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

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

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

Definition 6.2 A point $\mathbf{x}^*$ is a ***local strict maximum point*** of $f: \boldsymbol{S} \to \mathbb{R}$, $\boldsymbol{S} \subseteq \mathbb{R}^n$, if

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

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

Definition 6.3 A point $\mathbf{x}^*$ is a ***global maximum point*** of $f: \boldsymbol{S} \to \mathbb{R}$, $\boldsymbol{S} \subseteq \mathbb{R}^n$, if

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

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

Definition 6.4 A point $\mathbf{x}^*$ is a ***global strict maximum point*** of $f: \boldsymbol{S} \to \mathbb{R}$, $\boldsymbol{S} \subseteq \mathbb{R}^n$, if

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

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

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

$$y = f(x_1, x_2)$$
$$dy = f_1(x_1, x_2)dx_1 + f_2(x_1, x_2)dx_2.$$

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

Proof See Theorem 6.6 for the case of n variables case.❑

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 legibility,

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

The quadratic form $\mathbf{dx}^{\mathrm{T}}\mathbf{H}(x_1^*, x_2^*)\mathbf{dx} \leq 0$ 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 6.2 If (x_1^*, x_2^*) is a ***maximum point*** of a twice-differentiable function $f: \boldsymbol{S} \to \mathbb{R}$, $\boldsymbol{S} \subseteq \mathbb{R}^2$, 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) $f_{11}(x_1^*, x_2^*)f_{22}(x_1^*, x_2^*) - [f_{12}(x_1^*, x_2^*)]^2 \geq 0$.

Proof See Theorem 6.6 for the case of n variables case.❑

Theorem 6.3 If (x_1^*, x_2^*) is a ***minimum point*** of a twice-differentiable function $f: \boldsymbol{S} \to \mathbb{R}$, $\boldsymbol{S} \subseteq \mathbb{R}^2$, 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) $f_{11}(x_1^*, x_2^*)f_{22}(x_1^*, x_2^*) - [f_{12}(x_1^*, x_2^*)]^2 \geq 0$.

Proof See Theorem 6.6 and for the case of n variables case.❑

Example *Profit Maximization*. Let the production function be given by

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

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

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

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

$$\pi_L(L^*,K^*) = pf_L(L^*,K^*) - p_L = 0$$
$$\pi_K(L^*,K^*) = pf_K(L^*,K^*) - p_K = 0.$$

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

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

By the second-order necessary condition,

$$pf_{LL}(L^*,K^*) \leq 0$$
$$pf_{KK}(L^*,K^*) \leq 0$$
$$p^2 f_{LL}(L^*,K^*) f_{KK}(L^*,K^*) - (pf_{LK}(L^*,K^*))^2 \geq 0.$$

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 6.4 For a twice-differentiable function , $f: \boldsymbol{S} \to \mathbb{R}$, $\boldsymbol{S} \subseteq \mathbb{R}^2$, if

a) $f_1(x_1^*, x_2^*) = 0, f_2(x_1^*, x_2^*) = 0$
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***.

Proof See Theorem 6.6 and Theorem 6.8 for the case of n variables case.❑

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

Theorem 6.5 For a twice-differentiable function $f: \boldsymbol{S} \to \mathbb{R}$, $\boldsymbol{S} \subseteq \mathbb{R}^2$, if

a) $f_1(x_1^*, x_2^*) = 0, f_2(x_1^*, x_2^*) = 0$
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***.

Proof See Theorem 6.6 and Theorem 6.8 for the case of n variables case.❑

Example Use the sufficient conditions to find the point of maximum profit, when the production function is $q = f(L, K) = 5L^{\frac{1}{5}}K^{\frac{2}{5}}$, and $p_L = 4, p_K = 3$. The output price p is assumed to be 1. The profit function is then given by

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

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

$$L_0^{-\frac{4}{5}}K_0^{\frac{2}{5}} - 4 = 0$$
$$2L_0^{\frac{1}{5}}K_0^{-\frac{3}{5}} - 3 = 0$$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

6.6 Multivariable Optimization without Constraints

The necessary and sufficient conditions can be found by defining a composite function $g_{\mathbf{d}}: \mathbb{R} \to \mathbb{R}$, $g_{\mathbf{d}}(t) = f(\mathbf{x}^* + t\mathbf{d})$, where $\mathbf{d} \in \mathbb{R}^n$, and find its first- and second-order directional derivatives. The following establishes that the maximum point of f is closely related to the maximum point of $g_{\mathbf{d}}$.

