Essential Math

Yang Jiao (Joy)

The Graduate Institute of International and Development Studies, Geneva

Sets & Elementary Notation

Sets

A **set** is a collection of elements. For example,

$$\{1, 2, 3\}$$

$$\{A, B\} \text{ or } \{bottle, book\} \text{ or } \{3, computer, y\}$$

$$\{\{1, 1\}, \{1, 2\}, \{2, 1\}\} \text{ or } \{(1, 1), (1, 2), (2, 1)\}$$

Commonly used sets

- $\mathbb{N}$: the set of natural numbers $\{1, 2, 3, \ldots\}$
- $\mathbb{Z}$: the set of integers $\{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$
- $\mathbb{Q}$: the set of rational numbers
- $\mathbb{R}$: the set of real numbers
- $\mathbb{R}_+$: the set of non-negative real numbers
- $\mathbb{R}_{++}$: the set of positive real numbers

Elementary notations

$\in$ denotes "(is) element of". For example,

$$3 \in \mathbb{N}$$

$\notin$ denotes "(is) not element of". For example,

$$3.5 \notin \mathbb{N}$$

If we want to denote a generic element of a set, we write it as a variable. For example, we write "an (any) element in the set $\{1, 2, 3\}$" as

$$x \in \{1, 2, 3\}$$

Elementary notations

Example :

(a) We denote statement "each element of the set $\{1,2,3\}$ is smaller than 4" as

$$x < 4 \quad \forall x \in \{1,2,3\}$$

where $\forall$ means "for each".

(b) We denote "there exists an element in the set $\{1,2,3\}$ which is smaller than 3" as

$$\exists x \in \{1,2,3\} \text{ s.t. } x < 3$$

where $\exists$ means "there exists", s.t. means "such that". Similarly, we can write

$$\nexists x \in \{1,2,3\} \text{ s.t. } x > 4$$

Elementary notations

$A \subseteq B$ means "A is subset of B"

Every set is a subset of the set itself. $A \subseteq A$

$A \subset B$ means "A is subset of B and there are elements in B that are not in A", i.e. we write $A \subset B$ if $A \subseteq B$ but not $B \subseteq A$.

Example :

$$\mathbb{N} \subseteq \mathbb{Z}$$

or more precisely

$$\mathbb{N} \subset \mathbb{Z}$$

we can also write

$$x \in \mathbb{Z} \quad \forall x \in \mathbb{N}$$

Elementary notations

We can write $x \in \mathbb{Z} \quad \forall x \in \mathbb{N}$ also as

$$x \in \mathbb{N} \Rightarrow x \in \mathbb{Z}$$

$A \Rightarrow B$ reads as "if statement A holds then statement B holds", i.e. "statement A is a **sufficient** condition for statement B".

However, A is **not necessary** for B, we could have statement B even without statement A. For example, we can have $x \in \mathbb{Z}$ even for $x \notin \mathbb{N}$.

$A \Rightarrow B$ is equivalent to $\neg B \Rightarrow \neg A$, where $\neg$ denotes a negation and is read as "not".

Elementary notations

Example :

$$\text{run} \Rightarrow \text{red shirt}$$

It is equivalent to

$$\neg\text{red shirt} \Rightarrow \neg\text{run}$$

Elementary notations

If $A \Rightarrow B$ as well as $B \Rightarrow A$, then

$$A \Leftrightarrow B$$

This reads as "statement A holds iff (if and only if) statement B holds".
In other words, statement A is a **necessary and sufficient** condition for statement B (and vice versa), i.e. the two statements are equivalent.

Elementary notations

We can define a set based on another set. For example, another way to denote the set $\{1, 2, 3\}$ is to write

$$\{x \in \mathbb{N} : x < 4\} \quad \text{or} \quad \{x \in \mathbb{N} \text{ s.t. } x < 4\}$$

This is read as "all elements x in $\mathbb{N}$ s.t. x is smaller than 4".

Exercises

Exercise 1 : Define the following sets :

(a) all natural numbers divisible by three

(b) all (positive and negative) integers divisible by three

(c) all pairs $(1,1),(1,2),(1,3)$......

(d) all the intervals $[x, x+5)$, whereby x is a positive real number

Elementary notations

The empty set is denoted as $\emptyset$. For example, we have

$$\{x \in \mathbb{N} : x < 0\} = \emptyset$$

This is because $\nexists x \in \mathbb{N}$ s.t. $x < 0$.

