Lecture 6: Production Functions, Cost Minimization, and Lagrange Multipliers

6.1 Cost minimization and convex analysis

When there is a production function f for a single output producer with n inputs, the input requirement set for producing output level y is

$$V(y) = \{x \in \boldsymbol{R}^{\mathrm{n}} : f(x) \geqslant y\}.$$

The cost function for the producer facing wage vector $w = (w_1, \dots, w_n)$ is the support function

$$c(w, y) = \inf\{w \cdot x : f(x) \geqslant y\}.$$

The Support Function Theorem tells us that holding y fixed, c is concave in w, and if $\hat{x}$ is the unique cost minimizer, then

$$\frac{\partial c}{\partial w_i} = x_i^*$$

and when c is twice differentiable in w, the Hessian matrix

$$\left[\frac{\partial^2 c}{\partial w_i \partial w_j}\right] = \left[\frac{\partial x_i^*}{\partial w_j}\right] \quad \text{is symmetric and negative semidefinite.}$$

But the Support Function Theorem doesn't tell us that c is twice differentiable or how it depends on y. When the production function f is differentiable, we can use the Lagrange Multiplier Theorem to find additional results.

6.2 Classical Lagrange Multiplier Theorem

6.2.1 Definition *A* ***constrained optimization problem*** *is characterized by an* ***objective function*** *f and m* ***constraint functions****, $g_1, \dots, g_m$. The constraints take the form of either* ***equality constraints*** *($g_i(x) = 0$, $i = 1, \dots, m$) or* ***inequality constraints*** *($g_i(x) \geqslant 0$, $i = 1, \dots, m$).*

A point x^* is a ***constrained local maximizer*** *of f subject to the equality constraints $g_1(x) = 0$, $g_2(x) = 0$, ..., $g_m(x) = 0$ in some neighborhood W of x^* if x^* satisfies the constraints and also satisfies $f(x^*) \geqslant f(x)$ for all $x \in W$ that also satisfy the constraints.*

A ***constrained local minimizer*** *is defined similarly, and the case of inequality constraints is also dealt with as you should expect.*

Frequently, the true constraints are inequality constraints, but we can see that at an extremum, the will be satisfied as equalities, and we may write them as equality constraints.

Associated with such a problem is a function called the ***Lagrangean****:*

$$L(x; \lambda) = f(x) + \lambda \cdot g(x) = f(x) + \lambda_1 g_1(x) + \cdots + \lambda_m g_m(x).$$

The numbers λ_i are called ***Lagrange multipliers****.*

6.2.2 Lagrange Multiplier Theorem *Let $X \subset \boldsymbol{R}^{\mathrm{n}}$, and let $f, g_1, \ldots, g_m \colon X \to \boldsymbol{R}$ be continuous. Let x^* be an interior constrained local maximizer of f subject to $g(x) = 0$. Suppose f, $g_1, \ldots, g_m$ are differentiable at x^*, and that the Lagrange Constraint Qualification holds, that is, ${g_1}'(x^*), \ldots, {g_m}'(x^*)$ are linearly independent.*

Then there exist real numbers $\lambda_1^*, \ldots, \lambda_m^*$, such that

$$f'(x^*) + \sum_{i=1}^{m} \lambda_i^* {g_i}'(x^*) = 0.$$

6.2.3 Remark The way I wrote the Lagrangean above is the preferred way to write the Lagrangean for maximization. For minimization, the preferred way to write the Lagrangean is

$$L(x; \lambda) = f(x) - \lambda \cdot g(x) = f(x) - \lambda_1 g_1(x) + \cdots + \lambda_m g_m(x).$$

There is no need to do this unless you care bout the sign of the multipliers (and I do). Also, the constraint $g(x) = 0$ is the same as the constraint $-g(x) = 0$, so when deciding how to write the constraint, if there is a true inequality constraint $g(x) \geqslant 0$ that we know a priori must hold with equality, write the equality constraint as $g(x) = 0$. This will become clearer wen you look at the examples in what follows.

