiplier v_k, and sum over all $k = 1, \ldots, K$. Subtracting this sum from Eq. (5.46), we obtain

$$\Delta f(\bar{x}) - \sum_k v_k \, \Delta h_k(\bar{x}) = \left[\nabla f(\bar{x}) - \sum_k v_k \, \nabla h_k(\bar{x}) \right] \Delta x$$
$$+ \tfrac{1}{2} \, \Delta x^{\mathrm{T}} \left[\mathbf{H}_f - \sum_k v_k \mathbf{H}_k \right] \Delta x + O(\Delta x) \tag{5.48}$$

For $\bar{x} + \Delta x$ to be feasible,

$$\Delta h_k(\bar{x}) = 0 \tag{5.49}$$

Assuming that the constraint qualification is satisfied at $\bar{x}$, the Kuhn–Tucker necessary theorem implies that

$$\nabla f(\bar{x}) - \sum_k v_k \, \nabla h_k(\bar{x}) = 0 \tag{5.50}$$

Using Eqs. (5.49) and (5.50), Eq. (5.48) reduces to

$$\Delta f(\bar{x}) = \tfrac{1}{2} \, \Delta x^{\mathrm{T}} \left[\mathbf{H}_f - \sum_k v_k \mathbf{H}_k \right] \Delta x + O(\Delta x) \tag{5.51}$$

For $\bar{x}$ to be a local minimum, it is *necessary* that $\Delta f(\bar{x}) \geqslant 0$ for all feasible movement Δx around $\bar{x}$. Using Eqs. (5.49) and (5.51), the above condition implies that

$$\Delta x^{\mathrm{T}}\left[\mathbf{H}_f - \sum_k v_k \mathbf{H}_k\right] \Delta x > 0 \tag{5.52}$$

for all Δx satisfying

$$\Delta h_k(\bar{x}) = 0 \qquad \text{for } k = 1, \ldots, K \tag{5.53}$$

Using Eq. (5.47) and ignoring the second and higher order terms in Δx, Eq. (5.53) reduces to

$$\Delta h_k(\bar{x}) = \nabla h_k(\bar{x})\,\Delta x = 0$$

Thus assuming that the constraint qualification is satisfied at $\bar{x}$, the necessary conditions for $\bar{x}$ to be a local minimum are as follows:

1. There exists v_k, $k = 1, \ldots, K$, such that $(\bar{x}, v)$ is a Kuhn–Tucker point.
2. $\Delta x^{\mathrm{T}}[\mathbf{H}_f - \sum_k v_k \mathbf{H}_k]\,\Delta x \geqslant 0$ for all Δx satisfying

$$\nabla h_k(\bar{x})\,\Delta x = 0 \qquad \text{for } k = 1, \ldots, K$$

Similarly, the sufficient condition for $\bar{x}$ to be strict local minimum is given by

$$\Delta f(\bar{x}) > 0 \qquad \text{for all feasible } \Delta x \text{ around } \bar{x}$$

This implies that

$$\Delta x^{\mathrm{T}} \left[\mathbf{H}_f - \sum_k v_k \mathbf{H}_k \right] \Delta x > 0$$

for all Δx satisfying

$$\nabla h_k(\bar{x})\, \Delta x = 0 \qquad \text{for all } k = 1, \ldots, K \tag{5.54}$$

We shall now present the formal statements of second-order necessary and sufficient conditions for an NLP problem involving both equality and inequality constraints.

Theorem 5.6   Second-Order Necessity Theorem

Consider the NLP problem given by Problem P1. Let f, g, and h be twice-differentiable functions, and let x^* be feasible for the nonlinear program. Let the active constraint set at x^* be $I = \{j | g_j(x^*) = 0\}$. Furthermore, assume that $\nabla g_j(x^*)$ for $j \in I$ and $\nabla h_k(x^*)$ for $k = 1, 2, \ldots, K$ are linearly independent. Then the *necessary conditions* that x^* be a *local minimum* to the NLP problem are as follows:

1. There exists (u^*, v^*) such that (x^*, u^*, v^*) is a Kuhn–Tucker point.
2. For every vector $y_{(1 \times N)}$ satisfying

$$\nabla g_j(x^*)y = 0 \qquad \text{for } j \in I \tag{5.55}$$

$$\nabla h_k(x^*)y = 0 \qquad \text{for } k = 1, 2, \ldots, K \tag{5.56}$$

it follows that

$$y^{\mathrm{T}} \mathbf{H}_L(x^*, u^*, v^*) y \geqslant 0 \tag{5.57}$$

where

$$L(x, u, v) = f(x) - \sum_{j=1}^{J} u_j g_j(x) - \sum_{k=1}^{K} v_k h_k(x)$$

and $\mathbf{H}_L(x^*, u^*, v^*)$ is the Hessian matrix of the second partial derivatives of L with respect to x evaluated at (x^*, u^*, v^*).

