POLI 270 - Mathematical and Statistical Foundations
Prof. S. Saiegh
Fall 2010

Lecture Notes - Class 9

December 2, 2010

Optimization

To examine the uses of the derivative, we are going to finally move into some practical problems.

> Basically, we will use the notions of first-order critical value and the second derivative test to solve optimization problems.

We will look first at single variable optimization problems, and then use matrix algebra to examine multivariate problems.

One-Variable Optimization

Say we want to maximize an agent's utility, as in most microeconomics or game theory's problems.

> Then, we just need to specify our agent's utility function, and the maximize it.
>
> We already know that if f is a differentiable function that achieves a maximum at x^*, then its first derivative must be equal to zero at x^* and that its second derivative must be less than zero at x^*.

These conditions are usually known as the *first-order condition* and the *second-order condition*, respectively, and they can be expressed mathematically by:

$$f'(x^*) = 0$$
$$f''(x^*) \leq 0.$$

We should note, again, that these are *necessary* not sufficient conditions; a point in the domain that solves the maximization problem must satisfy the conditions, but there may be points that satisfy the conditions that do not solve the maximization problem.

> For a minimization problem, the first order condition is the same, but the second order condition becomes $f''(x^*) \geq 0$.

Example 1 *Consider the function* $f(x) = log\ x - ax$.

The F.O.C. for a maximum with respect to x is that

$$\frac{1}{x} - a = 0,$$

and the second-order condition is that

$$-\frac{1}{x^2} \leq 0.$$

We can see that the second-order condition is automatically satisfied, so it follows that the optimal value of x is $x^* = \frac{1}{a}$.

Now let $g(x) = u(x) - bx$.

In this case, we cannot solve explicitly for the x^* that maximizes this expression, but we know that it must satisfy the two conditions $u'(x^*) = b$ and $u''(x^*) \leq 0$.

Illustrations: Nash Equilibrium

Many optimization problems involve the optimal decision making by a single agent – a firm or a voter– in very simple environments.

Many situations, though, resemble more complex environments.

In particular, in many instances, the decisions of several individual agents produce a final outcome.

Game theory provides a useful analytical framework to study these situations.

The goal is to analyze the outcomes of interdependent decisions, moving away from a purely informal understanding of a situation to the formal statement of a game.

There are several ways of describing a game. In general, though, we need to characterize the set of players, the set of actions that each player can choose, and the players' preferences over possible outcomes.

The key element that defines a game, is that the reward of each player depends on the choices made by all others, not just his/her own decision.

Once we have our game, we can go ahead and "solve" it. In order to do that, we need a *solution concept.*

Our solution concept is simply a behavioral assumption about how players choose their actions, given the consequences of their actions, and their preferences over outcomes.

Let us concentrate on Nash Equilibrium, the most commonly used solution concept in Game Theory.

A **Nash equilibrium** is a situation where each player's action is such that she cannot do better by choosing a different action, given the actions chosen by the other players.

Example 2 *A synergistic relationship. Two individuals are involved in a common project. If both individuals devote more effort to the project, they are both better off. For any given effort of the other individual, the return to an individual's effort first increases, then decreases.*

Each player's set of actions is the set of effort levels (nonnegative numbers). Individual i's preferences are represented by the payoff function $u_i(a_i, a_{-i}) = a_i(c + a_{-i} - a_i)$, where a_i is i's effort level, a_{-i} is the other individual's effort level, and $c > 0$ is a constant. Formally,

Players, $N = \{1, 2\}$

Actions, $A_i = [0, \infty)$

$u_i(a_i, a_{-i}) = a_i(c + a_{-i} - a_i)$

To find the NE of this game, we can construct and analyze the players' best response functions,

$$B_i(a_{-i}) = \max_{a_i \in A_i} \{u_i(a_i, a_{-i})\}$$

Given a_{-i}, individual i's payoff is a quadratic function of a_i,

$$u_i(a_i, a_{-i}) = a_i c + a_i a_{-i} - a_i^2,$$

that is zero when $a_i = 0$ and when $a_i = c + a_{-i}$, and reaches a maximum in between.