6.3 Using the LMT

Since the LMT tells us what is true at the optimum, we can sometimes use the necessary conditions to pin down what the optimum is. For example, the Cobb–Douglas production function is given by

$$y = f(x) = \gamma x_1^{\alpha_1} \cdots x_n^{\alpha_n},$$

where each $\alpha_i > 0$, $i = 1, \ldots, n$. It is homogeneous of degree

$$\alpha = \sum_{i=1}^{n} \alpha_i.$$

This function was proposed by Charles Cobb and Paul Douglas [3] as a model for U.S. GDP, and it works surprisingly well empirically. When $\gamma = \alpha = 1$, it is a weighted geometric mean of the inputs.

To find the associated cost function we start by writing the Lagrangean for a minimum, where the true constraint is $f(x) - y \geqslant 0$, as

$$L(x;\lambda) = w \cdot x - \lambda(\gamma x_1^{\alpha_1} \cdots x_n^{\alpha_n} - y)$$

The first order conditions, using the binding constraint $y = \gamma x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ are:

$$\frac{\partial L}{\partial x_i} = w_i - \lambda \alpha_i \frac{y}{x_i} = 0 \quad i = 1, \dots, n.$$

So

$$x_i = \lambda \alpha_i \frac{y}{w_i} \quad i = 1, \dots, n. \tag{1}$$

But $y = \gamma {x_1}^{\alpha_1} \cdots {x_n}_n^{\alpha}$, so

$$y = \gamma \prod_{i=1}^n \left(\lambda \alpha_i \frac{y}{w_i}\right)^{\alpha_i} = \gamma \lambda^\alpha y^\alpha \prod_{i=1}^n \left(\frac{\alpha_i}{w_i}\right)^{\alpha_i}.$$

Solving this for λ gives

$$\begin{aligned}\hat{\lambda} &= \left[\gamma y^{\alpha-1} \prod_{i=1}^n \left(\frac{\alpha_i}{w_i}\right)^{\alpha_i}\right]^{-1/\alpha} \\ &= \gamma^{-1/\alpha} y^{(1-\alpha)/\alpha} \left(\prod_{i=1}^n \alpha_i^{-\alpha_i/\alpha}\right)\left(\prod_{i=1}^n w^{\alpha_i/\alpha}\right)\end{aligned}$$

To simplify notation a bit, set

$$\beta_i = \frac{\alpha_i}{\alpha},$$
$$b = \gamma^{\frac{-1}{\alpha}} \cdot \prod_i \alpha_i^{-\beta_i},$$

so

$$\hat{\lambda} = b y^{(1-\alpha)/\alpha} \prod_{i=1}^n w_i^{\beta_i}.$$

Substituting this for λ in (1) gives the conditional factor demands

$$\begin{aligned}\hat{x}_j(y,w) &= b y^{(1-\alpha)/\alpha} \prod_{i=1}^n w_i^{\beta_i} \alpha_j \frac{y}{w_j} \\ &= \frac{\alpha_j}{w_j} b y^{1/\alpha} \prod_{i=1}^n w_i^{\beta_i},\end{aligned}$$

for $j = 1, \dots, n$. So the cost function is

$$c(y,w) = \alpha b y^{1/\alpha} \prod_{i=1}^n w_i^{\beta_i},$$

which is a Cobb–Douglas function of ws.

Note that

$$\frac{\partial c(y,w)}{\partial y} = by^{(1-\alpha)/\alpha} \prod_{i=1}^{n} w_i^{\beta_i} = \hat{\lambda},$$

and

$$\frac{\partial c(y,w)}{\partial w_j} = \alpha \frac{\beta_j}{w_j} by^{(1-\alpha)/\alpha} \prod_{i=1}^{n} w_i^{\beta_i} = \hat{x}_j(y,w).$$

6.4 Second Order Conditions

The following result may be found in my on-line notes [2, Theorem 274].

