Chapter 8. Second-Order Conditions

Xiaoxiao Hu

April 6, 2020

8.A. Local and Global Maxima

- In Chapter 7, we have discussed the sufficient conditions for optimality, confined to the context of concave programming (or more braodly, quasi-concave programming).
- Especially, when F is concave and G is convex, the first-order conditions are sufficient for maximization.
- More accurately, the conditions are sufficient for a *global* maximum.
- That is, x^* satisfying the conditions does at least as well as *any* other feasible x.

Local and Global Maxima

- We obtain a *global* maximum in concave programming (quasi-concave programming) since the convexity (quasi-convexity) properties are defined *globally*.
- For example, recall the definition of convexity,

Definition 6.B.4 (Convex Function). *A function $f : \mathcal{S} \to \mathbb{R}$, defined on a convex set $\mathcal{S} \subset \mathbb{R}^N$, is convex if*

$$f(\alpha x^a + (1-\alpha)x^b) \leq \alpha f(x^a) + (1-\alpha)f(x^b), \tag{6.4}$$

for all x^a, $x^b \in \mathcal{S}$ and for all $\alpha \in [0,1]$.

Local and Global Maxima

- Similar requirements appear for concavity and quasi-convextiy (quasiconcavity).

- These properties ensure that the desired curvature is over the full domain and thus sufficient for a *global* maximum.

Local and Global Maxima

- The conclusions of a global maximiximum are ideal.
- However, in applications, we may not have functions that have the desired convexity property.

Local and Global Maxima

- In this chapter, we will focus on the curvature of the objective and constraint functions in a *small neighborhood* of the proposed optimum.
- The conditions are expressed in terms of the second-order derivatives of the functions at the point.
- Such conditions are sufficient for *local* optima – x^* satisfying the conditions does better than any other feasible x *in a sufficiently small neighborhood* of x^*.

Local and Global Maxima

- It is a useful property when global conditions are not met.

- Moreover, it has a valuable by-product: the second-order conditions play an instrumental role in determining the *comparative static* responses of the optimum choice variables x.

- We will discuss the comparative static result while we develop the theory of second-order conditions.

8.B. Unconstrained Maximization

- We will start with the simple cases of unconstrained maximization.

- First, consider the following unconstrained maximization problem with a scalar x:

$$\max_x F(x).$$

- Let x^* be a candidate for the optimum choice.

Unconstrained Maximization

- Expand F in a Taylor series around x^*:

$$F(x) = F(x^*) + F'(x^*)(x - x^*) + \frac{1}{2}F''(x^*)(x - x^*)^2 + ... \quad (8.1)$$

- The first-order necessary condition is $F'(x^*) = 0$.
- Then (8.1) becomes

$$\begin{aligned} F(x) &= F(x^*) + \frac{1}{2}F''(x^*)(x - x^*)^2 + ... \\ \Longrightarrow F(x) - F(x^*) &= \frac{1}{2}F''(x^*)(x - x^*)^2 + ... \end{aligned} \quad (8.2)$$

- For x sufficiently close to x^*, the quadratic term will dominate higher-order terms in the Taylor expansion.

Unconstrained Maximization

$$F(x) - F(x^*) = \frac{1}{2}F''(x^*)(x - x^*)^2 + ... \qquad (8.2)$$

for x in the small neighborhood of x^*.

(i) $F''(x^*) > 0 \implies F(x) - F(x^*) > 0 \implies F(x) > F(x^*)$.

- x^* will not be a maximum of $F(x)$ in the neighborhood.
- It will not be a maximum over the whole range of F.
- This argument gives a *second-order necessary* condition for x^* to yield a maximum, *local* or *global*:

$$F''(x^*) \leq 0. \qquad (8.3)$$

Unconstrained Maximization

$$F(x) - F(x^*) = \frac{1}{2}F''(x^*)(x - x^*)^2 + \ldots \tag{8.2}$$

for x in the small neighborhood of x^*.

(ii) $F''(x^*) < 0 \implies F(x) - F(x^*) < 0 \implies F(x) < F(x^*)$.

- In a small neighborhood of x^*, we will have $F(x^*) > F(x)$, irrespective of the signs of higher-order terms.
- Thus,
$$F''(x) < 0 \tag{8.4}$$
is a *second-order sufficient* condition for x^* to yield a *local* maximum.

Unconstrained Maximization

Note the differences between the weak inequality condition

$$F''(x^*) \leq 0 \tag{8.3}$$

and the strict inequality condition

$$F''(x) < 0 \tag{8.4}$$

(i) (8.3) is a *necessary* condition, while (8.4) is a *sufficient* condition.

(ii) (8.3) is a condition for both *local* and *global* maximum, while (8.4) is a condition only for *local* maximum.

Unconstrained Maximization

- A local maximum satisfying the second-order sufficient condition is called a *regular* maximum.
- If the maximum is "irregular", that is , if $F''(x) = 0$, then we have to look at the higher-order derivatives.

$$F(x) - F(x^*) = \frac{1}{3!} F'''(x^*)(x - x^*)^3 + \frac{1}{4!} F''''(x^*)(x - x^*)^4 + ...$$

- Then, $F'''(x^*) = 0$ is a necessary condition; $F'''(x^*) = 0$ and $F''''(x) < 0$ is a sufficient condition.
- We will focus on the *regular* maximum.

Comparative Statics

- Now suppose that the problem involves a parameter θ, that is, the objective function is $F(x, \theta)$.
- The first-order necessary conditions is

$$F_x(x^*, \theta) = 0. \tag{8.5}$$

(8.5) implicitly defines x^* as a function of θ.

Comparative Statics

- Totally differentiate the first-order condition,

$$F_x(x^*, \theta) = 0. \tag{8.5}$$

we have

$$\begin{aligned} & F_{xx}(x^*, \theta)\mathrm{d}x^* + F_{x\theta}(x^*, \theta)\mathrm{d}\theta = 0 \\ \text{or} \quad & \frac{\mathrm{d}x^*}{\mathrm{d}\theta} = -\frac{F_{x\theta}(x^*, \theta)}{F_{xx}(x^*, \theta)}. \end{aligned} \tag{8.6}$$

- At a *regular* maximum, $F_{xx}(x^*, \theta) < 0$, the sign of $\mathrm{d}x^*/\mathrm{d}\theta$ is the same as the sign of $F_{x\theta}(x^*, \theta)$.

An Economic Illustration

- Consider the following revenue maximization problem:

$$\max_x R(x,\theta) \equiv \max_x P(x,\theta)\cdot x,$$

  where x is the output and θ is a shift parameter; $P(x,\theta)$ is the inverse demand curve.

