10. Second Order Sufficient Conditions for a Maximum or Minimum

Before we can develop our second order conditions, we need to introduce the concept of a second order directional derivative. Suppose we are given a direction v and the first order directional derivative in this direction exists for all $(x_1, x_2, \ldots, x_N)$ in a neighbourhood; i.e., $D_v f(x_1, x_2, \ldots, x_N)$ exists. Now pick another direction $u \equiv (u_1, u_2, \ldots, u_2)$ where $u_1^2 + u_2^2 + \ldots + u_N^2 = 1$.

Definition: The second order directional derivative of f in the directions v and u evaluated at the point $x_1, x_2, \ldots, x_N$ is defined as the following limit if it exists:

(1) $$D_{vu}f(x_1, x_2, \ldots, x_N) \equiv \lim_{t \to 0} [D_v f(x_1 + tu_1, \ldots, x_N + tu_N) - D_v f(x_1, \ldots, x_N)] / t$$

[Geometric Interpretation?]

Note that if $v = e_i$ and $u = e_j$, then $D_{e_i e_j} f(x_1, \ldots, x_N) = f_{ij}(x_1, \ldots, x_N)$, the second order partial derivative of f with respect to x_i and x_j. For example, if $N = 2$, $v = (1, 0)$ and $u = (0, 1)$, then $D_{vu}f(x_1, x_2) = D_{e_1 e_2} f(x_1, x_2) = f_{12}(x_1, x_2) \equiv \partial^2 f(x_1, x_2)/\partial x_1 \partial x_2$. Recall that it is straightforward to compute second order partial derivatives using ordinary calculus rules. However, for general directions v and u, it is not easy to compute $D_{vu}f(x_1, \ldots, x_N)$.

The following theorem allows us to express a general second order directional derivative in terms of second order partial derivatives.

Second Order Directional Derivative Theorem. If the first and second order partial derivatives of f exist and are continuous in a neighbourhood around the point $x_1, \ldots, x_N$, then the second order directional derivative of f in the directions $v = (v_1, \ldots, v_N)$ and $u = (u_1, \ldots, u_N)$ evaluated at $x_1, \ldots, x_N$ exists and may be calculated as the following weighted sum of second order partial derivatives

(2) $$D_{vu}f(x_1, \ldots, x_N) = \sum_{i=1}^{N} \sum_{j=1}^{N} v_i f_{ij}(x_1, \ldots, x_N) u_j.$$

Proof: For simplicity, we shall prove only the case where $N = 2$. The general case follows in an analogous manner. By the definition of $D_{vu}f(x_1, x_2)$, we have:

$$D_{vu}f(x_1, x_2) = \lim_{t \to 0} [D_v f(x_1 + tu_1, x_2 + tu_2) - D_v f(x_1, x_2)] / t$$

$$= \lim_{t \to 0} [(v_1 f_1(x_1 + tu_1, x_2 + tu_2) + v_2 f_2(x_1 + tu_1, x_2 + tu_2)) - (v_1 f_1(x_1, x_2) + v_2 f_2(x_1, x_2)] / t$$

applying the First Order Directional Derivative Theorem to $D_v f(x_1 + tu_1, x_2 + tu_2)$ and $D_v f(x_1, x_2)$

$$= \lim_{t \to 0}[(v_1\{f_1(x_1 + tu_1, x_2 + tu_2) - f_1(x_1, x_2)\}$$

$$+ v_2\{f_2(x_1 + tu_1, x_2 + tu_2) - f_2(x_1, x_2)\}] / t$$

collecting terms involving v_1 and v_2

$$= \lim_{t \to 0} \sum_{i=1}^{2} v_i\{f_i(x_1 + tu_1, x_2 + tu_2) - f_i(x_1, x_2)\} / t$$

$$= \sum_{i=1}^{2} v_i D_u f_i(x_1, x_2)$$

by the definition of $D_u f_i$

$$= \sum_{i=1}^{2} v_i\{\sum_{j=1}^{2} u_j f_{ij}(x_1, x_2)\}$$

applying the First Order Directional Derivative Theorem to $D_u f_1$ and $D_u f_2$

$$= \sum_{i=1}^{2} \sum_{j=1}^{2} v_j f_{ij}(x_1, x_2) u_j \qquad \text{rearranging terms}$$

$$= v_1 f_{11}(x_1, x_2)u_1 + v_1 f_{12}(x_1, x_2) u_2 + v_2 f_{21}(x_1, x_2)u_1 + v_2 f_{22}(x_1, x_2)u_2$$

Q.E.D.