Elementary notations

The **Cartesian product** of $\{1, 2\}$ and $\{a, b\}$, denoted by

$$\{1, 2\} \times \{a, b\}$$

is the set $\{(1, a), (1, b), (2, a), (2, b)\}$

The set with elements (x, y) whereby x and y are both natural numbers :

$$\mathbb{N}^2 = \{(x, y) : x, y \in \mathbb{N}\}$$

The set of all three-dimensional vectors with elements that are real numbers (real-valued vectors) is denoted by $\mathbb{R}^3$.

The set of all real-valued vectors of length n is $\mathbb{R}^n$

Functions

Functions

A **function** is a mapping from one set to another. For example,

$$f(x) = x^2$$

maps $x \in \mathbb{R}$ into $x^2 \in \mathbb{R}$.

To refer generically to this function, we write f or $f(x)$ (or $f(z)$, the variable we assign to the argument is irrelevant).

To refer to the function evaluated at a specific point, we write e.g. $f(3)$, or we write more generically $f(x^*)$ for some particular $x^* \in \mathbb{R}$

Functions

Example :

(a) The function

$$g(x) = \begin{cases} 0 & \text{if } \ x < 1/6 \\ 1 & \text{otherwise (i.e. } \ x \geq 1/6) \end{cases}$$

maps $\mathbb{R}$ into $\{0, 1\}$.
A short way to write the same function is

$$g(x) = \mathbf{1}\{x < 1/6\}$$

whereby $\mathbf{1}\{\cdot\}$ is the indicator function (or indicator-operator). It returns a one if the condition inside the brackets is true and a zero otherwise.

Functions

Example :

(b) The factorial

$$f(x) = x!$$

maps $\mathbb{N}$ to $\mathbb{N}$.

(c) The function

$$f(x) = 1/x$$

is not defined for $x = 0$. It maps $\mathbb{R}\backslash\{0\}$ into $\mathbb{R}$

Functions

To generically denote some function that maps a set A into a set B, we write

$$f : A \to B$$

We call the first set A the **domain** of function g and the second set B its **codomain**.

Functions

Exponential function :

$$\exp\{x\} \text{ or } e^x$$

Natural logarithm :

$$\ln x \text{ or } \log x$$

Linear function :

$$f(x) = a + bx \text{ for } a, b \in \mathbb{R}$$

Quadratic function :

$$f(x) = a + bx + cx^2 \text{ for } a, b, c \in \mathbb{R}$$

We can define the set of quadratic functions as

$$\{f : f(x) = a + bx + cx^2, a, b, c \in \mathbb{R}, x \in \mathbb{R}\}$$

Monotonic functions

f is **strictly increasing** if $f(x_1) < f(x_2)$ whenever $x_1 < x_2$.

f is **strictly decreasing** if $f(x_1) > f(x_2)$ whenever $x_1 < x_2$.

In either case, f is said to be a **monotonic function**.

g is **non-increasing** if $g(x_1) \geq g(x_2)$ whenever $x_1 < x_2$.

g is **non-decreasing** if $g(x_1) \leq g(x_2)$ whenever $x_1 < x_2$.

In either case, g is said to be a **weakly monotonic function**.

Functions

f^{-1} is the **inverse** or **inverse-image** of f. For example,

$$\text{if } f(x) = 3x, \text{ then } f^{-1}(x) = x/3$$

$$\text{if } f(x) = x^2, \text{ then } f^{-1}(x) = \pm\sqrt{x}$$

In general, $f(f^{-1}(y)) = y$ and $f^{-1}(f(x)) = x$.

Functions

A function is said to be **injective** or **"one-to-one"** if for any single, unique element in the domain it returns a different, unique element in the codomain.

If the reverse is also true, then we speak of a **bijective** function.

Note that a function f being bijective is equivalent to f and f^{-1} both being injective.

A function can also be **surjective** (i.e. only f^{-1} is injective, not f), and it can be neither of these three definitions.

Functions

Example :

$f(x) = 3x$ is an injective function.

$f(x) = 3x$ is also bijective.

In contrast, $f(x) = x^2$ is only injective, not bijective.

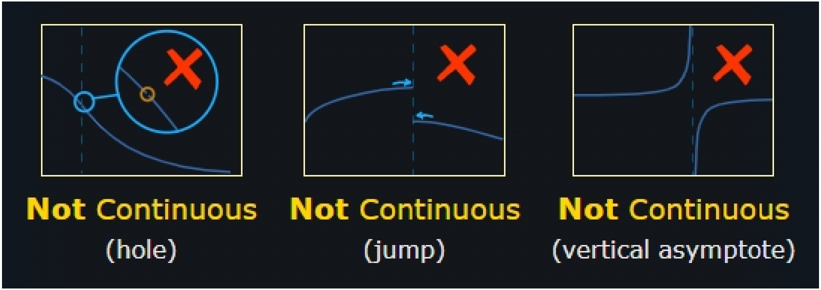
The exponential function and the natural logarithm as well as linear functions are bijective.