- Suppose $R_\theta(x,\theta) = P_\theta(x,\theta)\cdot x > 0$ for all x.
- That is, an increase in θ shifts the demand and the revenue curves upward.

An Economic Illustration

- By the first-order necessary condition,

$$R_x(x^*, \theta) = P_x(x^*, \theta) \cdot x^* + P(x^*, \theta) = 0. \tag{8.7}$$

- Totally differentiate (8.7), we have

$$\begin{aligned} R_{xx}(x^*, \theta)\mathrm{d}x^* + R_{x\theta}(x^*, \theta)\mathrm{d}\theta &= 0 \\ \Longrightarrow \frac{\mathrm{d}x^*}{\mathrm{d}\theta} &= -\frac{R_{x\theta}(x^*, \theta)}{R_{xx}(x^*, \theta)} \end{aligned} \tag{8.8}$$

- At a *regular* maximum, we have $R_{xx}(x^*, \theta) < 0$.
- Therefore, the sign of $\mathrm{d}x^*/\mathrm{d}\theta$ is the same as the sign of $R_{x\theta}(x^*, \theta)$.

An Economic Illustration

- Thus, if $R_{x\theta}(x^*, \theta) > 0$, an increase in θ will increase the revenue-maximizing output x^*.
- This is true if the increase in θ shifts the *marginal* revenue upward:

$$\frac{\mathrm{d}R_x(x, \theta)}{\mathrm{d}\theta} > 0.$$

An Economic Illustration

- Of course, it is perfectly possible that as $\theta \uparrow$,

  (i) the average revenue shifts up: $P_\theta(x,\theta) > 0$;

  (ii) the marginal revenue shifts down: $\mathrm{d}R_x(x,\theta)/\mathrm{d}\theta < 0$.

- What is needed is a twist that reduces the elasticity of demand ($E_d > 0$). To see this,

$$R_x(x,\theta) = P_x(x,\theta) + P(x,\theta) = P(x,\theta)\left[1 - \frac{1}{E_d}\right].$$

- If the marginal revenue does shift down , then a favorable shift of demand will cause output to fall.

More Choice Variables

- Let us turn to the case with a vector of choice variables.

- Now the Taylor expansion becomes

$$
\begin{aligned}
F(x) =& F(x^*) + F_x(x^*)(x - x^*) \\
& + \frac{1}{2}(x - x^*)^T F_{xx}(x^*)(x - x^*) + ... \qquad (8.9) \\
=& F(x^*) + \sum_{j=1}^{n} \left[F_j(x^*)(x_j - x_j^*) \right] \\
& + \frac{1}{2} \sum_{j=1}^{n} \sum_{k=1}^{n} F_{jk}(x_j - x_j^*)(x_k - x_k^*) + ...
\end{aligned}
$$

More Choice Variables

- The first-order necessary condition is $F_x(x^*) = 0$.
- Then (8.9) becomes

$$
\begin{aligned}
F(x) &= F(x^*) + \frac{1}{2}(x - x^*)^T F_{xx}(x^*)(x - x^*) + \dots \\
\Longrightarrow F(x) - F(x^*) &= \frac{1}{2}(x - x^*)^T F_{xx}(x^*)(x - x^*) + \dots
\end{aligned}
$$

More Choice Variables

- For x sufficiently close to x^*, the quadratic term dominates high-order terms. Therefore,

  (i) $(x - x^*)^T F_{xx}(x^*)(x - x^*) \leq 0$ is the second-order necessary condition for x^* to yield a local or global maximum;

  (ii) $(x - x^*)^T F_{xx}(x^*)(x - x^*) < 0$ is the second-order sufficient condition for x^* to yield a local maximum.

More Choice Variables

We will next link the second-order derivative test with the mathematical concepts of Negative (Semi-)Definiteness of matrices.

Negative (Semi-)Definite Matrix

Definition 8.B.1 (Negative Definite). *A symmetric* $N \times N$ *matrix* M *is negative definite if*

$$y^T M y < 0 \tag{8.10}$$

for all non-zero $y \in \mathbb{R}^N$*.*

Negative (Semi-)Definite Matrix

Definition 8.B.2 (Negative Semi-definite)**.** *A symmetric $N \times N$ matrix M is negative semi-definite if*

$$y^T M y \leq 0 \tag{8.11}$$

for all $y \in \mathbb{R}^N$.

Negative (Semi-)Definite Matrix

Example 8.B.1. $M = \begin{bmatrix} -2 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -2 \end{bmatrix}$ *is negative definite.*

Negative (Semi-)Definite Matrix

Example 8.B.1. $M = \begin{bmatrix} -2 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -2 \end{bmatrix}$ *is negative definite.*

Solution. *For any non-zero* $y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$, *we have*

$$y^T M y = -\left[y_1^2 + (y_1 - y_2)^2 + (y_2 - y_3)^2 + y_3^2\right] < 0.$$

Negative (Semi-)Definite Matrix

Example 8.B.2. $M = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$ *is negative semi-definite.*

Negative (Semi-)Definite Matrix

Example 8.B.2. $M = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$ *is negetive semi-definite.*

Solution. *For any* $y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$, *we have*

$$y^T M y = \begin{bmatrix} y_1 & y_2 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = -(y_1 + y_2)^2 \leq 0.$$

Negative (Semi-)Definite Matrix

- Note that a matrix M with all negative entries may not be negative definite.

- Example 8.B.3 illustrates the case where all entries in M is negative whereas M is not negative definite.

Negative (Semi-)Definite Matrix

Example 8.B.3. $M = \begin{bmatrix} -1 & -2 \\ -2 & -1 \end{bmatrix}$ *is not negative definite.*

Negative (Semi-)Definite Matrix

Example 8.B.3. $M = \begin{bmatrix} -1 & -2 \\ -2 & -1 \end{bmatrix}$ *is not negative definite.*

Solution. *For* $y = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ *we have*

$$y^T M y = \begin{bmatrix} -1 & 1 \end{bmatrix} \begin{bmatrix} -1 & -2 \\ -2 & -1 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = 2 > 0.$$

Positive (Semi-)Definite Matrix

Similarly, we could define positive (semi-)definite matrices analogously.

Definition 8.B.3 (Positive Definite)**.** *A symmetric $N \times N$ matrix M is positive definite if*

$$y^T M y > 0 \tag{8.12}$$

for all non-zero $y \in \mathbb{R}^N$.

Positive (Semi-)Definite Matrix

Definition 8.B.4 (Positive Semi-definite). *A symmetric $N \times N$ matrix M is positive semi-definite if*

$$y^T M y \geq 0 \tag{8.13}$$

for all $y \in \mathbb{R}^N$.

