

LECTURE 8: MULTIVARIABLE STATIC OPTIMIZATION

I. INTRODUCTION

- Many economic models are derived from the behavior of individuals, firms or policy makers who are trying to maximize some measure of welfare subject to constraints on their behavior. Optimization is, therefore, the most important mathematical concept in economics
- Our study of optimization encompasses three categorizations of problems. The first distinction is between **univariate** and **multivariate** optimization, i.e. finding extreme values of functions with one variable vs. finding extreme values of functions with many variables. Read Chapter 9 of Klein and bring yourself back up to speed quickly on the first and second order conditions for a solution to a univariate maximization problem.
- The second distinction is between **constrained** and **unconstrained** optimization. You should be familiar with both multivariate optimization from Econ 201. What we do here should add to that knowledge.
- The final distinction is between **static** and **dynamic** optimization, i.e. between one shot optimization decisions and decisions in which your current choice may affect a subsequent optimization decision.

II. UNIVARIATE OPTIMIZATION

- Optimization lies at the heart of economic behavior. Therefore, it is vital that we be able to figure out what the extreme values of a function are. Here, I will first derive the necessary first order (FOC) and sufficient second order (SOC) conditions for finding the extreme value(s) of a univariate function.
- I will be using differentials to find the extreme value(s) of a function. This will allow for an easier extension of this analysis from the univariate case to the multivariate case.

First Order (Necessary) Conditions

- Given a function of the form $y = f(x)$ we can calculate the differential as $dy = f'(x)dx$. If x^* is a candidate for an extreme point then the value of y should not change for any small change dx in the vicinity of x^* . In other words we should see that $dy \equiv f'(x^*)dx = 0$ for all dx in the vicinity of x^* . This only holds true when $f'(x^*) = 0$, which is the traditional first order condition
- Keep in mind that the FOC is only a necessary condition for an extreme point, not a sufficient one. It is possible that $f'(x^*) = 0$ even when x^* is neither a minimum nor a maximum. Consider the function $y = x^3$. Since $f'(x) = 3x^2$, we can show that $f'(0) = 0$. However, zero is neither a local minimum nor a local maximum of $f(x)$.

Second Order (Sufficient) Conditions

- The FOC does not tell us whether the extreme point is a minimum or a maximum. For that, we need to think about SOCs. To determine the nature of the extreme point, we look at the behavior of the differential around the extreme point. To do this we calculate the second order differential of y as $d^2y = d[f'(x)dx] = f''(x)dx^2$.
- If $y = f(x)$ has a local maximum at $x = x^*$ then not only will dy be zero at x^* , dy has to also be declining when x changes in the neighborhood of x^* . In other words in the vicinity of a peak we should have $d^2y \equiv f''(x^*)dx^2 < 0$. So the second order sufficient condition for a **local maximum** is that $f''(x^*) < 0$.
- If $y = f(x)$ has a local minimum at $x = x^*$ then not only will dy be zero at x^* , dy has to also be increasing when x changes in the neighborhood of x^* . In other words in the vicinity of a valley we should have $d^2y \equiv f''(x^*)dx^2 > 0$. So the second order sufficient condition for a **local minimum** is that $f''(x^*) > 0$.
- If x^* is neither a local maximum or a local minimum then the value of dy will be increasing when x changes in one direction and decreasing when x changes in the other direction, in the neighborhood of x^* . We have $d^2y \equiv f''(x^*)dx^2 = 0$.

Concavity, Convexity and Global Extrema

- We can also distinguish between local and global extreme points. If the function $f(x)$ is **concave**, then any extreme point is a **global maximum**. In other words if $f''(x) \geq 0$ everywhere, then x^* is a global minimum.
- Furthermore, if the function $f(x)$ is **strictly concave**, then any extreme point is a **unique global maximum**. In other words if $f''(x) > 0$ everywhere, then x^* is a unique global minimum.
- Similarly, if the function $f(x)$ is **convex**, then any extreme point is a **global maximum**. In other words if $f''(x) \leq 0$ everywhere, then x^* is a global maximum.
- Furthermore, if the function $f(x)$ is **strictly convex** then any extreme point is a **unique global maximum**. In other words if $f''(x) < 0$ everywhere, then x^* is a unique global maximum.

