APEC Math Review

Part 5 Multi-variable Calculus

Ling Yao

August, 2020

Basic linear algebra

A **scalar** is a single number.
A **vector** is a $k \times 1$ list of numbers.

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_k \end{pmatrix}$$

A **matrix** is a $k \times r$ rectangular array of numbers.

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1r} \\ a_{21} & a_{22} & \dots & a_{2r} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \dots & a_{kr} \end{bmatrix}$$

Transpose

$$\mathbf{a}' \text{ or } \mathbf{a}^T = \begin{pmatrix} a_1 & a_2 & \dots & a_k \end{pmatrix}$$

$$\mathbf{A}' \text{ or } \mathbf{A}^T = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{k1} \\ a_{12} & a_{22} & \dots & a_{k2} \\ \vdots & \vdots & & \vdots \\ a_{1r} & a_{2r} & \dots & a_{kr} \end{bmatrix}$$

Partition

A matrix can be partitioned into a column of row vectors or a row of column vectors:

$$\mathbf{A} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_r \end{bmatrix},$$

where $\mathbf{a}_j = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ki} \end{bmatrix}$ and $\alpha_i = \begin{bmatrix} a_{j1} & a_{j2} & \cdots & a_{jr} \end{bmatrix}$.

A matrix can also be partitioned into smaller matrices:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \dots & \mathbf{A}_{1r} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \dots & \mathbf{A}_{2r} \\ \vdots & \vdots & & \vdots \\ \mathbf{A}_{k1} & \mathbf{A}_{k2} & \dots & \mathbf{A}_{kr} \end{bmatrix}$$

Addition

The matrices **A** and **B** are of the same order, the sum is

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1r} + b_{1r} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2r} + b_{2r} \\ \vdots & \vdots & & \vdots \\ a_{k1} + b_{k1} & a_{k2} + b_{k2} & \dots & a_{kr} + b_{kr} \end{bmatrix}.$$

The matrix addition is commutative

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$$

and associative

$$\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}.$$

Also,

$$(\mathbf{A} + \mathbf{B})' = \mathbf{A}' + \mathbf{B}'.$$

Multiplication

If $\mathbf{A}$ is $k \times r$ and c is a real number, their product is

$$\mathbf{A}c = \begin{bmatrix} a_{11}c & a_{12}c & \dots & a_{1r}c \\ a_{21}c & a_{22}c & \dots & a_{2r}c \\ \vdots & \vdots & & \vdots \\ a_{k1}c & a_{k2}c & \dots & a_{kr}c \end{bmatrix}$$

If $\mathbf{a}$ and $\mathbf{b}$ are both $k \times 1$, their **inner product** is

$$\mathbf{a}'\mathbf{b} = a_1 b_1 + a_2 b_2 + \cdots + a_k b_k = \sum_{j=1}^{k} a_j b_j$$

$\mathbf{a}'\mathbf{b} = \mathbf{b}'\mathbf{a}$

$\mathbf{a}$ and $\mathbf{b}$ are **orthogonal** if $\mathbf{a}'\mathbf{b} = 0$.

Multiplication

If $\mathbf{A}$ is $k \times r$ and $\mathbf{B}$ is $r \times s$, $\mathbf{A}$ and $\mathbf{B}$ are **conformable** and their product is defined.

$$\mathbf{AB} = \begin{bmatrix} \mathbf{a}_1' \\ \mathbf{a}_2' \\ \vdots \\ \mathbf{a}_k' \end{bmatrix} \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_s \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1'\mathbf{b}_1 & \mathbf{a}_1'\mathbf{b}_2 & \dots & \mathbf{a}_1'\mathbf{b}_s \\ \mathbf{a}_2'\mathbf{b}_1 & \mathbf{a}_2'\mathbf{b}_2 & \dots & \mathbf{a}_2'\mathbf{b}_s \\ \vdots & \vdots & & \vdots \\ \mathbf{a}_k'\mathbf{b}_1 & \mathbf{a}_k'\mathbf{b}_2 & \dots & \mathbf{a}_k'\mathbf{b}_s \end{bmatrix}$$

The multiplication is not commutative

$$\mathbf{AB} \neq \mathbf{BA},$$

but is associative

$$\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$$

and distributive

$$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}.$$

Also,

$$(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'.$$

Special matrices

Square matrices

- A matrix is **square** if $k = r$.
- A square matrix is **symmetric** if $\mathbf{A} = \mathbf{A}'$.
- A square matrix is **diagonal** is the off-diagonal elements are all zero.
- A square matrix is **upper (lower) diagonal** if all elements below (above) the diagonal equal zero.
- An **identity matrix** is a diagonal matrix with ones on the diagonal.
- A square matrix $\mathbf{B}$ for which $\mathbf{B} \cdot \mathbf{B} = \mathbf{B}$ is **idempotent.**

Positive and negative definite matrices

A $k \times k$ real symmmetric matrix $\mathbf{A}$ is **positive definite** iif for all $\mathbf{c} \neq \mathbf{0}$ in $\mathbb{R}^k$, $\mathbf{c}'\mathbf{A}\mathbf{c} > 0$.

