

## Chapter 7

### Multivariable Unconstrained Optimization: Applications

**7.1 Competitive Firm Input Choices: Cobb-Douglas Technology** The profit function of a firm with Cobb-Douglas production function in a competitive product and inputs markets is given by

$$\max \pi = pL^\alpha K^\beta - wL - rK$$

First-Order Sufficient Condition:

$$\left. \begin{aligned} \frac{\partial \pi}{\partial L} = \alpha p L^{\alpha-1} K^{\beta} - w = 0 \\ \frac{\partial \pi}{\partial K} = \beta p L^{\alpha} K^{\beta-1} - r = 0 \end{aligned} \right\} \Rightarrow \begin{cases} VMP_L = w \\ VMP_K = r, \end{cases}$$

where  $VMP_L$  and  $VMP_K$  are the values of marginal product of labor and capital, respectively.

Second-Order Sufficient Condition: The Hessian at the critical point  $(L^*, K^*)$  is negative definite.

$$\mathbf{H}(L^*, K^*) = \begin{bmatrix} \alpha(\alpha-1)pL^{*\alpha-2}K^{*\beta} & \alpha\beta pL^{*\alpha-1}K^{*\beta-1} \\ \alpha\beta pL^{*\alpha-1}K^{*\beta-1} & \beta(\beta-1)pL^{*\alpha}K^{*\beta-2} \end{bmatrix}$$

Test the negative definiteness:

$$|\mathbf{H}_1| = \alpha(\alpha-1)pL^{*\alpha-2}K^{*\beta} < 0, \text{ if } 0 < \alpha < 1$$

$$|\mathbf{H}_2| = |\mathbf{H}| = \alpha(\alpha-1)pL^{*\alpha-2}K^{*\beta} \beta(\beta-1)pL^{*\alpha}K^{*\beta-2} - (\alpha\beta pL^{*\alpha-1}K^{*\beta-1})^2$$

$$= \alpha(\alpha-1)\beta(\beta-1)p^2L^{*2\alpha-2}K^{*2\beta-2} - \alpha^2\beta^2p^2L^{*2\alpha-2}K^{*2\beta-2} > 0$$

$$\Leftrightarrow (\alpha-1)(\beta-1) - \alpha\beta > 0, \text{ if } \beta > 0$$

$$\Leftrightarrow \alpha\beta - \alpha - \beta + 1 - \alpha\beta = 1 - \alpha - \beta > 0.$$

- That is,  $\alpha > 0$ ,  $\beta > 0$  and  $\alpha + \beta < 1$ . The production function has to be homogeneous of degree less than 1.
- That means the production function is decreasing returns to scale, and by Microeconomics, the marginal cost is positively sloped.

**Comparative Static Analysis:** The optimal solution  $(L^*, K^*)$  can be solved as functions of parameters  $\alpha$ ,  $\beta$ ,  $w$ ,  $r$ , and  $p$ . That is, assuming  $\alpha$ ,  $\beta$  unchanged, at a given particular  $(\bar{w}, \bar{r}, \bar{p})$ , the first-order sufficient conditions can be written as an implicit functions of  $(L^*, K^*)$  as follows.

$$\mathbf{f}(L^*, K^*; \bar{w}, \bar{r}, \bar{p}) = \begin{bmatrix} f^1(L^*, K^*; \bar{w}, \bar{r}, \bar{p}) \\ f^2(L^*, K^*; \bar{w}, \bar{r}, \bar{p}) \end{bmatrix} = \begin{bmatrix} \alpha \bar{p} L^{*\alpha-1} K^{*\beta} - \bar{w} \\ \beta \bar{p} L^{*\alpha} K^{*\beta-1} - \bar{r} \end{bmatrix} = \mathbf{0}$$