III. MULTIVARIATE OPTIMIZATION

First Order (Necessary) Conditions

- We can extend this analysis to thinking about multivariate optimization decision problems using differentials.
- Given a generic multivariate function of the form $z = f(x_1 \cdots x_n)$ we can calculate the differential of z as $dz = \sum_{i=1}^n f_i(x_1, x_2, \dots, x_n)dx_i$. At an extreme point $z^* = (x_1^*, x_2^*, \dots, x_n^*)$, we must have $dz \equiv \sum_{i=1}^n f_i(x_1^*, x_2^*, \dots, x_n^*)dx_i = 0$ or else we would be able to reach a higher or lower point in the immediate vicinity.

- Given this, the first order necessary conditions for an extreme point of f are that each of the partial derivatives of f , evaluated at the stationary point, must equal zero. In other words if $(x_1^*, x_2^*, \dots, x_n^*)$ is an extreme point, then the following must be true $f_i(x_1^*, x_2^*, \dots, x_n^*) = 0$, $\forall i = 1, \dots, n$

Example:

Consider the function $z = -x^2 + xy - y^2 + 3x$. The extreme points of z satisfy the first order conditions (FOCs)

$$\begin{aligned} z_x &\equiv -2x + y + 3 = 0 \\ z_y &\equiv x - 2y = 0 \end{aligned}$$

- The solution to this (which you can calculate by solving this simple 2×2 equation system) is $x = 2$ and $y = 1$. The extreme value of z is 3.
- Note that the FOC does not tell us whether the extreme value is a minimum or a maximum (or neither, i.e. a point of inflection). Hence, we need a second order condition to ascertain the true nature of any extremum of z .

Second Order (Sufficient) Conditions

- We will apply the same principle we used in the univariate case to a multivariate function. The required steps are much more complicated because the second order conditions do not involve a single variable any more. In fact, with n variables we have n^2 different second order derivatives ($n(n+1)/2$ of which are unique) to calculate.
- We can calculate the 2nd order differential of z , d^2z , as

$$\begin{aligned} d^2z &= \frac{\partial dz}{\partial x_1} dx_1 + \frac{\partial dz}{\partial x_2} dx_2 + \dots + \frac{\partial dz}{\partial x_n} dx_n \\ &= \frac{\partial (f_1 dx_1 + f_2 dx_2 + \dots + f_n dx_n)}{\partial x_1} dx_1 + \frac{\partial (f_1 dx_1 + f_2 dx_2 + \dots + f_n dx_n)}{\partial x_2} dx_2 \\ &\quad + \dots + \frac{\partial (f_1 dx_1 + f_2 dx_2 + \dots + f_n dx_n)}{\partial x_n} dx_n \\ &= (f_{11} dx_1^2 + f_{12} dx_1 dx_2 + \dots + f_{1n} dx_1 dx_n) + (f_{21} dx_2 dx_1 + f_{22} dx_2^2 + \dots + f_{2n} dx_2 dx_n) \\ &\quad + \dots + (f_{n1} dx_1 dx_n dx_1 + f_{n2} dx_n dx_2 + \dots + f_{nn} dx_n^2) \end{aligned}$$

- Since the order of differentiation is immaterial (Young's Theorem) $f_{ij} = f_{ji}$. We therefore get the following expression for the 2nd differential of z .

$$d^2z = \sum_{i=1}^n f_{ii} dx_i^2 + 2 \sum_{i=1}^n \sum_{j=i+1}^n f_{ij} dx_i dx_j$$

- The SOC can then be defined as follows: the function $z = f(x_1, x_2, \dots, x_n)$ has a local maximum at $x^* \equiv (x_1^*, x_2^*, \dots, x_n^*)$ if $d^2z < 0$ at x^* . In this case d^2z is said to be **negative definite**.