Indefinite Matrix

Remark. *A matrix that is not positive semi-definite and not negative semi-definite is called **indefinite**.*

Definiteness of Matrices

- There are various ways to check definiteness of matrices.
- In Examples 8.B.1, 8.B.2 and 8.B.3, we have used the definition to check the definiteness.
- Below, we will introduce the determinantal test for definiteness.

Principal Minor

Before discussing the general theorem, we need to learn some new concepts.

Definition 8.B.5 (Principal Submatrix and Principal Minor)**.** *Let M be a $N \times N$ matrix. A $k \times k$ submatrix of M formed by deleting $n-k$ rows and the same $n-k$ columns of M is called the k^{th} order* ***principal submatrix*** *of M. The determinant of a principal submatrix is called the k^{th} order* ***principal minor*** *of M.*

Principal Minor

Example 8.B.4.

For a general 3×3 *matrix* $M = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$.

1. *There is one* 3^{rd} *order principal minor, namely,* $\det M$*;*

Principal Minor: Example 8.B.4

2. There are three 2^{nd} order principal minors, namely,

a) $\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, formed by deleting the 3^{rd} row and column;

b) $\det \begin{bmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{bmatrix}$, formed by deleting the 2^{nd} row and column;

c) $\det \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}$, formed by deleting the 1^{st} row and column.

Principal Minor: Example 8.B.4

3. There are three 1^{st} order principal minors, namely,

   a) $\det\begin{bmatrix} a_{11} \end{bmatrix}$, formed by deleting the 2^{nd} and 3^{rd} rows and colomns;

   b) $\det\begin{bmatrix} a_{22} \end{bmatrix}$, formed by deleting the 1^{st} and 3^{rd} rows and colomns;

   c) $\det\begin{bmatrix} a_{33} \end{bmatrix}$, formed by deleting the 1^{st} and 2^{nd} rows and colomns.

Leading Principal Minor

Definition 8.B.6 (Leading Principal Submatrix and Leading Principal Minor)**.** *Let M be a $N \times N$ matrix. The k^{th} order prinipal submatrix of M obtained by deleting the last $n-k$ rows and column of M is called the k^{th} order **leading principal submatrix** of M; and its determinant is called the k^{th} order **leading principal minor** of M.*

Leading Principal Minor

Example 8.B.5. *For the 3×3 matrix in Example 8.B.4,*

1. *The* 3^{rd} *order leading principal minor is* $\det M$*;*
2. *The* 2^{nd} *order leading principal minor is* $\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$*;*
3. *The* 1^{st} *order leading principal minor is* $\det \begin{bmatrix} a_{11} \end{bmatrix}$*.*

Definiteness of Matrices

The following two theorems provide the algorithm for testing the definiteness of a symmetric matrix.

Theorem 8.1. *Let M be an $N \times N$ symmetric matrix. Then*

1. *M is positive definite if and only if all its leading principal minors are positive;*

2. *M is negative definite if and only if all its leading principal minors of odd order are negative; and all its leading principal minors of even order are positive.*

Definiteness of Matrices

Theorem 8.2. *Let M be an $N \times N$ symmetric matrix. Then*

1. *M is positive semi-definite if and only if all its principal minors are non-negative;*

2. *M is negative semi-definite if and only if all its principal minors of odd order are non-positive ; and all all its principal minors of even order are non-negative.*

Remark. *Please note that to check the semi-definiteness of matrices, we must unfortunately check not only the leading principal minors, but all principal minors.*

Definiteness of Matrices

- Returning to our maximization problem.
- We could rewrite the second-order conditions using the terminology of (semi-)definiteness of matrices.

(i) The second-order necessary condition: $F_{xx}(x^*)$ is negative semi-definite;

(ii) The second-order sufficient condition: $F_{xx}(x^*)$ is negative definite.

Remark. *The second-order partial derivative matrix, F_{xx}, is called Hessian Matrix.*

Concavity

- We would like to compare and contrast the second-order conditions with the property of concavity.

Proposition 7.A.1 (Concave Function)**.** *A differentiable function $f : \mathcal{S} \to \mathbb{R}$, defined on a convex set $\mathcal{S} \subset \mathbb{R}^N$, is concave if and only if*

$$f_x(x^a)(x^b - x^a) \geq f(x^b) - f(x^a), \tag{7.1}$$

for all $x^a, x^b \in \mathcal{S}$.

Concavity

For twice continusouly differentiable functions, this concavity property could be interpreted in terms of second-order derivatives.

Theorem 8.3. *The (twice continuously differentiable) function* $f : \mathcal{S} \to \mathbb{R}$ *is concave if and only if* f_{xx} *is negative semi-definite for every* $x \in \mathcal{S}$*. If* f_{xx} *is negative definite for every* $x \in \mathcal{S}$*, then the function is strictly concave.*

Concavity

- The link between *Concavity* and the second-order necessary condition is clear:

  (i) Concavity requires F_{xx} to be negative semi-definite for every x;

  (ii) Second-order necessary condition only requires F_{xx} to be negative semi-definite for x^*.

Concavity

- This is why the second-order conditions are useful: it is applicable to the functions that do not have the desired concavity property over their whole domain of definition.

- Of course, on the other hand, if the function do have the concavity property, it will satisfy the second-order necessary condition.

Concavity

The Remark below summarizes this observation.

Remark. *To apply the second-order conditions we derived in this chapter, the objective function need not be concave (defined globally). It only needs to be "concave" at the point* x^**:* $F_{xx}(x^*)$ *is negative semi-definite.*

Comparative Statics

- Similar to the scalar variable case, we could derive the comparative static result by

  1. totally differentiating first-order necessary condition;

  2. applying the second-order conditions.

- See Example 8.4 Part I for an application.

8.C. Constrained Optimization

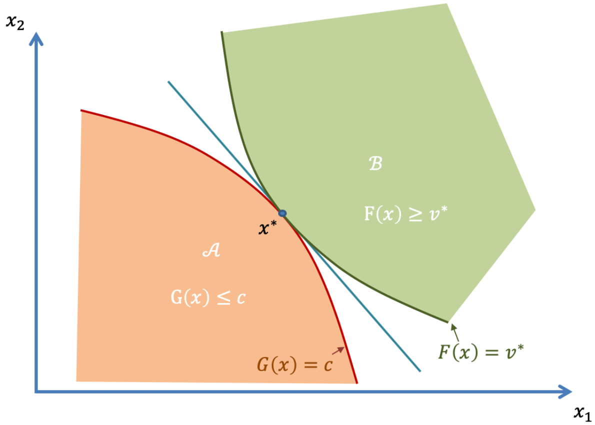
We will begin with the simplest case of two choice variables and one equality constraint.

$$\max_{x_1, x_2} F(x_1, x_2)$$

$$\text{s.t. } G(x_1, x_2) = c$$

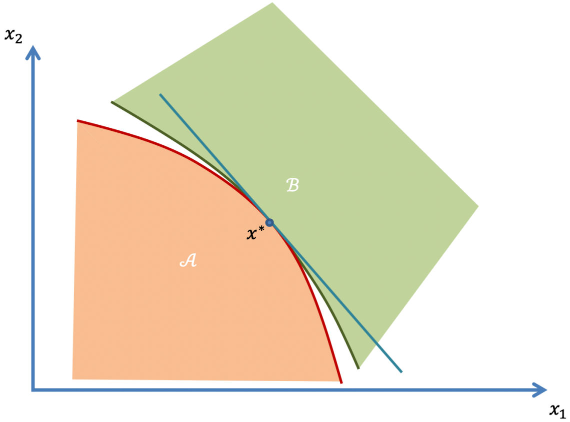
where F and G are increasing functions of their arguments.