The function $u_i(a_i, a_{-i})$ reaches a maximum at a_i^* if $u_i(a_i^*, a_{-i}) \geq u_i(a_i, a_{-i})$ for all $a_i \in A_i$.

The F.O.C. for a maximum with respect to a_i is that

$$\frac{\partial u_i}{\partial a_i} = 0, \text{ or } c + a_{-i} - 2a_i = 0$$

So, the optimal value of a_i is $a_i^* = \frac{1}{2}(c + a_{-i})$.

Now, in order to find the mutual best responses, we should solve the system of equations:

$$a_1^* = \frac{1}{2}(c + a_2^*)$$

$$a_2^* = \frac{1}{2}(c + a_1^*)$$

Substituting the second equation in the first we get

$$a_1^* = \frac{1}{2}(c + \frac{1}{2}(c + a_1^*))$$

$$a_1^* = c$$

Substituting this value into the second equation we get

$$a_2^* = \frac{1}{2}(c + c)$$

$$a_2^* = c$$

Therefore, this game has a unique NE, $(a_1^*, a_2^*) = (c, c)$.

Example 3 *Cournot's Model of Oligopoly. An oligopoly is characterized by market interactions with a small number of firms. The classic model of oligopoly is due to Cournot (1838). The modern study of oligopoly is grounded almost entirely on game theory.*

Suppose that there are two firms that produce an identical good.

Each firm must decide how much output to produce without knowing the production decision of the other firm.

If the firms produce a total of Q units of the good, the market price will then be $p(Q)$.

Given a production level q_i for firm i, the market price is then,

$$p(Q) \equiv p(q_i + q_{-i})$$

The function p is called the inverse demand function.

The cost to form i of producing q_i units of the good is $c_i(q_i)$.

Firm i's profit is equal to its revenue minus its cost:

$$\pi_i(q_i, q_{-i}) = p(q_i + q_{-i})q_i - c_i(q_i)$$

– We assume that for $p > 0$, if the firms' total output increases, then price decreases.

Firm i's action in this game is a choice of a given production level (output) and the payoff to firm i is its profits.

Each firm, thus, wants to maximize its profits. For example, firm 1's maximization problem is

$$\max_{q_1} \pi(q_1, q_2) = p(q_1 + q_2)q_1 - c_1(q_1)$$

Firm 1's profits depend on the amount of output chosen by Firm 2, and in order to make an informed decision Firm 1 must forecast Firm 2's output decision.

A Nash equilibrium of this games is a set of outputs (q_1^*, q_2^*) in which each firm is choosing its profit-maximizing output level.

Let us compute now the Nash equilibrium of Cournot's game using specific forms of the functions p and c_i:

- Let $p(Q) = \alpha - \beta Q$. Recall that $Q = q_1 + q_2$
- Let $c_i = 0$

Firm 1's profit is then, $\pi(q_1, q_2) = [\alpha - \beta(q_1 + q_2)]q_1$; and its best response function is,

$$B_1(q_2) = \max_{q_1 \in q}\{\pi_1(q_1, q_2)\}$$

We can take the first derivative of π_1 with respect to q_1, and the FOC is

$$\frac{\partial \pi_1}{\partial q_1}\{(\alpha - \beta q_2)q_1 - (\beta q_1)^2\} = \alpha - \beta q_2 - 2\beta q_1 = 0$$