Assuming that our function $f(x_1, \ldots, x_N)$ has continuous second order partial derivatives, we may now derive sufficient conditions for an interior local max or min by applying our old one dimensional conditions to all possible directions.

Thus we have the following:

Second Order Sufficient Conditions for f to attain a strict local max:

(3) $D_v f(x_1, \ldots, x_N) = 0$ for all directions v and

(4) $D_{vv} f(x_1, \ldots, x_N) = \sum_{i=1}^{N} \sum_{j=1}^{N} v_i f_{ij}(x_1, \ldots, x_N) v_j < 0$

for all v such that $v_1^2 + \ldots + v_N^2 = 1$

Second Order Sufficient Conditions for f to attain a strict local min at $x_1, \ldots, x_N$:

(5) $D_v f(x_1, \ldots, x_N) = 0$ for all directions v and

(6) $D_{vv}f(x_1, \ldots, x_N) = \sum_{i=1}^{N}\sum_{j=1}^{N} v_i f_{ij}(x_1, \ldots, x_N) v_j > 0$

for all v such that $v_1^2 + \ldots + v_N^2 = 1$.

We indicated in the previous section, that conditions (3) or (5) are equivalent to the following first order necessary conditions which can readily be verified:

(7) $f_i(x_1, \ldots, x_N) = 0, \qquad i = 1, \ldots, N$

However, if $N \geq 2$, then conditions (4) and (6) involve checking an infinite number of inequalities, which is not practical. Hence we need to convert conditions (4) and (6) into equivalent sets of conditions that can be checked. This task is accomplished in courses in matrix algebra and we will not attempt to do it here for the general case. However, we shall do the case where N = 2. We first, require a preliminary result. We state it for N = 2, but it is valid for a general N.

Young's Theorem. If f_1, f_2 and f_{12} exist and are continuous functions around the point (x_1, x_2), then the second order partial derivative $f_{21}(x_1, x_2)$ exists and equals

(8) $f_{21}(x_1, x_2) = f_{12}(x_1, x_2)$.

Proof: By the definition of $f_{21}(x_1, x_2)$, we have

$$f_{21}(x_1, x_2) \equiv \lim_{h \to 0} [f_2(x_1 + h, x_2) - f_2(x_1, x_2)] / h$$

$$= \lim_{h \to 0} h^{-1}[\lim_{k \to 0} \{f(x_1 + h, x_2 + k) - f(x_1 + h, x_2)\} / k + \lim_{k \to 0} \{f(x_1, x_2 + k) - f(x_1, x_2)\} / k].$$

by the definition of f_2 used two times

$$= \lim_{h \to 0} \lim_{k \to 0} h^{-1}k^{-1}[g(x_1 + h) - g(x_1)]$$

defining $g(t) \equiv f(t, x_2 + k) - f(t, x_2)$

$$= \lim_{h \to 0} \lim_{k \to 0} h^{-1}k^{-1}[g'(x_1^*)]$$

where x_1^* is between $x_1 + h$ and x_1, applying the Mean Value Theorem to g

$$= \lim_{h \to 0} \lim_{k \to 0} k^{-1}[f_1(x_1^*, x_2 + k) - f_1(x_1^*, x_2)]$$

cancelling h's and using the definition of g

$$= \lim_{h \to 0} f_{12}(x_1^*, x_2)$$

by the definition of f_{12}

$= f_{12}(x_1, x_2)$

using the continuity of f_{12}.

Q.E.D.