The Implicit Function Theorem applies here because

$$\nabla_{\begin{bmatrix} L \\ K \end{bmatrix}} \mathbf{f}(L^*, K^*; \bar{w}, \bar{r}, \bar{p}) = \mathbf{H}(L^*, K^*; \bar{w}, \bar{r}, \bar{p}),$$

is nonsingular because the Hessian is negative definite. We have

- a) there are functions  $L(w, r, p)$  and  $K(w, r, p)$  such that

$$\mathbf{f}(L(w, r, p), K(w, r, p); w, r, p) = \mathbf{0},$$

for  $|w - \bar{w}| < \varepsilon$ ,  $|r - \bar{r}| < \varepsilon$ , and  $|p - \bar{p}| < \varepsilon$  for some  $\varepsilon > 0$ .

- b)  $L(\bar{w}, \bar{r}, \bar{p}) = L^*$  and  $K(\bar{w}, \bar{r}, \bar{p}) = K^*$   
c) The gradient

$$\begin{aligned} \nabla_{\begin{bmatrix} w \\ r \\ p \end{bmatrix}} \begin{bmatrix} L(\bar{w}, \bar{r}, \bar{p}) \\ K(\bar{w}, \bar{r}, \bar{p}) \end{bmatrix} &= - \left[ \nabla_{\begin{bmatrix} L \\ K \end{bmatrix}} \mathbf{f}(L^*, K^*; \bar{w}, \bar{r}, \bar{p}) \right]^{-1} \nabla_{\begin{bmatrix} w \\ r \\ p \end{bmatrix}} \mathbf{f}(L^*, K^*; \bar{w}, \bar{r}, \bar{p}) \\ &= - \mathbf{H}(L^*, K^*; \bar{w}, \bar{r}, \bar{p})^{-1} \nabla_{\begin{bmatrix} w \\ r \\ p \end{bmatrix}} \mathbf{f}(L^*, K^*; \bar{w}, \bar{r}, \bar{p}) \\ &= - \begin{bmatrix} \alpha(\alpha-1) \bar{p} L^{*\alpha-2} K^{*\beta} & \alpha \beta \bar{p} L^{*\alpha-1} K^{*\beta-1} \\ \alpha \beta \bar{p} L^{*\alpha-1} K^{*\beta-1} & \beta(\beta-1) \bar{p} L^{*\alpha} K^{*\beta-2} \end{bmatrix}^{-1} \begin{bmatrix} -1 & 0 & \alpha L^{*\alpha-1} K^{*\beta} \\ 0 & -1 & \beta L^{*\alpha} K^{*\beta-1} \end{bmatrix} \end{aligned}$$

and by Cramer's Rule

$$\begin{aligned} \frac{\partial L(\bar{w}, \bar{r}, \bar{p})}{\partial w} &= - \frac{\begin{vmatrix} -1 & \alpha\beta\bar{p}L^{*\alpha-1}K^{*\beta-1} \\ 0 & \beta(\beta-1)\bar{p}L^{*\alpha}K^{*\beta-2} \end{vmatrix}}{\begin{vmatrix} \alpha(\alpha-1)\bar{p}L^{*\alpha-2}K^{*\beta} & \alpha\beta\bar{p}L^{*\alpha-1}K^{*\beta-1} \\ \alpha\beta\bar{p}L^{*\alpha-1}K^{*\beta-1} & \beta(\beta-1)\bar{p}L^{*\alpha}K^{*\beta-2} \end{vmatrix}} \\ &= \frac{\beta(\beta-1)\bar{p}L^{*\alpha}K^{*\beta-2}}{|\mathbf{H}|} < 0. \end{aligned}$$

We can similarly find all other partial derivatives.