A $k \times k$ real symmmetric matrix $\mathbf{A}$ is **negative definite** iif for all $\mathbf{c} \neq \mathbf{0}$ in $\mathbb{R}^k$, $\mathbf{c}'\mathbf{A}\mathbf{c} < 0$

$\mathbf{A}$ is positive or negative **semidefinite**, respectively, if the inequality is weak.

If a matrix is negative semidefinte, all the diagonal elements must be ≤ 0.

Exercise: matrix algebra

1. Let $\mathbf{A} = \begin{pmatrix} 2 & 3 & 1 \\ 0 & -1 & 2 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$, verify that $(\mathbf{BA})^{\mathsf{T}} = \mathbf{A}^{\mathsf{T}}\mathbf{B}^{\mathsf{T}}$.

2. Show that $\begin{pmatrix} 3 & 6 \\ -1 & -2 \end{pmatrix}$ is idempotent.

Partial derivatives

Let: $f : \mathbb{R}^n \to \mathbb{R}$. Then for each variable x_i at each point $\mathbf{x}^0 = (x_1^0, ..., x_n^0)$ in the domain of f,

$$\frac{\partial f}{\partial x_i}(x_1^0, ..., x_n^0) = \lim_{h \to 0} \frac{f(x_1^0, ..., x_i^0 + h, ..., x_n^0) - f(x_1^0, ..., x_i^0, ..., x_n^0)}{h},$$

if this limit exists.

Q: what is the difference among the notations d, ∂ and Δ?

d - derivative when there is only one variable

∂ - partial derivative

Δ - differential

Total derivatives

First order derivatives: the **Jacobian derivative** at $\mathbf{x}^*$

$$Df_{\mathbf{x}^*} = \left(\frac{\partial F}{\partial x_1}(\mathbf{x}^*) \cdots \frac{\partial F}{\partial x_n}(\mathbf{x}^*) \right).$$

The **gradient** at $\mathbf{x}^*$ is written in a column vector

$$\nabla F(\mathbf{x}^*) = \begin{pmatrix} \frac{\partial F}{\partial x_1}(\mathbf{x}^*) \\ \vdots \\ \frac{\partial F}{\partial x_n}(\mathbf{x}^*) \end{pmatrix}.$$

The gradient points in the direction at which F increases most rapidly.

Total derivatives

Second order derivatives: the **Hessian matrix**

$$D^2 f_{\mathbf{x}^*} \equiv \begin{bmatrix} \frac{\partial^2 f}{\partial {x_1}^2} & \frac{\partial^2 f}{\partial x_2 \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial {x_2}^2} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n} & \frac{\partial^2 f}{\partial x_2 \partial x_n} & \cdots & \frac{\partial^2 f}{\partial {x_n}^2} \end{bmatrix}$$

The Hessian matrix is symmetric by the Young's Theorem.

Approximation by differentials

The total differential of f at $\mathbf{x}^*$

$$df = \frac{\partial f}{\partial x_1}(\mathbf{x}^*)\Delta x_1 + \cdots + \frac{\partial f}{\partial x_n}(\mathbf{x}^*)\Delta x_n$$

approximates the actual change $\Delta f = f(\mathbf{x}^* + \Delta \mathbf{x}) - f(\mathbf{x}^*)$.

Application: Slutsky matrix

For a demand function $\mathbf{x}(\mathbf{p}, w)$, the total differential of each good x_i for $i = 1, ..., k$ is

$$dx_i = \frac{\partial x_i(\mathbf{p}, w)}{\partial p_1} dp_1 + \cdots + \frac{\partial x_i(\mathbf{p}, w)}{\partial p_k} dp_k + \frac{\partial x_i(\mathbf{p}, w)}{\partial w} dw$$

In vector notation it is (Glewwe, APEC 8001 lecture notes)

$$d\mathbf{x} = \frac{\partial \mathbf{x}(\mathbf{p}, w)}{\partial \mathbf{p}} d\mathbf{p} + \frac{\partial \mathbf{x}(\mathbf{p}, w)}{\partial w} dw$$

By the Walras' law $dw = \mathbf{x}(\mathbf{p}, w) \cdot d\mathbf{p}$

$$d\mathbf{x} = [\frac{\partial \mathbf{x}(\mathbf{p}, w)}{\partial \mathbf{p}} + \frac{\partial \mathbf{x}(\mathbf{p}, w)}{\partial w} \mathbf{x}^T(\mathbf{p}, w)] \cdot d\mathbf{p}$$

By the law of demand $d\mathbf{p} \cdot d\mathbf{x} \leq 0$

$$d\mathbf{p} \cdot [\frac{\partial \mathbf{x}(\mathbf{p}, w)}{\partial \mathbf{p}} + \frac{\partial \mathbf{x}(\mathbf{p}, w)}{\partial w} \mathbf{x}^T(\mathbf{p}, w)] \cdot d\mathbf{p} \leq 0$$

The expression in the square brackets is a $k \times k$ matrix known as the Slutsky matrix. It is negtative semidefinite.