Therefore, $B_1(q_2)$ is:

$$q_1^* = \frac{\alpha - \beta q_2}{2\beta}$$

and in the case of Firm 2, $\frac{\partial \pi_1}{\partial q_1} = 0$ implies that,

$$q_2^* = \frac{\alpha - \beta q_1}{2\beta}$$

To find the NE, we should solve:

$$q_1^* = \frac{\alpha - \beta q_2^*}{2\beta}$$

$$q_2^* = \frac{\alpha - \beta q_1^*}{2\beta}$$

Substituting the second equation into the first, we get

$$q_1^* = \frac{\alpha - \beta(\frac{\alpha - \beta q_1^*}{2\beta})}{2\beta}$$

And, ... (I will spare you the algebra)

$$q_1^* = \frac{\alpha}{3\beta}$$

Similarly,

$$q_2^* = \frac{\alpha}{3\beta}$$

So, the NE is $q_1^* = q_2^* = \frac{\alpha}{3\beta}$

Suppose now that $\alpha = 9$ and $\beta = 1$, then $q_1^* = q_2^* = 3$, $Q = 6$ and $\pi_i^* = 9$.

The envelope theorem.

Suppose that $f(x, a)$ is a function of both x and a.

We generally interpret a as being a parameter determined outside the problem being studied and x as the variable we wish to study.

Suppose that x is chosen to maximize the function.

For each different value of a there will typically be a different optimal choice of x.

In sufficiently regular cases, we will be able to write the function $x(a)$ that gives us the optimal choice of x for each different value of a.

For example, in some economic problem the choice variable might be the amount consumed or produced of some good while the parameter a will be a price.

We can also define the (optimal) value function, $M(a) = f(x(a), a)$.

This tells us what the optimized value of f is for different choices of a.

Example 4 *Recall from our first example that for $f(x, a) = log\ x - ax$, the optimal value of x is $x(a) = \frac{1}{a}$. Therefore, the value function for this problem is given by*

$$M(a) = log\ (\tfrac{1}{a}) - \tfrac{a}{a} = -log\ a - 1.$$

For the example $g(x, b) = u(x) - bx$, we have $M(b) = u(x(b)) - bx(b)$.

Many times we are interested in how the optimized value changes as the parameter a changes.

It turns out that there is a simple way to calculate this change.

By definition we have $M(a) = f(x(a), a)$.

Differentiating both sides of this identity, we have

$$\frac{dM(a)}{da} = \frac{\partial f(x(a),a)}{\partial x}\frac{\partial x(a)}{\partial a} + \frac{\partial f(x(a),a)}{\partial a}.$$

Since $x(a)$ is the choice of x that maximizes f, we know that

$$\frac{\partial f(x(a),a)}{\partial x} = 0.$$

Substituting this into the above expression, we have

$$\frac{dM(a)}{da} = \frac{\partial f(x(a),a)}{\partial a}.$$

A better way to write this is

$$\frac{dM(a)}{da} = \left.\frac{\partial f(x(a),a)}{\partial a}\right|_{x=x(a)}.$$

In this notation it is clear that the derivative is taken holding x fixed at the optimal value $x(a)$.

In words: the total derivative of the value function with respect to the parameter is equal to the partial derivative when the derivative is evaluated at the optimal choice.

This statement is the simplest form of the envelope theorem.

It is worthwhile thinking about why this happens.

When a changes there are two effects: the change in a directly affects f and the change in a affects x which in turn affects f.

But if x is chosen optimally, a small change in x has a zero effect on f, so the indirect effect drops out and only the direct effect is left.

Example 5 *Continuing with the $f(x,a) = \log x - ax$ example, we recall that $M(a) = -\log a - 1$. Hence $M'(a) = -\frac{1}{a}$. We can also see this using the envelope theorem; by direct calculation we see that $\frac{\partial f(x,a)}{\partial a} = -x$. Setting x equal to its optimal value, we have $\frac{\partial f(x,a)}{\partial a} = -\frac{1}{a} = M'(a)$.*

For the case where $M(b) = g(x(b), b) = u(x(b)) - bx(b)$, we have $M'(b) = -x(b)$.

Comparative statics.

Another question of interest in the social sciences is how the optimal choice changes as a parameter changes.

Analysis of this sort is known as **comparative statics** analysis or **sensitivity analysis**.