Theorem 6.6 A point $\mathbf{x}^*$ is a local maximum point of $f: \boldsymbol{S} \to \mathbb{R}$, $\boldsymbol{S} \subseteq \mathbb{R}^n$, if, and only if, $g_{\mathbf{d}}: \mathbb{R} \to \mathbb{R}$, $g_{\mathbf{d}}(t) = f(\mathbf{x}^* + t\mathbf{d})$, has a local maximum point at $t = 0$ for any $\mathbf{d} \in \mathbb{R}^n$.

Proof Note that f is not assumed to be differentiable. If $\mathbf{x}^*$ is a local maximum point of $f: \boldsymbol{S} \to \mathbb{R}$, $\boldsymbol{S} \subseteq \mathbb{R}^n$, there exists some $\varepsilon_0 > 0$, such that $f(\mathbf{x}^*) \geq f(\mathbf{x})$, for any $\|\mathbf{x} - \mathbf{x}^*\| < \varepsilon_0$. Suppose $t = 0$ is not a local maximum point for $g_{\mathbf{d}}$ for some $\mathbf{d} \in \mathbb{R}^n$. Then, for any $\varepsilon > 0$, there exists some $\hat{t}, -\varepsilon < \hat{t} < \varepsilon$, such that

$$g_{\mathbf{d}}(\hat{t}) > g_{\mathbf{d}}(0)$$

$$f(\mathbf{x}^* + \hat{t}\mathbf{d}) > f(\mathbf{x}^*),$$

Note that $\|\mathbf{x}^* + \hat{t}\mathbf{d} - \mathbf{x}^*\| = |\hat{t}|\|\mathbf{d}\|$. We then need to show that $|\hat{t}|\|\mathbf{d}\| < \varepsilon_0$. This can be easily done by choosing ε such that $\hat{t} < \varepsilon < \frac{\varepsilon_0}{\|\mathbf{d}\|}$ so that we have a contradiction.

If $t = 0$ is a local maximum point for $g_{\mathbf{d}}$ for any $\mathbf{d} \in \mathbb{R}^n$, there exists some $\varepsilon_0 > 0$, such that $g_{\mathbf{d}}(0) \geq g_{\mathbf{d}}(t)$, for any $t, -\varepsilon_0 < t < \varepsilon_0$. Suppose that $\mathbf{x}^*$ is not a local maximum point of $f: \boldsymbol{S} \to \mathbb{R}$, $\boldsymbol{S} \subseteq \mathbb{R}^n$. Then, for any $\varepsilon > 0$, there exists some $\hat{\mathbf{x}}$ such that $f(\mathbf{x}^*) < f(\hat{\mathbf{x}})$, $\|\hat{\mathbf{x}} - \mathbf{x}^*\| < \varepsilon$. Write $\mathbf{x}(\hat{t}) = \hat{\mathbf{x}} = \mathbf{x}^* + \hat{t}\mathbf{d}$, for some $\mathbf{d}$ such that $|\hat{t}| < \varepsilon_0$ so that we have a contradiction.

Choose $\mathbf{d} = z(\hat{\mathbf{x}} - \mathbf{x}^*)$, $z > \frac{\varepsilon}{\varepsilon_0\|\hat{\mathbf{x}} - \mathbf{x}^*\|}$. Then,

$$\|\hat{\mathbf{x}} - \mathbf{x}^*\| = \|\mathbf{x}^* + \hat{t}\mathbf{d} - \mathbf{x}^*\| = |\hat{t}|\|\mathbf{d}\| = |\hat{t}|z\|\hat{\mathbf{x}} - \mathbf{x}^*\| < \varepsilon$$
$$|\hat{t}| < \frac{\varepsilon}{z\|\hat{\mathbf{x}} - \mathbf{x}^*\|} < \varepsilon_0. \square$$

As a result of Theorem 6.6, we can apply the Necessary and Sufficient Conditions of single-variable function to obtain those of multi-variable functions.

