Lecture 7
Envelope Theorems, Bordered Hessians and Kuhn-Tucker Conditions

Eivind Eriksen

BI Norwegian School of Management
Department of Economics

October 15, 2010

Envelope theorems

In economic optimization problems, the objective functions that we try to maximize/minimize often depend on parameters, like prices. We want to find out how the optimal value is affected by changes in the parameters.

Example

Let $f(x; a) = -x^2 + 2ax + 4a^2$ be a function in one variable x that depends on a parameter a. For a given value of a, the stationary points of f is given by

$$\frac{\partial f}{\partial x} = -2x + 2a = 0 \quad \Leftrightarrow \quad x = a$$

and this is a (local and global) maximum point since $f(x; a)$ is concave considered as a function in x. We write $x^*(a) = a$ for the maximum point. The optimal value function $f^*(a) = f(x^*(a); a) = -a^2 + 2a^2 + 4a^2 = 5a^2$ gives the corresponding maximum value.

Envelope theorems: An example

Example (Continued)

The derivative of the value function is given by

$$\frac{\partial f^*}{\partial a} = \frac{\partial}{\partial a} f(x^*(a); a) = \frac{\partial}{\partial a} \left(5a^2\right) = 10a$$

On the other hand, we see that $f(x; a) = -x^2 + 2ax + 4a^2$ *gives*

$$\frac{\partial f}{\partial a} = 2x + 8a \quad \Rightarrow \quad \left(\frac{\partial f}{\partial a}\right)_{x = x^*(a)} = 2a + 8a = 10a$$

since $x^*(a) = a$.

The fact that these computations give the same result is not a coincidence, but a consequence of the envelope theorem for unconstrained optimization problems:

Envelope theorem for unconstrained maxima

Theorem

Let $f(\mathbf{x}; a)$ be a function in n variables $x_1, \dots, x_n$ that depends on a parameter a. For each value of a, let $\mathbf{x}^*(a)$ be a maximum or minimum point for $f(\mathbf{x}; a)$. Then

$$\frac{\partial}{\partial a} f(\mathbf{x}^*(a); a) = \left(\frac{\partial f}{\partial a} \right)_{\mathbf{x}=\mathbf{x}^*(a)}$$

The following example is a modification of Problem 3.1.2 in [FMEA]:

Example

A firm produces goods A and B. The price of A is 13*, and the price of B is p. The profit function is* $\pi(x, y) = 13x + py - C(x, y)$*, where*

$$C(x, y) = 0.04x^2 - 0.01xy + 0.01y^2 + 4x + 2y + 500$$

Determine the optimal value function $\pi^*(p)$*. Verify the envelope theorem.*

Envelope theorems: Another example

Solution

The profit function is $\pi(x, y) = 13x + py - C(x, y)$*, hence we compute*

$$\pi(x, y) = -0.04x^2 + 0.01xy - 0.01y^2 + 9x + (p - 2)y - 500$$

The first order conditions are

$$\pi_x = -0.08x + 0.01y + 9 = 0 \Rightarrow 8x - y = 900$$
$$\pi_y = 0.01x - 0.02y + p - 2 = 0 \Rightarrow x - 2y = 200 - 100p$$

This is a linear system with unique solution $x^* = \frac{1}{15}(1600 + 100p)$ *and* $y^* = \frac{1}{15}(-700 + 800p)$*. The Hessian* $\pi'' = \begin{pmatrix} -0.08 & 0.01 \\ 0.01 & -0.02 \end{pmatrix}$ *is negative definite since* $D_1 = -0.08 < 0$ *and* $D_2 = 0.0015 > 0$*. We conclude that* (x^*, y^*) *is a (local and global) maximum for* π*.*

Envelope theorems: Another example

Solution (Continued)

Hence the optimal value function $\pi^*(p) = \pi(x^*, y^*)$ *is given by*

$$\pi\left(\frac{1}{15}(1600+100p), \frac{1}{15}(-700+800p)\right) = \frac{80p^2 - 140p + 80}{3}$$

and its derivative is therefore

$$\frac{\partial}{\partial p}\pi(x^*, y^*) = \frac{160p - 140}{3}$$

On the other hand, the envelope theorem says that we can compute the derivative of the optimal value function as

$$\left(\frac{\partial \pi}{\partial p}\right)_{(x,y)=(x^*,y^*)} = y^* = \frac{1}{15}(-700+800p) = \frac{-140+160p}{3}$$