**HW** Baldani, p. 216, #8.2, 8.3, 8.4

**7.2 Competitive Firm Input Choices: General Production Technology** The same firm as in 7.1 but now with a generic production function  $f(L, K)$  will maximize the profit function

$$\max \pi = pf(L, K) - wL - rK.$$

First-Order Sufficient Condition:

$$\left. \begin{aligned} \frac{\partial \pi}{\partial L} = pf_L - w = 0 \\ \frac{\partial \pi}{\partial K} = pf_K - r = 0 \end{aligned} \right\} \Rightarrow \begin{cases} VMP_L = w \\ VMP_K = r \end{cases}$$

Second-Order Sufficient Condition:

$$\mathbf{H}(L^*, K^*) = \begin{bmatrix} pf_{LL} & pf_{LK} \\ pf_{KL} & pf_{KK} \end{bmatrix}$$

$$|\mathbf{H}_1| = pf_{LL} < 0$$

$$|\mathbf{H}_2| = |\mathbf{H}| = p^2(f_{LL}f_{KK} - f_{KL}^2) > 0$$

By Implicit Function Theorem:

$$\begin{aligned}
 \nabla_{\begin{bmatrix} w \\ r \\ p \end{bmatrix}} \begin{bmatrix} L(\bar{w}, \bar{r}, \bar{p}) \\ K(\bar{w}, \bar{r}, \bar{p}) \end{bmatrix} &= - \left[ \nabla_{\begin{bmatrix} L \\ K \end{bmatrix}} \mathbf{f}(L^*, K^*; \bar{w}, \bar{r}, \bar{p}) \right]^{-1} \nabla_{\begin{bmatrix} w \\ r \\ p \end{bmatrix}} \mathbf{f}(L^*, K^*; \bar{w}, \bar{r}, \bar{p}) \\
 &= -\mathbf{H}(L^*, K^*; \bar{w}, \bar{r}, \bar{p})^{-1} \nabla_{\begin{bmatrix} w \\ r \\ p \end{bmatrix}} \mathbf{f}(L^*, K^*; \bar{w}, \bar{r}, \bar{p}) \\
 &= - \begin{bmatrix} pf_{LL} & pf_{LK} \\ pf_{KL} & pf_{KK} \end{bmatrix}^{-1} \begin{bmatrix} -1 & 0 & f_L \\ 0 & -1 & f_K \end{bmatrix} \\
 \frac{\partial L(\bar{w}, \bar{r}, \bar{p})}{\partial w} &= - \frac{\begin{vmatrix} -1 & pf_{LK} \\ 0 & pf_{KK} \end{vmatrix}}{\begin{vmatrix} pf_{LL} & pf_{LK} \\ pf_{KL} & pf_{KK} \end{vmatrix}} \\
 &= \frac{pf_{KK}}{|\mathbf{H}|} < 0.
 \end{aligned}$$

- The demand for labor has a negative slope with respect to wage rate if the capital exhibits diminishing returns-- $f_{KK} < 0$  means the slope of the marginal product is negative.

**HW** Determine the signs of  $\frac{\partial L(\bar{w}, \bar{r}, \bar{p})}{\partial r}$  and

$$\frac{\partial L(\bar{w}, \bar{r}, \bar{p})}{\partial p}.$$

**HW** Baldani, p. 216, #8.5.

**7.3 Multiplant Firm** A firm with  $n$  plants.

$$TC_i(q_i) = C_i(q_i), \quad q_i = \text{quantity produced at plant } i.$$

$$TR = R(Q) = P(Q)Q$$

$$Q = \sum_{i=1}^n q_i$$

$$\pi(\mathbf{q}) = R(Q) - \sum_{i=1}^n C_i(q_i) = P(Q)Q - \sum_{i=1}^n C_i(q_i)$$

$$\pi_i(\mathbf{q}^*) = P'(Q^*)Q^* + P(Q^*) - C'_i(q_i^*) = 0, \quad i = 1, 2, \dots, n.$$

The last equality is the first-order sufficient condition and it implies that the marginal cost of each plant is equal to the marginal revenue, i.e.,

$$C'_1(q_1^*) = C'_2(q_2^*) = \dots = C'_n(q_n^*) = P'(Q^*)Q^* + P(Q^*) = R'(Q) = MR(Q)$$