Taylor series expansion in $\mathbb{R}^n$

Order one

$$F(\mathbf{a}+\mathbf{h}) = F(\mathbf{a}) + DF_{\mathbf{a}}\mathbf{h} + R_1(\mathbf{h};\mathbf{a})$$

Order two

$$F(\mathbf{a}+\mathbf{h}) = F(\mathbf{a}) + DF_{\mathbf{a}}\mathbf{h} + \frac{1}{2!}\mathbf{h}^T D^2 F_{\mathbf{a}}\mathbf{h} + R_2(\mathbf{h};\mathbf{a})$$

Order k

$$F(\mathbf{a}+\mathbf{h}) = F(\mathbf{a}) + DF_{\mathbf{a}}\mathbf{h} + \frac{1}{2!}\mathbf{h}^T DF_{\mathbf{a}}\mathbf{h} + \cdots + \frac{1}{k!}D^k F_{\mathbf{a}}(\mathbf{h},...,\mathbf{h}) + R_k(\mathbf{h};\mathbf{a})$$

Chain rule

If $\mathbf{x}(t) = \big(x_1(t), ..., x_n(t)\big)$ is a C^1 curve on an interval about t_0 and f is a C^1 function on a ball about $\mathbf{x}(t_0)$, then $g(t) \equiv f\big(x_1(t), ..., x_n(t)\big)$ is a C^1 function at t_0 and

$$\frac{dg}{dt}(t_0) = \frac{\partial f}{\partial x_1}\big(\mathbf{x}(t_0)\big)x_1'(t_0) \\ + \frac{\partial f}{\partial x_2}\big(\mathbf{x}(t_0)\big)x_2'(t_0) + \cdots + \frac{\partial f}{\partial x_n}\big(\mathbf{x}(t_0)\big)x_n'(t_0)$$

Exercise: Cournot and Engel aggregation

Previously we proved the Walras' law: $\mathbf{p} * \mathbf{x}(\mathbf{p}, w) = w$, show that

$$\sum_{n=1}^{N} p_n \frac{\partial x_n(\mathbf{p}, w)}{\partial p_k} + x_k(\mathbf{p}, w) = 0 \quad \textit{for } k = 1, ..., N,$$

(Cournot aggregation)

and

$$\sum_{n=1}^{N} p_n \frac{\partial x_n(\mathbf{p}, w)}{\partial w} = 1.$$

(Engel aggreagation)

Implicit function

Previously we have seen functions in this form:

$$e^y + xy - e = 0.$$

We can regard the left hand side as a two variables function, which equals to a constant:

$$G(x, y(x)) = c.$$

Differentiate with respect of x around x_0.

$$\frac{dG}{dx}(x_0, y(x_0)) * \frac{dx}{dx}(x_0) + \frac{dG}{dy}(x_0, y(x_0)) * \frac{dy}{dx}(x_0) = 0$$

$$\implies y'(x_0) = \frac{dy}{dx}(x_0) = -\frac{\frac{dG}{dx}(x_0, y_0)}{\frac{dG}{dy}(x_0, y_0)}$$

Implicit function theorem

Let $G(x_1, ..., x_k, y)$ be a C^1 function around the point $(x_1^*, ..., x_k^*, y^*)$. Suppose further that

$$G(x_1, ..., x_k, y) = c$$

and that

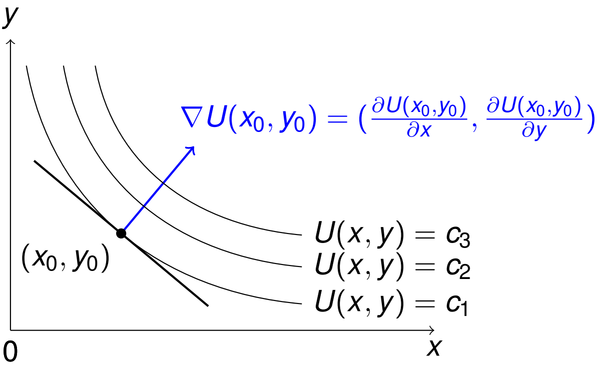
$$\frac{\partial G}{\partial y}(x_1^*, ..., x_k^*, y^*) \neq 0.$$

Then there is a C^1 function $y = y(x_1, ..., x_k)$ defined on an open ball B about $(x_1^*, ..., x_k^*)$ so that

1. $G(x_1, ..., x_k, y(x_1, ..., x_k)) = c$ for all $(x_1, ..., x_k) \in B$,
2. $y^* = y(x_1^*, ..., x_k^*)$, and
3. for each index i,

$$\frac{\partial y}{\partial x_i}(x_1^*, ..., x_k^*) = -\frac{\frac{\partial G}{\partial x_i}(x_1^*, ..., x_k^*, y^*)}{\frac{\partial G}{\partial y}(x_1^*, ..., x_k^*, y^*)}$$

Application: Marginal rate of substitution



The slope of the indifference curve at (x_0, y_0)

$$\frac{dy}{dx}(x_0) \equiv MRS(x_0, y_0) = -\frac{\frac{\partial U(x_0,y_0)}{\partial x}}{\frac{\partial U(x_0,y_0)}{\partial y}}$$

The vector $(1, MRS)$ is perpendicular to the gradient vector.

$$(1, MRS(x_0, y_0)) \cdot \nabla U(x_0, y_0) = 0$$

Exercise: Implicit function and gradient

(Simon and Blume, Exercise 15.13) A firm uses x hours of unskilled labor and y hours of skilled labor each day to produce $Q(x, y) = 60x^{2/3}y^{1/3}$ units of output per day. It currently employs 64 hours of unskilled labor and 27 hours of skilled labor.