Envelope theorem for constrained maxima

Theorem

Let $f(\mathbf{x}; a), g_1(\mathbf{x}; a), \dots, g_m(\mathbf{x}; a)$ be functions in n variables $x_1, \dots, x_n$ that depend on the parameter a. For a fixed value of a, consider the following Lagrange problem: Maximize/minimize $f(\mathbf{x}; a)$ subject to the constraints $g_1(\mathbf{x}; a) = \cdots = g_m(\mathbf{x}; a) = 0$. Let $\mathbf{x}^*(a)$ be a solution to the Lagrange problem, and let $\lambda^*(a) = \lambda_1^*(a), \dots, \lambda_m^*(a)$ be the corresponding Lagrange multipliers. If the NDCQ condition holds, then we have

$$\frac{\partial}{\partial a} f(\mathbf{x}^*(a); a) = \left(\frac{\partial \mathcal{L}}{\partial a} \right)_{\mathbf{x}=\mathbf{x}^*(a), \lambda=\lambda^*(a)}$$

Notice that any equality constraint can be re-written in the form used in the theorem, since

$$g_j(\mathbf{x}; a) = b_j \Leftrightarrow g_j(\mathbf{x}; a) - b_j = 0$$

Interpretation of Lagrange multipliers

In a Lagrange problem, the Lagrange function has the form

$$\mathcal{L}(\mathbf{x}, \lambda) = f(\mathbf{x}) - \lambda_1(g_1(\mathbf{x}) - b_1) - \dots \lambda_m(g_m(\mathbf{x}) - b_m)$$

Hence we see that the partial derivative with respect to the parameter $a = b_j$ is given by

$$\frac{\partial \mathcal{L}}{\partial b_j} = \lambda_j$$

By the envelope theorem for constrained maxima, this gives that

$$\frac{\partial f(\mathbf{x}^*)}{\partial b_j} = \lambda_j^*$$

where $\mathbf{x}^*$ is the solution to the Lagrange problem, $\lambda_1^*, \dots, \lambda_m^*$ are the corresponding Lagrange multipliers, and $f(\mathbf{x}^*)$ is the optimal value function.

Envelope theorems: A constrained example

Example

Consider the following Lagrange problem: Maximize $f(x,y) = x + 3y$ *subject to* $g(x,y) = x^2 + ay^2 = 10$. *When* $a = 1$, *we found earlier that* $\mathbf{x}^*(1) = (1,3)$ *is a solution, with Lagrange multiplier* $\lambda^*(1) = 1/2$ *and maximum value* $f^*(1) = f(\mathbf{x}^*(1)) = f(1,3) = 10$. *Use the envelope theorem to estimate the maximum value* $f^*(1.01)$ *when* $a = 1.01$, *and check this by computing the optimal value function* $f^*(a)$.

Solution

The NDCQ condition is satisfied when $a \neq 0$, *and the Lagrangian is given by*

$$\mathcal{L} = x + 3y - \lambda(x^2 + ay^2 - 10)$$

Envelope theorems: A constrained example

Solution (Continued)

By the envelope theorem, we have that

$$\left(\frac{\partial f^*(a)}{\partial a}\right)_{a=1} = \left(-\lambda y^2\right)_{\mathbf{x}=(1,3),\lambda=1/2} = -\frac{9}{2}$$

An estimate for $f^*(1.01)$ is therefore given by

$$f^*(1.01) \simeq f^*(1) + 0.01 \cdot \left(\frac{\partial f^*(a)}{\partial a}\right)_{a=1} = 10 - 0.045 = 9.955$$

To find an exact expression for $f^*(a)$, we solve the first order conditions:

$$\frac{\partial \mathcal{L}}{\partial x} = 1 - \lambda \cdot 2x = 0 \Rightarrow x = \frac{1}{2\lambda}$$
$$\frac{\partial \mathcal{L}}{\partial y} = 3 - \lambda \cdot 2ay = 0 \Rightarrow y = \frac{3}{2a\lambda}$$

Envelope theorems: A constrained example

Solution (Continued)

We substitute these values into the constraint $x^2 + ay^2 = 10$, *and get*

$$(\frac{1}{2\lambda})^2 + a(\frac{3}{2a\lambda})^2 = 10 \quad \Leftrightarrow \quad \frac{a+9}{4a\lambda^2} = 10$$

This gives $\lambda = \pm\sqrt{\frac{a+9}{40a}}$ *when* $a > 0$ *or* $a < -9$. *Substitution gives solutions for* $x^*(a)$, $y^*(a)$ *and* $f^*(a)$ *(see Lecture Notes for details). For* $a = 1.01$, *this gives* $x^*(1.01) \simeq 1.0045$, $y^*(1.01) \simeq 2.9836$ *and* $f^*(1.01) \simeq 9.9553$.