6.4.1 Theorem (Necessary Second Order Conditions for a Maximum) *Let $U \subset \boldsymbol{R}^{\mathrm{n}}$ and let $x^* \in \operatorname{int} U$. Let $f, g_1, \ldots, g_m \colon U \to \boldsymbol{R}$ be C^2, and suppose x^* is a local constrained maximizer of f subject to $g(x) = 0$. Define the Lagrangean $L(x,\lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x)$. Assume that ${g_1}'(x^*), \ldots, {g_m}'(x^*)$ are linearly independent, so the conclusion of the Lagrange Multiplier Theorem holds, that is, there are $\lambda_1^*, \ldots, \lambda_m^*$ satisfying the first order conditions*

$$L'_x(x^*,\lambda^*) = f'(x^*) + \sum_{i=1}^{m} \lambda_i^* {g_i}'(x^*) = 0.$$

Then

$$\sum_{i=1}^{n}\sum_{j=1}^{n} D_{ij}L(x^*,\lambda^*) v_i v_j \leqslant 0,$$

for all $v \neq 0$ satisfying ${g_i}'(x^*) \cdot v = 0$, $i = 1, \ldots, m$.

Since minimizing f is the same as maximizing $-f$, we do not need any new results for minimization, but there a few things worth pointing out.

The Lagrangean for maximizing $-f$ subject to $g_i = 0$, $i = 1, \ldots, m$ is

$$-f(x) + \sum_{i=1}^{m} \lambda_i g_i(x),$$

The second order condition for maximizing $-f$ is that

$$\sum_{i=1}^{n}\sum_{j=1}^{n} \left(-D_{ij} f(x^*) + \sum_{i=1}^{m} \lambda^* D_{ij} g(x^*) \right) v_i v_j \leqslant 0,$$

for all $v \neq 0$ satisfying ${g_i}'(x^*) \cdot v = 0$, $i = 1, \ldots, m$. This can be rewritten as

$$\sum_{i=1}^{n}\sum_{j=1}^{n} \left(D_{ij} f(x^*) - \sum_{i=1}^{m} \lambda^* D_{ij} g(x^*) \right) v_i v_j \geqslant 0,$$

which explains why I prefer to write the Lagrangean for a minimization problem as

$$L(x, \lambda) = f(x) - \sum_{i=1}^{m} \lambda_i g_i(x).$$

The first order conditions will be exactly the same. For the second order conditions we have the following.

6.4.2 Theorem (Necessary Second Order Conditions for a Minimum) *Let $U \subset \boldsymbol{R}^{\mathrm{n}}$ and let $x^* \in \operatorname{int} U$. Let $f, g_1, \ldots, g_m \colon U \to \boldsymbol{R}$ be C^2, and suppose x^* is a local constrained minimizer of f subject to $g(x) = 0$. Define the Lagrangean*

$$L(x, \lambda) = f(x) - \sum_{i=1}^{m} \lambda_i g_i(x).$$

Assume that ${g_1}'(x^*), \ldots, {g_m}'(x^*)$ are linearly independent, so the conclusion of the Lagrange Multiplier Theorem holds, that is, there are $\lambda_1^*, \ldots, \lambda_m^*$ satisfying the first order conditions

$$L'_x(x^*, \lambda^*) = f'(x^*) - \sum_{i=1}^{m} \lambda_i^* {g_i}'(x^*) = 0.$$

Then

$$\sum_{i=1}^{n} \sum_{j=1}^{n} D_{ij} L(x^*, \lambda^*) v_i v_j \geqslant 0,$$

for all $v \neq 0$ satisfying ${g_i}'(x^*) \cdot v = 0$, $i = 1, \ldots, m$.