1. In what direction (expressed in a vector) should it change (x, y) if it wants to increase output most rapidly?
2. The firms is planning to hire an additional 1.5 hour of skilled labor. Use the calculus to estimate the corresponding change in unskilled labor that should keep its output at its current level.

Homogeneous functions

For any scalar k, a real-valued function $f(x_1, ..., x_n)$ is **homogeneous of degree** k if

$$f(tx_1, ..., tx_n) = t^k f(x_1, ..., x_n) \textit{ for all } x_1, ..., x_n \textit{ and all } t > 0.$$

Exercise: homogeneous functions

(M.W.G Exercise 2.E.1)
Suppose there are 3 goods, the demand function $\mathbf{x}(\mathbf{p}, w)$ is defined by

$$x_1(\mathbf{p}, w) = \frac{p_2}{p_1 + p_2 + p_3} \frac{w}{p_1}$$
$$x_2(\mathbf{p}, w) = \frac{p_3}{p_1 + p_2 + p_3} \frac{w}{p_2}$$
$$x_3(\mathbf{p}, w) = \frac{\beta p_1}{p_1 + p_2 + p_3} \frac{w}{p_3}$$

Does this demand function satisfy homogeneity of degree 0?

Property

The partial derivatives of a function $f(x_1, ..., x_n)$ homogeneous of degree k are homogeneous of degree $k - 1$.

Proof:

By definition

$$f(tx_1, ..., tx_n) = t^k f(x_1, ..., x_n).$$

Take the first derivative *w.r.t.* x_1.

$$\frac{\partial f(tx_1, ..., tx_n)}{\partial tx_1} \cdot t = t^k \frac{\partial f(x_1, ..., x_n)}{\partial x_1}$$

$$\frac{\partial f(tx_1, ..., tx_n)}{\partial tx_1} \cdot = t^{k-1} \frac{\partial f(x_1, ..., x_n)}{\partial x_1}$$

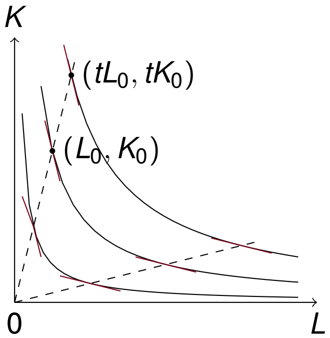
Applications

1. The demand function is *H.O.D.* 0 by nature - no money illusion

$$\mathbf{x}(tp_1, ..., tp_n, tw) = t^0\mathbf{x}(p_1, ..., p_n, w) = \mathbf{x}(p_1, ..., p_n, w)$$

2. A production function $Q = f(x_1, ..., x_n)$ which is homogeneous of degree k exhibits
   - constant returns to scale if $k = 1$
   - increasing returns to scale if $k > 1$
   - decreasing returns to scale if $k < 1$
3. Homogeneous utility function

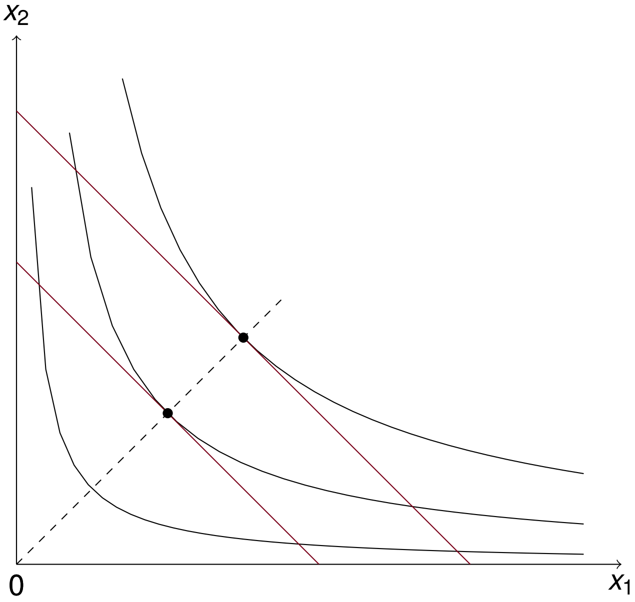
Application: marginal rate of technical substitution constant along a ray



$$MRTS(L_0, K_0) = \left| -\frac{\frac{\partial f(L_0,K_0)}{\partial L}}{\frac{\partial f(L_0,K_0)}{\partial K}} \right|$$

$$MRTS(tL_0, tK_0) = \left| -\frac{\frac{\partial f(tL_0,tK_0)}{\partial L}}{\frac{\partial f(tL_0,tK_0)}{\partial K}} \right| = \frac{t^{k-1}\frac{\partial f(L_0,K_0)}{\partial L}}{t^{k-1}\frac{\partial f(L_0,K_0)}{\partial K}} = MRTS(L_0, K_0)$$

Application: homogeneous utility function



Exercise: homogeneous function

(MWG exercise 2.E.4) Show that income elasticity of demand of a good ($\frac{\partial x_i(\mathbf{p},w)}{\partial w} \cdot \frac{w}{x_i}$) is 1 if the demand function $x_i(\mathbf{p}, w)$ homogeneous of degree 1 in w.