Constrained Optimization



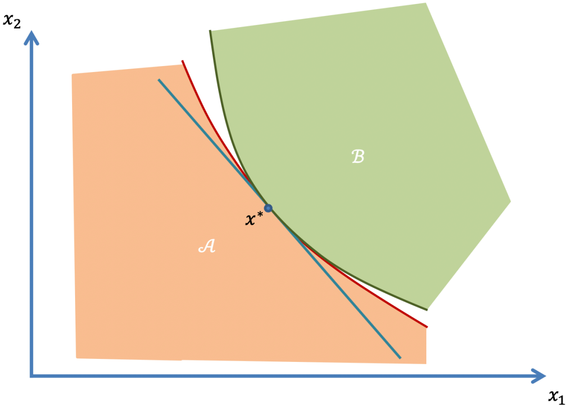
(a)

Constrained Optimization



(b)

Constrained Optimization



(c)

Constrained Optimization

- We have seen these figures in Chapter 6 and mentioned that the *relative* curvature of F and G matters for maximization: the contour of F should be more convex than that of G.

- To express the idea algebraically, we think of x_2 as a function of x_1 along the contour of F and G, and find the second-order derivative of this function.

Constrained Optimization

- For F, the function of the contour is $F(x_1, x_2) = v$.

- Total differentiation gives

$$\begin{aligned} & F_1(x_1, x_2)\mathrm{d}x_1 + F_2(x_1, x_2)\mathrm{d}x_2 = 0 \\ \Longrightarrow & \frac{\mathrm{d}x_2}{\mathrm{d}x_1} = -\frac{F_1(x_1, x_2)}{F_2(x_1, x_2)}. \end{aligned} \tag{8.14}$$

- To obtain the curvature, we differentiate (8.14) with respect x_1 (remember we think of x_2 as a function of x_1):

$$
\begin{aligned}
\frac{\mathrm{d}^2 x_2}{\mathrm{d}x_1^2} &= -\frac{\mathrm{d}}{\mathrm{d}x_1}\left[\frac{F_1(x_1, x_2(x_1))}{F_2(x_1, x_2(x_1))}\right] \\
&= -\frac{F_2\left[F_{11} + F_{12}\frac{\mathrm{d}x_2}{\mathrm{d}x_1}\right] - F_1\left[F_{21} + F_{22}\frac{\mathrm{d}x_2}{\mathrm{d}x_1}\right]}{{F_2}^2} \\
&= -\frac{F_2\left[F_{11} - F_{12}\frac{F_1}{F_2}\right] - F_1\left[F_{21} - F_{22}\frac{F_1}{F_2}\right]}{{F_2}^2} \\
&\underbrace{=}_{F_{12}=F_{21}} -\frac{{F_2}^2 F_{11} - 2F_1F_2F_{12} + {F_1}^2 F_{22}}{{F_2}^3}.
\end{aligned}
$$

Remark. *The symmetry of the second derivative matrix follows from the Schwarz's theorem: if* F *has continuous second partial derivative at* a*, then,* $\frac{\partial^2 f(a)}{\partial x_i \partial x_j} = \frac{\partial^2 f(a)}{\partial x_j \partial x_i}$.

Constrained Optimization

A similar expression could be derived for the second-order derivative along the constraint curve:

$$\frac{\mathrm{d}^2 x_2}{\mathrm{d}x_1^2} = -\frac{{G_2}^2 G_{11} - 2G_1 G_2 G_{12} + {G_1}^2 G_{22}}{{G_2}^3}.$$

Constrained Optimization

The second-order sufficient condition for x^* to be a local optimum is that $\mathrm{d}^2 x_2/dx_1{}^2$ along the F contour should be greater than that along the G contour:

$$
\begin{aligned}
& -\frac{F_2{}^2 F_{11} - 2F_1 F_2 F_{12} + F_1{}^2 F_{22}}{F_2{}^3} > -\frac{G_2{}^2 G_{11} - 2G_1 G_2 G_{12} + G_1{}^2 G_{22}}{G_2{}^3} \\
\underbrace{\Longrightarrow}_{\text{FOC: } F_j = \lambda G_j} & -\frac{\lambda^2 G_2{}^2 F_{11} - 2\lambda G_1 \lambda G_2 F_{12} + \lambda^2 G_1{}^2 F_{22}}{\lambda^3 G_2{}^3} > -\frac{G_2{}^2 G_{11} - 2G_1 G_2 G_{12} + G_1{}^2 G_{22}}{G_2{}^3} \\
\underbrace{\Longrightarrow}_{G_j > 0, \lambda > 0} & G_2{}^2 \left(F_{11} - \lambda G_{11}\right) - 2G_1 G_2 (F_{12} - \lambda G_{12}) + G_1{}^2 (F_{22} - \lambda G_{22}) < 0,
\end{aligned}
$$

evaluated at x^*.

This is more neatly expressed in matrix notation:

$$\det \begin{bmatrix} 0 & -G_1 & -G_2 \\ -G_1 & F_{11} - \lambda G_{11} & F_{12} - \lambda G_{12} \\ -G_2 & F_{21} - \lambda G_{21} & F_{22} - \lambda G_{22} \end{bmatrix} > 0, \tag{8.15}$$

evaluated at x^*.

Generalization to more variables and more constraints

- Next, we provide without proof the conditions for the general problem with n choice variables and m equation constraints ($m < n$).
- Similar to the matrix notation in (8.15), we form the partitioned matrix:

$$\begin{bmatrix} 0 & -G_x \\ -{G_x}^T & F_{xx} - \lambda G_{xx} \end{bmatrix}, \tag{8.16}$$

evaluated at x^*.