The basic calculation goes as follows:

We know that the optimal choice function $x(a)$ must satisfy the condition

$$\frac{\partial f(x(a),a)}{\partial x} \equiv 0.$$

Differentiating both sides of this identity,

$$\frac{\partial^2 f(x(a),a)}{\partial x^2}\frac{dx(a)}{da} + \frac{\partial^2 f(x(a),a)}{\partial x \partial a} \equiv 0.$$

Solving for $\frac{dx(a)}{da}$, we have

$$\frac{dx(a)}{da} = -\frac{\frac{\partial^2 f(x(a),a)}{\partial x \partial a}}{\frac{\partial^2 f(x(a),a)}{\partial x^2}}$$

We know that the denominator of this expression is negative due to the second-order conditions for maximization.

Noting the minus sign preceding the fraction, we can conclude that

$$\text{sign}\frac{dx(a)}{da} = \text{sign}\frac{\partial^2 f(x(a),a)}{\partial x \partial a}.$$

Hence, the sign of the derivative of the optimal choice with respect to the parameter depends only on the second cross-partial of the objective function with respect to x and a.

The nice feature of this is that we don't actually have to repeat this calculation every time; we can simply use the information about the cross- partial.

Example 6 *If $f(x,a) = log\ x - ax$, we already know that $x(a) = \frac{1}{a}$. By direct calculation $x'(a) < 0$. But we could have seen this without solving the maximization problem simply by observing that*

$$\frac{\partial^2 f(x(a),a)}{\partial x \partial a} = -1 < 0$$

In the optimization problem with the objective function $g(x,b) = u(x) - bx$, we can see that

$$\text{sign } x'(b) = \text{sign } (-1) < 0$$

This is a remarkable example: we know almost nothing about the shape of the function$u(x)$, and yet we are able to determine how the optimal choice must change as the parameter changes simply by using the properties of the *form* of the objective function.

For minimization problems, all that changes is the sign of the denominator.

Since the second-order condition for minimization implies that the second derivative with respect to the choice variable is positive, we see that the sign of the derivative of the choice variable with respect to the parameter is the opposite of the sign of the cross-partial derivative.

Multivariate Maximization

Comparative Statics.

Suppose that $f(x_1, x_2)$, and that both x_1 and x_2 are functions of a (i.e. $x_1(a)$ and $x_2(a)$). Just as before, we may want to determine how the optimal choice changes as the parameter a changes.

We know that the optimal choices have to satisfy the first-order conditions

$$\frac{\partial f(x_1(a), x_2(a), a)}{\partial x_1} = 0$$

$$\frac{\partial f(x_1(a), x_2(a), a)}{\partial x_2} = 0.$$

Differentiating these two expressions with respect to a, we have

$$\frac{\partial^2 f}{\partial x_1^2}\frac{\partial x_1}{\partial a} + \frac{\partial^2 f}{\partial x_1 \partial x_2}\frac{\partial x_2}{\partial a} + \frac{\partial^2 f}{\partial x_1 \partial a} = 0$$

$$\frac{\partial^2 f}{\partial x_2 \partial x_1}\frac{\partial x_1}{\partial a} + \frac{\partial^2 f}{\partial x_2^2}\frac{\partial x_2}{\partial a} + \frac{\partial^2 f}{\partial x_2 \partial a} = 0$$

This is more conveniently written in matrix form as

$$\begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix} \begin{bmatrix} \frac{\partial x_1}{\partial a} \\ \frac{\partial x_2}{\partial a} \end{bmatrix} = \begin{bmatrix} -\frac{\partial^2 f}{\partial x_1 \partial a} \\ -\frac{\partial^2 f}{\partial x_2 \partial a} \end{bmatrix}.$$

If the matrix on the LHS of this expression is invertible, we can solve this system of equations to get

$$\begin{bmatrix} \frac{\partial x_1}{\partial a} \\ \frac{\partial x_2}{\partial a} \end{bmatrix} = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix}^{-1} \begin{bmatrix} -\frac{\partial^2 f}{\partial x_1 \partial a} \\ -\frac{\partial^2 f}{\partial x_2 \partial a} \end{bmatrix}.$$

Rather than invert the matrix, it is often easier to use **Cramer's rule** to solve the system of equations for $\frac{\partial x_1}{\partial a}$ and $\frac{\partial x_2}{\partial a}$.