Functions

A function f is continuous if

$$\lim_{x \to c} f(x) = f(c) \quad \forall c$$

x and c have to be points in the domain of f.



Essentially, a continuous function is one that can be drawn on a graph without lifting the pencil off the paper.

Exercises

Exercise 2 : Tell if the following functions are continuous :

(a) $f(x) = x^2$

(b) $f(x) = |x|$

(c) $f(x) = 1\{x < 1/6\}$

(d) $f(x) = 1/x$

Differentiation

The **first-order derivative** of a function $f : \mathbb{R} \to \mathbb{R}$ is defined as

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

We write the derivative f' also as $f^{(1)}$, $\frac{\partial f(x)}{\partial x}$ or $\frac{df(x)}{dx}$ (when the function is with scalar argument).

Note that f' is itself also a function. To then evaluate it at a point x^*, we write $f'(x^*), f^{(1)}(x^*)$, $\left.\frac{\partial f(x)}{\partial x}\right|_{x=x^*}$ or $\left.\frac{df(x)}{dx}\right|_{x=x^*}$

Differentiation

Let $a, c \in \mathbb{R}$ and $b \in \mathbb{R}_+$, some useful derivatives :

- $$\frac{\partial}{\partial x}\left[c + ax^b\right] = abx^{b-1}$$
- $$\frac{\partial}{\partial x}\log x = \frac{1}{x}$$
- $$\frac{\partial}{\partial x}\exp\{x\} = \exp\{x\}$$

Differentiation

Rules of differentiation :

- $$\frac{\partial}{\partial x} af(x) = af'(x)$$
- Combination rule
$$\frac{\partial}{\partial x}[f(x) + g(x)] = f'(x) + g'(x)$$
- Product rule
$$\frac{\partial}{\partial x} f(x)g(x) = f'(x)g(x) + f(x)g'(x)$$
- Quotient rule
$$\frac{\partial}{\partial x} \frac{f(x)}{g(x)} = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}$$
- Composite function rule
$$\frac{\partial}{\partial x} f(g(x)) = f'(g(x))g'(x) = \left.\frac{\partial f(z)}{\partial z}\right|_{z=g(x)} g'(x)$$

Exercises

Exercise 3 : derive the following functions :

(a) $f(x) = 2x^4 - 7x^{-1}$

(b) $f(x) = \log x - 3x^2$

(c) $f(x) = (x^4 - 3x^2)(5x + 1)$

(d) $f(x) = \log(x)\left(3x^2\right)$

(e) $f(x) = (x^2 + 1)/(2x^3 + 1)$

(f) $f(x) = \frac{2x^3 - 5}{\sqrt{x^2 + 1}}$

(g) $f(x) = \log\left(3x^2\right)$

Differentiation

A function f is said to be **differentiable** if its derivative is defined for all points in its domain.

A function f is said to be **continuously differentiable** if it is differentiable and its derivative is a continuous function.

Every differentiable function is continuous, but continuous function is not necessarily differentiable. For example, $f(x) = |x|$ is continuous but not differentiable at $x = 0$.

Differentiation

We can also take higher-order derivatives of a function by repeatedly taking derivatives.

The **second order derivative** of function f is

$$f''(x) = \frac{\partial f'(x)}{\partial x}$$

denoted as $f^{(2)}(x)$ or $\frac{\partial^2 f(x)}{\partial^2 x}$.

The ***k*th order derivative** of f is denoted as $f^{(k)}$ or $\frac{\partial^k f(x)}{\partial^k x}$.

Differentiation

For functions with multiple arguments, we can take derivatives w.r.t. (with respect to) different arguments.

For example, for

$$f(x, y) = x^3 + e^y$$

we have

$$\frac{\partial f}{\partial x} = 3x^2$$

and

$$\frac{\partial f}{\partial y} = e^y$$

Integration

For a function defined on a discrete domain, like $\{1, 2, \ldots, n\}$, we could simply sum up the function evaluated at all the different x in the domain :

$$\sum_{x \in \{1,2,\ldots,n\}} f(x) = \sum_{x=1}^{n} f(x)$$

For example, for $f(x) = x^2$ defined on $\{1, 2, \ldots, n\}$, we get

$$1^2 + 2^2 + \ldots + n^2 = n(n+1)(2n+1)/6$$

Integration

In contrast, we cannot do so for a function defined on a continuous domain, like $\mathbb{R}$.

