

Chapter 6

Multivariable Optimization without Constraints

6.1 Definitions of Extreme Points Similar to the definitions of extreme points of single-variable functions, we have

Definition A point $\mathbf{x}^*$ is a *local maximum point* of $f: S \rightarrow \mathbb{R}$, $S \subseteq \mathbb{R}^n$, if

$$f(\mathbf{x}^*) \geq f(\mathbf{x}),$$

for any $\mathbf{x} \in S$ such that $\|\mathbf{x} - \mathbf{x}^*\| < \varepsilon$ for some $\varepsilon > 0$.

Definition A point $\mathbf{x}^*$ is a *local strict maximum point* of $f: S \rightarrow \mathbb{R}$, $S \subseteq \mathbb{R}^n$, if

$$f(\mathbf{x}^*) > f(\mathbf{x}),$$

for any $\mathbf{x} \in S$, $\mathbf{x} \neq \mathbf{x}^*$, such that $\|\mathbf{x} - \mathbf{x}^*\| < \varepsilon$ for some $\varepsilon > 0$.

Definition A point $\mathbf{x}^*$ is a *global maximum point* of $f: S \rightarrow \mathbb{R}$, $S \subseteq \mathbb{R}^n$, if

$$f(\mathbf{x}^*) \geq f(\mathbf{x}),$$

for any $\mathbf{x} \in S$.

Definition A point $\mathbf{x}^*$ is a *global strict maximum point* of $f: S \rightarrow \mathbb{R}$, $S \subseteq \mathbb{R}^n$, if

$$f(\mathbf{x}^*) > f(\mathbf{x}),$$

for any $\mathbf{x} \in S$, $\mathbf{x} \neq \mathbf{x}^*$.

6.2 2-Variable Optimization A 2-variable differentiable function and its total differential are given by

$$y = f(x_1, x_2)$$

$$dy = f_1 dx_1 + f_2 dx_2.$$

At extreme point (x_1^*, x_2^*) , minimum or maximum, for any arbitrarily small changes in x_1 and x_2 , in terms of dx_1 and dx_2 , there is no change in y , i.e.,

$$dy = f_1(x_1^*, x_2^*)dx_1 + f_2(x_1^*, x_2^*)dx_2 = 0.$$

6.3 First-Order Necessary Condition From the last equation, we can conclude that

Theorem If (x_1^*, x_2^*) is an extreme point of a differentiable function $f: S \rightarrow \mathbf{R}$, $S \subseteq \mathbf{R}^2$, then $f_1(x_1^*, x_2^*) = 0$ and $f_2(x_1^*, x_2^*) = 0$.

At maximum point (x_1^*, x_2^*) , the change of the change must be non-positive, i.e.,

$$dy^2 \leq 0,$$

where for any dx_1 and dx_2 , with the argument (x_1^*, x_2^*) is suppressed for brevity and ligibility,

$$\begin{aligned} dy^2 &= d(f_1(x_1^*, x_2^*)dx_1 + f_2(x_1^*, x_2^*)dx_2) \\ &= (f_{11}dx_1 + f_{12}dx_2)dx_1 + (f_{21}dx_1 + f_{22}dx_2)dx_2 \\ &= f_{11}dx_1^2 + 2f_{12}dx_1dx_2 + f_{22}dx_2^2 \\ &= \begin{bmatrix} dx_1 & dx_2 \end{bmatrix} \begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix} \begin{bmatrix} dx_1 \\ dx_2 \end{bmatrix} \\ &= \mathbf{dx}^T \mathbf{H}(x_1^*, x_2^*) \mathbf{dx} \leq 0. \end{aligned}$$

The quadratic form $\mathbf{dx}^T \mathbf{H}(x_1^*, x_2^*) \mathbf{dx} \leq 0$ for any vector $\mathbf{dx}$ means that the symmetric matrix $\mathbf{H}$ is negative semi-definite. We have the second-order necessary condition as follows.