For example, if we want to solve for $\frac{\partial x_1}{\partial a}$, we can apply Cramer's rule to express this derivative as the ratio of two determinants:

$$\frac{\partial x_1}{\partial a} = \frac{\begin{vmatrix} -\frac{\partial^2 f}{\partial x_1 \partial a} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ -\frac{\partial^2 f}{\partial x_2 \partial a} & \frac{\partial^2 f}{\partial x_2^2} \end{vmatrix}}{\begin{vmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{vmatrix}}.$$

By the second-order condition for maximization, the matrix in the denominator of this expression is a negative semidefinite matrix.

Matrix algebra tells us that this matrix must have a positive determinant.

Therefore, the sign of $\frac{\partial x_1}{\partial a}$ is simply the sign of the determinant in the numerator.

Example 7 *Let* $f(x_1, x_2, a_1, a_2) = u_1(x_1) + u_2(x_2) - a_1 x_1 - a_2 x_2$.

The first-order conditions for maximizing f are

$$u_1'(x_1^*) - a_1 = 0$$
$$u_2'(x_2^*) - a_2 = 0$$

The second order condition is that the matrix

$$\mathbf{H} = \begin{bmatrix} u_1''(x_1^*) & 0 \\ 0 & u_2''(x_2^*) \end{bmatrix}$$

is negative semidefinite.

Since a negative semidefinite matrix must have diagonal terms that are less than or equal to zero, it follows that $u_1''(x_1^*) \leq 0$ and $u_2''(x_2^*) \leq 0$.

The maximized value function is given by

$$M(a_1, a_2) \equiv \max_{x1,x2} u_1(x_1) + u_2(x_2) - a_1 x_1 - a_2 x_2,$$

and a simple calculation using the envelope theorem shows that

$$\frac{\partial M}{\partial a_1} = -x_1^*$$
$$\frac{\partial M}{\partial a_2} = -x_2^*$$

The comparative statics calculation immediately above shows that

$$sign\,\frac{\partial x_1}{\partial a_1} = sign\begin{vmatrix} 1 & 0 \\ 0 & u_2''(x_2^*) \end{vmatrix}.$$

Carrying out the calculation of the determinant,

$$sign\,\frac{\partial x_1}{\partial a_1} \leq 0.$$

Note that we can determine how the choice variable responds to changes in the parameter without knowing anything about the explicit functional form of u_1 or u_2; we only have to know the *structure* of the objective function –in this case, that it is **additively separable**.

Constrained Optimization

I will use demand theory in economics to explain and then apply the concept of constrained optimization.

Utility Maximization

Demand theory is based on the premise that consumers maximize utility subject to a budget constraint. Utility is assumed to be an increasing function of the quantities of goods consumed, but marginal utility is assumed to decrease with consumption.

Suppose there are only two goods, x and y available to the consumer. Let m be the fixed amount o money available to her, and let p_x and p_y be the prices of the two goods.

Suppose that the consumer maximizes her utility subject to the constraint that all income is spent on the two goods. The problem of utility maximization can then be written as

$$max\ u(x, y) \tag{1}$$

subject to

$$p_x x + p_y y = m \tag{2}$$

To determine the individual consumer's demand for the two goods, we choose those values of x and y that maximize (1) subject to (2) (to simplify our life, we assume that the utility function is continuous and that goods are infinitely divisible).