For such functions, we can compute the **integral** e.g. between the points $x = a$ and $x = b$:

$$\int_a^b f(x)dx$$

Integration

Proposition :

$$\int_a^b g'(x)dx = g(b) - g(a) = [g(x)]_a^b$$

To find $\int_a^b f(x)dx$, we have to find a **primitive** g of f such that $f = g'$; the desired integral is then equal to $g(b) - g(a)$.

Example :

$$\int_1^2 3x^2 dx = \left(x^3 + c\right)\Big|_{x=2} - \left(x^3 + c\right)\Big|_{x=1} = 8 - 1 = 7$$

Note that the constant c cancels out.

Integration

Sometimes we leave out the bounds of the integration because we want to integrate a function over its whole domain.

For example, rather than writing $\int_{-\infty}^{\infty} f(x)dx = \int_{\mathbb{R}} f(x)dx$ we could simply write $\int f(x)dx$.

Or because we want to compute the primitive of the integral function.

For example, we have $\int 3x^2 dx = x^3 + \mathbf{C}$

Integration

Integration by parts

$$\int_a^b f(x)g'(x)dx = [f(x)g(x)]_a^b - \int_a^b f'(x)g(x)dx$$

Vectors & Matrices

Vectors & Matrices

Scalar is a single number.

Vectors and **matrices** are ordered lists of several scalars, arranged in a rectangular way.

A vector with n components is called an ***n*-vector**.

A matrix with m rows and n columns is called an $m \times n$ **matrix**.

Note that a vector is just a $k \times 1$ matrix and a scalar is a 1×1 matrix.

Vectors & Matrices

The i th **component** of the vector a is denoted by a_i.

The **element** in the i th row and j th column of a matrix A is called the (i, j) element of A, denoted as a_{ij}.

For example,

$$v = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \quad M = \begin{bmatrix} 2 & 2 & 7 \\ 4 & 9 & 3 \end{bmatrix}$$

v is 2×1 vector and M is 2×3 matrix.

$v_1 = 3$ and $v_2 = 1$

$M_{12} = 2$ and $M_{23} = 3$

Vectors & Matrices

A vector is a **column-vector** if it has just one column (i.e. it is arranged vertically)

A vector is a **row-vector** if it has just one row (i.e. it is arranged horizontally).

We will by default always use column-vectors.

Elementary operations

Transpose flips the elements of a vector or matrix.

$$v = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \quad v' = \begin{bmatrix} 3 & 1 \end{bmatrix}$$

$$M = \begin{bmatrix} 2 & 2 & 7 \\ 4 & 9 & 3 \end{bmatrix}, \quad M' = \begin{bmatrix} 2 & 4 \\ 2 & 9 \\ 7 & 3 \end{bmatrix}$$

$$s = 1, \quad s' = s$$

Elementary operations

Addition If two matrices (vectors) have the **same dimensions**, we can add them together by adding together the corresponding elements in the two matrices (vectors).

$$\begin{bmatrix} 3 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 2 & 7 \\ 4 & 9 & 3 \end{bmatrix} + \begin{bmatrix} 1 & -3 & 2 \\ 5 & -3 & 6 \end{bmatrix} = \begin{bmatrix} 3 & -1 & 9 \\ 9 & 6 & 9 \end{bmatrix}$$

Elementary operations

Multiply by a scalar Multiplying a matrix (vector) by a scalar just involves multiplying each element by this scalar.

$$s\begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 3s \\ s \end{bmatrix}$$

$$s\begin{bmatrix} 2 & 2 & 7 \\ 4 & 9 & 3 \end{bmatrix} = \begin{bmatrix} 2s & 2s & 7s \\ 4s & 9s & 3s \end{bmatrix}$$

Elementary operations

Rules of addition : for matrices (or vectors) A, B, C and scalars λ and μ

- $A + B = B + A$
- $(A + B) + C = A + (B + C)$
- $\lambda(A + B) = \lambda A + \lambda B$
- $(\lambda + \mu)A = \lambda A + \mu A$
- $\lambda(\mu A) = \mu(\lambda A) = (\lambda\mu)A$

Elementary operations

Multiply row vector and column vector We can multiply a $1 \times k$ vector with a $k \times 1$ vector **(the inner dimensions k have to be the same)**.