6.4 Second-Order Necessary Condition

Theorem If (x_1^*, x_2^*) is a *maximum point* of a twice-differentiable function $f: S \rightarrow \mathbf{R}$, $S \subseteq \mathbf{R}^2$, then $\mathbf{H}(x_1^*, x_2^*)$ is *negative semi-definite*. That means,

- a) $f_{11}(x_1^*, x_2^*) \leq 0, f_{22}(x_1^*, x_2^*) \leq 0$, and
- b) $f_{11}(x_1^*, x_2^*)f_{22}(x_1^*, x_2^*) - [f_{12}(x_1^*, x_2^*)]^2 \geq 0$.

Theorem If (x_1^*, x_2^*) is a *minimum point* of a twice-differentiable function $f: S \rightarrow \mathbf{R}$, $S \subseteq \mathbf{R}^2$, then $\mathbf{H}(x_1^*, x_2^*)$ is *positive semi-definite*. That means,

- a) $f_{11}(x_1^*, x_2^*) \geq 0, f_{22}(x_1^*, x_2^*) \geq 0$, and
- b) $f_{11}(x_1^*, x_2^*)f_{22}(x_1^*, x_2^*) - [f_{12}(x_1^*, x_2^*)]^2 \geq 0$.

Example Profit Maximization. Let the production function be given by

$$q = f(K, L),$$

each finished good is sold at p bahts. The price of capital and labor is fixed at p_K and p_L , respectively. The profit function is thus a function of both capital K and labor L ,

$$\pi(K, L) = pf(K, L) - (p_K K + p_L L).$$

If (K^*, L^*) is the profit maximizing (locally) point, by the first order necessary condition,

$$\begin{aligned}\pi_K(K^*, L^*) &= pf_K(K^*, L^*) - p_K = 0 \\ \pi_L(K^*, L^*) &= pf_L(K^*, L^*) - p_L = 0.\end{aligned}$$

Thus, at the maximum point, the *value of the marginal product* of each input must equal its price. Since $p_K, p_L > 0$, this is equivalent to

$$\frac{f_K(K^*, L^*)}{f_L(K^*, L^*)} = \frac{p_K}{p_L}.$$

By the second-order necessary condition,

$$\begin{aligned} pf_{KK}(K^*, L^*) &\leq 0 \\ pf_{LL}(K^*, L^*) &\leq 0 \\ p^2 f_{KK}(K^*, L^*) f_{LL}(K^*, L^*) - (pf_{KL}(K^*, L^*))^2 &\geq 0. \end{aligned}$$

6.5 Sufficient Conditions These are the conditions that, when satisfied, guarantee that a point is a maximum or minimum point of a twice-differentiable function. However, these conditions may not identify every maximum and minimum points.

Theorem For a twice-differentiable function $f: S \rightarrow \mathbf{R}$, $S \subseteq \mathbf{R}^2$, if

- a) $f_1(x_1^*, x_2^*) = 0, f_2(x_1^*, x_2^*) = 0$, and
b) $f_{11}(x_1^*, x_2^*) < 0$ and $f_{11}(x_1^*, x_2^*) f_{22}(x_1^*, x_2^*) - [f_{12}(x_1^*, x_2^*)]^2 > 0$,
then (x_1^*, x_2^*) is a **local maximum point**.

- Both conditions must be satisfied.
- Part (b) means the Hessian is negative definite.
- The sufficient conditions actually find a local **strict** maximum point.

Theorem For a twice-differentiable function $f: S \rightarrow \mathbf{R}$, $S \subseteq \mathbf{R}^2$, if

- a) $f_1(x_1^*, x_2^*) = 0, f_2(x_1^*, x_2^*) = 0$, and
b) $f_{11}(x_1^*, x_2^*) > 0$ and $f_{11}(x_1^*, x_2^*) f_{22}(x_1^*, x_2^*) - [f_{12}(x_1^*, x_2^*)]^2 > 0$,
then (x_1^*, x_2^*) is a **local minimum point**.