Bordered Hessians

The bordered Hessian is a second-order condition for local maxima and minima in Lagrange problems. We consider the simplest case, where the objective function $f(\mathbf{x})$ is a function in two variables and there is one constraint of the form $g(\mathbf{x}) = b$. In this case, the bordered Hessian is the determinant

$$B = \begin{vmatrix} 0 & g_1' & g_2' \\ g_1' & \mathcal{L}_{11}'' & \mathcal{L}_{12}'' \\ g_2' & \mathcal{L}_{21}'' & \mathcal{L}_{22}'' \end{vmatrix}$$

Example

Find the bordered Hessian for the following local Lagrange problem: Find local maxima/minima for $f(x_1, x_2) = x_1 + 3x_2$ *subject to the constraint* $g(x_1, x_2) = x_1^2 + x_2^2 = 10$.

Bordered Hessians: An example

Solution

The Lagrangian is $\mathcal{L} = x_1 + 3x_2 - \lambda(x_1^2 + x_2^2 - 10)$. *We compute the bordered Hessian*

$$B = \begin{vmatrix} 0 & 2x_1 & 2x_2 \\ 2x_1 & -2\lambda & 0 \\ 2x_2 & 0 & -2\lambda \end{vmatrix} = -2x_1(-4x_1\lambda) + 2x_2(4x_2\lambda) = 8\lambda(x_1^2 + x_2^2)$$

and since $x_1^2 + x_2^2 = 10$ *by the constraint, we get* $B = 80\lambda$. *We solved the first order conditions and the constraint earlier, and found the two solutions* $(x_1, x_2, \lambda) = (1, 3, 1/2)$ *and* $(x_1, x_2, \lambda) = (-1, -3, -1/2)$. *So the bordered Hessian is* $B = 40$ *in* $\mathbf{x} = (1, 3)$, *and* $B = -40$ *in* $\mathbf{x} = (-1, -3)$. *Using the following theorem, we see that* $(1, 3)$ *is a local maximum and that* $(-1, -3)$ *is a local minimum for* $f(x_1, x_2)$ *subject to* $x_1^2 + x_2^2 = 10$.

Bordered Hessian Theorem

Theorem

Consider the following local Lagrange problem: Find local maxima/minima for $f(x_1, x_2)$ subject to $g(x_1, x_2) = b$. Assume that $\mathbf{x}^* = (x_1^*, x_2^*)$ satisfy the constraint $g(x_1^*, x_2^*) = b$ and that $(x_1^*, x_2^*, \lambda^*)$ satisfy the first order conditions for some Lagrange multiplier λ^*. Then we have:

1. *If the bordered Hessian $B(x_1^*, x_2^*, \lambda^*) < 0$, then (x_1^*, x_2^*) is a local minima for $f(\mathbf{x})$ subject to $g(\mathbf{x}) = b$.*
2. *If the bordered Hessian $B(x_1^*, x_2^*, \lambda^*) > 0$, then (x_1^*, x_2^*) is a local maxima for $f(\mathbf{x})$ subject to $g(\mathbf{x}) = b$.*

Optimization problems with inequality constraints

We consider the following optimization problem with inequality constraints:

Optimization problem with inequality constraints

Maximize/minimize $f(\mathbf{x})$ subject to the inequality constraints $g_1(\mathbf{x}) \leq b_1$, $g_2(\mathbf{x}) \leq b_2, \ldots, g_m(\mathbf{x}) \leq b_m$.

In this problem, f and $g_1, \ldots, g_m$ are function in n variables $x_1, x_2, \ldots, x_n$ and $b_1, b_2, \ldots, b_m$ are constants.

Example (Problem 8.9)

Maximize the function $f(x_1, x_2) = x_1^2 + x_2^2 + x_2 - 1$ *subject to* $g(x_1, x_2) = x_1^2 + x_2^2 \leq 1$.

To solve this constrained optimization problem with inequality constraints, we must use a variation of the Lagrange method.

Kuhn-Tucker conditions

Definition

Just as in the case of equality constraints, the Lagrangian is given by

$$\mathcal{L}(\mathbf{x}, \lambda) = f(\mathbf{x}) - \lambda_1(g_1(\mathbf{x}) - b_1) - \lambda_2(g_2(\mathbf{x}) - b_2) - \cdots - \lambda_m(g_m(\mathbf{x}) - b_m)$$

In the case of inequality constraints, we solve the Kuhn-Tucker conditions in additions to the inequalities $g_1(\mathbf{x}) \leq b_1, \ldots, g_m(\mathbf{x}) \leq b_m$. The Kuhn-Tucker conditions for maximum consist of the first order conditions

$$\frac{\partial \mathcal{L}}{\partial x_1} = 0, \quad \frac{\partial \mathcal{L}}{\partial x_2} = 0, \quad \frac{\partial \mathcal{L}}{\partial x_3} = 0, \quad \ldots, \quad \frac{\partial \mathcal{L}}{\partial x_n} = 0$$

and the complementary slackness conditions given by

$$\lambda_j \geq 0 \text{ for } j = 1, 2, \ldots, m \text{ and } \lambda_j = 0 \text{ whenever } g_j(\mathbf{x}) < b_j$$

When we solve the Kuhn-Tucker conditions together with the inequality constraints $g_1(\mathbf{x}) \leq b_1, \ldots, g_m(\mathbf{x}) \leq b_m$, we obtain candidates for maximum.