The result has the dimensions 1×1, i.e. it is a scalar.

$$\begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} = ac + bd$$

Elementary operations

Multiply two matrices We can multiply a $n \times k$ matrix A with a $k \times m$ matrix B **(the inner dimensions k have to be the same)**.

The result AB is $n \times m$, and the elements in row i and column j of AB is the sum of the product of the elements of the row i of matrix A by the column j of matrix B.

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} e & f & g \\ h & i & j \end{bmatrix} = \begin{bmatrix} ae + bh & af + bi & ag + bj \\ ce + dh & cf + di & cg + dj \end{bmatrix}$$

The order by which two matrices are multiplied matters : here BA is undefined.

Exercises

Exercise 4 : compute the product of A and B

$$A = \begin{pmatrix} 3 & 1 & 4 \\ 2 & 0 & 5 \end{pmatrix} \quad B = \begin{pmatrix} 3 & 1 & 4 & 2 \\ 1 & 0 & 6 & 3 \\ 2 & 1 & 3 & 4 \end{pmatrix}$$

Elementary operations

Multiply column vector and row vector We can multiply a $n \times 1$ vector with a $1 \times m$ vector.

The result is an $n \times m$ matrix.

$$\begin{bmatrix} a \\ b \end{bmatrix} \begin{bmatrix} c & d & e \end{bmatrix} = \begin{bmatrix} ac & ad & ae \\ bc & bd & be \end{bmatrix}$$

Elementary operations

Rules of matrix multiplication : for $n \times m$ matrix A $m \times n$ matrix B and $n \times n$ matrix C

- $(AB)C = A(BC)$
- $(A' + B)C = A'C + BC, \quad C(A + B') = CA + CB'$
- $(\lambda A)B = \lambda(AB) = A(\lambda B)$

In general

$$AB \neq BA$$

If $m \neq n$, AB is $m \times m$ and BA is $n \times n$.

If $m = n$, AB may or may not be equal to BA.

Elementary operations

Combining transposition and addition or multiplication, respectively, we have the rules

- $(\alpha A)' = \alpha A'$
- $(AB)' = B'A'$
- $(ABC)' = C'B'A'$
- $(A + B)' = A' + B'$

Exercises

Exercise 5

(a) Consider the matrices

$$A = \begin{pmatrix} -1 & 0 \\ 2 & 3 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 2 \\ 3 & 0 \end{pmatrix}$$

Find and compare AB and BA.

(b) Consider the matrices

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 0 & 1 \end{pmatrix} \quad B = \begin{pmatrix} 4 & 3 \\ 2 & 1 \end{pmatrix} \quad C = \begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix}$$

Show $(AB)C = A(BC)$.

Why we need matrix?

Matrix notation is useful due to its efficiency.

Example :

(a) We can write

$$v_1 b_1 + v_2 b_2 + \ldots + v_k b_k = \sum_{j=1}^{k} v_j b_j$$

compactly as

$$v'b$$

where $v = (v_1, v_2, \ldots, v_k)'$ and $b = (b_1, b_2, \ldots, b_k)'$ are both $k \times 1$ vectors.

Why we need matrix ?

Example :

(b) We can write the three equations

$$x_i' b = 0 \text{ for } i = 1, 2, 3$$

compactly as

$$Xb = 0$$

where

$$X = \begin{bmatrix} x_1' \\ x_2' \\ x_3' \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1k} \\ x_{21} & x_{22} & \dots & x_{2k} \\ x_{31} & x_{32} & \dots & x_{3k} \end{bmatrix}$$

is a $3 \times k$ matrix that stacks the vectors x_1, x_2 and x_3 along rows and 0 is a 3×1 vector of zeros.

Why we need matrix ?

Example :

(c) Note that

$$v'v = v_1^2 + v_2^2 + v_3^2$$

computes the sum of squares of the elements in vector v.

As a result, $v'v \geq 0$ for any vector v.

Why we need matrix ?

Example :

(d)

$$x_i x_i' = \begin{bmatrix} x_{i1} \\ x_{i2} \\ \dots \\ \dots \\ x_{ik} \end{bmatrix} \begin{bmatrix} x_{i1} & x_{i2} & \dots & x_{ik} \end{bmatrix} = \begin{bmatrix} x_{i1}^2 & x_{i1}x_{i2} & \dots & x_{i1}x_{ik} \\ x_{i2}x_{i1} & x_{i2}^2 & \dots & x_{i2}x_{ik} \\ \vdots & \vdots & \ddots & \vdots \\ x_{ik}x_{i1} & x_{ik}x_{i2} & \dots & x_{ik}^2 \end{bmatrix}$$

gives a $k \times k$ matrix. We can write

$$\sum_{i=1}^{3} x_i x_i'$$

compactly as

$$X'X = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} x_1' \\ x_2' \\ x_3' \end{bmatrix}$$

Some specific matrices

A matrix full of zeros is called a **zero matrix** and will be denoted by $0_{n\times m}$.