- The top left partition 0 is $m \times m$; the bottom right partition $F_{xx} - \lambda G_{xx}$ is $n \times n$; and G_x is $m \times n$.

Generalization to more variables and more constraints

Remark. *The matrix*

$$\begin{bmatrix} 0 & -G_x \\ -{G_x}^T & F_{xx} - \lambda G_{xx} \end{bmatrix}$$

is called Bordered Hessian Matrix.

Generalization to more variables and more constraints

- For second-order sufficient condition, we need to look at $n-m$ of the bordered Hessian's leading principal minors.

- Intuitively, we can think of the m constraints as reducing the problem to one with $n-m$ free variables.

- For example, the maximization problem: $\max_{x,y,z} x+y^2+z$ subject to $x+y+z=1$ can be reduced to $\max_{x,y} x+y^2+(1-x-y)$ with no constraint.

Generalization to more variables and more constraints

- The smallest minor we consider consisting of the truncated first $2m+1$ rows and columns, the next consisting of the truncated first $2m+2$ rows and columns, and so on, with the last being the determinant of the entire bordered Hessian.

- A sufficient condition for a local maximum of F is that the smallest minor has the same sign as $(-1)^{m+1}$ and that the rest of the principal minors alternate in sign.

Generalization to more variables and more constraints

The result is summarized in Theorem 8.4 below.

Theorem 8.4 (Second-order Sufficient Condition for Constrained Maximization Problem). *If the last $n-m$ leading principal minors of the bordered Hessian matrix at the proposed optimum x^* is such that the smallest minor (the $(2m+1)^{th}$ minor) has the same sign as $(-1)^{m+1}$ and the rest of the principal minors alternate in sign, then x^* is the local maximum of the constrained maximization problem.*

Generalization to more variables and more constraints

It is easy to check that (8.15) satisfies the sufficient condition for a local maximum for the two-variable one-constraint case:

1. For the two-variable one-constraint case ($n = 2, m = 1$), we need to look at $n - m = 1$ leading principal minors. Therefore, we only need to compute the determinant of the border Hessian.

2. The sign requirement for maximum is

   $(-1)^{m+1} = (-1)^2 > 0.$

Generalization to more variables and more constraints

Example 8.C.1. *Consider the maximization problem with three variables* ($n = 3$) *and two constraints* ($m = 2$)*:*

$$\max_{x,y,z} F(x,y,z) \equiv z$$

$$s.t. \;\; G^1(x,y,z) \equiv x + y + z = 12$$

$$G^2(x,y,z) \equiv x^2 + y^2 - z = 0$$

More variables and more constraints: Example 8.C.1

- *The Lagrangian is*

$$\mathcal{L}(x, y, z, \lambda, \mu) = z + \lambda(12 - x - y - z) + \mu(-x^2 - y^2 + z).$$

- *The first-order necessary conditions are*

$$\partial\mathcal{L}/\partial x = -\lambda - 2\mu x = 0$$

$$\partial\mathcal{L}/\partial y = -\lambda - 2\mu y = 0$$

$$\partial\mathcal{L}/\partial z = 1 - \lambda + \mu = 0$$

$$\partial\mathcal{L}/\partial\lambda = 12 - x - y - z = 0$$

$$\partial\mathcal{L}/\partial\mu = -x^2 - y^2 + z = 0$$

More variables and more constraints: Example 8.C.1

- *The stationary points are* $(x^*, y^*, z^*, \lambda, \mu) = (2, 2, 8, \frac{4}{5}, -\frac{1}{5})$ *and* $(-3, -3, 18, \frac{6}{5}, \frac{1}{5})$.
- *The bordered Hessian matrix is*

$$\begin{bmatrix} 0 & 0 & -G_x^1 & -G_y^1 & -G_z^1 \\ 0 & 0 & -G_x^2 & -G_y^2 & -G_z^2 \\ -G_x^1 & -G_x^2 & \mathcal{L}_{11} & \mathcal{L}_{12} & \mathcal{L}_{13} \\ -G_y^1 & -G_y^2 & \mathcal{L}_{21} & \mathcal{L}_{22} & \mathcal{L}_{23} \\ -G_z^1 & -G_z^2 & \mathcal{L}_{31} & \mathcal{L}_{32} & \mathcal{L}_{33} \end{bmatrix} = \begin{bmatrix} 0 & 0 & -1 & -1 & -1 \\ 0 & 0 & -2x & -2y & 1 \\ -1 & -2x & -2\mu & 0 & 0 \\ -1 & -2y & 0 & -2\mu & 0 \\ -1 & 1 & 0 & 0 & 0 \end{bmatrix}$$

More variables and more constraints: Example 8.C.1

- *We need to check* $n - m = 1$ *leading principal minors, i.e., we only need to check the determinant of the bordered Hessian.*
- *For local maximum, the sign requirement is* $(-1)^{m+1} = (-1)^3 < 0$.

More variables and more constraints: Example 8.C.1

1. 1^{st} *proposed optimum:* $(x^*, y^*, z^*, \lambda, \mu) = (2, 2, 8, \frac{4}{5}, -\frac{1}{5})$

   - *The determinant of the bordered Hessian is* $20 > 0$.

2. 2^{nd} *proposed optimum:* $(x^*, y^*, z^*, \lambda, \mu) = (-3, -3, 18, \frac{6}{5}, \frac{1}{5})$

   - *The determinant of the bordered Hessian is* $-20 < 0$.

Therefore, the 2^{nd} *proposed optimum* $(x^*, y^*, z^*, \lambda, \mu) = (-3, -3, 18, \frac{6}{5}, \frac{1}{5})$ *is a local maximum.*

Generalization to more variables and more constraints

Example 8.C.2. *Consider the following maximization problem with three variables* ($n = 3$) *and one constraint* ($m = 1$)*:*

$$\max_{x,y,z} F(x, y, z) \equiv x + y + z$$

$$s.t. \;\; G^1(x, y, z) \equiv x^2 + y^2 + z^2 = 3$$

More variables and more constraints: Example 8.C.2

- *The Lagrangian is*

$$\mathcal{L}(x, y, z, \lambda) = x + y + z + \lambda(3 - x^2 - y^2 - z^2).$$

- *The first-order necessary conditions are*

$$\partial\mathcal{L}/\partial x = 1 - 2\lambda x = 0$$

$$\partial\mathcal{L}/\partial y = 1 - 2\lambda y = 0$$

$$\partial\mathcal{L}/\partial z = 1 - 2\lambda z = 0$$

$$\partial\mathcal{L}/\partial \lambda = 3 - x^2 - y^2 - z^2 = 0$$

More variables and more constraints: Example 8.C.2

- *The stationary points are* $(x^*, y^*, z^*, \lambda) = (-1, -1, -1, -\frac{1}{2})$ *and* $(1, 1, 1, \frac{1}{2})$.