- The function $z = f(x_1, x_2, \dots, x_n)$ has a local minimum at $x^* \equiv (x_1^*, x_2^*, \dots, x_n^*)$ if $d^2z > 0$ at x^* . In this case d^2z is said to be **positive definite**.
- The extreme point is neither a minimum nor a maximum (a saddle point, the multivariable analog to a point of inflection) if $d^2z > 0$ for some values and $d^2z < 0$ for others near x^* .
- The intuition is as follows. Suppose we have found a prospective candidate for an extreme point, i.e. some point $z^* = f(x_1^*, x_2^*, \dots, x_n^*)$ at which $dz^* = 0$. Suppose we move away from z^* in all directions and find that dz becomes negative. Since dz was zero at z^* and < 0 in the vicinity of z^* , we know that z^* is a local maximum. In other words, if dz is declining in the neighborhood of z^* , given that it is zero at z^* , then z^* is a peak. We can also express this as follows: **if $d^2z < 0$ at z^* then z^* is a local maximum**.
- Conversely, suppose we move away from z^* in all directions and find that dz becomes positive. Since dz was zero at z^* and > 0 in the vicinity of z^* , we know that z^* is a local minimum. In other words, if dz is increasing in the neighborhood of z^* , given that it is zero at z^* , then z^* is a valley. We can also express this as follows: **if $d^2z > 0$ at z^* then z^* is a local minimum**.
- If dz could be either increasing or decreasing in value depending on which direction we move in, given that it equaled zero at z^* , then z^* is said to be a saddle-point (the multivariable analog to a point of inflection). We can also express this as follows: **if $d^2z <> 0$ at z^* (depending on which direction we move in) then z^* is a saddle-point**.
- To check the second order condition, we need to calculate

$$d^2z \equiv \sum_{i=1}^n f_{ii} dx_i^2 + 2 \sum_{i=1}^n \sum_{j=i+1}^n f_{ij} dx_i dx_j$$

and check its value in the vicinity of $(x_1^*, x_2^*, \dots, x_n^*)$. This can be pretty tedious so there is an easier way to test the SOC by using a special matrix called the Hessian matrix.

- Given a function $z = f(x_1, x_2, \dots, x_n)$, the **Hessian matrix** associated with that function is

$$H = \begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1n} \\ f_{21} & f_{22} & \cdots & f_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ f_{n1} & f_{n2} & \cdots & f_{nn} \end{bmatrix}$$

- From this Hessian matrix, we can construct a sequence of determinants known as **Principal Minors**. These are distinct from the minors that we talked about earlier. The n principal minors of H are

$$|H_1| = |f_{11}|, |H_2| = \begin{vmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{vmatrix}, |H_3| = \begin{vmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{vmatrix}, \dots, |H_n| = \begin{vmatrix} f_{11} & f_{12} & \cdots & f_{1n} \\ f_{21} & f_{22} & \cdots & f_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ f_{n1} & f_{n2} & \cdots & f_{nn} \end{vmatrix}$$

- Basically the j th principal minor is the determinant of the matrix constructed by taking the first j rows and j columns of the Hessian. The SOC for multivariable static optimization is (the proof is very nasty, come see me if you are masochistic)

- z^* is a **local maximum** (d^2z is **negative definite**) if $|H_1| < 0, |H_2| > 0, |H_3| < 0, \dots, (-1)^n |H_n| > 0$ i.e. the principal minors, evaluated at $x = x^*$, are of alternating signs with the 1st one being negative.
- z^* is a **local minimum** (d^2z is **positive definite**) if $|H_1| > 0, |H_2| > 0, |H_3| > 0, \dots, |H_n| > 0$, i.e. the principal minors, evaluated at $x = x^*$, are all positive.