Example Use the sufficient conditions to find the point of maximum profit, when the production function is

$q = f(K, L) = 5K^{\frac{2}{5}}L^{\frac{1}{5}}$, and $p_K = 3$, $p_L = 4$. The output price p is assumed to be 1. The profit function is then given by

$$\pi(K, L) = 5K^{\frac{2}{5}}L^{\frac{1}{5}} - 3K - 4L.$$

The partial derivatives with respect to each input at the maximum point are given by

$$\begin{aligned} 2K_0^{-\frac{3}{5}}L_0^{\frac{1}{5}} - 3 &= 0 \\ K_0^{\frac{2}{5}}L_0^{-\frac{4}{5}} - 4 &= 0. \end{aligned}$$

The **critical point** is $(K_0, L_0) = \left(\frac{2}{9}, \frac{1}{12}\right)$. Verify that the second-order sufficient conditions

$$\begin{aligned} -\frac{6}{5}K_0^{-\frac{8}{5}}L_0^{\frac{1}{5}} &< 0, \text{ and} \\ \left(-\frac{6}{5}K_0^{-\frac{8}{5}}L_0^{\frac{1}{5}}\right)\left(-\frac{4}{5}K_0^{\frac{2}{5}}L_0^{-\frac{9}{5}}\right) - \left(\frac{2}{5}K_0^{-\frac{3}{5}}L_0^{-\frac{4}{5}}\right)^2 &> 0, \end{aligned}$$

are satisfied. Then conclude that the point $(K_0, L_0) = \left(\frac{2}{9}, \frac{1}{12}\right)$ is a local maximum point.

- If all the prices p, p_K and p_L are increased by 5%, will the point $(K_0, L_0) = \left(\frac{2}{9}, \frac{1}{12}\right)$ still be a local maximum point?

HW With production function $q = f(K, L) = 5K^2L$, recalculate the maximum point.

Example Find the input levels that maximize the profit when the production function is given by

$$q = f(K, L) = 5L + K^{\frac{1}{2}}L^{\frac{3}{4}},$$

and $p = 2$, $p_K = 4$, and $p_L = 16$. The profit function is then given by

$$\pi(K, L) = 2 \left(5L + K^{\frac{1}{2}} L^{\frac{3}{4}} \right) - 4K - 16L .$$

The partial derivatives with respect to each input at the maximum point are given by

$$\begin{aligned} K_0^{-\frac{1}{2}} L_0^{\frac{3}{4}} - 4 &= 0 \\ 10 + \frac{3}{4} K_0^{\frac{1}{2}} L_0^{-\frac{1}{4}} - 16 &= 0. \end{aligned}$$

The **critical point** is $(K_0, L_0) = (256, 256)$. However, the second-order sufficient conditions are not satisfied because

$$\begin{aligned} \pi_{KK}(K_0, L_0) &= -\frac{1}{2} K_0^{-\frac{3}{2}} L_0^{\frac{3}{4}} < 0 \\ \pi_{KK}(K_0, L_0) \pi_{LL}(K_0, L_0) - [\pi_{KL}(K_0, L_0)]^2 &= \left(-\frac{1}{2} K_0^{-\frac{3}{2}} L_0^{\frac{3}{4}} \right) \left(-\frac{3}{8} K_0^{\frac{1}{2}} L_0^{-\frac{5}{4}} \right) - \left(\frac{3}{4} K_0^{-\frac{1}{2}} L_0^{-\frac{1}{4}} \right)^2 \\ &= \frac{3}{16} L_0^{-\frac{1}{2}} K_0^{-1} - \frac{9}{16} L_0^{-\frac{1}{2}} K_0^{-1} \\ &= -\frac{6}{16} L_0^{-\frac{1}{2}} K_0^{-1} < 0. \end{aligned}$$

We cannot conclude that the critical point found is a maximum point.

HW Baldani, p. 193, #7.1 (a,b,c), 7.2 (a,b,c)

6.6 Multivariable Optimization without Constraints

With the same exposition as in the 2-variable case, we can discuss the necessary and sufficient conditions of the extreme points of functions with n variables as follows. A differentiable function with n variables and its differential is given by

$$y = f(\mathbf{x}) = f(x_1, x_2, \dots, x_n)$$

$$dy = \nabla f(\mathbf{x})^T \mathbf{dx} = \sum_{j=1}^n f_j(\mathbf{x}) dx_j$$

At extreme point $\mathbf{x}^* = (x_1^*, x_2^*, \dots, x_n^*)$, minimum or maximum, for any arbitrarily small changes in terms of $\mathbf{dx}$ there is no change in y , i.e.,

$$dy = \nabla f(\mathbf{x}^*)^T \mathbf{dx} = \sum_{j=1}^n f_j(\mathbf{x}^*) dx_j = 0.$$