- *The bordered Hessian matrix is*

$$\begin{bmatrix} 0 & -G_x^1 & -G_y^1 & -G_z^1 \\ -G_x^1 & \mathcal{L}_{11} & \mathcal{L}_{12} & \mathcal{L}_{13} \\ -G_y^1 & \mathcal{L}_{21} & \mathcal{L}_{22} & \mathcal{L}_{23} \\ -G_z^1 & \mathcal{L}_{31} & \mathcal{L}_{32} & \mathcal{L}_{33} \end{bmatrix} = \begin{bmatrix} 0 & -2x & -2y & -2z \\ -2x & -2\lambda & 0 & 0 \\ -2y & 0 & -2\lambda & 0 \\ -2z & 0 & 0 & -2\lambda \end{bmatrix}$$

More variables and more constraints: Example 8.C.2

- *We need to check* $n - m = 2$ *leading principal minors, i.e., the* 3^{rd} *order and the entire bordered Hessian.*
- *For local maximum, the sign requirement is* $(-1)^{m+1} = (-1)^2 > 0$ *for the* 3^{rd} *order leading principal minor and* < 0 *for the entired bordered Hessian.*

More variables and more constraints: Example 8.C.2

1. *The first proposed optimum:* $(x^*, y^*, z^*, \lambda) = (-1, -1, -1, -\frac{1}{2})$
   - *The* 3^{rd} *order leading principal minor is* $-8 < 0$*;*
   - *The determinant of the bordered Hessian is* $-12 < 0$*.*
2. *The second proposed optimum:* $(x^*, y^*, z^*, \lambda) = (1, 1, 1, \frac{1}{2})$
   - *The* 3^{rd} *order leading principal minor is* $8 > 0$*;*
   - *The determinant of the bordered Hessian is* $-12 < 0$*.*

Thus, the 2^{nd} *proposed optimum* $(x^*, y^*, z^*, \lambda) = (1, 1, 1, \frac{1}{2})$ *is a local maximum.*

Comparative Statics

- For the constrained maximization problem, we could derive the comparative static results by

  (i) totally differentiating the first-order necessary condition and the constrained equations;

  (ii) applying the second-order conditions.

- See Example 8.4 Part II for an application.

Inequality Constraints

Finally, we consider the maximization problem

$$\max_x F(x)$$
$$\text{s.t. } G(x) \leq c.$$

- After applying the Kuhn-Tucker first-order necessary conditions and solving for the stationary points, we know which constraints are binding and which are not in those candidate optima.

Inequality Constraints

- It seems that for each stationary point, we could treat the binding constraints as the equality constraints and simply ignore the slack constraints.
- The intuition is correct in general, but there is one tricky point: it is possible that the inequality constraint is binding but at the same time its corresponding Lagrange multiplier is equal to 0.
- These inequality constraints are *degenerate inequality constraints*.

Inequality Constraints

- The conclusion is that to check the second-order sufficient condition, we should only keep the binding constraints with strictly positive corresponding Lagrange multipliers.

- In other words, we form the bordered Hessian Matrix using only the constraints with strictly positive Lagrange multipliers and then apply Theorem 8.4.

8.D. Envelope Properties

In Chapter 5, we established the envelope property of the maximum value function:

$$V(\theta) = \max_x \{F(x, \theta) \mid G(x) \le c\}.$$

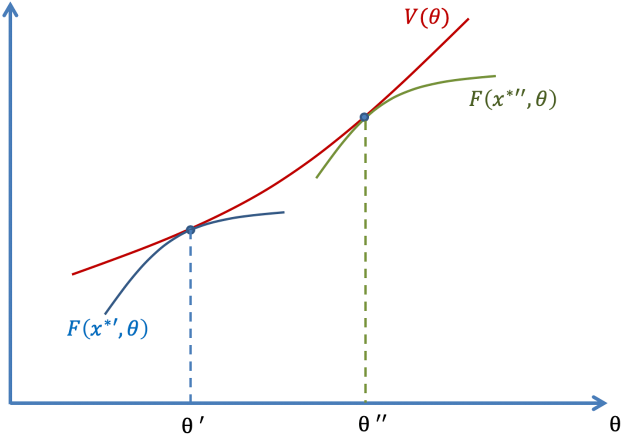
- $V(\theta)$ is the upper envelope of the family of functions $F(x, \theta)$ in each of which x is held fixed.

Envelope Properties

- Subsequently, we have considered the more general problem of short-run and long-run maximum value functions, where the vector of choice variables x is partitioned into subvectors (y, z) and z is helpd fixed in the short-run.

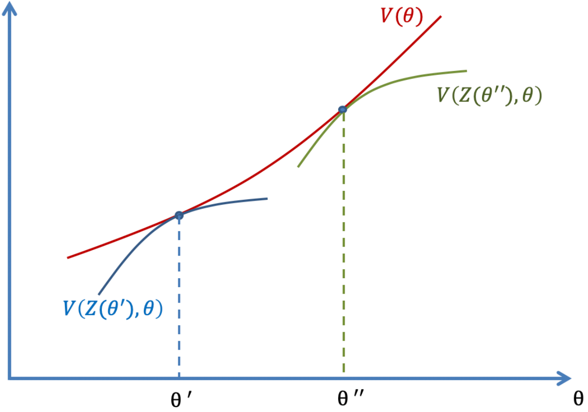
- $V(\theta)$, the long-run optimum value function, is the upper envelope of the family of value functions $V(z, \theta)$, the short-run maximum value functions.

Envelope Properties



(a) Envelope Theorem

Envelope Properties



(b) Short-run and Long-run Curves

Envelope Properties

- We have also mentioned the curvature properties of the envelopes.
- In Figure (a), V is more convex than each F.
- In Figure (b), $V(\theta)$ is more convex than $V(z, \theta)$.
- That is, the fewer variables are held fixed, the more convex should the maximum value function be.
- This second-order envelope property is the subject of this section.