Example 5.7 [5]

$$\text{Minimize} \quad f(x) = (x_1 - 1)^2 + x_2^2$$
$$\text{Subject to} \quad g_1(x) = -x_1 + x_2^2 \geqslant 0$$

Suppose we want to verify whether $x^* = (0, 0)$ is optimal.

Solution

$$\nabla f(x) = [2(x_1 - 1), 2x_2]$$

$$\nabla g_1(x) = (-1, 2x_2) \qquad I = \{1\}$$

Since $\nabla g_1(x^*) = (-1, 0)$ is linearly independent, the constraint qualification is satisfied at x^*. The first-order KTCs are given by

$$2(x_1 - 1) + u_1 = 0$$

$$2x_2 - 2x_2 u_1 = 0$$

$$u_1(-x_1 + x_2^2) = 0$$

$$u_1 \geqslant 0$$

Here $x^* = (0, 0)$ and $u_1^* = 2$ satisfy the above conditions. Hence, $(x^*, u^*) = (0, 0, 2)$ is Kuhn–Tucker point and x^* satisfies the first-order necessary conditions of optimality by Theorem 5.1. In other words, we do not know whether or not (0, 0) is an optimal solution to the NLP problem!

Let us now apply the second-order necessary conditions to test whether (0, 0) is a local minimum to the NLP problem. The first part of Theorem 5.6 is already satisfied, since $(x^*, u^*) = (0, 0, 2)$ is a Kuhn–Tucker point. To prove the second-order conditions, compute

$$\mathbf{H}_L(x, u) = \begin{bmatrix} 2 & 0 \\ 0 & 2 - 2u_1 \end{bmatrix}$$

At (x^*, u^*),

$$\mathbf{H}_L(x^*, u^*) = \begin{bmatrix} 2 & 0 \\ 0 & -4 \end{bmatrix}$$

We have to verify whether

$$y^{\mathrm{T}} \begin{bmatrix} 2 & 0 \\ 0 & -4 \end{bmatrix} y \geqslant 0$$

for all y satisfying

$$\nabla g_1(x^*) y = 0 \qquad \text{or} \qquad (-1, 0) \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = 0$$

In other words, we need to consider only vectors (y_1, y_2) of the form $(0, y_2)$ to satisfy Eq. (5.57). Now,

$$(0,\ y_2)\begin{bmatrix} 2 & 0 \\ 0 & -4 \end{bmatrix}\begin{bmatrix} 0 \\ y_2 \end{bmatrix} = -4y_2^2 < 0 \qquad \text{for all } y_2 \neq 0$$

Thus, $x^* = (0, 0)$ does not satisfy the second-order necessary conditions, and hence its is *not* a local minimum for the NLP problem.

Theorem 5.7 Second-Order Sufficiency Theorem

Sufficient conditions that a point x^* is a *strict local minimum* of the NLP problem P1, where f, g_j, and h_k are twice-differentiable functions, are as follows:

(i) There exists (u^*, v^*) such that (x^*, u^*, v^*) is a Kuhn–Tucker point.

(ii) For every nonzero vector $y_{(1\times N)}$ satisfying

$$\nabla g_j(x^*)y = 0 \qquad j \in I_1 = \{j | g_j(x^*) = 0,\ u_j^* > 0\} \tag{5.58}$$

$$\nabla g_j(x^*)y \geqslant 0 \qquad j \in I_2 = \{j | g_j(x^*) = 0,\ u_j^* = 0\} \tag{5.59}$$

$$\nabla h_k(x^*)y = 0 \qquad k = 1, 2, \ldots, K \tag{5.60}$$

$$y \neq 0$$

it follows that

$$y^{\mathrm{T}}\mathbf{H}_L(x^*, u^*, v^*)y > 0 \tag{5.61}$$

Note: $I_1 \cup I_2 = I$, the set of all active constraints at x^*.