6.7 First-Order Necessary Condition From the last equation, we can conclude that

Theorem If $\mathbf{x}^*$ is an extreme point of a differentiable function $f: S \rightarrow \mathbf{R}$, $S \subseteq \mathbf{R}^n$, then

$$\nabla f(\mathbf{x}^*) = \begin{bmatrix} f_1(\mathbf{x}^*) \\ f_2(\mathbf{x}^*) \\ \vdots \\ f_n(\mathbf{x}^*) \end{bmatrix} = \mathbf{0}.$$

At maximum point $\mathbf{x}^*$, the change of the change must be non-positive, i.e.,

$$dy^2 \leq 0,$$

where for any $\mathbf{dx}$,

$$\begin{aligned} dy^2 &= d\left(\nabla f(\mathbf{x}^*)^T \mathbf{dx}\right) = d\left(\sum_{j=1}^n f_j(\mathbf{x}^*) dx_j\right) \\ &= (f_{11} dx_1 + f_{12} dx_2 + \dots + f_{1n} dx_n) dx_1 \\ &\quad + (f_{21} dx_1 + f_{22} dx_2 + \dots + f_{2n} dx_n) dx_2 \\ &\quad + \dots + (f_{n1} dx_1 + f_{n2} dx_2 + \dots + f_{nn} dx_n) dx_n \\ &= \mathbf{dx}^T \mathbf{H}(\mathbf{x}^*) \mathbf{dx} \leq 0. \end{aligned}$$

That means when $\mathbf{x}^*$ is found, moving away from it in any direction $d\mathbf{x}$ infinitesimally will not increase the value of y .

The quadratic form $d\mathbf{x}^T \mathbf{H}(\mathbf{x}^*) d\mathbf{x} \leq 0$ for any vector $d\mathbf{x}$ means that the symmetric matrix $\mathbf{H}$ is negative semi-definite. (See Simon and Blume [1994] Theorem 16.2, page 383, for the test of semi-definiteness of a square matrix). We have the second-order necessary condition as follows.

6.8 Second-Order Necessary Condition

Theorem If $\mathbf{x}^*$ is a *maximum point* of a twice-differentiable function $f: S \rightarrow \mathbf{R}$, $S \subseteq \mathbf{R}^n$, then $\mathbf{H}(\mathbf{x}^*)$ is *negative semi-definite*.

Theorem If $\mathbf{x}^*$ is a *minimum point* of a twice-differentiable function $f: S \rightarrow \mathbf{R}$, $S \subseteq \mathbf{R}^n$, then $\mathbf{H}(\mathbf{x}^*)$ is *positive semi-definite*.

6.9 Sufficient Conditions These are the conditions that, when satisfied, guarantee that a point is a maximum or minimum point of a twice-differentiable function. However, these conditions may not identify every maximum and minimum points.

Theorem For a twice-differentiable function $f, f: S \rightarrow \mathbf{R}$, $S \subseteq \mathbf{R}^n$, if

- a) $\nabla f(\mathbf{x}^*) = \mathbf{0}$ and
- b) $\mathbf{H}(\mathbf{x}^*)$ is *negative* definite,

then $\mathbf{x}^*$ is a *local maximum point*.

Theorem For a twice-differentiable function $f, f: S \rightarrow \mathbf{R}$, $S \subseteq \mathbf{R}^n$, if

- a) $\nabla f(\mathbf{x}^*) = \mathbf{0}$ and
- b) $\mathbf{H}(\mathbf{x}^*)$ is *positive* definite,

then $\mathbf{x}^*$ is a *local minimum point*.

- Both conditions must be satisfied.

- The sufficient conditions actually find a local *strict* maximum (minimum) point.

6.10 Test of Definiteness of the Hessian

Definition The submatrix $\mathbf{H}_k$ of $\mathbf{H}(\mathbf{x}^*)$ is called the *leading principal submatrix* of order k if it is the matrix $\mathbf{H}(\mathbf{x}^*)$ with the last $n - k$ rows and columns deleted.