6.7 First-Order and Second-Order Necessary Condition

Theorem 6.7 If $\mathbf{x}^*$ is an extreme point of a twice-differentiable function $f: \boldsymbol{S} \to \mathbb{R}, \boldsymbol{S} \subseteq \mathbb{R}^n$**,** then

1) $\nabla f(\mathbf{x}^*) = \mathbf{0}$,
2) $\mathbf{d}^{\mathrm{T}}\,\mathbf{H}(\mathbf{x}^*)\mathbf{d} \leq 0$, for any $\mathbf{d} \in \mathbb{R}^n$.

Proof Define composite function $g_{\mathbf{d}}: \mathbb{R} \to \mathbb{R}$, $g_{\mathbf{d}}(t) = f(\mathbf{x}^* + t\mathbf{d})$, where $\mathbf{d} \in \mathbb{R}^n$. The function $g_{\mathbf{d}}$ attains local maximum value at $t = 0$ for each direction $\mathbf{d} \in \mathbb{R}^n$. By Theorem 2.5, for any $\mathbf{d} \in \mathbb{R}^n$

1) $g'_{\mathbf{d}}(0) = \nabla f(\mathbf{x}^*)^{\mathrm{T}}\mathbf{d} = 0 \Leftrightarrow \nabla f(\mathbf{x}^*) = \mathbf{0}$.
2) $g''_{\mathbf{d}}(0) = \mathbf{d}^{\mathrm{T}}\nabla^2 f(\mathbf{x}^*)\mathbf{d} = \mathbf{d}^{\mathrm{T}}\mathbf{H}(\mathbf{x}^*)\mathbf{d} \leq 0. \square$

This means that when $\mathbf{x}^*$ is a local maximum point, moving away infinitestimally from it in any direction $\mathbf{d}$ will not increase the value of f, and the curvature of the graph is concave, which means the Hessian is negative semi-definite according to the following definition.

Definition 6.5 A square matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ is ***negative (positive) semi-definite*** if for any $\mathbf{d} \in \mathbb{R}^n$, $\mathbf{d}^{\mathrm{T}}\mathbf{A}\mathbf{d} \leq (\geq) 0$.

For the Sufficient Conditions, we will need the following definiteness definition.

Definition 6.6 A square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is ***negative (positive) definite*** if for any $\mathbf{d} \in \mathbb{R}^n, \mathbf{d} \neq \mathbf{0}$, $\mathbf{d}^{\mathrm{T}}\mathbf{A}\mathbf{d} < (>) 0$.

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 of multivariable. However, these conditions may not identify every maximum and minimum points.

Theorem 6.8 For a twice-differentiable function $f: \boldsymbol{S} \to \mathbb{R}$, $\boldsymbol{S} \subseteq \mathbb{R}^n$, if at any $\mathbf{x}^* \in \boldsymbol{S}$,

1) $\nabla f(\mathbf{x}^*) = \mathbf{0}$, and
2) $\mathbf{H}(\mathbf{x}^*)$ is negative (positive) definite,

then $\mathbf{x}^*$ is a ***local maximum (minimum) point*** of f.

Proof Since $\mathbf{x}^*$ is a local maximum (minimum) point of f if, and only if, the composite function $g_{\mathbf{d}}: \mathbb{R} \to \mathbb{R}$, $g_{\mathbf{d}}(t) = f(\mathbf{x}^* + t\mathbf{d})$, has a local maximum (minimum) at $t = 0$, for any $\mathbf{d} \in \mathbb{R}^n$. By Theorem 2.6, the function $g_{\mathbf{d}}$ has a local maximum at $t = 0$, if

1) $g'_{\mathbf{d}}(0) = \nabla f(\mathbf{x}^*)^{\mathrm{T}}\mathbf{d} = 0 \Leftrightarrow \nabla f(\mathbf{x}^*) = \mathbf{0}$, and
2) $g''_{\mathbf{d}}(0) = \mathbf{d}^{\mathrm{T}}\nabla^2 f(\mathbf{x}^*)\mathbf{d} = \mathbf{d}^{\mathrm{T}}\mathbf{H}(\mathbf{x}^*)\mathbf{d} < 0$.

