ECON 186 Class Notes: Optimization Part 2

Jijian Fan

Hessians

- The Hessian matrix is a matrix of all partial derivatives of a function.
- Given the function $f(x_1, x_2, ..., x_n)$, where all partial derivatives exist and are continuous, the Hessian of f is

$$H(f) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

- Given the quadratic form $d^2 z = f_{xx} dx^2 + 2f_{xy} dxdy + f_{yy} dy^2$, the Hessian determinant (sometimes called the Hessian) is

$$|H| = \begin{vmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{vmatrix}$$

Examples

- Is $q = 5u^2 + 3uv + 2v^2$ either positive or negative definite?
  - The discriminant of q is $\begin{vmatrix} 5 & 1.5 \\ 1.5 & 2 \end{vmatrix}$, with first leading principal minor $|5| > 0$ and second leading principal minor $\begin{vmatrix} 5 & 1.5 \\ 1.5 & 2 \end{vmatrix} = 10 - 2.25 = 7.75 > 0$
  - So, q is positive definite.
- Given $f_{xx} = -2$, $f_{xy} = 1$, and $f_{yy} = -1$ at a certain point on a function $z = f(x, y)$, does d^2z have a definite sign at that point?
  - The discriminant of the quadratic form d^2z is $\begin{vmatrix} -2 & 1 \\ 1 & -1 \end{vmatrix}$, which has leading principal minors $-2 < 0$ and $\begin{vmatrix} -2 & 1 \\ 1 & -1 \end{vmatrix} = 1 > 0$, so d^2z is negative definite, which means the point in question is a local maximum.

Three-variable Quadratic Forms

- Similar conditions can analogously be obtained for a function of three or more variables. Consider a quadratic form q with three variables u_1, u_2, and u_3. Then:

$$q(u_1, u_2, u_3) = d_{11}(u_1^2) + d_{12}(u_1 u_2) + d_{13}(u_1 u_3) + d_{21}(u_2 u_1) + d_{22}(u_2^2)$$

$$+d_{23}(u_2 u_3) + d_{31}(u_3 u_1) + d_{32}(u_3 u_2) + d_{33}(u_3^2) = \sum_{i=1}^{3} \sum_{j=1}^{3} d_{ij} u_i u_j$$

$$= \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix} \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \equiv u' D u$$

Three-variable Quadratic Forms

- Now, there are three leading principal minors:

$$|D_1| \equiv d_{11} \qquad |D_2| \equiv \begin{vmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{vmatrix} \qquad |D_3| \equiv \begin{vmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{vmatrix}$$

- The sufficient condition for positive definiteness (local minimum) is that $|D_1| > 0$, $|D_2| > 0$, and $|D_3| > 0$.
- The sufficient condition for negative definiteness (local maximum) is that $|D_1| < 0$, $|D_2| > 0$, and $|D_3| < 0$.

Examples

- Find and classify the critical points of the function $f(x,y,z) = x^2 + y^2 + 7z^2 + xy + 3yz$.
  - $f_x = 2x + y$, $f_y = 2y + x + 3z$, $f_z = 14z + 3y$. It is easy to see that the only critical point is $(0,0,0)$.
  - $f_{xx} = 2$, $f_{yy} = 2$, $f_{zz} = 14$, $f_{xy} = f_{yx} = 1$, $f_{xz} = f_{zx} = 0$, $f_{yz} = f_{zy} = 3$. We then compute the Hessian:

$$\begin{vmatrix} f_{xx} & f_{yx} & f_{zx} \\ f_{yx} & f_{yy} & f_{yz} \\ f_{zx} & f_{zy} & f_{zz} \end{vmatrix} = \begin{vmatrix} 2 & 1 & 0 \\ 1 & 2 & 3 \\ 0 & 3 & 14 \end{vmatrix}$$