Definition $|\mathbf{H}_k|$ is called the *leading principal minor* of order k .

Theorem $\mathbf{H}(\mathbf{x}^*)$ is *positive* definite if $|\mathbf{H}_k| > 0$, $k = 1, \dots, n$.

Theorem $\mathbf{H}(\mathbf{x}^*)$ is *negative* definite if $(-1)^k |\mathbf{H}_k| > 0$, $k = 1, \dots, n$.

HW Baldani, p. 194, #7.4 (a,b,c,d)

6.11 Concavity, Convexity and Optimization

The sufficient conditions can only guarantee that a point is a local maximum or minimum. We need additional assumption to obtain a global extreme point. One of such assumption is the concavity or convexity of the function. This is similar to the concavity and convexity of the single-variable functions as discussed in Chapter 2.

Definition A function $f: S \rightarrow \mathbf{R}$, $S \subseteq \mathbf{R}^n$, is a *concave* (convex) function, if

$$f(\lambda \mathbf{x}^1 + (1 - \lambda) \mathbf{x}^2) \geq \lambda f(\mathbf{x}^1) + (1 - \lambda) f(\mathbf{x}^2),$$

$$(\leq)$$

for any $\mathbf{x}^1, \mathbf{x}^2 \in S$, and $0 \leq \lambda \leq 1$.

Definition A function $f: S \rightarrow \mathbf{R}$, $S \subseteq \mathbf{R}^n$, is a *strictly concave* (convex) function, if

$$f(\lambda \mathbf{x}^1 + (1-\lambda)\mathbf{x}^2) > \lambda f(\mathbf{x}^1) + (1-\lambda)f(\mathbf{x}^2),$$

$$(<)$$

for any $\mathbf{x}^1, \mathbf{x}^2 \in \mathcal{S}$, $\mathbf{x}^1 \neq \mathbf{x}^2$, and $0 < \lambda < 1$.

HW Baldani, p. 194, #7.6, 7.7, 7.8

Theorem If f is concave, then

$$\mathbf{x}^* \text{ is local max} \Leftrightarrow \mathbf{x}^* \text{ is global max.}$$

Theorem If f is concave and differentiable, then

$$\nabla f(\mathbf{x}^*) = \mathbf{0} \Leftrightarrow \mathbf{x}^* \text{ is global max.}$$

Theorem If $f: \mathcal{S} \rightarrow \mathbf{R}$, $\mathcal{S} \subseteq \mathbf{R}^n$, is twice differentiable, then

f is concave $\Leftrightarrow \mathbf{H}(\mathbf{x})$ negative semi-definite, $\mathbf{x} \in \mathcal{S}$.

f is strictly concave $\Leftrightarrow \mathbf{H}(\mathbf{x})$ negative definite, $\mathbf{x} \in \mathcal{S}$.

Theorem If a matrix is positive (negative) definite, it is nonsingular.

6.12 Comparative Statics Analysis

Consider $y = f(\mathbf{x}; \mathbf{c})$, with a vector of parameters $\mathbf{c} \in \mathbf{R}^k$.

Suppose that the optimal solution is found by the first-order sufficient condition at some specific value of $\mathbf{c}_0$

$$\nabla_{\mathbf{x}} f(\mathbf{x}^*; \mathbf{c}_0) = \mathbf{0}.$$

This can be considered as a system of n implicit functions with n endogenous variables $\mathbf{x}$ and k exogeneous variables $\mathbf{c}$.

If the optimal solution is found by the sufficient conditions,

$\mathbf{H}(\mathbf{x}^*; \mathbf{c}_0)$ is either positive or negative definite and thus

nonsingular. Thus the Implicit Function Theorem applies because

$$\nabla_{\mathbf{x}}^2 f(\mathbf{x}^*; \mathbf{c}_0) = \nabla_{\mathbf{x}} (\nabla_{\mathbf{x}} f(\mathbf{x}^*; \mathbf{c}_0)) = \mathbf{H}(\mathbf{x}^*; \mathbf{c}_0).$$