For local minimum, the inequality in (2) is of the reverse direction.

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

6.10 Test of Definiteness of the Hessian

See Simon and Blume [1994] Theorem 16.2, page 383, for the test of semi-definiteness of a square matrix. Only the test for definiteness is discussed here.

Definition 6.7 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 6.8 $|\mathbf{H}_k|$ is called the ***leading principal minor*** of order k.

Theorem 6.9 $\mathbf{H}(\mathbf{x}^*)$ is ***positive*** definite if $|\mathbf{H}_k| > 0$, $k = 1, 2, \ldots, n$.

Theorem 6.10 $\mathbf{H}(\mathbf{x}^*)$ is ***negative*** definite if $(-1)^k|\mathbf{H}_k| > 0$, $k = 1,2,\dots,n$.

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 assumptions 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 6.9 A function $f: \boldsymbol{S} \to \mathbb{R}$, $\boldsymbol{S} \subseteq \mathbb{R}^n$, is a ***concave*** (convex) function, if

$$f(\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) \geq \lambda f(\mathbf{x}_1) + (1-\lambda)f(\mathbf{x}_2)$$

for any $\mathbf{x}_1, \mathbf{x}_2 \in \boldsymbol{S}$, and $0 \leq \lambda \leq 1$.

Definition 6.10 A function $f: \boldsymbol{S} \to \mathbb{R}$, $\boldsymbol{S} \subseteq \mathbb{R}^n$, 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 $\mathbf{x}_1, \mathbf{x}_2 \in \boldsymbol{S}$, $\mathbf{x}_1 \neq \mathbf{x}_2$, and $0 < \lambda < 1$.

Definitions for convex and strictly convex functions are defined exactly the same with the direction of inequality reversed.

We will have the following results similar to the case of single variable in Chapter 2. We can extend the results of Chapter 2 to multivariable concave (convex) function by first proving the following Theorem 6.10. First we need the following definition.

Definition 6.11 Given a function $f: \boldsymbol{S} \to \mathbb{R}^n$, $\boldsymbol{S} \subseteq \mathbb{R}$, for any $\mathbf{x}, \mathbf{y} \in \boldsymbol{S}$, define $g_{\mathbf{xy}}: [0,1] \to \mathbb{R}$, where

$$g_{\mathbf{xy}}(t) = f(t\mathbf{x} + (1-t)\mathbf{y}).$$

• Note that $f(t\mathbf{x} + (1-t)\mathbf{y}) = f(\mathbf{y} - t(\mathbf{y} - \mathbf{x}))$

Theorem 6.10 (Simon & Blume, Theorem 21.1) A function $f: \boldsymbol{S} \to \mathbb{R}$, $\boldsymbol{S}$ is a convex set, $\boldsymbol{S} \subseteq \mathbb{R}^n$, is concave (convex) if, and only if, for any $\mathbf{x}, \mathbf{y} \in \boldsymbol{S}$, $g_{\mathbf{xy}}$ is.

Proof First, assume that $g_{\mathbf{xy}}$ is concave for any $\mathbf{x}, \mathbf{y} \in \boldsymbol{S}$. Note that

$g_{\mathbf{xy}}(0) = f(\mathbf{y})$,
$g_{\mathbf{xy}}(1) = f(\mathbf{x})$, and
$g_{\mathbf{xy}}(\lambda) = f(\mathbf{x} + (1-\lambda)(\mathbf{y}-\mathbf{x})) = f(\lambda\mathbf{x} + (1-\lambda)\mathbf{y})$.