  - $|D_1| = 2 > 0$, $|D_2| = \begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} = 4 - 1 = 3 > 0$, $|D_3| = \begin{vmatrix} 2 & 1 & 0 \\ 1 & 2 & 3 \\ 0 & 3 & 14 \end{vmatrix} =$ $2\begin{vmatrix} 2 & 3 \\ 3 & 14 \end{vmatrix} - 1\begin{vmatrix} 1 & 3 \\ 0 & 14 \end{vmatrix} + 0\begin{vmatrix} 1 & 2 \\ 0 & 3 \end{vmatrix} = 2(28-9) - 14 = 24 > 0$
  - So, since the Hessian is positive definite, the only critical point $(0,0,0)$ is a local minimum.

Examples

- Find the extreme values of $f(x_1, x_2, x_3) = z = -x_1^3 + 3x_1x_3 + 2x_2 - x_2^2 - 3x_3^2$
  - $f_1 = -3x_1^2 + 3x_3,\ f_2 = 2 - 2x_2,\ f_3 = 3x_1 - 6x_3$
  - So, we have a system of three equations:

$$-3x_1^2 + 3x_3 = 0 \rightarrow x_3 = x_1^2 \rightarrow x_3 = \left(\tfrac{1}{2}\right)^2 = \tfrac{1}{4}$$
$$2 - 2x_2 = 0 \rightarrow x_2 = 1$$
$$3x_1 - 6x_3 = 0 \rightarrow x_1 = 2x_3 \rightarrow x_1 = 2x_1^2 \rightarrow x_1 = \tfrac{1}{2}$$

  - Additionally, since $x_1 - 2x_3 = 0$, $(0, 1, 0)$ must also be a solution, so the two roots are $(0, 1, 0)$ and $\left(\frac{1}{2}, 1, \frac{1}{4}\right)$.

Examples

- $f_{11} = -6x_1,\ f_{22} = -2,\ f_{33} = -6,\ f_{12} = f_{21} = 0,\ f_{13} = f_{31} = 3,$ $f_{23} = f_{32} = 0$
- So, the Hessian is
$$\begin{vmatrix} -6x_1 & 0 & 3 \\ 0 & -2 & 0 \\ 3 & 0 & -6 \end{vmatrix}$$
- $|D_1|(0,1,0) = 0$, so we already know that the point $(0,1,0)$ is indefinite, and in fact is not an extremum at all.
- $|D_1|(\frac{1}{2},1,\frac{1}{4}) = -3 < 0$, $|D_2|(\frac{1}{2},1,\frac{1}{4}) = \begin{vmatrix} -3 & 0 \\ 0 & -2 \end{vmatrix} = 6 > 0$,
$$|D_3|(\tfrac{1}{2},1,\tfrac{1}{4}) = \begin{vmatrix} -3 & 0 & 3 \\ 0 & -2 & 0 \\ 3 & 0 & -6 \end{vmatrix} = -3\begin{vmatrix} -2 & 0 \\ 0 & -6 \end{vmatrix} + 3\begin{vmatrix} 0 & -2 \\ 3 & 0 \end{vmatrix} =$$
$-36 + 18 = -18 < 0$
- The Hessian is negative definite, so the point $\left(\frac{1}{2},1,\frac{1}{4}\right)$ is a maximum.

Profit Maximization Example

- Consider a competitive firm with the following profit function:

$$\pi = R - C = PQ - wL - rK$$

- where P is price, Q is output, L is labor, K is capital, w is wage, r is the rental rate of capital. Since the firm is in a competitive market, P, w, and r are exogenous, while L, K, and Q are endogenous. However, Q is also function of K and L via the Cobb-Douglas production function

$$Q = Q(K, L) = L^{\alpha} K^{\beta}$$

- Assume that there are decreasing returns to scale where $\alpha = \beta < \frac{1}{2}$. Substituting in, the objective function becomes

$$\pi(K, L) = PL^{\alpha} K^{\alpha} - wL - rK$$

Profit Maximization Example

- First order conditions:

$$\frac{\partial \pi}{\partial L} = P\alpha L^{\alpha-1}K^{\alpha} - w = 0 \to K = \left(\frac{w}{P\alpha}L^{1-\alpha}\right)^{\frac{1}{\alpha}}$$

$$\frac{\partial \pi}{\partial K} = P\alpha L^{\alpha}K^{\alpha-1} - r = 0$$

Profit Maximization Example

- Before we continue, let's make sure that these equations for L and K do actually give us a maximum.