Envelope Properties

- Following the same notation of Chapter 5, let $Z(\theta)$ be the long-run optimum value of z.
- Then, the long-run and short-run value coincide at $Z(\theta)$:

$$V(\theta) = V(Z(\theta), \theta). \tag{8.17}$$

- Besides, two curves are tangential at $Z(\theta)$:

$$V_\theta(\theta) = V_\theta(Z(\theta), \theta). \tag{8.18}$$

Envelope Properties

- Now consider a deviation from θ to θ', we have

$$V(Z(\theta), \theta') \leq V(Z(\theta'), \theta') = V(\theta').$$

- Expand $V(Z(\theta), \theta')$ and $V(\theta')$ around θ in Taylor series:

$$\begin{aligned} & V(Z(\theta), \theta) + V_\theta(Z(\theta), \theta)(\theta' - \theta) + \frac{1}{2} V_{\theta\theta}(Z(\theta), \theta)(\theta' - \theta)^2 + ... \\ \leq & V(\theta) + V_\theta(\theta)(\theta' - \theta) + \frac{1}{2} V_{\theta\theta}(\theta' - \theta)^2 + ... \end{aligned} \tag{8.19}$$

Envelope Properties

By the first-order envelope properties

$$V(\theta) = V(Z(\theta), \theta) \tag{8.17}$$

$$\text{and } V_\theta(\theta) = V_\theta(Z(\theta), \theta) \tag{8.18}$$

Equation (8.19)

$$V(Z(\theta), \theta) + V_\theta(Z(\theta), \theta)(\theta' - \theta) + \frac{1}{2}V_{\theta\theta}(Z(\theta), \theta)(\theta' - \theta)^2 + \ldots$$

$$\leq V(\theta) + V_\theta(\theta)(\theta' - \theta) + \frac{1}{2}V_{\theta\theta}(\theta' - \theta)^2 + \ldots \tag{8.19}$$

becomes $$(V_{\theta\theta}(Z(\theta), \theta) - V_{\theta\theta}(\theta))(\theta' - \theta)^2 + \ldots \leq 0. \tag{8.20}$$

Envelope Properties

- Consider θ' sufficiently close to θ, the quadratic term in the expansion would dominate the rest of the terms.
- For the inequality

$$(V_{\theta\theta} - V_{\theta\theta})(Z(\theta), \theta)(\theta' - \theta)^2 + ... \leq 0 \tag{8.20}$$

to hold, a necessary condition is

$$V_{\theta\theta}(Z(\theta), \theta) \leq V_{\theta\theta}(\theta). \tag{8.21}$$

- This proves that the long-run maximum value function is at least as convex as the short-run value function at the point where the two are tangent.

Envelope Properties

For suitably "regular" maxima, we have a strict inequality in (8.21).

8.E. Examples

Example 8.1: Consumer Theory

Consider the consumer's expenditure minimization problem:

$$\min_x px \tag{EMP}$$

$$\text{s.t. } u(x) \geq u.$$

Example 8.1: Consumer Theory

- In Example 5.2, we define the consumer's expenditure function $E(p, u)$ as the minimum value to the expenditure minimization problem (EMP) above.
- We denote the optimum quantity as the compensate demand function $C(p, u)$.
- The envelope property implies:

$$C(p, u) = E_p(p, u). \tag{8.22}$$

Example 8.1: Consumer Theory

- In Example 6.2, we showed that the expenditure function $E(p, u)$ is concave in p.
- Now by Theorem 8.3, we know that it means that $E_{pp}(p, u)$ is negative semi-definite.
- Differentiating

$$C(p, u) = E_p(p, u) \tag{8.22}$$

with respect to p:

$$C_p(p, u) = E_{pp}(p, u). \tag{8.23}$$

Example 8.1: Consumer Theory

$$C_p(p, u) = E_{pp}(p, u). \tag{8.23}$$

(i) Because the second derivative matrix $E_{pp}(p, u)$ is symmetric by Schwarz's theorem, $C_p(p, u)$ is symmetric:

$$\frac{\partial C^j}{\partial p_k} = \frac{\partial C^k}{\partial p_j} = E_{jk}.$$

This is the symmetry of substitution effects of price changes.

Example 8.1: Consumer Theory

$$C_p(p,u) = E_{pp}(p,u). \tag{8.23}$$

(ii)
- $E_{pp}(p,u)$ is negative semi-definite.
- That is, $y^T E_{pp}(p,u) y \leq 0$ for all $y \in \mathbb{R}^n$.
- In particular, we could choose $y = e^j$
- Then $e^{j^T} E_{pp}(p,u) e^j = E_{jj} \leq 0 \implies \dfrac{\partial C^j}{\partial p_j} \leq 0. \quad (8.24)$
- This is true for all j.
- Therefore, the own substitution effects of price changes are non-positive.

Example 8.1: Consumer Theory

- The second result follows even more simply from the very concept of maximum.
- For interested students, please refer to the textbook or lecture notes.

Example 8.2: The LeChatelier Samuelson Principal

- Consider the consumer's expenditure minimization problem (EMP) again.
- Now, we focus on second-order envelope properties.
- Consider a change in p_1 and compare the following two situations:

  (i) Quantities of all goods are free to change optimally;

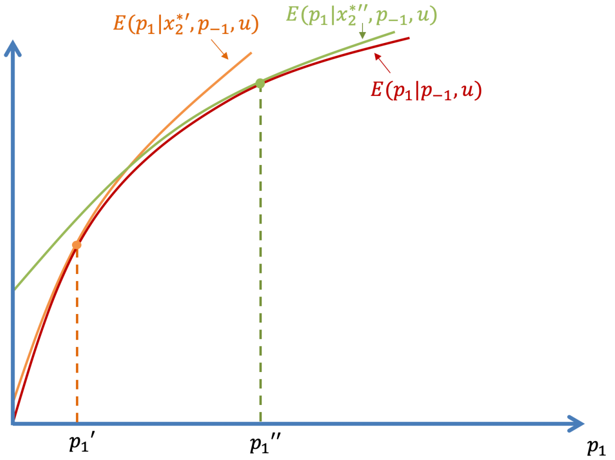
  (ii) Quantity x_2 must be kept fixed at its initially optimal level.

Example 8.2: The LeChatelier Samuelson Principal

- Let $E(p_1 \mid p_{-1}, u)$ denotes the expenditure function in situation (i) and $E(p_1 \mid x_2, p_{-1}, u)$ denotes the expenditure function in situation (ii) where x_2 must be kept fixed.

- Let $C(p_1 \mid p_{-1}, u)$ and $C(p_1 \mid x_2, p_{-1}, u)$ be the corresponding compensated demand.