Thus,

$$\begin{aligned} f(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) = g_{\mathbf{xy}}(\lambda) &= g_{\mathbf{xy}}(\lambda 1 + (1-\lambda)0) \\ &\geq \lambda g_{\mathbf{xy}}(1) + (1-\lambda)g_{\mathbf{xy}}(0) \\ &= \lambda f(\mathbf{x}) + (1-\lambda)f(\mathbf{y}). \end{aligned}$$

Now, if f is concave, we have to show that $g_{\mathbf{xy}}$ is concave for any $\mathbf{x}, \mathbf{y} \in \boldsymbol{S}$. For any $s_1, s_2, \lambda \in [0,1]$,

$$\begin{aligned} & g_{\mathbf{xy}}(\lambda s_1 + (1-\lambda)s_2) \\ &= f((\lambda s_1 + (1-\lambda)s_2)\mathbf{x} + \big(1 - (\lambda s_1 + (1-\lambda)s_2)\big)\mathbf{y}) \\ &= f((\lambda s_1 + (1-\lambda)s_2)\mathbf{x} + \big(\lambda + (1-\lambda) - (\lambda s_1 + (1-\lambda)s_2)\big)\mathbf{y}) \\ &= f((\lambda s_1 + (1-\lambda)s_2)\mathbf{x} + (\lambda(1-s_1) + (1-\lambda) - (1-\lambda)s_2)\mathbf{y}) \\ &= f((\lambda s_1 + (1-\lambda)s_2)\mathbf{x} + (\lambda(1-s_1) + (1-\lambda)(1-s_2))\mathbf{y}) \\ &= f(\lambda(s_1\mathbf{x} + (1-s_1))\mathbf{y}) + (1-\lambda)(s_2\mathbf{x} + (1-s_2)\mathbf{y})) \\ &\geq \lambda f(s_1\mathbf{x} + (1-s_1))\mathbf{y}) + \ (1-\lambda)f(s_2\mathbf{x} + (1-s_2)\mathbf{y}) \\ &= \lambda g_{\mathbf{xy}}(s_1) + (1-\lambda)g_{\mathbf{xy}}(s_2). \end{aligned}$$

Thus, $g_{\mathbf{xy}}$ is concave.❑

We can now use results from Chapter 2 to obtain the analogous results for the multivariable function.

Theorem 6.11 Let $f: \boldsymbol{S} \to \mathbb{R}^n$, $\boldsymbol{S} \subseteq \mathbb{R}$, be a differentiable function. The function f is concave (convex) if, and only if,

$$f(\mathbf{y}) - f(\mathbf{x}) \leq \nabla f(\mathbf{x})^{\mathrm{T}}(\mathbf{y} - \mathbf{x}),$$
$$(\geq)$$

for any $\mathbf{x}, \mathbf{y} \in \boldsymbol{S}$.

Proof We show only for the case of concave function. If we define function $g_{\mathbf{xy}}$ slightly differently from the previous theorem, namely

$$g_{\mathbf{xy}}(t) = f\big((1-t)\mathbf{x} + t\mathbf{y}\big) = f\big(\mathbf{x} + t(\mathbf{y} - \mathbf{x})\big).$$

then by the previous theorem, $f: \boldsymbol{S} \to \mathbb{R}^n$, $\boldsymbol{S} \subseteq \mathbb{R}$, is a differentiable and concave if, and only if $g_{\mathbf{xy}}$ is, and by Theorem 2.8,

$$\begin{aligned} &\Leftrightarrow g_{\mathbf{xy}}(1) - g_{\mathbf{xy}}(0) \leq g'_{\mathbf{xy}}(0)(1-0) \\ &\Leftrightarrow f(\mathbf{y}) - f(\mathbf{x}) \leq \nabla f(\mathbf{x})^{\mathrm{T}}(\mathbf{y} - \mathbf{x})❑ \end{aligned}$$

Similarly, with the same argument we can write for strict concavity and convexity

Theorem 6.12 Let $f: \boldsymbol{S} \to \mathbb{R}^n$, $\boldsymbol{S} \subseteq \mathbb{R}$, be a differentiable function. The function f is strictly concave (convex) if, and only if,

$$f(\mathbf{y}) - f(\mathbf{x}) \underset{(>)}{<} \nabla f(\mathbf{x})^{\mathrm{T}}(\mathbf{y} - \mathbf{x}),$$

for any $\mathbf{x}, \mathbf{y} \in \boldsymbol{S}, \mathbf{x} \neq \mathbf{y}$.

Problem. Prove this Theorem 6.12. The proof is similar to that of Theorem 6.11 but with strict inequality.