Necessary condition

Theorem

Assume that $\mathbf{x}^* = (x_1^*, \dots, x_N^*)$ solves the optimization problem with inequality constraints. If the NDCQ condition holds at $\mathbf{x}^*$, then there are unique Lagrange multipliers $\lambda_1, \dots, \lambda_m$ such that $(x_1^*, \dots, x_n^*, \lambda_1, \dots, \lambda_m)$ satisfy the Kuhn-Tucker conditions.

Given a point $\mathbf{x}^*$ satisfying the constraints, the NDCQ condition holds if the rows in the matrix

$$\begin{pmatrix} \frac{\partial g_1}{\partial x_1}(\mathbf{x}^*) & \frac{\partial g_1}{\partial x_2}(\mathbf{x}^*) & \dots & \frac{\partial g_1}{\partial x_n}(\mathbf{x}^*) \\ \frac{\partial g_2}{\partial x_1}(\mathbf{x}^*) & \frac{\partial g_2}{\partial x_2}(\mathbf{x}^*) & \dots & \frac{\partial g_2}{\partial x_n}(\mathbf{x}^*) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_m}{\partial x_1}(\mathbf{x}^*) & \frac{\partial g_m}{\partial x_2}(\mathbf{x}^*) & \dots & \frac{\partial g_m}{\partial x_n}(\mathbf{x}^*) \end{pmatrix}$$

corresponding to constraints where $g_j(\mathbf{x}^*) = b_j$ are linearly independent.

Kuhn-Tucker conditions: An example

Solution (Problem 8.9)

The Lagrangian is $\mathcal{L} = x_1^2 + x_2^2 + x_2 - 1 - \lambda(x_1^2 + x_2^2 - 1)$*, so the first order conditions are*

$$2x_1 - \lambda(2x_1) = 0 \Rightarrow 2x_1(1-\lambda) = 0$$
$$2x_2 + 1 - \lambda(2x_2) = 0 \Rightarrow 2x_2(1-\lambda) = -1$$

From the first equation, we get $x_1 = 0$ *or* $\lambda = 1$*. But* $\lambda = 1$ *is not possible by the second equation, so* $x_1 = 0$*. The second equation gives* $x_2 = \frac{-1}{2(1-\lambda)}$ *since* $\lambda \neq 1$*. The complementary slackness conditions are* $\lambda \geq 0$ *and* $\lambda = 0$ *if* $x_1^2 + x_2^2 < 1$*. We get two cases to consider.* **Case 1**: $x_1^2 + x_2^2 < 1, \lambda = 0$*. In this case,* $x_2 = -1/2$ *by the equation above, and this satisfy the inequality. So the point* $(x_1, x_2, \lambda) = (0, -1/2, 0)$ *is a candidate for maximality.* **Case 2**: $x_1^2 + x_2^2 = 1, \lambda \geq 0$*. Since* $x_1 = 0$*, we get* $x_2 = \pm 1$*.*

Kuhn-Tucker conditions: An example

Solution (Problem 8.9 Continued)

We solve for λ in each case, and check that $\lambda \geq 0$. We get two candidates for maximality, $(x_1, x_2, \lambda) = (0, 1, 3/2)$ and $(x_1, x_2, \lambda) = (0, -1, 1/2)$. We compute the values, and get

$$f(0, -1/2) = -1.25$$
$$f(0, 1) = 1$$
$$f(0, -1) = -1$$

We must check that the NDCQ condition holds. The matrix is $(2x_1 \quad 2x_2)$. If $x_1^2 + x_2^2 < 1$, the NDCQ condition is empty. If $x_1^2 + x_2^2 = 1$, the NDCQ condition is that $(2x_1 \quad 2x_2)$ has rank one, and this is satisfied. By the extreme value theorem, the function f has a maximum on the closed and bounded set given by $x_1^2 + x_2^2 \leq 1$ (a circular disk with radius one), and therefore $(x_1, x_2) = (0, 1)$ is a maximum point.

Kuhn-Tucker conditions for minima

General principle: A minimum for $f(\mathbf{x})$ is a maximum for $-f(\mathbf{x})$. Using this principle, we can write down Kuhn-Tucker conditions for minima:

Kuhn-Tucker conditions for minima

There are Kuhn-Tucker conditions for minima in a similar way as for maxima. The only difference is that the complementary slackness conditions are

$$\lambda_j \leq 0 \text{ for } j = 1, 2, \ldots, m \text{ and } \lambda_j = 0 \text{ whenever } g_j(\mathbf{x}) < b_j$$

in the case of minima.