Euler's theorem

Let $f(\mathbf{x})$ be a C^1 homogeneous function of degree k in $\mathbb{R}^n_+$. Then for all $\mathbf{x}$,

$$x_1 \frac{\partial f}{\partial x_1}(\mathbf{x}) + x_2 \frac{\partial f}{\partial x_2}(\mathbf{x}) +, \cdots, + x_n \frac{\partial f}{\partial x_n}(\mathbf{x}) = kf(\mathbf{x})$$

or, in gradient notation,

$$\mathbf{x} \cdot \nabla f(\mathbf{x}) = kf(\mathbf{x}).$$

Proof:
By definition

$$f(tx_1, ..., tx_n) = t^k f(x_1, ..., x_n).$$

Take derivative *w.r.t.* t on both sides:

$$x_1 \frac{\partial f}{\partial x_1}(t\mathbf{x}) + x_2 \frac{\partial f}{\partial x_2}(t\mathbf{x}) +, \cdots, + x_n \frac{\partial f}{\partial x_n}(t\mathbf{x}) = kt^{k-1} f(\mathbf{x})$$

Last, set $t = 1$.

Euler's theorem

Conversely, suppose that $f(\mathbf{x})$ is a C^1 in $\mathbb{R}^n_+$. Suppose that

$$x_1 \frac{\partial f}{\partial x_1}(\mathbf{x}) + x_2 \frac{\partial f}{\partial x_2}(\mathbf{x}) +, \cdots, + x_n \frac{\partial f}{\partial x_n}(\mathbf{x}) = kf(\mathbf{x})$$

for all $x \in \mathbb{R}^n_+$. Then f is homogeneous of degree k.

Homogenize functions

If we want to impose a homogeneity of degree k on a function $F(\mathbf{x}, z)$, we can invent a new function f such that

$$z^k f(\frac{1}{z}\mathbf{x}) = F(\mathbf{x}, z).$$

For example, we want to estimate a Cobb-Douglas production function

$$F(L, K) = \gamma L^{\alpha} K^{\beta}.$$

Being homogeneous of degree 1 in the inputs means $\alpha + \beta = 1$.

$$F(L, K) = \gamma L^{1-\beta} K^{\beta} = L \cdot \gamma (K/L)^{\beta}.$$

To impose this restriction, we can define a new variable $k = \frac{K}{L}$ and estimate

$$f(k) = F(L, K)/L = \gamma k^{\beta}.$$

Exercise

Verify the Euler's theorem using the Cobb-Douglas function

$$f(x_1, x_2) = Ax_1^{\alpha} x_2^{\beta},$$

where $\alpha + \beta = 1$

Example: translog cost function (Greene 10.3.2)

In the cost minimization problem we can solve for the optimal input demand function $x_m(Q, \mathbf{p})$, where Q is the production level and $\mathbf{p}$ is the input prices vector. The cost function is given by

$$C = \sum_{m=1}^{M} p_m x_m(Q, \mathbf{p}) = C(Q, \mathbf{p})$$

If the production technology is constant returns to scale, then we can write the cost function in terms of per unit cost

$$c(\mathbf{p}) = C/Q$$

Since the cost function is homogeneous of degree 1 in input prices, by the Euler's theorem,

$$\frac{\partial c(\mathbf{p})}{\partial p_1} p_1 + \frac{\partial c(\mathbf{p})}{\partial p_2} p_2 + \cdots + \frac{\partial c(\mathbf{p})}{\partial p_m} p_m = c(\mathbf{p})$$

Example: translog cost function (Greene 10.3.2)

continued...

Divide both sides by $c(\mathbf{p})$

$$\frac{\partial c(\mathbf{p})}{\partial p_1}\frac{p_1}{c(\mathbf{p})} + \frac{\partial c(\mathbf{p})}{\partial p_2}\frac{p_2}{c(\mathbf{p})} + \cdots + \frac{\partial c(\mathbf{p})}{\partial p_m}\frac{p_m}{c(\mathbf{p})} = 1$$

$$\Leftrightarrow \sum_{m=1}^{M} \frac{\partial ln\, c(\mathbf{p})}{\partial ln\, p_m} = 1$$

Example: translog cost function (Greene 10.3.2)

continued...
By expanding $ln\, c(\mathbf{p})$ in a second-order Taylor series expansion about the point $ln(\mathbf{p}) = \mathbf{0}$, we obtain

$$ln\, c \approx \beta_0 + \sum_{m=1}^{M} \left(\frac{\partial ln\, c}{\partial ln\, p_m}\right) ln\, p_m + \frac{1}{2} \sum_{m=1}^{M} \sum_{n=1}^{M} \left(\frac{\partial^2 ln\, c}{\partial ln\, p_m\, \partial ln\, p_n}\right) ln\, p_m\, ln\, p_n$$