5.8 GENERALIZED LAGRANGE MULTIPLIER METHOD

The usefulness of the Lagrange multiplier method for solving constrained optimization problems is not limited to differentiable functions. Many engineering problems may involve discontinuous or nondifferentiable functions to be optimized. Everett [4] generalized the Lagrange multiplier method presented earlier to handle such problems.

Consider the NLP problem

$$\begin{array}{lll} \text{Minimize} & f(x) & \\ \text{Subject to} & \underline{g_j(x) \geqslant b_j} & \text{for } j = 1, 2, \ldots, J \\ & x \in S & \end{array}$$

where S is a subset of R^N, imposing additional restrictions on the variables x_j (e.g., <u>S may be a discrete set</u>).

Everett's generalized Lagrangian function corresponding to the NLP problem is given by

$$E(x; \lambda) = f(x) - \sum_{j=1}^{J} \lambda_j g_j(x) \tag{5.62}$$

where the λ_j's are nonnegative multipliers.

Suppose the unconstrained minimum of $E(x; \lambda)$ over all $x \in S$ is attained at the point $\bar{x}$ for a fixed value of $\lambda = \bar{\lambda}$. Then, Everett [4] proved that $\bar{x}$ is an optimal solution to the following mathematical program:

$$\begin{aligned} &\text{Minimize} && f(x) \\ &\text{Subject to} && g_j(x) \geqslant g_j(\bar{x}) \qquad j = 1, \ldots, J \\ & && x \in S \end{aligned}$$

Hence, to solve the original NLP problem, it is sufficient to find nonnegative multipliers λ^* (called Everett's multipliers) such that the unconstrained minimum of $E(x; \lambda^*)$ over all $x \in S$ occurs at the point x^* such that

$$g_j(x^*) = b_j \qquad \text{for } j = 1, \ldots, J \tag{5.63}$$

We call this *Everett's condition.*

Inequality=> Equality

Theorem 5.8

Let $\lambda^{(1)}$ and $\lambda^{(2)}$ be nonnegative vectors such that

$$\lambda_i^{(2)} > \lambda_i^{(1)} \qquad \text{and} \qquad \lambda_j^{(1)} = \lambda_j^{(2)} \qquad \text{for all } j \neq i$$

If $\bar{x}^{(1)}$ and $\bar{x}^{(2)}$ minimize $E(x; \lambda)$ given by Eq. (5.62), then

$$g_i[\bar{x}^{(1)}] \geqslant g_i[\bar{x}^{(2)}]$$

$$E(x; \lambda) = f(x) - \sum_{j=1}^{J} \lambda_j g_j(x) \tag{5.62}$$

Example 5.9

$$\begin{aligned}
&\text{Minimize} \quad f(x) = x_1^2 + x_2^2 \\
&\text{Subject to} \quad g_1(x) = 2x_1 + x_2 \geqslant 2
\end{aligned}$$

Everett's function is given by

$$E(x;\, \lambda) = x_1^2 + x_2^2 - \lambda(2x_1 + x_2)$$

We begin Everett's method with $\lambda = \lambda_1 = 0$. The unconstrained minimum of $E(x;\, 0)$ occurs at the point $\overline{x}^{(1)} = (0, 0)$. Since $g_1(\overline{x}^{(1)}) = 0$, which is less than 2, we increase λ to increase $g_1(x)$. Choose $\lambda = \lambda_2 = 1$. The unconstrained minimum of $E(x;\, 1)$ occurs at the point $\overline{x}^{(2)} = (1, \frac{1}{2})$, $g(\overline{x}^{(2)}) = 2 + \frac{1}{2} > 2$. Hence, λ has to be decreased to get a solution that reduces the constraint value. The remaining steps are shown in Table 5.1. Note that the value of λ in each step is simply the midpoint of the two previous λ's, since we know that the optimal λ is between 0 and 1. The convergence is achieved at step 8.