Now consider conditions (6) when N = 2. Using Young's Theorem, these conditions become:

(9) $D_{vv}f(x_1, x_2) = v_1^2 f_{11}(x_1, x_2) + 2v_1v_2f_{12}(x_1, x_2) + v_2^2 f_{22}(x_1, x_2) > 0$

for all v_1 and v_2 such that $v_1^2 + v_2^2 = 1$.

Our problem is to reduce the infinite number of conditions in (9) down to a finite number of conditions.

Suppose $v_1 = 0$. Since $v_1^2 + v_2^2 = 1$, we have $v_2^2 = 1$ and (9) reduces to

(10) $f_{22}(x_1, x_2) > 0$.

Now let us fix $v_1 \neq 0$ and look at the right hand side of (9) as a function of v_2. We want to know if

(11) $g(v_2) \equiv v_1^2 f_{11}^* + 2v_1v_2 f_{12}^* + v_2^2 f_{22}^* > 0$

where $f_{ij}^* \equiv f_{ij}(x_1, x_2)$. Let us try to minimize g with respect to v_2:

$$g'(v_2) = 2v_1f_{12}^* + 2 v_2 f_{22}^* \overset{set}{=} 0 \implies v_2^* = -v_1 f_{12}^* / f_{22}^*$$
$$g''(v_2^*) = 2f_{22}^*$$
$$> 0 \qquad \text{using (10).}$$

Hence $v_2^* =$ globally minimizes $g(v_2)$. Thus if $g(v_2^*) > 0$, then $g(v_2) \geq g(v_2^*) > 0$ for all v_2. Now calculate $g(v_2^*)$:

$$g(v_2^*) = v_1^2 f_{11}^* + 2v_1(- v_1f_{12}^* / f_{22}^*)f_{12}^* + v_1^2 f_{12}^{*\,2}/f_{22}^* = v_1^2 [f_{11}^* - f_{12}^{*\,2}/f_{22}^*].$$

Thus necessary and sufficient conditions for (9) to be true are (10) and $f_{11}^* - f_{12}^{*\,2} / f_{22}^* > 0$ which in view of (10) is equivalent to:

(12) $f_{11}(x_1, x_2)\, f_{22}(x_1, x_2) - [f_{12}(x_1, x_2)]^2 > 0.$

Thus the infinite number of inequalities in (9) are equivalent to the two conditions (10) and (12).

Suppose Conditions (9) are satisfied. Then set $v_2 = 0$ and (9) reduces to:

(13) $f_{11}(x_1, x_2) > 0.$

Problem: Show that conditions (10) and (12) are equivalent to conditions (12) and (13).

Now consider conditions (4) when N = 2. We may modify our analysis on the previous page and show that (4) is equivalent to the following two conditions:

(14) $f_{11}(x_1, x_2) < 0$ and

(15) $f_{11}(x_1, x_2)\, f_{22}(x_1, x_2) - [f_{12}(x_1, x_2)]^2 > 0.$

Condition (14) may be replaced by

(16) $f_{22}(x_1, x_2) < 0.$

(Another way to derive (14) and (15) from (13) and (12) is to observe that maximizing $f(x_1, x_2)$ is equivalent to minimizing $-f(x_1, x_2)$. Thus we require $-f_{11}(x_1, x_2) > 0$ and $[-f_{11}(x_1, x_2)][-f_{22}(x_1, x_2)] - [-f_{12}(x_1, x_2)]^2 > 0$ and these two inequalities are equivalent to (14) and (15).)

Example 1: Maximize $f(x_1, x_2) \equiv 2x_1 + 2x_2 - x_1^2 + x_1x_2 - x_2^2$.