By the Implicit Function Theorem, there exists a differentiable function $\mathbf{x} = \mathbf{x}(\mathbf{c})$ and $\varepsilon > 0$ such that

a) $\nabla_{\mathbf{x}} f(\mathbf{x}(\mathbf{c}); \mathbf{c}) = \mathbf{0}$, for $\|\mathbf{c} - \mathbf{c}_0\| < \varepsilon$,

b) $\mathbf{x}(\mathbf{c}_0) = \mathbf{x}^*$, and

c) the gradient

$$\begin{aligned}\nabla \mathbf{x}(\mathbf{c}_0) &= -\left[\nabla_{\mathbf{x}} \left(\nabla_{\mathbf{x}} f(\mathbf{x}^*; \mathbf{c}_0)\right)\right]^{-1} \nabla_{\mathbf{c}} \left(\nabla_{\mathbf{x}} f(\mathbf{x}^*; \mathbf{c}_0)\right) \\ &= -\left[\nabla_{\mathbf{x}}^2 f(\mathbf{x}^*; \mathbf{c}_0)\right]^{-1} \nabla_{\mathbf{c}} \left(\nabla_{\mathbf{x}} f(\mathbf{x}^*; \mathbf{c}_0)\right) \\ &= -\left[\mathbf{H}(\mathbf{x}^*; \mathbf{c}_0)\right]^{-1} \nabla_{\mathbf{c}} \left(\nabla_{\mathbf{x}} f(\mathbf{x}^*; \mathbf{c}_0)\right).\end{aligned}$$

Example Suppose $y = f(x_1, x_2, x_3; c)$ is a function of three decision variables and one parameter c . By the first-order condition at $(x_1^*, x_2^*, x_3^*; c_0)$ as the system of implicit functions

$$f_1(x_1^*, x_2^*, x_3^*; c_0) = 0$$

$$f_2(x_1^*, x_2^*, x_3^*; c_0) = 0$$

$$f_3(x_1^*, x_2^*, x_3^*; c_0) = 0$$

There exists a function $\mathbf{x} = \mathbf{x}(c)$ such that $\nabla_{\mathbf{x}} f(\mathbf{x}(c); c) = \mathbf{0}$ for $|c - c_0| < \varepsilon$, $\mathbf{x}^* = \mathbf{x}(c_0)$ and

$$\begin{aligned}\mathbf{x}'(c_0) &= -\left[\mathbf{H}(\mathbf{x}^*; c_0)\right]^{-1} \nabla_{\mathbf{c}} \left(\nabla_{\mathbf{x}} f(\mathbf{x}^*; c_0)\right) \\ \begin{bmatrix} x_1'(c_0) \\ x_2'(c_0) \\ x_3'(c_0) \end{bmatrix} &= -\begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix}^{-1} \begin{bmatrix} f_{1c} \\ f_{2c} \\ f_{3c} \end{bmatrix},\end{aligned}$$

where all second-order and cross-partial derivatives are evaluated at $(x_1^*, x_2^*, x_3^*; c_0)$.

By Cramer's Rule,

$$x'_3(c_0) = - \frac{\begin{vmatrix} f_{11} & f_{12} & f_{1c} \\ f_{21} & f_{22} & f_{2c} \\ f_{31} & f_{32} & f_{3c} \end{vmatrix}}{\begin{vmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{vmatrix}},$$

whose sign is difficult to determine without explicit computation.

If f_1 and f_2 does not involve c so that $f_{1c} = f_{2c} = 0$ and then

$$x'_3(c_0) = - \frac{\begin{vmatrix} f_{11} & f_{12} & 0 \\ f_{21} & f_{22} & 0 \\ f_{31} & f_{32} & f_{3c} \end{vmatrix}}{\begin{vmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{vmatrix}} = - \frac{f_{3c} |\mathbf{H}_2|}{|\mathbf{H}|}.$$

Thus if $\mathbf{x}^*$ is determined to be a local maximum point by the 1st and 2nd-order sufficient conditions, $-\frac{|\mathbf{H}_2|}{|\mathbf{H}|} > 0$ and thus

$$\text{sign}[x'_3(c_0)] = \text{sign}[f_{3c}].$$

HW Baldani, p. 194, #7.9 (a,b,c)