Table 5.1 Everett's Method for Example 5.9

Step t	λ_t	$\bar{x}^{(t)} = (\bar{x}_1^{(t)}, \bar{x}_2^{(t)})$	$g_1(\bar{x}^{(t)})$	Constraint Violation
1	0	$\bar{x}^{(1)} = (0, 0)$	0	<2
2	1	$\bar{x}^{(2)} = (1, 0.5)$	2.5	>2
3	0.5	$\bar{x}^{(3)} = (0.5, 0.25)$	1.25	<2
4	0.75	$\bar{x}^{(4)} = (0.75, 0.375)$	1.875	<2
5	0.88	$\bar{x}^{(5)} = (0.88, 0.44)$	2.2	>2
6	0.82	$\bar{x}^{(6)} = (0.82, 0.41)$	2.05	>2
7	0.78	$\bar{x}^{(7)} = (0.78, 0.39)$	1.95	<2
8	0.8	$\bar{x}^{(8)} = (0.8, 0.4)$	2	2

Bipartite method

5.9 GENERALIZATION OF CONVEX FUNCTIONS

Definition: Pseudoconvex Function

A differentiable function $f(x)$ defined on an open convex set S is *pseudoconvex* on S if and only if for all $x^{(1)}, x^{(2)} \in S$

$$\nabla f(x^{(1)})(x^{(2)} - x^{(1)}) \geqslant 0 \Rightarrow f(x^{(2)}) \geqslant f(x^{(1)})$$

Remarks

1. $f(x)$ is pseudoconcave if $-f(x)$ is pseudoconvex.
2. Every convex function is also pseudoconvex, but a pseudoconvex function may not be convex.

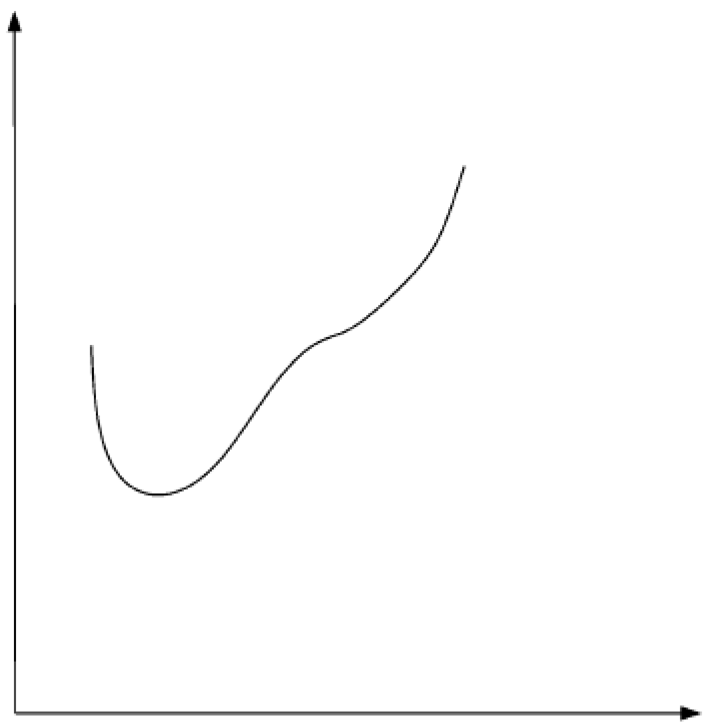


Figure 5.2. Pseudoconvex function.

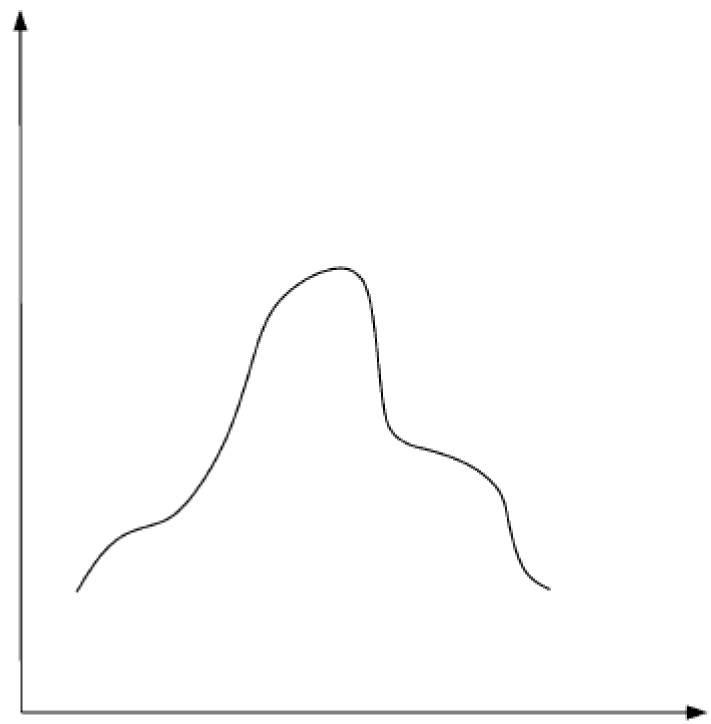


Figure 5.3. Pseudoconcave function.