$$|H| = \begin{vmatrix} \pi_{LL} & \pi_{LK} \\ \pi_{KL} & \pi_{KK} \end{vmatrix} = \begin{vmatrix} P\alpha(\alpha-1)L^{\alpha-2}K^{\alpha} & P\alpha^2 L^{\alpha-1}K^{\alpha-1} \\ P\alpha^2 L^{\alpha-1}K^{\alpha-1} & P\alpha(\alpha-1)L^{\alpha}K^{\alpha-2} \end{vmatrix}$$

$$= P^2\alpha^2(\alpha-1)^2 L^{2\alpha-2}K^{2\alpha-2} - P^2\alpha^4 L^{2\alpha-2}K^{2\alpha-2}$$

$$= P^2\alpha^2(\alpha^2-2\alpha+1)L^{2\alpha-2}K^{2\alpha-2} - P^2\alpha^4 L^{2\alpha-2}K^{2\alpha-2}$$

$$= P^2\alpha^2 L^{2\alpha-2}K^{2\alpha-2}(1-2\alpha) + P^2\alpha^4 L^{2\alpha-2}K^{2\alpha-2} - P^2\alpha^4 L^{2\alpha-2}K^{2\alpha-2}$$

$$= P^2\alpha^2 L^{2\alpha-2}K^{2\alpha-2}(1-2\alpha)$$

- $|H_1| = P\alpha(\alpha-1)L^{\alpha-2}K^{\alpha} < 0$ and $|H| > 0$, so the hessian is negative definite, so L and K as defined by the FOC's represents the optimal quantities that will maximize profit.

Profit Maximization Example

- Plugging in the FOC for L into the FOC for K, we get

$$P\alpha L^{\alpha}K^{\alpha-1} - r = P\alpha L^{\alpha}\left[\left(\frac{w}{P\alpha}L^{1-\alpha}\right)^{\frac{1}{\alpha}}\right]^{\alpha-1} - r$$

$$= 0 \rightarrow P\alpha L^{\alpha}\left[\left(\frac{w}{P\alpha}\right)^{\frac{1}{\alpha}}L^{\frac{1-\alpha}{\alpha}}\right]^{\alpha-1} = P\alpha\left(\frac{w}{P\alpha}\right)^{\frac{\alpha-1}{\alpha}}L^{\frac{(1-\alpha)(\alpha-1)}{\alpha}+\alpha} - r$$

$$= P^{-\frac{\alpha-1}{\alpha}+1}\alpha^{-\frac{\alpha-1}{\alpha}+1}L^{\frac{-\alpha^2+2\alpha-1+\alpha^2}{\alpha}}w^{\frac{\alpha-1}{\alpha}} - r$$

$$= (P\alpha)^{\frac{1}{\alpha}}L^{\frac{2\alpha-1}{\alpha}}w^{\frac{\alpha-1}{\alpha}} - r = 0 \rightarrow (P\alpha)^{\frac{1}{\alpha}}L^{\frac{2\alpha-1}{\alpha}}w^{\frac{\alpha-1}{\alpha}} = r \rightarrow$$

$$(P\alpha)^{\frac{1}{\alpha}}w^{\frac{\alpha-1}{\alpha}}r^{-1} = L^{-\frac{2\alpha-1}{\alpha}} = L^{\frac{1-2\alpha}{\alpha}} \rightarrow L^{*} = \left(P\alpha w^{\alpha-1}r^{-\alpha}\right)^{\frac{1}{1-2\alpha}}$$