The Hessian is given by

$$\mathbf{H}(\mathbf{q}^*) = \begin{bmatrix} R'' - C''_1 & R'' & \dots & R'' \\ R'' & R'' - C''_2 & \dots & \vdots \\ \vdots & \dots & \ddots & R'' \\ R'' & \dots & R'' & R'' - C''_n \end{bmatrix},$$

where  $R'' = P''(Q^*)Q^* + 2P'(Q^*)$ . Test of negative definiteness:

$$\begin{aligned} |\mathbf{H}_1| &= R'' - C''_1 < 0 \\ |\mathbf{H}_2| &= (R'' - C''_1)(R'' - C''_2) - R''^2 = C''_1 C''_2 - R''(C''_1 + C''_2) > 0, \text{ and} \\ (-1)^i |\mathbf{H}_i| &> 0, \quad i = 3, 4, \dots, n. \end{aligned}$$

For perfect competition,  $R'' = 0$  and

$$\mathbf{H}(\mathbf{q}^*) = \begin{bmatrix} -C''_1 & 0 & \dots & 0 \\ 0 & -C''_2 & \dots & \vdots \\ \vdots & \dots & \ddots & 0 \\ 0 & \dots & 0 & -C''_n \end{bmatrix},$$

which is negative definite if  $C''_i > 0$ ,  $i = 1, 2, \dots, n$ .

For monopoly, this is not necessarily true if the plants have increasing return to scale ( $C''_i < 0$ ).

**HW** Show that the Hessian

$$\mathbf{H}(\mathbf{q}^*) = \begin{bmatrix} R'' - C''_1 & R'' & \dots & R'' \\ R'' & R'' - C''_2 & \dots & \vdots \\ \vdots & \dots & \ddots & R'' \\ R'' & \dots & R'' & R'' - C''_n \end{bmatrix}$$

is not necessarily negative definite if  $C_i'' < 0$ ,  $i = 1, 2, \dots, n$ . (Hint: Show for  $n = 2$ )

**HW** Show that the Hessian  $\mathbf{H}(\mathbf{q}^*)$  above is negative definite if  $R'' < 0$  and  $C_i'' > 0$ ,  $i = 1, 2, \dots, n$ .

**Solution:** If we assume instead that  $R'' > 0$  and  $C_i'' > 0$  then we can equivalently show that

$$\mathbf{H}(\mathbf{q}^*) = \begin{bmatrix} R'' + C_1'' & R'' & \cdots & R'' \\ R'' & R'' + C_2'' & \cdots & \vdots \\ \vdots & \cdots & \ddots & R'' \\ R'' & \cdots & R'' & R'' + C_n'' \end{bmatrix}$$

is positive definite. We will prove by induction. If  $n = 2$ , then  $|\mathbf{H}_1| = R'' + C_1'' > 0$  and

$$|\mathbf{H}_2| = (R'' + C_1'')(R'' + C_2'') - R''^2 = C_1''C_2'' + R''(C_1'' + C_2'') > 0.$$

Note that  $|\mathbf{H}_2| > 0$  even when  $C_1'' = 0$ . We can then state the induction hypothesis that  $|\mathbf{H}(\mathbf{q}^*)_{n \times n}| > 0$  when  $R'' > 0$ ,  $C_1'' \geq 0$  and  $C_i'' > 0$ ,  $i = 2, 3, \dots, n$ . We need to prove

$$\left| \mathbf{H}(\mathbf{q}^*)_{(n+1) \times (n+1)} \right| = \begin{vmatrix} R'' + C_1'' & R'' & \cdots & R'' \\ R'' & R'' + C_2'' & \cdots & \vdots \\ \vdots & \cdots & \ddots & R'' \\ R'' & \cdots & R'' & R'' + C_{n+1}'' \end{vmatrix} > 0.$$