If we treat the derivatives as coefficients, then the cost function becomes

$$\begin{aligned} ln\, c &= \beta_0 + \beta_1 ln\, p_1 + \cdots + \beta_M ln\, p_M \\ &\quad + \delta_{11}\left(\frac{1}{2} ln^2 p_1\right) + \delta_{12} ln\, p_1\, ln\, p_2 + \delta_{22}\left(\frac{1}{2} ln^2 p_2\right) + \cdots + \delta_{MM}\left(\frac{1}{2} ln^2 p_M\right) \end{aligned}$$

Example: translog cost function (Greene 10.3.2)

continued...
Take derivative w.r.t. each of the p_m on both sides we can obtain a system of equations

$$\frac{\partial \ln c}{\partial \ln p_1} = \beta_1 + \delta_{11} \ln p_1 + \delta_{12} \ln p_2 + \cdots + \delta_{1M} \ln p_M$$

$$\frac{\partial \ln c}{\partial \ln p_2} = \beta_2 + \delta_{21} \ln p_1 + \delta_{22} \ln p_2 + \cdots + \delta_{2M} \ln p_M$$

$$\vdots$$

$$\frac{\partial \ln c}{\partial \ln p_M} = \beta_M + \delta_{M1} \ln p_1 + \delta_{M2} \ln p_2 + \cdots + \delta_{MM} \ln p_M$$

Example: translog cost function (Greene 10.3.2)

continued...

The matrix $\begin{pmatrix} \delta_{11} & \delta_{12} & \cdots & \delta_{1M} \\ \delta_{21} & \delta_{22} & \cdots & \delta_{2M} \\ \vdots & & & \\ \delta_{1M} & \delta_{M2} & \cdots & \delta_{MM} \end{pmatrix}$ is symmetric by Young's theorem.

The homogeneity of the cost function implies $\sum_{m=1}^{M} \frac{\partial ln\, c(\mathbf{p})}{\partial ln\, p_m} = 1$

Combined, they imply that

$$\delta_{mn} = \delta_{nm}$$

$$\sum_{m=1}^{M} \beta_m = 1$$

$$\sum_{m=1}^{M} \delta_{mn} = \sum_{n=1}^{M} \delta_{nm} = 0$$

Example: translog cost function (Greene 10.3.2)

continued...
We can impose the homogeneity by substituting the δ_{1M} in the first equation with $-\delta_{11} - \delta_{12} - \cdots - \delta_{1(M-1)}$ and do the same operation for each equation. Then we will end up with a system equation

$$\frac{\partial ln\, c}{\partial ln\, p_1} = \beta_1 + \delta_{11} ln \frac{p_1}{p_M} + \delta_{12} ln \frac{p_2}{p_M} + \cdots + \delta_{1(M-1)} ln \frac{p_{M-1}}{p_M}$$

$$\frac{\partial ln\, c}{\partial ln\, p_2} = \beta_1 + \delta_{21} ln \frac{p_1}{p_M} + \delta_{22} ln \frac{p_2}{p_M} + \cdots + \delta_{1(M-1)} ln \frac{p_{M-1}}{p_M}$$

$$\vdots$$

$$\frac{\partial ln\, c}{\partial ln\, p_M} = \beta_1 + \delta_{M1} ln \frac{p_1}{p_M} + \delta_{M2} ln \frac{p_2}{p_M} + \cdots + \delta_{M(M-1)} ln \frac{p_{M-1}}{p_M}$$

We can also drop one of the equations.

Example: translog cost function (Greene 10.3.2)

continued...
Turns out $\frac{\partial ln\, c}{\partial ln\, p_m}$ equals to the cost-minimizing factor cost share of input m.

$$\begin{aligned} s_m^* &= \frac{p_m}{C} x_m^* \\ &= \frac{p_m}{C} \frac{\partial C(Q, \mathbf{p})}{\partial p_m} \qquad\qquad (Shephard's\ lemma) \\ &= \frac{\partial ln\, C(Q, \mathbf{p})}{\partial ln\, p_m} \\ &= \frac{\partial ln\, c(\mathbf{p})}{\partial ln\, p_m} \end{aligned}$$

The last equality sign is based on constant returns to scale:
$ln\, C(Q, \mathbf{p}) = ln\, Q + ln\, c(\mathbf{p})$.