Theorem 6.13 (Simon and Blume [1994], Theorem 21.5, page 513) If $f: \boldsymbol{S} \to \mathbb{R}$, $\boldsymbol{S}$ is a convex set, $\boldsymbol{S} \subseteq \mathbb{R}^n$, is twice differentiable, then
a) f is concave $\Leftrightarrow$For any $\mathbf{x} \in \boldsymbol{S}$, $\mathbf{d}^{\mathrm{T}}\mathbf{H}(\mathbf{x})\mathbf{d} \leq 0$, $\mathbf{d} \in \mathbb{R}^n$,
b) For any $\mathbf{x} \in \boldsymbol{S}$, $\mathbf{d}^{\mathrm{T}}\mathbf{H}(\mathbf{x})\mathbf{d} < 0$, $\mathbf{d} \in \mathbb{R}^n$, $\mathbf{d} \neq \mathbf{0} \Longrightarrow f$ is strictly concave

Proof (a) ($\Longrightarrow$) Let f be concave over a convex set $\boldsymbol{S}$. We have to show that $\mathbf{d}^{\mathrm{T}}\mathbf{H}(\mathbf{x})\mathbf{d} \leq 0$, $\mathbf{d} \in \mathbb{R}^n$. For any $\mathbf{x} \in \boldsymbol{S}$, since f is differentiable at $\mathbf{x}$, there exists some sufficiently small $t_0 > 0$, such that $\mathbf{y_d} = \mathbf{x} + t_0\mathbf{d} \in \boldsymbol{S}$, for any $\mathbf{d} \in \mathbb{R}^n$. Since f is concave, $g_{\mathbf{xy_d}}$ is concave, where

$$g_{\mathbf{xy_d}}(t) = f\big((1-t)\mathbf{x} + t\mathbf{y_d}\big) = f\big(\mathbf{x} + t(\mathbf{y_d} - \mathbf{x})\big),$$

and $g''_{\mathbf{xy_d}}(0) \leq 0$. Thus

$$\begin{aligned} 0 \geq g''_{\mathbf{xy_d}}(0) &= (\mathbf{y_d} - \mathbf{x})^{\mathrm{T}}\mathbf{H}(\mathbf{x})(\mathbf{y_d} - \mathbf{x}) \\ &= t_0\mathbf{d}^{\mathrm{T}}\mathbf{H}(\mathbf{x})t_0\mathbf{d} \\ &= t_0^2\mathbf{d}^{\mathrm{T}}\mathbf{H}(\mathbf{x})\mathbf{d}. \end{aligned}$$

Thus, we have $\mathbf{d}^{\mathrm{T}}\mathbf{H}(\mathbf{x})\mathbf{d} \leq 0$, for any $\mathbf{x} \in \boldsymbol{S}$.

($\Longleftarrow$) We need to show that if, for any $\mathbf{x} \in \boldsymbol{S}$, $\mathbf{d}^{\mathrm{T}}\mathbf{H}(\mathbf{x})\mathbf{d} \leq 0$, $\mathbf{d} \in \mathbb{R}^n$, then f is concave. For any $\mathbf{y}, \mathbf{z} \in \boldsymbol{S}$, define $g_{\mathbf{yz}}: [0,1] \to \mathbb{R}$, where

$$g_{\mathbf{yz}}(t) = f(t\mathbf{y} + (1-t)\mathbf{z}) = f(\mathbf{z} + t(\mathbf{y} - \mathbf{z})).$$

If $g_{\mathbf{yz}}(t)$ is concave for any $\mathbf{y}, \mathbf{z} \in \boldsymbol{S}$, then so is f. Since f is twice differentiable so is $g_{\mathbf{yz}}$. By Theorem 2.12, we can show $g_{\mathbf{yz}}$ is concave by showing that $g''_{\mathbf{yz}}(t) \leq 0$, for any $t \in [0,1]$.