By a property of determinant,

$$\begin{aligned}
 \left| \mathbf{H}(\mathbf{q}^*)_{(n+1) \times (n+1)} \right| &= \begin{vmatrix} C_1'' & R'' & \cdots & \cdots & R'' \\ -C_2'' & R'' + C_2'' & \cdots & \cdots & \vdots \\ 0 & R'' & \ddots & \cdots & \vdots \\ \vdots & \cdots & \cdots & R'' + C_{n+1}'' & R'' \\ 0 & R'' & \cdots & R'' & R'' + C_{n+1}'' \end{vmatrix} \\
 &= C_1'' \begin{vmatrix} R'' + C_2'' & R'' & \cdots & R'' \\ R'' & R'' + C_3'' & \cdots & \vdots \\ \vdots & \cdots & \ddots & R'' \\ R'' & \cdots & R'' & R'' + C_{n+1}'' \end{vmatrix} \\
 &\quad + (-1)^{2+1} (-C_2'') \begin{vmatrix} R'' & R'' & \cdots & R'' \\ R'' & R'' + C_3'' & \cdots & \vdots \\ \vdots & \cdots & \ddots & R'' \\ R'' & \cdots & R'' & R'' + C_{n+1}'' \end{vmatrix},
 \end{aligned}$$

which is positive by the induction hypothesis.

- If we assume the functional forms of TR and TC explicitly with parameters, we can perform the sensitivity analysis using the Implicit Function Theorem.

**HW** Baldani, p. 217, #8.9

**7.4 Multi-Market Monopoly** A monopoly has two separate markets with similar demands  $D_1(q_1) = \alpha P(q_1)$  and  $D_2(q_2) = P(q_2)$ , where  $q_1$  and  $q_2$  are quantities sold in the two markets. Assume that  $\alpha > 1$  so that market 1 is more important. The total revenues earned are:

$$\begin{aligned}
 R_1(q_1) &= \alpha P(q_1) q_1 \\
 R_2(q_1) &= P(q_2) q_2.
 \end{aligned}$$

The total cost of output  $Q = q_1 + q_2$  is

$$TC(Q) = C(Q) + tq_2,$$

where  $t$  is the extra cost per unit to sell in market 2. The profit function is thus

$$\begin{aligned}\pi(q_1, q_2) &= R_1(q_1) + R_2(q_2) - C(Q) - tq_2 \\ &= \alpha P(q_1)q_1 + P(q_2)q_2 - C(Q) - tq_2.\end{aligned}$$

First-order Sufficient Condition:

$$\begin{aligned}\pi_1(q_1^*, q_2^*) &= R_1'(q_1^*) - C'(q_1^* + q_2^*) \\ &= \alpha P'(q_1^*)q_1^* + \alpha P(q_1^*) - C'(q_1^* + q_2^*) = 0 \\ \pi_2(q_1^*, q_2^*) &= R_2'(q_2^*) - C'(q_1^* + q_2^*) - t \\ &= P'(q_2^*)q_2^* + P(q_2^*) - C'(q_1^* + q_2^*) - t = 0.\end{aligned}$$

We have  $q_1^* > q_2^*$ . (Why?)

Second-order Sufficient Condition: The Hessian is given by

$$\mathbf{H}(q_1^*, q_2^*) = \begin{bmatrix} \alpha R_1'' - C'' & -C'' \\ -C'' & R_2'' - C'' \end{bmatrix},$$

where

$$\begin{aligned}\alpha R_1'' &= \alpha P''(q_1^*)q_1^* + 2\alpha P'(q_1^*) \\ R_2'' &= P''(q_2^*)q_2^* + 2P'(q_2^*).\end{aligned}$$

Test of negative definiteness of the Hessian: If  $R_1'' < 0$ ,  $R_2'' < 0$  and  $C'' > 0$

$$\begin{aligned}|\mathbf{H}_1(q_1^*, q_2^*)| &= \alpha R_1'' - C'' < 0 \\ |\mathbf{H}(q_1^*, q_2^*)| &= (\alpha R_1'' - C'')(R_2'' - C'') - C''^2 = \alpha R_1''R_2'' - \alpha R_1''C'' - R_2''C'' > 0.\end{aligned}$$