$$\nabla f(x^{(1)})(x^{(2)} - x^{(1)}) \geqslant 0 \Rightarrow f(x^{(2)}) \geqslant f(x^{(1)})$$

Definition: Strictly Quasi-Convex Function

A function $f(x)$ defined on a convex set S is *strictly quasi-convex* on S if and only if

$$f(\lambda x^{(1)} + (1 - \lambda)x^{(2)}) < \max[f(x^{(1)}), f(x^{(2)})]$$

$$\text{for all } x^{(1)}, x^{(2)} \in S \qquad 0 < \lambda < 1 \qquad f(x^{(1)}) \neq f(x^{(2)})$$

Definition: Quasi-Convex Function

A function $f(x)$ defined on a convex set S is *quasi-convex* on S if and only if

$$f(\lambda x^{(1)} + (1 - \lambda)x^{(2)}) \leqslant \max[f(x^{(1)}), f(x^{(2)})]$$

$$\text{for all } x^{(1)}, x^{(2)} \in S \qquad 0 \leqslant \lambda \leqslant 1$$

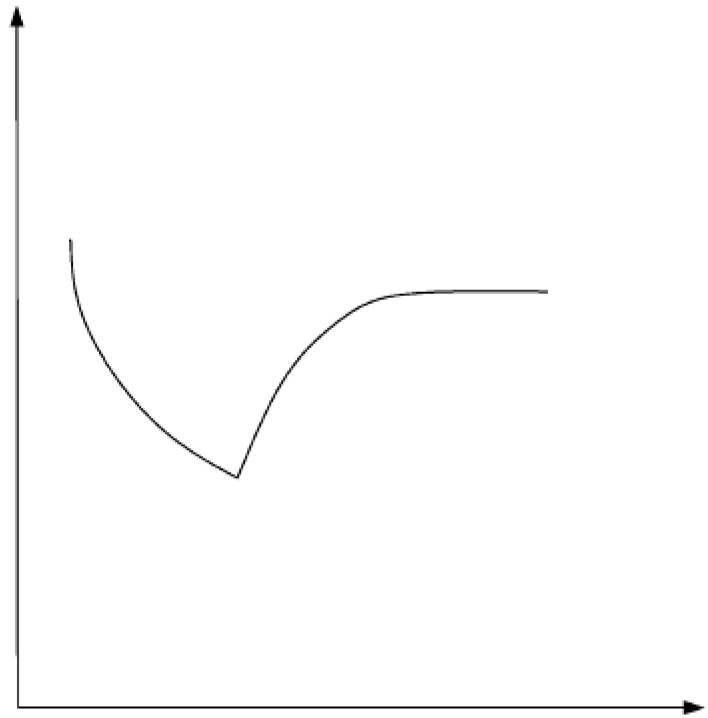


Figure 5.4. Quasi-convex function.

$$f(\lambda x^{(1)} + (1 - \lambda)x^{(2)}) \leqslant \max[f(x^{(1)}), f(x^{(2)})]$$

$$\text{for all } x^{(1)}, x^{(2)} \in S \qquad 0 \leqslant \lambda \leqslant 1$$