Take the first- and second-order derivatives of $g_{\mathbf{yz}}(t)$, we have

$$g'_{\mathbf{yz}}(t) = \nabla f(t\mathbf{y} + (1-t)\mathbf{z})^{\mathrm{T}}(\mathbf{y} - \mathbf{z})$$

$$g''_{\mathbf{yz}}(t) = (\mathbf{y}-\mathbf{z})^{\mathrm{T}}\mathbf{H}(t\mathbf{y}+(1-t)\mathbf{z})(\mathbf{y}-\mathbf{z})$$

Since $\mathbf{d}^{\mathrm{T}}\mathbf{H}(\mathbf{x})\mathbf{d} \leq 0$ for any $\mathbf{d}^{\mathrm{T}}\mathbf{H}(\mathbf{x})\mathbf{d} \leq 0$ and $\mathbf{d} \in \mathbb{R}^n$, then $g''_{\mathbf{yz}}(t) \leq 0$. By Theorem 6.10, $g_{\mathbf{yz}}$ is concave if, and only if, f is.❑

Problem Prove (b) of Theorem 6.13 above. Note that the implication arrow is only in one direction.

This means that if the function f is twice differentiable, it is concave if and only if the Hessian is negative semi-definite everywhere. And the function is strictly concave if the Hessian is negative definite everywhere.

Theorem 6.14 If a matrix is positive (negative) definite, it is nonsingular.

Problem Prove Theorem 6.14

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

Theorem 6.15 Let $f: \boldsymbol{S} \to \mathbb{R}, \boldsymbol{S} \subseteq \mathbb{R}^n$,
a) If f is concave, a local maximum point $\mathbf{x}^*$ is also a global one.
b) If f is strictly concave, a local maximum point $\mathbf{x}^*$ is also a strictly global one.

Proof (a) We can first apply Theorem 6.6 that $\mathbf{x}^*$ is a local max for f if and only if the local max for $g: \mathbb{R} \to \mathbb{R}$, $g(t) = f(\mathbf{x}^* + t\mathbf{d})$, at $t^* = 0$ for any $\mathbf{d} \in \mathbb{R}^n$. Thus, by concavity of g, $t^* = 0$ is a global max for any $\mathbf{d} \in \mathbb{R}^n$, and it follows that $\mathbf{x}^*$ is the global max for f.❑

Theorem 6.16 Let $f: \boldsymbol{S} \to \mathbb{R}, \boldsymbol{S} \subseteq \mathbb{R}^n$, be differentiable.

a) If f is concave and $\nabla f(\mathbf{x}^*) = \mathbf{0}, \mathbf{x}^* \in \boldsymbol{S}$, then $\mathbf{x}^*$is a global maximum of f.
b) If f is strictly concave and $\nabla f(\mathbf{x}^*) = \mathbf{0}, \mathbf{x}^* \in \boldsymbol{S}$, then $\mathbf{x}^*$is a strict global maximum of f.

Proof We prove only for (a). By Theorem 6.11, for any $\mathbf{y} \in \boldsymbol{S}$

$$f(\mathbf{y}) - f(\mathbf{x}^*) \leq \nabla f(\mathbf{x}^*)^{\mathrm{T}}(\mathbf{y}-\mathbf{x}) = 0.$$

Thus, $f(\mathbf{x}^*) \geq f(\mathbf{y})$, and $\mathbf{x}^*$is a global maximum of f.❑

Problem Prove (b) of Theorem 6.16.

6.12 Comparative Statics Analysis

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

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

By the Implicit Function Theorem (Theorem 4.8), 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

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

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

$$\begin{aligned} f_1(x_1^*, x_2^*, x_3^*; c_0) &= 0 \\ f_2(x_1^*, x_2^*, x_3^*; c_0) &= 0 \\ f_3(x_1^*, x_2^*, x_3^*; c_0) &= 0 \end{aligned}$$

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

$$\mathbf{x}'(c_0) = -[\mathbf{H}(\mathbf{x}^*; c_0)]^{-1} \nabla_c\left(\nabla_{\mathbf{x}} f(\mathbf{x}^*; c_0)\right)$$

$$\begin{bmatrix} x_1'(c_0) \\ x_2'(c_0) \\ x_3'(c_0) \end{bmatrix} = \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix} \begin{bmatrix} f_{1c} \\ f_{2c} \\ f_{3c} \end{bmatrix}$$

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 $sign[x_3'(c_0)] = sign[f_{3c}]$.

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