**Comparative Static Analysis:** Write the first-order condition as implicit functions.

$$\nabla_{\mathbf{q}} \pi(q_1^*, q_2^*; \alpha, t) = \begin{bmatrix} \pi_1(q_1^*, q_2^*; \alpha, t) \\ \pi_2(q_1^*, q_2^*; \alpha, t) \end{bmatrix} = \begin{bmatrix} \alpha R_1'(q_1^*) - C'(q_1^* + q_2^*) \\ R_2'(q_2^*) - C'(q_1^* + q_2^*) - t \end{bmatrix} = \mathbf{0}$$

Assuming  $\nabla_{\mathbf{q}}^2 \pi(q_1^*, q_2^*; \alpha, t) = \mathbf{H}(q_1^*, q_2^*; \alpha, t)$  being a nonsingular matrix, the Implicit Function Theorem yields

$$\begin{aligned} \nabla_{\begin{bmatrix} \alpha \\ t \end{bmatrix}} \mathbf{q}^* &= -\nabla_{\mathbf{q}}^2 \pi(q_1^*, q_2^*; \alpha, t)^{-1} \nabla_{\begin{bmatrix} \alpha \\ t \end{bmatrix}} (\nabla_{\mathbf{q}} \pi(q_1^*, q_2^*; \alpha, t)) \\ &= -\mathbf{H}(q_1^*, q_2^*)^{-1} \begin{bmatrix} R_1'(q_1^*) & 0 \\ 0 & -1 \end{bmatrix} \\ &= -\begin{bmatrix} \alpha R_1'' - C'' & -C'' \\ -C'' & R_2'' - C'' \end{bmatrix}^{-1} \begin{bmatrix} R_1'(q_1^*) & 0 \\ 0 & -1 \end{bmatrix}. \end{aligned}$$

By Cramer's Rule, if  $R_1'(q_1^*) > 0$ ,  $C'' > 0$

$$\begin{aligned} \frac{\partial q_1^*}{\partial \alpha} &= -\frac{\begin{vmatrix} R_1'(q_1^*) & -C'' \\ 0 & R_2'' - C'' \end{vmatrix}}{\begin{vmatrix} \mathbf{H}(q_1^*, q_2^*) \end{vmatrix}} = -\frac{R_1'(q_1^*)(R_2'' - C'')}{\begin{vmatrix} \mathbf{H}(q_1^*, q_2^*) \end{vmatrix}} > 0, \\ \frac{\partial q_1^*}{\partial t} &= -\frac{\begin{vmatrix} 0 & -C'' \\ -1 & R_2'' - C'' \end{vmatrix}}{\begin{vmatrix} \mathbf{H}(q_1^*, q_2^*) \end{vmatrix}} = \frac{C''}{\begin{vmatrix} \mathbf{H}(q_1^*, q_2^*) \end{vmatrix}} > 0. \end{aligned}$$

Thus,

$$\begin{aligned} dq_1^* &= \frac{\partial q_1^*}{\partial \alpha} d\alpha + \frac{\partial q_1^*}{\partial t} dt \\ &= \frac{-R_1'(q_1^*)(R_2'' - C'')d\alpha + C''dt}{\begin{vmatrix} \mathbf{H}(q_1^*, q_2^*) \end{vmatrix}}, \end{aligned}$$

and similarly,

$$\begin{aligned} dq_2^* &= \frac{\partial q_2^*}{\partial \alpha} d\alpha + \frac{\partial q_2^*}{\partial t} dt \\ &= \frac{-R_1'(q_1^*)C''d\alpha + (\alpha R_1'' - C'')dt}{\begin{vmatrix} \mathbf{H}(q_1^*, q_2^*) \end{vmatrix}}. \end{aligned}$$