Example 8.2: The LeChatelier Samuelson Principal



Example 8.2: The LeChatelier Samuelson Principal

Envelope properties of the curves:

1. First-order envelope property shows that the curves will be tangential at the point where x_2 is at its optimal value;
2. Second-order envelope property shows that $E(p_1 \mid p_{-1}, u)$ is more concave than $E(p_1 \mid x_2^{*\prime}, p_{-1}, u)$ and $E(p_1 \mid x_2^{*\prime\prime}, p_{-1}, u)$:

$$E_{p_1 p_1}(p_1 \mid p_{-1}, u) \leq E_{p_1 p_1}(p_1 \mid x_2^{*\prime}, p_{-1}, u)$$

$$\text{and } E_{p_1 p_1}(p_1 \mid p_{-1}, u) \leq E_{p_1 p_1}(p_1 \mid x_2^{*\prime\prime}, p_{-1}, u).$$

Example 8.2: The LeChatelier Samuelson Principal

- We know from (8.23) in Example 8.1 that

$$C^1_{p_1}(p_1 \mid p_{-1}, u) = E_{p_1 p_1}(p_1 \mid p_{-1}, u)$$

$$C^1_{p_1}(p_1 \mid x_2, p_{-1}, u) = E_{p_1 p_1}(p_1 \mid x_2, p_{-1}, u)$$

- Therefore, $\quad C^1_{p_1}(p_1 \mid p_{-1}, u) \leq C^1_{p_1}(p_1 \mid x_2, p_{-1}, u)$

$$\underbrace{\Longrightarrow}_{C^1_{p_1}(p_1|p_{-1},u)\leq 0, C^1_{p_1}(p_1|x_2,p_{-1},u)\leq 0} \left|C^1_{p_1}(p_1 \mid p_{-1}, u)\right| \geq \left|C^1_{p_1}(p_1 \mid x_2, p_{-1}, u)\right|$$

i.e.,

$$\left|\frac{\partial x_1}{\partial p_1}\right|_{x_2 \text{ free}} \geq \left|\frac{\partial x_1}{\partial p_1}\right|_{x_2 \text{ fixed}} \tag{8.25}$$

Example 8.2: The LeChatelier Samuelson Principal

- Fixing quantity of some other good 2 makes the compensated demand for good 1 less responsive to its own price.
- This is true irrespective of whether good 1 and good 2 are substitutes or complements.
- This is known as the LeChatelier Samuelson Principle.

Example 8.4: Use of Second-order Conditions (Part I)

- Consider a firm that buys a vector x of inputs at prices w, produced output $y = f(x)$, and sells it for revenue $R(y)$.

- The firm's profit maximization problem is

$$\max_x F(x, w) \equiv \max_x R(f(x)) - wx,$$

where w is a row vector of input prices.

Example 8.4: Use of Second-order Conditions (Part I)

- First-order necessary condition gives

$$F_x(x^*, w) = R'(f(x^*))f_x(x^*) - w = 0. \qquad (8.35)$$

- Totally differentiate (8.35), we have

$$F_{xx}(x^*, w)\mathrm{d}x^* + F_{xw}(x^*, w)\mathrm{d}w^T = 0$$

$$\Longrightarrow \mathrm{d}x^* = -F_{xx}(x^*, w)^{-1}F_{xw}(x^*, w)\mathrm{d}w^T. \qquad (8.36)$$

- From the functional form of F, we have $F_{xw}(x^*, w) = -I$.
- Plugging it into (8.36), we have

$$\mathrm{d}x^* = F_{xx}(x^*, w)^{-1}\mathrm{d}w^T \implies \mathrm{d}w\mathrm{d}x^* = \mathrm{d}wF_{xx}(x^*, w)^{-1}\mathrm{d}w^T.$$

Example 8.4: Use of Second-order Conditions (Part I)

- By the second-order necessary condition, $F_{xx}(x^*)$ is negative semi-definite.
- The inverse of a negative semi-definite matrix is also negative semi-definite.
- So $\mathrm{d}w\mathrm{d}x^* = \mathrm{d}wF_{xx}(x^*, w)^{-1}\mathrm{d}w^T \leq 0$.
- If the maximum is "regular", that is, the second-order sufficient condition is satisfied, then

$$\mathrm{d}w\mathrm{d}x^* < 0.$$

Example 8.4: Use of Second-order Conditions (Part II)

- Consider the consumer's utility maximization problem:

$$\max_{x} U(x)$$

$$\text{s.t. } px = I.$$

- The first-order necessary condition is

$$U_x(x^*) - \lambda p = 0. \tag{8.37}$$

Example 8.4: Use of Second-order Conditions (Part II)

- We want to find pure substitution effect of a price change.

- So, for the price change $\mathrm{d}p$, we compensate the consumer $\mathrm{d}I = x^{*T}\mathrm{d}p^T$.

- Under such compensation, the initial optimal bundle x^* is still affordable, i.e., x^* satisfies the new budget constraint:

$$(p + \mathrm{d}p)x^* = px^* + x^{*T}\mathrm{d}p^T = I + \mathrm{d}I.$$

Example 8.4: Use of Second-order Conditions (Part II)

- The optimal choice x^* and the Lagrange multiplier λ change as p changes.
- Totally differentiate (8.37) gives

$$U_{xx}(x^*)\mathrm{d}x^* - p^T\mathrm{d}\lambda - \lambda\mathrm{d}p^T = 0. \tag{8.38}$$

- Totally differentiate the budget constraint gives

$$p\mathrm{d}x^* + {x^*}^T\mathrm{d}p^T = \mathrm{d}I = {x^*}^T\mathrm{d}p^T \implies p\mathrm{d}x^* = 0. \tag{8.39}$$

Example 8.4: Use of Second-order Conditions (Part II)

- Combining (8.38) and (8.39), and organizing them in matrix form, we have

$$\begin{bmatrix} 0 & -p \\ -p^T & U_{xx} \end{bmatrix} \begin{bmatrix} \mathrm{d}\lambda \\ \mathrm{d}x^* \end{bmatrix} = \begin{bmatrix} 0 \\ \lambda \mathrm{d}p^T \end{bmatrix} \tag{8.40}$$

- The bordered Hessian is $\begin{bmatrix} 0 & -p \\ -p^T & U_{xx} \end{bmatrix}$.

Example 8.4: Use of Second-order Conditions (Part II)

- If the maximum is "regular" (2^{nd} order sufficient condition is satisfied), then bordered Hessian is negative definite.
- In particular, $\begin{bmatrix} \mathrm{d}\lambda & \mathrm{d}x^{*T} \end{bmatrix} \begin{bmatrix} 0 & -p \\ -p^T & U_{xx} \end{bmatrix} \begin{bmatrix} \mathrm{d}\lambda \\ \mathrm{d}x^* \end{bmatrix} < 0.$
- Combining it with (8.40), we have

$$\begin{bmatrix} \mathrm{d}\lambda & \mathrm{d}x^{*T} \end{bmatrix} \underbrace{\begin{bmatrix} 0 \\ \lambda \mathrm{d}p^T \end{bmatrix}}_{(8.40)} = \lambda \mathrm{d}p \mathrm{d}x^* < 0 \underbrace{\Longrightarrow}_{\lambda > 0} \mathrm{d}p \mathrm{d}x^* < 0.$$