6.5 Envelope Theorem for Constrained Extrema

6.5.1 Theorem (Envelope Theorem for Constrained Maximization) *Let $X \subset \boldsymbol{R}^{\mathrm{n}}$ and $P \subset \boldsymbol{R}^{\ell}$ be open, and let $f, g_1, \ldots, g_m \colon X \times P \to \boldsymbol{R}$ be C^1. For each $p \in P$, let $x^*(p)$ be an interior constrained local maximizer of $f(x,p)$ subject to $g(x,p) = 0$. Define the Lagrangean*

$$L(x, \lambda; p) = f(x, p) + \sum_{i=1}^{m} \lambda_i g_i(x, p),$$

and assume that the conclusion of the Lagrange Multiplier Theorem holds for each p, that is, there exist real numbers $\lambda_1^*(p), \ldots, \lambda_m^*(p)$, such that the first order conditions

$$\frac{\partial L\Big(x^*(p), \lambda^*(p), p\Big)}{\partial x} = f'_x\Big(x^*(p), p\Big) + \sum_{i=1}^{m} \lambda_i^*(p) {g_i}'_x\Big(x^*(p), p\Big) = 0$$

are satisfied. Assume that $x^* \colon P \to X$ and $\lambda^* \colon P \to \boldsymbol{R}^{\mathrm{m}}$ are C^1. Set

$$V(p) = f\left(x^*(p), p\right).$$

Then V is C^1 and

$$\frac{\partial V(p)}{\partial p_j} = \frac{\partial L\Big(x^*(p), \lambda^*(p), p\Big)}{\partial p_j} = \frac{\partial f(x^*, p)}{\partial p_j} + \sum_{i=1}^m \lambda_i^*(p) \frac{\partial g_i(x^*, p)}{\partial p_j}.$$

Proof: Clearly V is C^1 as the composition of C^1 functions. Since x^* satisfies the constraints, we have

$$V(p) = f\left(x^*(p), p\right) = f\left(x^*(p), p\right) + \sum_{i=1}^m \lambda_i^*(p) g_i(x^*, p).$$

Therefore by the chain rule,

$$\begin{aligned}
\frac{\partial V(p)}{\partial p_j} &= \left(\sum_{k=1}^n \frac{\partial f(x^*, p)}{\partial x_k} \frac{\partial x^{*k}}{\partial p_j}\right) + \frac{\partial f(x^*, p)}{\partial p_j} \\
&\quad + \sum_{i=1}^m \left\{ \frac{\partial \lambda_i^*(p)}{\partial p_j} g_i(x^*, p) + \lambda^*(p) \left[\left(\sum_{k=1}^n \frac{\partial g_i(x^*, p)}{\partial x_k} \frac{\partial x^{*k}}{\partial p_j} \right) + \frac{\partial g_i(x^*, p)}{\partial p_j} \right] \right\} \\
&= \frac{\partial f(x^*, p)}{\partial p_j} + \sum_{i=1}^m \lambda_i^*(p) \frac{\partial g_i(x^*, p)}{\partial p_j} \\
&\quad + \sum_{i=1}^m \frac{\partial \lambda_i^*(p)}{\partial p_j} g_i(x^*, p) \qquad (2) \\
&\quad + \sum_{k=1}^n \left(\frac{\partial f(x^*, p)}{\partial x_k} + \sum_{i=1}^m \lambda^*(p) \frac{\partial g_i(x^*, p)}{\partial x_k} \right) \frac{\partial x^{*k}}{\partial p_j}. \qquad (3)
\end{aligned}$$

The theorem now follows from the fact that both terms (2) and (3) are zero. Term (2) is zero since each g_i is zero as x^* satisfies the constraints, and term (3) is zero, as the first order conditions imply that each $\frac{\partial f(x^*,p)}{\partial x_k} + \sum_{i=1}^m \lambda^*(p) \frac{\partial g_i(x^*,p)}{\partial x_k} = 0$. ∎

6.5.2 Theorem (Envelope Theorem for Minimization) *Let $X \subset \boldsymbol{R}^n$ and $P \subset \boldsymbol{R}^\ell$ be open, and let $f, g_1, \dots, g_m \colon X \times P \to \boldsymbol{R}$ be C^1. For each $p \in P$, let $x^*(p)$ be an interior constrained local maximizer of $f(x, p)$ subject to $g(x, p) = 0$. Define the Lagrangean*

$$L(x, \lambda; p) = f(x, p) - \sum_{i=1}^m \lambda_i g_i(x, p),$$

and assume that the conclusion of the Lagrange Multiplier Theorem holds for each p, that is, there exist real numbers $\lambda_1^*(p), \dots, \lambda_m^*(p)$, such that the first order conditions