$f_1(x_1, x_2) = 2 - 2x_1 + x_2 \overset{set}{=} 0$

$f_2(x_1, x_2) = 2 + x_1 - 2x_2 \overset{set}{=} 0$

Solution is $x_1^* = x_2^* = 2.$

$f_{11}(x_1, x_2) = -2$, $f_{12}(x_1, x_2) = 1$

$f_{21}(x_1, x_2) = 1$, $f_{22}(x_1, x_2) = -2$

$f_{11}(x_1^*, x_2^*) = -2 < 0$

$f_{11}(x_1^*, x_2^*)\, f_{22}(^*_1, x_2^*) - [f_{12}(x_1^*, x_2^*])^2$

$= [-2]\,[-2] - 1^2$

$= 4 - 1$

$= 3 > 0$

Therefore the second order sufficient conditions for a local max are satisfied at $x_1^* = 2$, $x_2^* = 2$. Since there is only one solution to the first order conditions and $f(x_1, x_2)$ becomes large and negative as $x_1^2 + x_2^2$ becomes large, we conclude that our local max is also the global max.

Example 2: Two Input Cobb-Douglas Production Function

Suppose that a competitive firm utilizes positive amounts of two inputs, x_1 and x_2, in order to produce units of a single output y. The technology of the firm may be summarized by means of a *production function* f; i.e., $y = f(x_1, x_2)$ denotes the maximum amount that can be produced in a certain period of time using x_1 units of input 1 and x_2 units of input 2. Suppose that the firm can sell units of output at the fixed price $p_0 > 0$ and can purchase units of input 1 and 2 at the fixed prices p_1 and p_2. Then the firm's *constrained profit maximization problem* is

(17) $\max_{y,\, x_1,\, x_2}\{p_0 y - p_1 x_1 - p_2 x_2 :\ y = f(x_1, x_2)\}.$

Problem (17) is a constrained profit maximization problem involving 3 decision variables, y, x_1, x_2; 3 exogenous variables, p_0, p_1, p_2; and one (production function) constraint. Let us substitute the constraint function into the objective function and reduce (17) into the following *unconstrained profit maximization problem* involving the two decision variables, x_1 and x_2:

(18) $\max_{x_1,\, x_2}\{p_0 f(x_1, x_2) - p_1 x_1 - p_2 x_2\}.$

We cannot solve (18) until we are given a concrete functional form for the production function f. Suppose that

(19) $y = f(x_1, x_2) \equiv \alpha x_1^{\beta_1} x_2^{\beta_2}$

where the technological parameters satisfy the following restrictions

(20) $\alpha > 0,\ \beta_1 > 0,\ \beta_2 > 0,\ \beta_1 + \beta_2 < 1.$

Substituting (19) into (18), we find that the first order necessary conditions for (18) are:

(21) $\beta_1 p_0 \alpha x_1^{\beta_1 - 1} x_2^{\beta_2} - p_1 \overset{set}{=} 0;$

(22) $\beta_2 p_0 \alpha x_1^{\beta_1} x_2^{\beta_2 - 1} - p_2 \overset{set}{=} 0.$

Multiply (21) by x_1, multiply (22) by x_2. We get:

(23) $\Box_1 p_0 \Box x_1 \; x_1^{\Box_1} x_2^{\Box_2} = p_1 x_1$ and

(24) $\Box_2 p_0 \Box x_2 \; x_1^{\Box_1} x_2^{\Box_2} = p_2 x_2$.

Now divide (23) by (24) and simplify. We get

(25) $x_1 = \Box_1 \Box_2^{\Box 1} p_1^{-1} \; p_2 x_2$.

Substitute (25) into (24) and simplify. We get

(26) $x_2^* = [a\Box_2 \Box_1^{\Box_1} \Box_2^{\Box\Box_1}]^{1/(1-\Box_1-\Box_2)} p_0^{1/(1-\Box_1-\Box_2)} p_1^{-\Box_1/(1-\Box_1-\Box_2)} p_2^{(\Box_1-1)/(1-\Box_1-\Box_2)}$

$\equiv k_2 p_0^{1/(1-\Box_1-\Box_2)} p_1^{-\Box_1/(1-\Box_1-\Box_2)} p_2^{(\Box_1-1)/(1-\Box_1-\Box_2)}$

$\equiv D_2(p_0, p_1, p_2, a, \Box_1, \Box_2)$

Now substitute (26) into (25) and get:

(27) $x_1^* = [\Box_1 \Box_2^{\Box 1} k_2] p_0^{1/(1-\Box_1-\Box_2)} p_1^{(\Box_2-1)(1-\Box_1-\Box_2)} p_2^{\Box\Box_2/(1-\Box_1-\Box_2)}$

$\equiv k_1 p_0^{1/(1-\Box_1-\Box_2)} p_1^{(\Box_2-1)/(1-\Box_1-\Box_2)} p_2^{\Box\Box_2/(1-\Box_1-\Box_2)}$

$\equiv D_1(p_0, p_1, p_2, a, \Box_1, \Box_2)$.