Example:

In the case $z = -x^2 + xy - y^2 + 3x$, we showed that the stationary points of z were $x = 2$ and $y = 1$. The Hessian is

$$H = \begin{bmatrix} z_{xx} & z_{xy} \\ z_{yx} & z_{yy} \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}$$

The principal minors are $|H_1| = -2, |H_2| = \begin{vmatrix} -2 & 1 \\ 1 & -2 \end{vmatrix} = 3$. Given the alternating signs of the principal minors, we can conclude that z^* is a maximum.

Concavity, Convexity and Global Extrema

- The final piece of theory relates to calculating the concavity and convexity of a multivariate function. Recall that in the univariate case, we said that a function $f(x)$ was said to be concave if the following identity holds for any two points x_1 and x_2 in the function's domain, where $0 < \lambda < 1$

$$f(\lambda x_1 + (1 - \lambda)x_2) \geq \lambda f(x_1) + (1 - \lambda)f(x_2)$$

- Similarly, a function $f(x)$ was said to be convex if the following identity holds.

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2)$$

- The same definition holds in the multivariate case, except that x is now a vector (x_1, x_2, \dots, x_n) instead of a single variable. Calculating whether a multivariate function is concave or convex is difficult. We can't draw a graph very easily and it is hard to think about points of a function lying above planes which would be the multivariate analog to how we visually identified whether a univariate function was concave or not.
- We can use the Hessian and the principal minors to test for concavity or convexity of a function. The definitions are as follows:
 - A function $z = f(x_1, x_2, \dots, x_n)$ is said to be **concave** if d^2z is *everywhere negative semi-definite*, i.e. if $|H_1| \leq 0, |H_2| \geq 0, |H_3| \leq 0, \dots, (-1)^n |H_n| \geq 0$
 - A function $z = f(x_1, x_2, \dots, x_n)$ is said to be **convex** if d^2z is *everywhere positive semi-definite*, i.e. if $|H_1| \geq 0, |H_2| \geq 0, |H_3| \geq 0, \dots, |H_n| \geq 0$
- A concave function has a global maximum, but it is not necessarily unique, i.e. it may have more than 1 peak. A convex function has a global minimum, but it is not necessarily unique so it may have more than 1 low point.

- Note that the distinction from the SOC discussed earlier is that the principal minors exhibit the behavior at **all points** in the domain, not just in the vicinity of the extreme point.
- A strictly concave function has a unique global maximum, so the extreme point we find using the FOCs is the unique maximum of the function.
- A strictly convex function has a unique global minimum, so the extreme point we find using the FOCs is the unique minimum of the function.
- A function $z = f(x_1, x_2, \dots, x_n)$ is said to be **strictly concave** if d^2z is *everywhere* **negative definite**, i.e. if $|H_1| < 0, |H_2| > 0, |H_3| < 0, \dots, (-1)^n |H_n| < 0$
- A function $z = f(x_1, x_2, \dots, x_n)$ is said to be **strictly convex** if d^2z is *everywhere* **positive definite**, i.e. if $|H_1| > 0, |H_2| > 0, |H_3| > 0, \dots, |H_n| > 0$

IV. APPLICATIONS

Profit Maximization by a Multi-Product Firm

- Consider a firm that sells two goods A and B, in two perfectly competitive markets. Assume that the cost function of the firm is given by $C(Q_A, Q_B) = 3Q_A^2 + 2Q_B^2 + 4Q_AQ_B$.
- Since the prices of the goods are exogenous to the firm, the profit function can be written as $\Pi(Q_A, Q_B) = P_AQ_A + P_BQ_B - (3Q_A^2 + 2Q_B^2 + 4Q_AQ_B)$
- The producer's decision is $\max_{Q_A, Q_B} \Pi(Q_A, Q_B)$ and the FOC are

$$\begin{aligned} P_A - 6Q_A - 4Q_B &= 0 \\ P_B - 4Q_B - 4Q_A &= 0 \end{aligned}$$