$$\frac{\partial L\Big(x^*(p), \lambda^*(p), p\Big)}{\partial x} = f_x'\Big(x^*(p), p\Big) - \sum_{i=1}^m \lambda_i^*(p) g_{i\,x}'\Big(x^*(p), p\Big) = 0$$

are satisfied. Assume that $x^*\colon P \to X$ *and* $\lambda^*\colon P \to \boldsymbol{R}^{\mathrm{m}}$ *are* C^1*. Set*

$$V(p) = f\left(x^*(p), p\right).$$

Then V *is* C^1 *and*

$$\frac{\partial V(p)}{\partial p_j} = \frac{\partial L\Big(x^*(p), \lambda^*(p), p\Big)}{\partial p_j} = \frac{\partial f(x^*, p)}{\partial p_j} - \sum_{i=1}^{m} \lambda_i^*(p) \frac{\partial g_i(x^*, p)}{\partial p_j}.$$

The proof is the same as that of Theorem 6.5.1.

6.6 The Envelope Theorem and Cost Minimization

minimize $\sum_i w_i x_i$ subject to $f(x_1, \dots, x_n) = y$

$$L(x, \lambda; y, w) = \sum_i w_i x_i - \lambda\left(f(x_1, \dots, x_n) - y\right)$$

$$c(y, w) = \sum_i w_i \hat{x}_i(y, w)$$

By the Envelope Theorem,

$$\frac{\partial c}{\partial y} = \left.\frac{\partial L}{\partial y}\right|_{\substack{x=\hat{x}(y,w)\\ \lambda=\hat{\lambda}(y,w)}} = \hat{\lambda}$$

The Lagrange multiplier is the marginal cost. Also,

$$\frac{\partial c}{\partial w_i} = \left.\frac{\partial L}{\partial w_i}\right|_{\substack{x=\hat{x}(y,w)\\ \lambda=\hat{\lambda}(y,w)}} = \hat{x}_i$$

6.7 Appendix: Quadratic forms under constraint

The following results used to be well known (e.g., Mann [6], Debreu [4], Samuelson [8, pp. 378–379], Quirk [7, pp. 22–25], and Diewert and Woodland [5, Appendix, Lemma 3]) and may be found in my one-line notes [1].

6.7.1 Theorem ([1], Theorem 7) *Suppose* A *is an* $n \times n$ *symmetric matrix and* $b^1, \dots, b^m \in \boldsymbol{R}^{\mathrm{n}}$ *are linearly independent. Assume*

$$x'Ax < 0 \text{ for all nonzero } x \text{ satisfying } B'x = 0\ ,$$

where B *is the* $n \times m$ *matrix whose* j^{th} *column is* b^j*. (Here* B' *denotes the transpose of* B*.) Then:*

1. *The matrix*

$$\left[\begin{array}{c|c} A & B \\ \hline B' & 0 \end{array}\right]$$

is invertible.

2. *Write*

$$\left[\begin{array}{c|c} A & B \\ \hline B' & 0 \end{array}\right]^{-1} = \left[\begin{array}{c|c} C & D \\ \hline D' & E \end{array}\right].$$

Then C is negative semidefinite of rank $n-m$, with $Cx = 0$ if and only if x is a linear combination of $b^1, \dots, b^m$.

6.7.2 Theorem ([1], Theorem 8) *Suppose A is an $n \times n$ symmetric matrix and $b^1, \dots, b^m \in \boldsymbol{R}^{\mathrm{n}}$ are linearly independent. Assume*

$$x'Ax \leqslant 0 \text{ for all nonzero } x \text{ satisfying } B'x = 0\ ,$$

where B is the $n \times m$ matrix whose j^{th} column is b^j. Suppose also that the matrix

$$\left[\begin{array}{c|c} A & B \\ \hline B' & 0 \end{array}\right]$$

is invertible. Then

$$x'Ax < 0$$

for all nonzero x satisfying $B'x = 0$.