Profit Maximization Example

- Similarly, we can find that $K^* = \left(P\alpha r^{\alpha-1} w^{-\alpha}\right)^{\frac{1}{1-2\alpha}}$
- Then, we can find the optimal quantity expressed only as a function of the exogenous parameters:

$$Q^* = (L^*)^\alpha (K^*)^\alpha = \left(P\alpha w^{\alpha-1} r^{-\alpha}\right)^{\frac{\alpha}{1-2\alpha}} \left(P\alpha r^{\alpha-1} w^{-\alpha}\right)^{\frac{\alpha}{1-2\alpha}} =$$

$$= \left(\frac{\alpha^2 P^2}{wr}\right)^{\frac{\alpha}{1-2\alpha}}$$

Constrained Optimization

- Up to this point, we have considered only problems of unconstrained optimization, that is, where an economic entity chooses the values of some variables to optimize a dependent variable with no restriction.
- However, consider a firm which seeks to maximizes profits with the production of two goods, but faces a production quota where $Q_1 + Q_2 = 950$. In this case, the choice variables are not only simultaneous, but also dependent. The solving of this problem is called constrained optimization.
- As another example, consider that a consumer wants to maximize their utility, given by
$$U = x_1 x_2 + 2x_1$$
- However, the consumer does not have an infinite amount of money, so they cannot buy an infinite amount of goods as would maximize their utility. Instead, the individual only has \$60 to spend and x_1 costs \$4 and x_2 costs \$2, so their budget constraint is
$$4x_1 + 2x_2 = 60$$

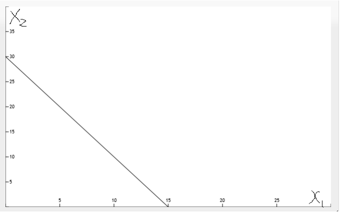
Constrained Optimization

- So, the individual's optimization problem can be stated as

$$max\ U = x_1 x_2 + 2x_1 \qquad subject\ to$$

$$4x_1 + 2x_2 = 60$$

- We call this constraint a budget constraint and it restricts the domain of the utility function, and as a result, the range of the objective function.
- In an unconstrained setting, x_1 and x_2 could take any value ≥ 0, but now the pair (x_1, x_2) must lie on the budget line.



Constrained Optimization

- The first method of solving constrained optimization is that of substitution. In the above example, we can take the budget constraint and find:

$$x_2 = \frac{60 - 4x_1}{2} = 30 - 2x_1$$

- Plug into the utility function to get:

$$U = x_1(30 - 2x_1) + 2x_1 = 32x_1 - 2x_1^2$$

$$\frac{\partial U}{\partial x_1} = 32 - 4x_1 = 0 \rightarrow x_1^* = 8$$

$$x_2 = 30 - 2(8) = 14$$

$$U^* = 8(14) + 2(8) = 128$$

- Also, we can easily see that $\frac{dU^2}{dx_1^2} = -4 < 0$, so $x_1^* = 8$ represents a constrained maximum of U.

Constrained Optimization

- Another method, which is generally much more useful, especially for more complex and more than one constraint is called the Lagrange-multiplier method.
- The Lagrangian function for the previous example is:

$$L = x_1 x_2 + 2x_1 + \lambda \left(60 - 4x_1 - 2x_2\right)$$

- λ is called the Lagrange multiplier (which we will discuss later). To solve the Lagrangian, we treat λ as a choice variable, so that the derivative with respect to λ will automatically satisfy the constraint. The first order conditions are:

$$\frac{\partial L}{\partial x_1} = x_2 + 2 - 4\lambda = 0 \rightarrow \lambda = \frac{x_2 + 2}{4}$$