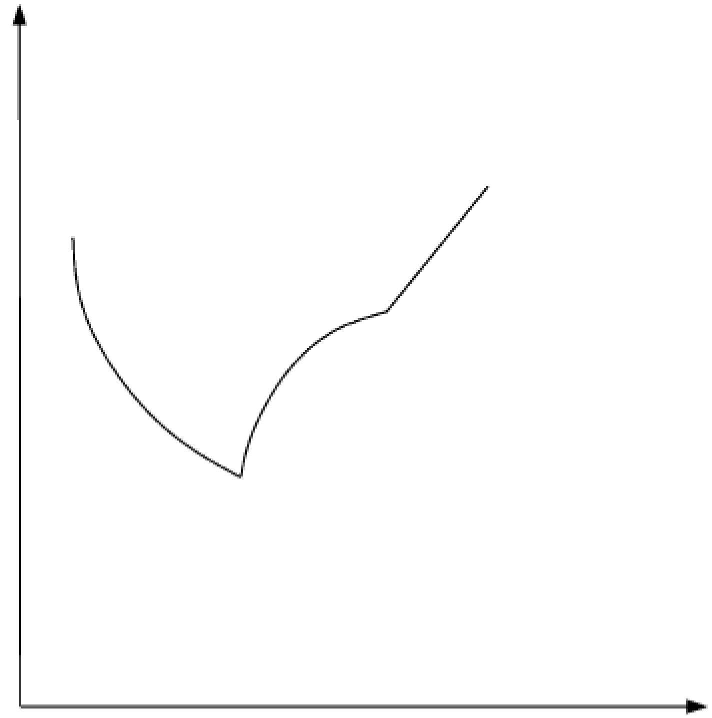


Figure 5.5. Strictly quasi-convex function.

$$f(\lambda x^{(1)} + (1 - \lambda)x^{(2)}) < \max[f(x^{(1)}), f(x^{(2)})]$$

$$\text{for all } x^{(1)}, x^{(2)} \in S \qquad 0 < \lambda < 1 \qquad f(x^{(1)}) \neq f(x^{(2)})$$

Remarks

1. A pseudoconvex function is also a quasi-convex function. But a quasi-convex function may not be pseudoconvex.
2. A strictly quasi-convex function need not necessarily be quasi-convex unless $f(x)$ is assumed to be *continuous* on the convex set of S.

Example 5.10

$$f(x) = \begin{cases} x & \text{for } x \leqslant 0 \\ 0 & \text{for } 0 \leqslant x \leqslant 1 \\ x - 1 & \text{for } x \geqslant 1 \end{cases}$$

The above function is both quasi-convex and quasi-concave but neither strictly quasi-convex nor strictly quasi-concave.

Example 5.11

$$f(x) = x^3 \qquad \text{for } x \in \mathbb{R}$$

The above function is both strictly quasi-convex and strictly quasi-concave, but it is neither quasi-convex nor quasi-concave because of the inflection point at $x = 0$.

Theorem 5.9

Let f be pseudoconvex function defined on an (open) convex set S. If $\nabla f(\bar{x}) = 0$, then $\bar{x}$ minimizes $f(x)$ over all $x \in S$.

Theorem 5.10

Let f be a strictly quasi-convex function defined on a convex set S. Then a local minimum of f is also a global minimum.

Theorem 5.11 Generalization of Kuhn–Tucker Sufficient Optimality Theorem

Consider the NLP problem

$$\begin{aligned} &\text{Minimize} && f(x) \\ &\text{Subject to} && g_j(x) \geqslant 0 \qquad \text{for } j = 1, 2, \ldots, J \\ & && h_k(x) = 0 \qquad \text{for } k = 1, 2, \ldots, K \\ & && x = (x_1, x_2, \ldots, x_N) \end{aligned}$$

KTP is as follows: Find x, u, and v such that

$$\nabla f(x) - \sum_j u_j \nabla g_j(x) - \sum_k v_k \nabla h_k(x) = 0$$

$$u_j \geqslant 0 \qquad v_k \text{ unrestricted in sign}$$

$$u_j g_j(x) = 0 \qquad \text{for all } j$$

$$g_j(x) \geqslant 0 \qquad h_k(x) = 0$$

Let $f(x)$ be pseudoconvex, g_j be quasi-concave, and h_k be both quasi-convex and quasi-concave. If $(\overline{x}, \overline{u}, \overline{v})$ solves the KTP, then $\overline{x}$ solves the NLP problem.

Summary

1. Necessary and sufficient conditions of optimality for constrained optimization problems
2. Lagrangian optimality conditions (equality and inequality constraints)
3. Kuhn-Tucker optimality conditions (first-order conditions, involving gradients)
4. Optimal => KTC : functions are differentiable + constraints qualification
5. KTC => Optimal :
   Objective function- convex
   Inequality constrains- concave
   Equality constrains – linear
6. Saddlepoint conditions applicable – functions are not differentiable
7. Second order necessary conditions – functions are twice differentiable
   Second order sufficient conditions – do not need convexity of functions and linearity of equality constraints
8. Pseudo-convex and Quasi-convex - relaxation of convexity