The Consumer's Optimum

To solve the constrained optimization problem given by equations (1) and (2), we can use the method of Lagrange multipliers. We first write the "Lagrangian" for the problem. To do this, we rewrite the constraint as $p_x x + p_y y - m = 0$. The Lagrangian is then

$$\mathcal{L} = u(x, y) - \lambda(p_x x + p_y y - m) \tag{3}$$

If we choose values of x and y that satisfy the budget constraint, then the second term in (3) will be zero, and maximizing $\mathcal{L}$ will be equivalent to maximizing $u(x, y)$. By differentiating $\mathcal{L}$ with respect to x, y and λ and then equating the derivatives to zero, we obtain the necessary conditions for a maximum:

$$\begin{aligned} \frac{\partial u(x,y)}{\partial x} - \lambda p_x = 0 \\ \frac{\partial u(x,y)}{\partial y} - \lambda p_y = 0 \\ p_x x + p_y y - m = 0 \end{aligned} \tag{4}$$

The third condition is the original budget constraint. The first two conditions of (4) tell us that each good will be consumed up to the point at which the marginal utility from consumption is a multiple (λ) of the price of the good. To see the implication of this, we combine the first two conditions to obtain the *equal marginal principle*:

$$\lambda = \frac{\frac{\partial u(x,y)}{\partial x}}{p_x} = \frac{\frac{\partial u(x,y)}{\partial y}}{p_y} \tag{5}$$

In other words, the marginal utility of each good divided by its price is the same. To be optimizing, the consumer must be getting the same utility from the last dollar spent by consuming either x or y. Were this not the case, consuming more of one one good and less of the other would increase utility.

To characterize the individual's optimum in more detail, we can rewrite the information in (5) to obtain

$$\frac{\frac{\partial u(x,y)}{\partial x}}{\frac{\partial u(x,y)}{\partial y}} = \frac{p_x}{p_y} \tag{6}$$

The fraction on the left is the *marginal rate of substitution* between good x and y, and the fraction on the right might be called the *economic rate of substitution* between goods x and y.

In general, the three equations in (4) can be solved to determine the three unknowns x, y, and λ as a function of the two prices and income. Substitution for λ then allows us to solve for the demands of each of the two goods in terms of income and the prices of the two commodities. This can be seen in terms of an example:

Example 8 *A frequently used utility function is the Cobb-Douglas utility function, which can be represented in two forms:*

$$u(x,y) = x^{\alpha} y^{1-\alpha}$$

and

$$u(x,y) = \alpha \ log(x) + (1-\alpha) \ log(y)$$

The two forms are equivalent for the purposes of demand theory because they both yield the identical demand functions for goods x and y.

I will derive the demand functions for the second form. To find the demand functions for x and y, given the usual budget constraint, we first write the Lagrangian

$$\mathcal{L} = \alpha \ log(x) + (1-\alpha) \ log(y) - \lambda(p_x x + p_y y - m)$$

Now differentiating with respect to x, y, and λ, and setting the derivatives equal to zero, we obtain

$$\frac{\partial \mathcal{L}}{\partial x} = \frac{\alpha}{x} - \lambda p_x = 0$$
$$\frac{\partial \mathcal{L}}{\partial y} = \frac{(1-\alpha)}{y} - \lambda p_y = 0$$
$$\frac{\partial \mathcal{L}}{\partial \lambda} = p_x x + p_y y - m = 0$$

The first two conditions imply that

$$\begin{aligned} p_x x &= \frac{\alpha}{\lambda} \\ p_y y &= \frac{(1-\alpha)}{\lambda} \end{aligned} \tag{7}$$

Combining these with the last condition (the budget constraint) gives:

$$\frac{\alpha}{\lambda} + \frac{(1-\alpha)}{\lambda} - m = 0,$$

or

$$\lambda = \frac{1}{m}.$$

Now we can substitute this expression for λ back into (7) to obtain the demand functions:

$$x = \frac{\alpha m}{p_x},$$

and

$$y = \frac{(1-\alpha)m}{p_y}.$$

In this example, the demand for each good depends only on the price of that good and on income, and not on the price of the other good. Thus, the cross-price elasticities of demand are 0.