Example: translog cost function (Greene 10.3.2)

continued...
Therefore the empirical model is set up as

$$s_1^* = \beta_1 + \delta_{11} ln\,\frac{p_1}{p_M} + \delta_{12} ln\,\frac{p_2}{p_M} + \cdots + \delta_{1(M-1)} ln\,\frac{p_{M-1}}{p_M}$$

$$s_2^* = \beta_1 + \delta_{21} ln\,\frac{p_1}{p_M} + \delta_{22} ln\,\frac{p_2}{p_M} + \cdots + \delta_{2(M-1)} ln\,\frac{p_{M-1}}{p_M}$$

$$\vdots$$

$$s_{(M-1)}^* = \beta_1 + \delta_{(M-1)1} ln\,\frac{p_1}{p_M} + \delta_{(M-1)2} ln\,\frac{p_2}{p_M} + \cdots + \delta_{(M-1)(M-1)} ln\,\frac{p_{M-1}}{p_M}$$

Monotonic transformation

Let I be an interval on the real line. Then $g : I \to \mathbb{R}$ is a **monotonic transformation** of I if g is a strictly increasing function on I. Furthermore, if g is a monotonic transformation and u is a real-valued function of n variables, then we say that

$$g \circ u : \mathbf{x} \mapsto g(u(\mathbf{x}))$$

is a **monotonic transformation of u.**

Examples on a monotonic transformation of the utility function $U(x, y) = xy$:

- $3xy$
- $(xy)^2 + 1$
- $(xy)^3 + xy$
- e^{xy}

Cardinal v.s. ordinal

Preferences are ordinal, but homogeneity is a cardinal property.

e.g. both $U(x, y) = xy$ and $U(x, y) = (xy)^3 + xy$ can represent the same preference, but the latter is not homogeneous.

But homogeneity is great because it leads to some nice properties we discussed before.

So we need to define a class of functions with all the nice properties that homogeneous functions have.

Homothetic functions

A function $v : \mathbb{R}^n_+ \to \mathbb{R}$ is called **homothetic** if it is a monotone transformation of a homogeneous function, that is if there is a monotonic transformation $z \mapsto g(z)$ and a homogeneous function $u : \mathbb{R}^n_+ \to \mathbb{R}$ such that $v(\mathbf{x}) = g(u(\mathbf{x}))$ for all $\mathbf{x}$ in the domain.

Homothetic functions

Let $u : \mathbb{R}^n_+ \to \mathbb{R}$ be a strictly monotonic function. Then, u is homothetic if and only if for all $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n_+$,

$$u(\mathbf{x}) \geq u(\mathbf{y}) \Leftrightarrow u(\alpha\mathbf{x}) \geq u(\alpha\mathbf{y}) \textit{ for all } \alpha > 0.$$

Let u be a C^1 function on $\mathbb{R}^n_+$. If u is homothetic, then the slope of the tangent planes to the level sets of u are constant along rays from the origin; in other words, for every i, j and for every $\mathbf{x} \in \mathbb{R}^n_+$,

$$\frac{\frac{\partial u}{\partial x_i}(t\mathbf{x})}{\frac{\partial u}{\partial x_j}(t\mathbf{x})} = \frac{\frac{\partial u}{\partial x_i}(\mathbf{x})}{\frac{\partial u}{\partial x_j}(\mathbf{x})} \textit{ for all } t > 0.$$

Single-variable and multi-variable concavity

Let f be a real-valued function defined on the convex subset U of $\mathbb{R}^n$. Then f is **concave** iif for every $\mathbf{x^1}, \mathbf{x^2} \in U$ the function $g_{\mathbf{x^1},\mathbf{x^2}}(t) \equiv f(t\mathbf{x^2} + (1-t)\mathbf{x^1})$ is concave on $\{t \in \mathbb{R} \mid t\mathbf{x^2} + (1-t)\mathbf{x^1} \in U\}$.

Single-variable and multi-variable concavity

Proof:

Suppose $g_{\mathbf{x^1},\mathbf{x^2}}(t)$ is concave.
Note that $g_{\mathbf{x^1},\mathbf{x^2}}(0) = f(\mathbf{x^1})$ and $g_{\mathbf{x^1},\mathbf{x^2}}(1) = f(\mathbf{x^2})$.

$$\begin{aligned} f(t\mathbf{x^2} + (1-t)\mathbf{x^1}) &= g_{\mathbf{x^1},\mathbf{x^2}}(t) \\ &= g_{\mathbf{x^1},\mathbf{x^2}}(t \cdot 1 + (1-t) \cdot 0) \\ &\geq t g_{\mathbf{x^1},\mathbf{x^2}}(1) + (1-t) g_{\mathbf{x^1},\mathbf{x^2}}(0) \\ &= t f(\mathbf{x^2}) + (1-t) f(\mathbf{x^1}) \end{aligned}$$

So f is concave.

Single-variable and multi-variable concavity

continued...

Suppose f is concave. Choose $s_1, s_2 \in \mathbb{R}$ and $t \in [0,1]$. Note that $g_{\mathbf{x^1},\mathbf{x^2}}(s_1) = f(s_1\mathbf{x^2} + (1-s_1)\mathbf{x^1})$ and $g_{\mathbf{x^1},\mathbf{x^2}}(s_2) = g_{\mathbf{x^1},\mathbf{x^2}}(s_2)$.

$$\begin{aligned}
& g_{\mathbf{x^1},\mathbf{x^2}}(ts_1 + (1-t)s_2) \\
& = f((ts_1 + (1-t)s_2)\mathbf{x^2} + (1 - (ts_1 + (1-t)s_2))\mathbf{x^1}) \\
& = f(t(s_1\mathbf{x^2} + (1-s_1)\mathbf{x^1}) + (1-t)(s_2\mathbf{x^2} + (1-s_2)\mathbf{x^1}) \\
& \geq tf(s_1\mathbf{x^2} + (1-s_1)\mathbf{x^1}) + (1-t)f((s_2\mathbf{x^2} + (1-s_2)\mathbf{x^1})) \\
& = tg_{\mathbf{x^1},\mathbf{x^2}}(s_1) + (1-t)g_{\mathbf{x^1},\mathbf{x^2}}(s_2)
\end{aligned}$$