The x_1 and x_2 solutions to (18) defined by (27) and (26) are the firm's system of *profit maximizing input demand functions*, D_1 and D_2. These two functions tell us how the firm's demands will vary as the output and input prices p_0, p_1 and p_2 change. The demand functions also depend on the technological parameters $\Box$, $\Box_1$ and $\Box_2$ that appeared in the production function (19).

To obtain the firm's output supply function S, substitute (26) and (27) into (19) and obtain:

(28) $y^* = [\Box \, k_1^{\Box_1} k_2^{\Box_2}] p_0^{(\Box_1+\Box_2)/(1-\Box_1-\Box_2)} p_1^{-\Box_1/(1-\Box_1-\Box_2)} p_2^{-\Box_2/(1-\Box_1-\Box_2)}$

$\equiv k_0 p_0^{(\Box_1+\Box_2)/(1-\Box_1-\Box_2)} p_1^{-\Box_1/(1-\Box_1-\Box_2)} p_2^{-\Box_2/(1-\Box_1-\Box_2)}$

$\equiv S(p_0, p_1, p_2, \Box, \Box_1, \Box_2)$.

This example illustrates how the firm's profit maximization problem generates output supply functions and input demand functions.

Question: if we were given data on the firm's output in period t, y^t, the output price p_0^t and the input prices p_1^t and p_2^t for period t for t = 1, . . ., T, how could we use these data to obtain estimates for the firm's technological parameters α, β_1 and β_2? *Hint*: take the logarithm of both sides of (28) and recall example 2 in section 3 above.

We are not quite through with the above example: we have not checked whether our solution defined by (26) and (27) satisfies the second order sufficient conditions for a local maximum. We first calculate the second order partial derivatives $f_{ij}^* \equiv f_{ij}(x_1^*, x_2^*)$ for i, j = 1, 2:

$$f_{11}^* = p_0\alpha\beta_1(\beta_1 - 1)x_1^{*\beta_1-2}x_2^{*\beta_2}\,;\; f_{12}^* = p_0\alpha\beta_1\beta_2\, x_1^{*\beta_1-1}x_2^{*\beta_2-1} = f_{21}^*$$

$$f_{22}^* = p_0\alpha\beta_2(\beta_2 - 1)x_1^{*\beta_1}\, x_1^{*\beta_2-1}$$

By (20) $\beta_1 < 1$ and $\beta_2 < 1$ so that

(29) $\quad f_{11}^* < 0$ and $f_{22}^* < 0$.

(30) $\quad f_{11}^*\, f_{22}^* - (f_{12}^*)^2 = [p_0\alpha\, x_1^{*\beta_1-1}x_2^{*\beta_2-1})]^2[\beta_1(\beta_1 - 1)\beta_2(\beta_2 - 1) - (\beta_1\beta_2)^2]$

$$= [p_0\alpha\, x_1^{*\beta_1-1}x_2^{*\beta_2-1})]^2\beta_1\beta_2[1 - \beta_1 - \beta_2]$$

$$> 0 \qquad\qquad \text{using (20).}$$

Thus the second order sufficient conditions for a local max are satisfied and our solution functions defined by (26) and (27) are valid.

11. The Lagrange Multiplier Technique

In economics, constrained maximization problems are often solved using the Lagrange multiplier technique. Thus it is necessary for us to explain what is is.