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Properties :

- $A + 0 = 0 + A = A$
- $A - A = 0$
- $0 - A = -A$
- $A0 = 0 \quad ; \quad 0A = 0$

Some specific matrices

A matrix M is said to be **symmetric** if $M' = M$.

The matrix $X'X$ is symmetric.

Some specific matrices

A matrix M is **square** if it has the same number of rows and columns.

Square matrices can be raised to powers ; e.g. we can compute $M^2 = MM$ or $M^c = MM^{c-1}$.

Some specific matrices

A square matrix D is **diagonal** if it has non-zero elements only along the diagonal. For example, the $k \times k$ diagonal matrix

$$S = \begin{bmatrix} s_1 & 0 & \dots & 0 \\ 0 & s_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & s_k \end{bmatrix}$$

which we can also write as

$$S = \operatorname{diag}(s_1, s_2, \dots, s_k)$$

For a diagonal matrix S we have

$$S^c = \operatorname{diag}(s_1^c, s_2^c, \dots, s_k^c)$$

Some specific matrices

The **identity** matrix is a diagonal matrix which has just ones on its diagonal.

We denote a $k \times k$ identity matrix as $I_k = \text{diag}\,(1, 1, \ldots, 1)$

Properties :

- $I^c = I$
- $AI_m = A$ and $I_n A = A$

Some specific matrices

A square matrix T_1 is **lower-triangular** if it has non-zero elements only on and below its diagonal.

$$T_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 4 & 0 & 0 \\ -1 & 3 & 5 & 0 \\ 9 & -4 & 5 & 1 \end{bmatrix}$$

A square matrix T_2 is **upper-triangular** if it has non-zero elements only on and above its diagonal.

$$T_2 = \begin{bmatrix} 1 & 4 & 1 & -4 \\ 0 & 4 & 5 & 0 \\ 0 & 0 & 5 & 6 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

Transposing a lower-triangular matrix gives an upper-triangular matrix, and vice versa.

Exercises

Exercise 6 :

Let $A = \begin{pmatrix} 1 & -1 \\ 1 & 3 \end{pmatrix}$. Calculate A^2 and A^3.

Exercise 7 :

Calculate AB when

$$A = \begin{pmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{pmatrix}, \quad B = \begin{pmatrix} 3 & -1 & 6 \\ 0 & 2 & 1 \\ 0 & 0 & -5 \end{pmatrix}$$

What general result about upper triangular matrices does your answer suggest ?
What is the corresponding result for lower triangular matrices ?

Linear dependence

The vectors $v_1, v_2, \ldots, v_k$ are said to be **linearly dependent** if we can find scalars $a_1, a_2, \ldots, a_k$ - whereby at least one of them has to be non-zero - such that

$$a_1 v_1 + a_2 v_2 + \ldots + a_k v_k = 0.$$

Otherwise they are **linearly independent**.

Linear dependence implies that we can write at least one of the vectors (the one multiplied by a non-zero scalar) - w.l.o.g. the first vector, v_1 - as a linear combination of the others :

$$v_1 = \frac{-a_2}{a_1} v_2 + \ldots + \frac{-a_k}{a_1} v_k$$

Testing for linear dependence

Example : Consider the vectors

$$a = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}, b = \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix}, c = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$$

Suppose there are scalars α, β, γ such that $\alpha a + \beta b + \gamma c = 0$.

We have

$$2\alpha + 4\beta + \gamma = 0 \quad (1)$$

$$\alpha + \beta + \gamma = 0 \quad (2)$$

$$2\alpha + 3\beta + 2\gamma = 0 \quad (3)$$

and therefore $\gamma = \beta = \alpha = 0$, so the vectors a, b and c are linearly independent.

Exercises

Exercise 8 : Suppose

$$\mathbf{x} = \begin{bmatrix} 3 \\ 2q \\ 6 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} p+2 \\ -5 \\ 3r \end{bmatrix}$$

If $\mathbf{x} = 2\mathbf{y}$, find p, q and r.

Exercise 9 :

Which of the following sets of vectors are linearly dependent ?

$$(a) \quad \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad (b) \quad \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

$$(c) \quad \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}, \quad \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix}$$

Rank

The **rank** of a matrix M is the number of linearly independent columns (or rows) of M.

