Econ 205 - Slides from Lecture 14

Joel Sobel

September 10, 2010

Theorem ("Lagrange Multipliers")

Theorem

If x^* *solves* $\max f(x)$ *subject to* $G(x) = 0$ *then there exists* λ *such that* $Df(x^*) = \lambda DG(x^*)$*.*

It is standard to write the equation

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

The λ_i are often called Lagrange Multipliers.

Comments

- FOC replace a max problem.
- λ have economic interpretation.
- Write the constraints of the original problem are $G(x) = b$.
- What happens to the solution of the problem as b changes?
- Let $V(b) \equiv f(x^*(b))$ where $x^*(b)$ solves: $\max f(x)$ subject to $G(x) = b$.
- Assume that you can solve problem at a point (say $b = 0$).

- If regularity condition holds, then IFT guarantees that we can solve $G(x) = b$ for W as a function of u and b in a neighborhood of $(u, b) = (u^*, 0)$.
- If $x^*(b) = (u^*(b), w^*(b))$ satisfies the FOC, then we can differentiate $V(b)$ to obtain: $DV(0) =$

  $$((D_u f(x^*) + D_w f(x^*)) (D_u W(x^*, 0)) Du^*(0) + D_w f(x^*) D_b W(x^*, 0).$$

- The first two terms on RHS are 0 by FOC.
- . The final term on the right is the contribution that comes from the fact that W may change with b.
- $$D_w f(x^*) D_b W(x^*, 0) = D_w f(x^*) (D_w G(x^*))^{-1} \mathbf{I} = \lambda,$$

  where the last equation is just the definition of λ. Hence we have the following equality constrained envelope theorem:

  $$DV(\mathbf{0}) = \lambda. \quad (1)$$

Interpretation

How does the value of the problem change if the right-hand side of the ith constraint increases from 0?
The answer is λ_i.

The Kuhn-Tucker Theorem

Inequality constraints: $S = \{x : g_i(x) \leq b_i, i \in I\}$.

- Includes equalities as a special case (because $g_i(x) = b_i$ is equivalent to $-g_i(x) \leq -b_i$ and $g_i(x) \leq b_i$.
- We know how to solve problems without constraints.
- We know how to solve problems with equality constraints.
- If you have inequality constraints and you think that x^* is an optimum, then (locally) the constraints will divide into the group that are satisfied as equations (binding or active constraints) and the others.
- The first group we treat as we would in an equality constrained problem.
- The second group we treat as in an inequality constrained problem.

The First-Order Condition

- Expect a first-order condition like:

$$\nabla f(x^*) = \sum_{i \in J} \lambda_i \nabla g_i(x^*) \tag{2}$$

  where J is the set of binding constraints.
- Alternative form:

$$\nabla f(x^*) = \sum_{i \in I} \lambda_i \nabla g_i(x^*) \tag{3}$$

  and

$$\lambda_i (b_i - g_i(x^*)) = 0 \text{ for all } i \in I. \tag{4}$$

Interpretation

- Equation (3) differs from equation (2) because it includes all constraints (not just the binding ones).
- Equation (3) and (4) are equivalent to (2).
- Equation (4) can only be satisfied (for a given i) if either $\lambda_i = 0$ or $g_i(x^*) = 0$.
- It says a multiplier associated with non-binding constraints is zero.

Complementary Slackness

- Equation (4) is called a complementary slackness condition. (Constraints that do not bind are said to have slack.)
- Lagrange multipliers are "prices" or values of resources represented in each constraint.
- It is useful to think of multiplier λ_i as the amount that value would increase if the ith constraint were relaxed by a unit or, alternatively, how much the person solving the problem would pay to have an "extra" unit to allocate on the ith constraint.
- If a constraint does not bind, it should be the case that this value is zero (why would you pay to relax a constraint that is already relaxed?).
- If you are willing to pay to gain more resource ($\lambda_i = 0$) it must be the case that the constraint is binding. It is possible (but rare) to have both $\lambda_i = 0$ and $g_i(x^*) = 0$. This happens when you have "just enough" of the ith resource for your purposes, but cannot use anymore.

Summary

Theorem

Suppose x^* *solves:* $\max f(x)$ *subject to* $x \in S$ *when* S *is defined by inequalities* $S = \{x : g_i(x) \leq b_i\}$ *and* f *and* g_i *are differentiable. There exists* λ_k^* *such that*

$$\nabla f(x^*) + \sum_{i \in I} \lambda_i^* \nabla \left(g_i(x^*) - b_i\right) = 0 \tag{5}$$

and

$$\lambda_i^* \left(g_i(x^*) - b_i\right) = 0 \textit{ for all } i. \tag{6}$$

1. Appropriate generalization of unconstrained and equality constrained cases.
2. Multipliers have interpretation: $D_i f(x^*) = \lambda_i$
3. We need an additional "regularity" condition to get this result in general. (Supplementary notes have details.)

Second-Order Conditions

- First-order conditions for optimality do not distinguish local maxima from local minima from points that are neither.
- Second-order conditions do provide useful information.
- Equality constraints: you derive first-order conditions by requiring that the objective function be maximized in all directions that will satisfy the constraints of the problem. The second-order conditions must hold for exactly these directions.