Suppose f and b are differentiable functions of two variables x_1 and x_2 and we wish to solve the following constrained maximization problem (e.g., recall example 1 in section 3):

(1) $\max_{x_1,x_2}\{f(x_1, x_2) : b(x_1, x_2) = 0\}$.

Suppose that the maximum to (1) occurs at a point x_1^*, x_2^* in the interior of the domain of both functions. Suppose also that $f_2(x_1^*, x_2^*) \neq 0$ and $b_2(x_1^*, x_2^*) \neq 0$.

Consider the indifference or level curve of f through the point (x_1^*, x_2^*). This is the set $\{x_1, x_2 : f(x_1, x_2) = f(x_1^*, x_2^*)\}$. Consider also the constraint curve, $\{x_1, x_2 : b(x_1, x_2) = 0\}$. Obviously, if x_1^*, x_2^* solves (1), then x_1^*, x_2^* is on the constraint curve; i.e.,

(2) $b(x_1^*, x_2^*) = 0$.

From elementary geometrical considerations, it can be seen that if x_1^*, x_2^* solves (1), then the indifference curve of f through x_1^*, x_2^* must be tangent to the constraint curve and the point of tangency occurs at x_1^*, x_2^*. Thus the slopes of the two curves must be equal. From formula (17) in section 6, the slope of the indifference curve through x_1^*, x_2^* is $-f_1(x_1^*, x_2^*)/f_2(x_1^*, x_2^*)$ while the slope of the constraint function at x_1^*, x_2^* is $-b_1(x_1^*, x_2^*)/b_2(x_1^*, x_2^*)$. Equating these two slopes yields the equation

(3) $-f_1(x_1^*, x_2^*)/f_2(x_1^*, x_2^*) = -b_1(x_1^*, x_2^*)/b_2(x_1^*, x_2^*)$

Now rearrange (3) to yield the following equation:

(4) $-f_1(x_1^*, x_2^*)/b_1(x_1^*, x_2^*) = -f_2(x_1^*, x_2^*)/b_2(x_1^*, x_2^*)$
$\equiv \lambda^*$

where we have defined the common ratio in (4) to be the number λ^*. Now we may rearrange equations (4) to yield the following two equations involving λ^*:

(5) $f_1(x_1^*, x_2^*) + \lambda^* b_1(x_1^*, x_2^*) = 0$

(6) $f_2(x_1^*, x_2^*) + \lambda^* b_2(x_1^*, x_2^*) = 0.$

Equations (2), (5) and (6) may be regarded as three equations in the three unknowns x_1^*, x_2^* and λ^*. These are *Lagrange's first order necessary conditions* for x_1^*, x_2^* to solve the constrained maximization problem (1).

These conditions may be obtained in a simple way by defining the *Lagrangian* $L(x_1, x_2, \lambda)$ by:

(7) $L(x_1, x_2, \lambda) \equiv f(x_1, x_2) + \lambda b(x_1, x_2).$

It can be verified that equations (5), (6) and (2) are equivalent to the following first order conditions:

(8) $\partial L(x_1^*, x_2^*, \lambda^*)/\partial x_1 = 0;\ \partial L(x_1^*, x_2^*, \lambda^*)/\partial x_2 = 0;\ \partial L(x_1^*, x_2^*, \lambda^*)/\partial \lambda = 0$

The second order conditions for the Lagrangian technique are too complex for us to develop here. In practice, when using the Lagrange multiplier technique for solving (1), one simply hopes that the point x_1^*, x_2^* found by solving (8) is the desired maximum.

Question: Suppose (1) was a minimization problem instead of a maximization problem. How could we adapt the above technique?

Problem:

Determine whether the following functions have any local minimums or maximums. Check the relevant second order conditions. The domain of definition for each function is two dimensional space.

(i) $f(x_1,x_2) \equiv x_1^2 + x_2^2 - 2x_1 - 2x_2$;
(ii) $f(x_1,x_2) \equiv -x_1^2 + x_1x_2 - x_2^2 + x_1 - x_2$;
(iii) $f(x_1,x_2) \equiv x_1^2 - 2x_1x_2 + x_2^2$.