Rank tells whether we can reduce M to a matrix containing less vectors in its columns without losing relevant information.

For an $n \times m$ matrix, if $n \leq m$, it can have at most rank n; if $n \geq m$, it can have at most rank m.

If a matrix has the largest possible rank, we say the matrix has **full rank**. Otherwise it is **rank-deficient**.

It is not necessary to know how to manually compute a matrix's rank (same for inverse, determinant, eigenvalues and eigenvectors we will see later). Instead, we can use computer softwares to compute it numerically.

Inverse

Let M be a **square** $k \times k$ matrix. If it has full rank, then it is **invertible**.

The **inverse** M^{-1} is also $k \times k$, and we have

$$MM^{-1} = M^{-1}M = I$$

. Some useful results :

- $$(cM)^{-1} = c^{-1}M^{-1}$$

- $$S = \operatorname{diag}(s_1, s_2, \ldots, s_k) \quad S^{-1} = \operatorname{diag}\left(s_1^{-1}, s_2^{-1}, \ldots, s_k^{-1}\right)$$

- $$I^{-1} = I$$

Inverse

Combining inversion and transposition or multiplication, respectively, we have the rules

- $$\left(A^{-1}\right)' = \left(A'\right)^{-1}$$
- $$(AB)^{-1} = B^{-1}A^{-1}$$

However, $(A+B)^{-1}$ cannot (in general) be simplified any further.

Inverse

What is the point of inverting a matrix ?

Consider a system of linear equations $AX = b$ where

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

We have

$$\begin{aligned} & Ax = b \\ \Rightarrow \quad & A^{-1}Ax = A^{-1}b \\ \Rightarrow \quad & x = A^{-1}B \end{aligned}$$

If A is invertible, the unique solution to $AX = b$ is $x = A^{-1}B$.

Determinant

For a **square** $k \times k$ matrix M, we can compute its **determinant**, denoted by $|M|$ or $\det(M)$.

The determinant is a scalar computed from the elements of M, and it is non-zero iff the matrix M has full rank (i.e. is invertible).

Some useful results :

- $$S = \operatorname{diag}(s_1, s_2, \ldots, s_k) \quad |S| = \prod_{j=1}^{k} s_j$$
- $$|cM| = c^k |M|$$

Eigenvalues & Eigenvectors

Let M be a **square** $k \times k$ matrix. An **eigenvalue** λ and an **eigenvector** v of M satisfy the equation

$$Mv = \lambda v$$

For a $k \times k$ matrix M, we can find k such eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$ with corresponding eigenvectors $v_1, v_2, \ldots, v_k$.

The determinant of M is the product of its eigenvalues :

$$|M| = \prod_{j=1}^{k} \lambda_j$$

Eigenvalues & Eigenvectors

Eigenvalues are useful because they can be used to efficiently compute powers of a square matrix.

We can write

$$M = Q \Lambda Q^{-1}$$

where $\Lambda = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_k)$ is a diagonal matrix containing the eigenvalues of M, and Q is a $k \times k$ matrix stacking the eigenvectors along columns.

As a result, we have

$$M^c = Q \Lambda^c Q^{-1}$$

where $\Lambda^c = \operatorname{diag}(\lambda_1^c, \lambda_2^c, \ldots, \lambda_k^c)$

Eigenvalues & Eigenvectors

If all the eigenvalues of M are smaller than one in absolute value, then

$$\lim_{h \to \infty} M^h = 0$$

From

$$M^c = Q \Lambda^c Q^{-1}$$

we know the eigenvalues of M^c are $\lambda_1^c, \lambda_2^c, \ldots, \lambda_k^c$.

From

$$M^{-1} = Q \Lambda^{-1} Q^{-1}$$

we know the eigenvalues of M^{-1} are $\lambda_1^{-1}, \lambda_2^{-1}, \ldots, \lambda_k^{-1}$.

Revisiting functions

In its most general form, a real-valued function maps a d-dimensional domain into a c-dimensional codomain, $f : \mathbb{R}^d \to \mathbb{R}^c$

By using the vectors and matrices notations, we can write functions compactly.

