Math Review 2022

08-1202022

Mwaso Mnensa

Basic matrix algebra

- A scalar is a single number
- A vector is a list of number (a quantity having direction0
- $\boldsymbol{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_k \end{pmatrix}$
- A matrix is a 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 represented as having been divided into blocks, such as columns or rows or smaller matrices. These are partitions.

- $\boldsymbol{A} = (\ \alpha_2 \quad \ldots\ldots \quad \alpha_r) \quad where\ \alpha_i = \begin{pmatrix} a_{i1} \\ a_{i2} \\ \vdots \\ a_{ik} \end{pmatrix}$

- $= \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \end{pmatrix} where\ \alpha_j = (\alpha_{1j} \quad \alpha_{2j} \cdots \quad \alpha_{kj})$

Addition

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

- **A+B** = **B+A**   (commutative)
- **A+(B+C)** = **(A+B) + C**   (associative)
- $(\boldsymbol{A}+\boldsymbol{B})^{\boldsymbol{T}}=\boldsymbol{A}^{\boldsymbol{T}}+\boldsymbol{B}^{\boldsymbol{T}}$

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 $\mathbf{a}'\mathbf{b} = 0$.

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,

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

Square 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

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 $\mathbf{x}^0 = (x_1^0, \ldots, x_n^0)$ in the domain of f,

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

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

Example

- *Example*. Given :

$$f\left(x_1, x_2\right) = x_1^2 + 3x_1x_2 - x_2^2$$

- Compute the partial derivative with respect to each element:

Total derivative

- The <u>total derivative</u> of a function $f : R^n \rightarrow R$ tells us how f changes as we allow x_1 through x_n to change *simultaneously.* The total differential f at x can be approximated as follows:

- $df = \frac{\partial f(\boldsymbol{x})}{\partial x_1} dx_1 + \frac{\partial f(\boldsymbol{x})}{\partial x_2} dx_2 \ldots\ldots + \frac{\partial f(\boldsymbol{x})}{\partial x_n} dx_n$

The Jacobian

- For $f: R^n \to R$

- $J = \frac{df}{dx} = \left[\frac{\partial f}{\partial x_1}, \dots \dots, \frac{\partial f}{\partial x_n}\right]$

- $for \;\; f: R^n \to R^m$

- $J = \frac{d\boldsymbol{f}}{dx} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}, \dots \dots, \frac{\partial f_1}{\partial x_n} \\ \vdots \\ \frac{\partial f_m}{\partial x_1}, \dots \dots, \frac{\partial f_m}{\partial x_n} \end{bmatrix}$

The gradient

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.

Hessian matrix

- This is a matrix of second order derivatives of a function

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

They are symmetric about the diagonal (implication of Young's theorem)

Taylor series

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

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

Chain rule

If $\mathbf{x}(t) = (x_1(t), \dots, x_n(t))$ 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 $g(t) \equiv f(x_1(t), \dots, x_n(t))$ is a C^1 function at t_0 and

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

Implicit functions

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.

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

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

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

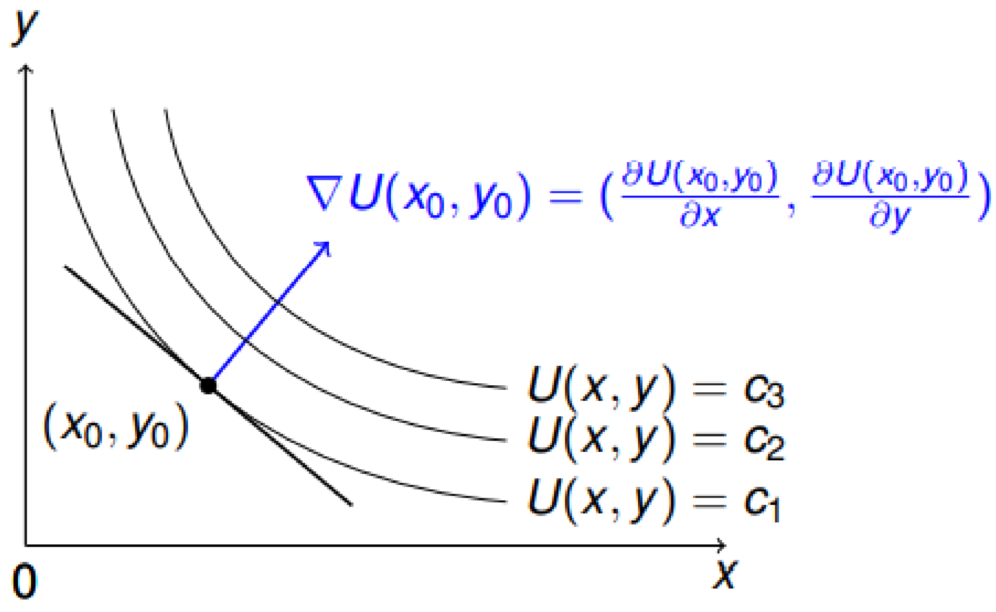
$$\frac{\partial G}{\partial y}(x_1^*, ..., x_k^*, y^*) \neq 0.$$

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,

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

MRS



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

(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.

Homogenous functions

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

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

Exercise

(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 of HODk function

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.

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

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

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

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 $t = 1$.

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.

Homogenizing a function

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 $\alpha + \beta = 1$. To impose this restriction, we can define a new variable $k = \frac{K}{L}$ and estimate

$$f(k) = \gamma k^{\beta}.$$

Then recover the production function by

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

Exercise

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

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

where $\alpha + \beta = 1$

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 vs. 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.