$$\frac{\partial L}{\partial x_2} = x_1 - 2\lambda = 0 \rightarrow \lambda = \frac{x_1}{2}$$

$$\frac{\partial L}{\partial \lambda} = 60 - 4x_1 - 2x_2 = 0$$

Constrained Optimization

$$\lambda = \lambda \to \frac{x_2+2}{4} = \frac{x_1}{2} \equiv \textit{Marginal rate of substitution} \to x_1 = \frac{x_2+2}{2}$$

$$60 - 4\left(\frac{x_2+2}{2}\right) - 2x_2 = 0 \to 60 - 4x_2 - 4 = 0 \to 4x_2 = 56 \to x_2^* = 14$$

$$60 - 4x_1 - 28 = 0 \to 4x_1 = 32 \to x_1^* = 8$$

Lagrangian Method

- Given an objective function

$$z = f(x, y)$$

- subject to

$$g(x, y) = c$$

- we can write the Lagrangian function as

$$L = f(x, y) + \lambda [c - g(x, y)]$$

- Then, the first order conditions are

$$L_\lambda : c - g(x, y) = 0$$

$$L_x : f_x - \lambda g_x = 0$$

$$L_y : f_y - \lambda g_y = 0$$

Lagrangian Example

- A firm's production function is $y = \sqrt{x} + \sqrt{z}$ and input prices are w_x and w_z. Find the quantities of x and z that minimize cost subject to the production function.

$$L = w_x x + w_z z - \lambda(\sqrt{x} + \sqrt{z} - y)$$

$$\frac{\partial L}{\partial x} : w_x - \frac{1}{2}\lambda x^{-\frac{1}{2}} = 0 \rightarrow w_x = \frac{1}{2}\lambda x^{-\frac{1}{2}} \rightarrow \lambda = 2 w_x x^{\frac{1}{2}}$$

$$\frac{\partial L}{\partial z} : w_z - \frac{1}{2}\lambda z^{-\frac{1}{2}} = 0 \rightarrow w_z = \frac{1}{2}\lambda z^{-\frac{1}{2}} \rightarrow \lambda = 2 w_z z^{\frac{1}{2}}$$

$$2 w_x x^{\frac{1}{2}} = 2 w_z z^{\frac{1}{2}} \rightarrow x^{\frac{1}{2}} = z^{\frac{1}{2}} \frac{w_z}{w_x} \rightarrow x = z \frac{w_z^2}{w_x^2}$$

Lagrangian Example

$$y = \sqrt{x} + \sqrt{z} = z^{\frac{1}{2}} \frac{w_z}{w_x} + z^{\frac{1}{2}} = \frac{z^{\frac{1}{2}} w_z + z^{\frac{1}{2}} w_x}{w_x} = \frac{z^{\frac{1}{2}} (w_x + w_z)}{w_x}$$

$$\rightarrow z^{\frac{1}{2}} = \frac{w_x y}{w_x + w_z} \rightarrow z^* = \frac{w_x^2 y^2}{(w_x + w_z)^2}$$

Plugging back into the marginal rate of technical substitution,

$$x^* = \left(\frac{w_x^2 y^2}{(w_x + w_z)^2} \right) \frac{w_z^2}{w_x^2} = \frac{w_z^2 y^2}{(w_x + w_z)^2}$$

Lagrangian Multiplier Interpretation

- The optimal value of L depends on $\lambda^*(c)$, $x^*(c)$, $y^*(c)$, so

$$L^* = f(x^*, y^*) + \lambda^* [c - g(x^*, y^*)]$$

$$\frac{dL^*}{dc} = (f_x - \lambda^* g_x) \frac{dx^*}{dc} + (f_y - \lambda^* g_y) \frac{dy^*}{dc} + [c - g(x^*, y^*)] \frac{d\lambda^*}{dc} + \lambda^*$$

- However, the first order conditions tell us that $c = g(x^*, y^*)$, $f_x = \lambda^* g_x$, and $f_y = \lambda^* g_y$, so the first three terms on the right hand side drop out and we are left with

$$\frac{dL^*}{dc} = \lambda^*$$

- So, the value of the Lagrange multiplier at the solution of the problem is a measure of the effect of a change in the constraint via the parameter c on the optimal value of the objective function.