A matrix A is **positive definite under the orthogonality constraints $b^1, \dots, b^m$** if it is symmetric and

$$x \cdot Ax > 0 \quad \text{for all} \quad x \neq 0 \quad \text{satisfying} \quad b^j \cdot x = 0, \quad j = 1, \dots, m.$$

For brevity, when the vectors $b^1, \dots, b^m$ are understood, we often say simply that A is **positive definite under constraint**. The notions of negative definiteness and semidefiniteness under constraint are defined in the obvious analogous way. Notice that we can replace $b^1, \dots, b^m$ by any basis for the span of $b^1, \dots, b^m$, so without loss of generality we may assume that $b^1, \dots, b^m$ are linearly independent, or even orthonormal.

6.7.3 Theorem ([1], Theorem 10) *Let A be an $n \times n$ symmetric matrix and let $\{b^1, \dots, b^m\}$ be linearly independent.*

1. *A is positive definite under the orthogonality constraints $b^1, \dots, b^m$ if and only if*

$$(-1)^m \begin{vmatrix} a_{11} & \dots & a_{1r} & b^1_1 & \dots & b^m_1 \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{r,1} & \dots & a_{rr} & b^1_r & \dots & b^m_r \\ b^1_1 & \dots & b^1_r & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ b^m_1 & \dots & b^m_r & 0 & \dots & 0 \end{vmatrix} > 0$$

for $r = m+1, \ldots, n$. That is, if and only if every r^{th}-order NW bordered principal minor has sign $(-1)^m$ for $r > m$.

2. *A is negative definite under the orthogonality constraints $b^1, \ldots, b^m$ if and only if*

$$(-1)^r \begin{vmatrix} a_{11} & \ldots & a_{1r} & b_1^1 & \ldots & b_1^m \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{r,1} & \ldots & a_{rr} & b_r^1 & \ldots & b_r^m \\ b_1^1 & \ldots & b_r^1 & 0 & \ldots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ b_1^m & \ldots & b_r^m & 0 & \ldots & 0 \end{vmatrix} > 0$$

for $r = m+1, \ldots, n$. That is, if and only if every r^{th}-order NW bordered principal minor has sign $(-1)^r$ for $r > m$.

Note that for positive definiteness under constraint all the NW bordered principal minors of order greater than m have the same sign, the sign depending on the number of constraints. For negative definiteness the NW bordered principal minors alternate in sign. For the case of one constraint ($m = 1$) if A is positive definite under constraint, then these minors are negative. Again with one constraint, if A is negative definite under constraint, then the minors of even order are positive and of odd order are negative.

References

[1] K. C. Border. 2001. More than you wanted to know about quadratic forms. On-line note.
http://www.its.caltech.edu/~kcborder/Notes/QuadraticForms.pdf

[2] ——— . 2015. Miscellaneous notes on optimization theory and related topics: Or a dessert course in optimization. On-line note.
http://www.its.caltech.edu/~kcborder/Notes/Maximization.pdf

[3] C. W. Cobb and P. H. Douglas. 1928. A theory of production. *American Economic Review* 18(1):139–165. Supplement, Papers and Proceedings of the Fortieth Annual Meeting of the American Economic Association
http://www.jstor.org/stable/1811556

[4] G. Debreu. 1952. Definite and semidefinite quadratic forms. *Econometrica* 20(2):295–300. http://www.jstor.org/stable/1907852

[5] W. E. Diewert and A. D. Woodland. 1977. Frank Knight's theorem in linear programming revisited. *Econometrica* 45(2):375–398.
http://www.jstor.org/stable/1911216

[6] H. B. Mann. 1943. Quadratic forms with linear constraints. *American Mathematical Monthly* 50(7):430–433. Reprinted in [**?**].
http://www.jstor.org/stable/2303666

[7] J. P. Quirk. 1976. *Intermediate microeconomics: Mathematical notes*. Chicago: Science Research Associates.

[8] P. A. Samuelson. 1965. *Foundations of economic analysis*. New York: Athenaeum. Reprint of the 1947 edition published by Harvard University Press.