We have

$$dQ^* = dq_1^* + dq_2^* = \frac{-R_1'(q_1^*)R_2''d\alpha + \alpha R_1''dt}{|\mathbf{H}(q_1^*, q_2^*)|}.$$

HW Baldani, p. 217, #8.10, 8.11

## 7.5 Statistical Estimation: Linear Regression

Recall matrix differentiation,

$$\begin{aligned}\nabla \mathbf{c}^T \mathbf{x} &= \mathbf{c} \\ \nabla \alpha \mathbf{x}^T \mathbf{x} &= 2\alpha \mathbf{x} \\ \nabla \mathbf{x}^T \mathbf{A} \mathbf{x} &= 2\mathbf{A} \mathbf{x},\end{aligned}$$

where  $\mathbf{A}$  is a symmetric matrix.

Linear regression model: *Least Squares* The dependent variable  $y$  is determined linearly by the independent variables  $x_j, j = 1, 2, \dots, k$ , with some random error  $\varepsilon$ .

Suppose there are  $n$  observation, we have for  $i = 1, 2, \dots, n$ ,

$$y_i = \beta_0 + \beta_1 x_i^1 + \beta_2 x_i^2 + \dots + \beta_k x_i^k + \varepsilon_i,$$

where  $\beta_j, j = 0, 1, 2, \dots, k$ , are unknown parameters. In matrix form,

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

where  $\mathbf{y}, \boldsymbol{\varepsilon} \in \mathbf{R}^n$ ,  $\boldsymbol{\beta} \in \mathbf{R}^{k+1}$ , and  $\mathbf{X} \in \mathbf{R}^{n \times (k+1)}$ . If we estimate  $\boldsymbol{\beta}$  by some  $\mathbf{b} \in \mathbf{R}^{k+1}$ , the estimate of  $\mathbf{y}$  is thus, and the error of estimation of  $\mathbf{y}$  is  $\hat{\mathbf{y}} = \mathbf{X}\mathbf{b}$ . The determination of  $\mathbf{b}$  by the Least Squares Method is the choice of  $\mathbf{b}$  such that the sum of squares of the error of estimation  $SSR = \mathbf{e}^T \mathbf{e} = \sum_{i=1}^n e_i^2$  is minimized. The minimization is thus given by

$$\begin{aligned}\min_{\mathbf{b} \in \mathbb{R}^{k+1}} f(\mathbf{b}) &= \mathbf{e}^T \mathbf{e} = (\mathbf{y} - \hat{\mathbf{y}})^T (\mathbf{y} - \hat{\mathbf{y}}) \\ &= (\mathbf{y} - \mathbf{X}\mathbf{b})^T (\mathbf{y} - \mathbf{X}\mathbf{b}) \\ &= \mathbf{y}^T \mathbf{y} - 2\mathbf{y}^T \mathbf{X}\mathbf{b} + \mathbf{b}^T \mathbf{X}^T \mathbf{X}\mathbf{b}.\end{aligned}$$

By the first-order sufficient condition, the critical solution is the solution that

$$\nabla f(\hat{\mathbf{b}}) = -2\mathbf{X}^T \mathbf{y} + 2\mathbf{X}^T \mathbf{X}\hat{\mathbf{b}} = \mathbf{0},$$

and thus the critical solution is given by

$$\hat{\mathbf{b}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y},$$

and the second-order sufficient condition requires that

$$\nabla^2 f(\mathbf{b}) = 2\mathbf{X}^T \mathbf{X},$$

be positive definite. The square symmetric matrix  $\mathbf{X}^T \mathbf{X}$  is always positive semidefinite. Why? Now, if it is also positive definite, the solution, as given by the first-order sufficient condition is uniquely defined because  $(\mathbf{X}^T \mathbf{X})^{-1}$  exists (Why?), is a strict local minimum point.

Can we say that is a strict global minimum?

**HW** What will happen if one of the independent variable is just a linear combination of some of the other independent variable? For example, what if  $x^1 = 1.5x^2$ ?