Lagrangian-Method with Multiple Constraints

- The Lagrange-multiplier method is equally applicable when there is more than one constraint, we just need a Lagrange-multiplier for each constraint.
- Consider the function $f(x_1, x_2, ..., x_n)$ subject to two constraints: $g(x_1, x_2, ..., x_n) = c$ and $h(x_1, x_2, ..., x_n) = d$. Then, the Lagrangian function can be written as:

$$L = f(x_1, x_2, ..., x_n) + \lambda [c - g(x_1, x_2, ..., x_n)] + \mu [d - h(x_1, x_2, ..., x_n)]$$

- Then, the first-order conditions will consist of the following $(n+2)$ simultaneous equations:

$$L_\lambda = c - g(x_1, x_2, ..., x_n) = 0$$

$$L_\mu = d - h(x_1, x_2, ..., x_n) = 0$$

$$L_i = f_i - \lambda g_i - \mu h_i = 0 \qquad (i = 1, 2, ..., n)$$

Multi-Constraint Lagrangian Example

- Find the maximum and minimum of $f(x,y,z) = 4y - 2z$ subject to $2x - y - z = 2$ and $x^2 + y^2 = 1$. The Lagrangian function is:

$$L = 4y - 2z + \lambda(2 - 2x + y + z) + \mu\left(1 - x^2 - y^2\right)$$

- The first order conditions are:

$$L_\lambda : 2 - 2x + y + z = 0 \tag{1}$$

$$L_\mu : 1 - x^2 - y^2 = 0 \tag{2}$$

$$L_x : -2\lambda - 2x\mu = 0 \tag{3}$$

$$L_y : 4 + \lambda - 2y\mu = 0 \tag{4}$$

$$L_z : -2 + \lambda = 0 \tag{5}$$

Multi-Constraint Lagrangian Example

$$(5) \to \lambda = 2$$

- Plug in $\lambda = 2$ to (3) and (4):

$$(-3) \to -2(2) - 2x\mu = 0 \to -2x\mu = 4 \to x = -\frac{2}{\mu} \tag{6}$$

$$(4) \to 4 + 2 - 2y\mu = 0 \to 6 = 2y\mu \to y = \frac{3}{\mu} \tag{7}$$

- Plug in (6) and (7) to (2):

$$1 - \left(-\frac{2}{\mu}\right)^2 - \left(\frac{3}{\mu}\right)^2 = 0 \to 1 = \frac{13}{\mu^2} \to \mu = \pm\sqrt{13}$$

Multi-Constraint Lagrangian Example

- So there are two possible solutions, where $\mu = \sqrt{13}$ and $\mu = -\sqrt{13}$.
- Case 1: $\mu = \sqrt{13}$
  - Plugging back in to (6), (7), and then (2), we get $x = -\frac{2}{\sqrt{13}}$, $y = \frac{3}{\sqrt{13}}$, and $0 = 2 + 2\left(\frac{2}{\sqrt{13}}\right) + \frac{3}{\sqrt{13}} + z \rightarrow z = -2 - \frac{7}{\sqrt{13}}$
- Case 2: $\mu = -\sqrt{13}$
  - Plugging back in to (6), (7), and then (2), we get $x = \frac{2}{\sqrt{13}}$, $y = -\frac{3}{\sqrt{13}}$, and $0 = 2 - 2\left(\frac{2}{\sqrt{13}}\right) - \frac{3}{\sqrt{13}} + z \rightarrow z = -2 + \frac{7}{\sqrt{13}}$
- These are both potential optimum. To confirm, we must use the second order conditions we will learn in the next lecture.