- Rewriting this in matrix form gives us

$$\begin{bmatrix} 6 & 4 \\ 4 & 4 \end{bmatrix} \begin{bmatrix} Q_A \\ Q_B \end{bmatrix} = \begin{bmatrix} P_A \\ P_B \end{bmatrix}$$

- Using Cramer's Rule, we can show that the solutions are

$$Q_A = \frac{1}{|A|} \begin{vmatrix} P_A & 4 \\ P_B & 4 \end{vmatrix} = \frac{1}{8} (4P_A - 4P_B), Q_B = \frac{1}{|A|} \begin{vmatrix} 6 & P_A \\ 4 & P_B \end{vmatrix} = \frac{1}{8} (6P_B - 4P_A)$$

- We can check the SOC by calculating the Hessian as $H = \begin{bmatrix} -6 & -4 \\ -4 & -4 \end{bmatrix}$, which in turn implies that the principal minors are $|H_1| = -6$ and $|H_2| = \begin{vmatrix} -6 & -4 \\ -4 & -4 \end{vmatrix} = 8$.
- Since the profit function is strictly concave everywhere, we conclude that there is a unique global maximum at $Q_A = \left[\frac{P_A - P_B}{2} \right], Q_B = \left[\frac{3P_B - 2P_A}{4} \right]$

Price Discrimination

- Consider a slightly different variation. This time we are dealing with a firm that sells a single good in two different markets, and needs to decide how much of that good to sell in each market. An example might be an airline which sells a single product, e.g. seats for air travel, in two markets: business fliers and leisure fliers.
- We will also use a general cost function and revenue function with parameters rather than specific numeric values. Finally, we will assume that the market is in fact a monopoly (so Air France may be the more appropriate example instead of Delta Airlines).
- Assume that the revenue functions of the firm in the two markets are $R(Q_A)$ and $R(Q_B)$. The total cost function is given by $C(Q)$ where $Q = Q_A + Q_B$. Also assume that the cost function is convex in Q and that the revenue functions are concave.
- The profit function of the airline can be written as $\Pi(Q_A, Q_B) = R_A(Q_A) + R_B(Q_B) - C(Q)$
- The producer's maximization decision is $\max_{Q_A, Q_B} \Pi(Q_A, Q_B)$ and the FOC are

$$R'_A(Q_A) - C'(Q) \frac{\partial Q}{\partial Q_A} \equiv R'_A(Q_A) - C'(Q) = 0$$

$$R'_B(Q_B) - C'(Q) \frac{\partial Q}{\partial Q_B} \equiv R'_B(Q_B) - C'(Q) = 0$$

- The first order condition state that marginal revenue in each market equal marginal cost of production, which is as we would expect with any maximizing decision.
- We can check the SOC by calculating the Hessian as $H = \begin{bmatrix} R''_A(Q_A) - C''(Q) & -C''(Q) \\ -C''(Q) & R''_B(Q_B) - C''(Q) \end{bmatrix}$, which in turn implies that the principal minors are $|H_1| = R''_A(Q_A) - C''(Q)$ and $|H_2| = \begin{vmatrix} R''_A(Q_A) - C''(Q) & -C''(Q) \\ -C''(Q) & R''_B(Q_B) - C''(Q) \end{vmatrix} = R''_A R''_B - C''(R''_A + R''_B)$.
- Given the assumptions that the cost function was strictly convex, and revenue was strictly concave we know that $C'' > 0$ and $R''_A, R''_B < 0$ everywhere.
- Then the sign of $|H_1|$ will be negative and the sign of $|H_2|$ is positive. Thus if we have a strictly concave revenue function and a strictly convex cost function we would get a unique global solution to our maximization problem.
- If we had a concave revenue function and a convex cost function we would get a global solution to our maximization problem, that may not be unique.