- If x^* is a maximum and $x^* + tv$ satisfies the constraints of the problem for t near zero, then $h(t) = f(x^* + tv)$ has a critical point at $t = 0$ and $h^{''}(0) \leq 0$.
- $h^{''}(0) = v^t D^2 f(x^*) v$, so $h^{''}(0) < 0$ if and only if a quadratic form is negative.
- If the problem is unconstrained, then the second-order conditions require that the matrix of second derivatives be negative semi-definite.
- In general, this condition need only apply in the directions consistent with the constraints.
- There is a theory of "Boardered Hessians" that allows you to use some insights from the theory of quadratic forms to classify when quadratic form restricted to a set of directions will be positive definite.

$$\max_{x_1, x_2} x_1^2 - x_2$$

$$\text{subject to} \quad x_1 - x_2 = 0$$

The constraint says $x_1 = x_2$ so the problem is

$$\max_{x_1} x_1^2 - x_1$$

and we can solve this as a one-variable unconstrained problem.

$$\max \log(x_1 + x_2) \text{ subject to: } x_1 + x_2 \leq 5, x_1, x_2 \geq 0.$$

- log is a strictly increasing function, so the problem is equivalent to solving $\max(x_1 + x_2)$ subject to the same constraints.
- Since the objective function is increasing in both variables, the constraint $x_1 + x_2 \leq 5$ must bind.
- Hence the problem is equivalent to maximizing $x_1 + x_2$ subject to $x_1 + x_2 = 5$ and $x_1, x_2 \geq 0$.
- If $x_1 \in [0, 5]$ any pair $(x_1, 5 - x_1)$ solves the problem.

Solving this using KT is possible, but not easy.

$$\max \log(1+x_1) - x_2^2 \text{ subject to: } x_1 + 2x_2 \leq 3, x_1, x_2 \geq 0.$$

The first-order conditions are

$$\frac{1}{1+x_1} = \lambda_1 - \lambda_2 \quad (7)$$

$$-2x_2^2 = 2\lambda_1 - \lambda_3 \quad (8)$$

$$\lambda_1 (3 - x_1 - 2x_2) = 0 \quad (9)$$

$$\lambda_2 x_1 = 0 \quad (10)$$

$$\lambda_3 x_2 = 0 \quad (11)$$

$3 = x_1 + 2x_2$ and $x_1 > 0$ in the solution. [Why?]
Hence:

- the third constraint is satisfied,
- $\lambda_2 = 0$ by the fourth constraint,
- and the first constraint gives: $\lambda_1 = 1/(1+x_1)$
- and the second constraint implies that $\lambda_3 > 0$.
- by the final constraint that $x_2 = 0$
- $x_1 = 3$ and $\lambda_1 = 1/4$.

Simple Allocation

$$\max x_1 \cdots x_n \text{ subject to } \sum_{i=1}^{k} x_i = K.$$

This problem has an equality constraint. If $v = x_1 \cdots x_n$, then we can write the first-order conditions as $v = \lambda x_i$ for all i. Hence $kv = \lambda K$ and therefore each $x_i = K/k$.

A Famous Inequality

Solve:

$$\max x_1^2 x_2^2 \cdots x_n^2 \text{ subject to } \sum_{i=1}^{n} x_i^2 = 1$$

is to set $x_n = (1/n)^n$ for all n.
Solution must satisfy

$$2\frac{f(x)}{x_i} = 2\lambda x_i,$$

which implies that the x_i are independent of i.
[Why is this a maximum and not a minimum?]
Given any n positive numbers $a_1, \ldots, a_n$, let

$$x_i = \frac{a_i^{1/2}}{(a_1 + \cdots + a_n)^{1/2}}, \text{ for } i = 1, \ldots, n.$$

It follows that $\sum_{i=1}^{n} x_i^2 = 1$ and so

$$\left(\frac{a_1 \cdots a_n}{(a_1 + \cdots + a_n)^n} \right)^{1/n} \leq \frac{1}{n}$$

Hype

It is possible to use techniques from constrained optimization to deduce other important results (the triangle inequality, the optimality of least squares, . . .).

Cobb-Douglas

Canonical consumer theory example.
Cobb-Douglas Utility function:

$$f(x) = x_1^{a_1} \cdots x_n^{a_n},$$

where the coefficients are nonnegative and sum to one.
Consumer problem:

$$\max f(x) \text{ subject to } p \cdot x \leq w$$

where $p \geq 0$, $p \neq 0$ is the vector of prices and $w > 0$ is wealth. Since the function f is increasing in its arguments, the budget constraint must hold as an equation.

The first-order conditions can be written

$$a_i f(x)/x_i = \lambda p_i$$

or

$$a_i f(x) = \lambda p_i x_i.$$

Summing both sides of this equation and using the fact that the a_i sum to one yields:

$$f(x) = \lambda p \cdot x = \lambda w.$$

It follows that

$$x_i = \frac{w a_i}{p_i} \text{ and } \lambda = \left(\frac{a_1}{p_1}\right)^{a_1} \cdots (a_n p_n)^{a_n}.$$

In Cobb-Douglas example, you have explicit formulas for solution and value function.
You can manipulate these to confirm generate properties (homogeneity and envelope theorem).