So g is concave. $\square$

Calculus criterion for concavity

Let f be a C^1 function on a convex subset U of $\mathbb{R}^n$. Then f is **concave** on U iif for all $\mathbf{x^1}, \mathbf{x^2} \in U$

$$\begin{aligned} f(\mathbf{x^2}) - f(\mathbf{x^1}) &\leq Df(\mathbf{x^1})(\mathbf{x^2} - \mathbf{x^1}) \\ &= \frac{\partial f}{\partial x_1}(\mathbf{x})(x_1^2 - x_1^1) + \cdots + \frac{\partial f}{\partial x_n}(\mathbf{x})(x_n^2 - x_n^1) \end{aligned}$$

Similarly, f is **convex** iif

$$f(\mathbf{x^2}) - f(\mathbf{x^1}) \geq Df(\mathbf{x^1})(\mathbf{x^2} - \mathbf{x^1}).$$

Calculus criterion for concavity

Proof:
Let $\mathbf{x^1}$ and $\mathbf{x^2}$ be arbitrary points in U. Let

$$\begin{aligned} g_{\mathbf{x^1},\mathbf{x^2}}(t) &\equiv f(t\mathbf{x^2} + (1-t)\mathbf{x^1}) \\ &= f(\mathbf{x^1} + t(\mathbf{x^2} - \mathbf{x^1})) \end{aligned}$$

By the Chain Rule,

$$g'_{\mathbf{x^1},\mathbf{x^2}}(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x^1} + t(\mathbf{x^2} - \mathbf{x^1}))(x_i^2 - x_i^1)$$

and

$$g'_{\mathbf{x^1},\mathbf{x^2}}(0) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x^1})(x_i^2 - x_i^1) = Df(\mathbf{x^1})(x^2 - x^1)$$

We have proved that f is concave iif $g_{\mathbf{x^1},\mathbf{x^2}}(t)$ is concave, which means

$$\begin{aligned} & g_{\mathbf{x^1},\mathbf{x^2}}(1) - g_{\mathbf{x^1},\mathbf{x^2}}(0) \leq g'_{\mathbf{x^1},\mathbf{x^2}}(0) \times (1-0) \\ \Leftrightarrow\ & f(\mathbf{x^2}) - f(\mathbf{x^1}) \leq Df(\mathbf{x^1})(\mathbf{x^2} - \mathbf{x^1}) \qquad \square \end{aligned}$$

Second derivatives and convexity

Let f be a C^2 function on an open convex subset D of $\mathbb{R}^n$. Then f is **concave** on D iif the Hessian matrix $D^2 f(\mathbf{x})$ is **negative semidefinite** for all $\mathbf{x} \in D$. f is **convex** on D iif the Hessian matrix $D^2 f(\mathbf{x})$ is **positive semidefinite** for all $\mathbf{x} \in D$.

Second derivatives and convexity

Show that $D^2 f(\mathbf{x})$ is negative semidefinite if f is concave.

Proof: I. Suppose f is concave, then $g(t) = f(\mathbf{x} + t\mathbf{z})$ is concave. From part 4 we know

$$g''(t) \leq 0$$

$$g(t) \leq g(t_0) + g'(t_0)(t - t_0)$$

Take derivative w.r.t. t on both sides of $g(t) = f(\mathbf{x} + t\mathbf{z})$:

$$g'(t) = \nabla f(\mathbf{x} + t\mathbf{z})\mathbf{z}$$

$$g''(t) = \mathbf{z}^T D^2 f(\mathbf{x} + t\mathbf{z})\mathbf{z} \leq 0$$

Set $t = 0$

$$\mathbf{z}^T D^2 f(\mathbf{x})\mathbf{z} \leq 0$$

Calculus criterion for quasiconcavity

Previously we defined quasiconcavity as ”for all $\mathbf{x}, \mathbf{y}$ in a convex set D in $\mathbb{R}^n$ and all $t \in [0, 1]$,

$$f(t\mathbf{x} + (1 - t)\mathbf{y}) \geq min\{f(\mathbf{x}), f(\mathbf{y})\}" .$$

Alternatively, f is quasiconcave on D iif

$$f(\mathbf{y}) \geq f(\mathbf{x}) \textit{ implies that } Df(\mathbf{x})(\mathbf{y} - \mathbf{x}) \geq 0$$

As an exercise, prove it.

- Every concave(convex) function is quasiconcave (quasiconvex).
- Any monotonic transformation of a concave(convex) function is a quasiconcave(quasiconvex) function.

Exercise

(Simon & Blume Example 21.10) Consider the CES function

$$Q(x_1, x_2) = (a_1 x_1^{\gamma} + a_2 x_2^{\gamma})^{1/\gamma}, \quad \textit{where}\; 0 < \gamma < 1$$

Is it quasiconcave?