Example (a) :

Function mapping $\mathbb{R}^2$ into $\mathbb{R}$

$$f(x, y) = 2x + y$$

can be written as

$$f(v) = \begin{bmatrix} 2 & 1 \end{bmatrix} v$$

where $v = (x, y)'$

Revisiting functions

Example (b) :

Function mapping $\mathbb{R}^2$ into $\mathbb{R}_+$

$$f\left(b_1, b_2\right) = \left(y_1 - x_{11}b_1 - x_{12}b_2\right)^2 + \left(y_2 - x_{21}b_1 - x_{22}b_2\right)^2 + \left(y_3 - x_{31}b_1 -\right.$$

can be written as

$$f(b) = \left(y_1 - x_1'b\right)^2 + \left(y_2 - x_2'b\right)^2 + \left(y_3 - x_3'b\right)^2 = \sum_{i=1}^{3}\left(y_i - x_i'b\right)^2$$

where $b = (b_1, b_2)'$. And even more compactly

$$f(b) = (Y - Xb)'(Y - Xb)$$

where

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}, \quad X = \begin{bmatrix} x_1' \\ x_2' \\ x_3' \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \end{bmatrix}$$

Exercises

Exercise 10 : Tell the dimension of the following functions' domain and codomain :

(a) $f(x) = \begin{bmatrix} x^2 \\ 2x \end{bmatrix}$

(b) $f(x) = \begin{bmatrix} x^2 & x^3 \\ 2x & 1 \end{bmatrix}$

(c) $f(v) = 2vv'$ where v is $n \times 1$ vector

(d) $f(M) = a'M$ where a is $n \times 1$ vector and M is $n \times m$ matrix

(e) $f(M) = a'Mb$ where a is $n \times 1$, b is $m \times 1$ and M is $n \times m$

Revisiting differention

Gradiant vector ∇f : the vector of partial derivatives of a scalar valued function $f : \mathbb{R}^N \to \mathbb{R}$

For example, let

$$f(v) = \begin{bmatrix} 2 & 1 \end{bmatrix} v, \text{ where } v = (v_1, v_2)'$$

The gradiant vector of f

$$\frac{\partial f}{\partial v} = \begin{bmatrix} \partial f / \partial v_1 \\ \partial f / \partial v_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Revisiting differention

Let

$$g(v) = \begin{bmatrix} 2v_1 + v_2 \\ v_1^2 + \log v_2 \\ -v_2^3 \end{bmatrix}$$

The derivative of each of the three scalar-outputs of g w.r.t. each of the two scalar-arguments

$$\frac{\partial g}{\partial v'} = \begin{bmatrix} \partial g_1/\partial v_1 & \partial g_1/\partial v_2 \\ \partial g_2/\partial v_1 & \partial g_2/\partial v_2 \\ \partial g_3/\partial v_1 & \partial g_3/\partial v_2 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 2v_1 & v_2^{-1} \\ 0 & -3v_2^2 \end{bmatrix}$$

$g(v)$ already returns a column-vector as its output, we have to take the derivatives w.r.t. v', i.e. go from left to right, rather than w.r.t. v, because we cannot go from top to bottom.

Revisiting differention

Hessian matrix : the matrix of second-order derivatives of a function $f : \mathbb{R}^k \to \mathbb{R}$

$$\frac{\partial^2 f}{\partial v \partial v'} = \frac{\partial}{\partial v'}\frac{\partial f}{\partial v} = \begin{bmatrix} \frac{\partial^2 f}{\partial v_1 \partial v_1} & \frac{\partial^2 f}{\partial v_1 \partial v_2} & \cdots & \frac{\partial^2 f}{\partial v_1 \partial v_k} \\ \frac{\partial^2 f}{\partial v_2 \partial v_1} & \frac{\partial^2 f}{\partial v_2 \partial v_2} & \cdots & \frac{\partial^2 f}{\partial v_2 \partial v_k} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial v_k \partial v_1} & \frac{\partial^2 f}{\partial v_k \partial v_2} & \cdots & \frac{\partial^2 f}{\partial v_k \partial v_k} \end{bmatrix}$$

Revisiting differention

Let

$$
\begin{aligned}
f(M) &= a' M b \\
&= \begin{bmatrix} a_1 & a_2 \end{bmatrix} \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \\
&= \begin{bmatrix} a_1 M_{11} + a_2 M_{21} & a_1 M_{12} + a_2 M_{22} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \\
&= b_1 a_1 M_{11} + b_1 a_2 M_{21} + b_2 a_1 M_{12} + b_2 a_2 M_{22}
\end{aligned}
$$

The first-order derivatives

$$
f' = \frac{\partial f}{\partial M} = \begin{bmatrix} \partial f / \partial M_{11} & \partial f / \partial M_{12} \\ \partial f / \partial M_{21} & \partial f / \partial M_{22} \end{bmatrix} = \begin{bmatrix} b_1 a_1 & b_2 a_1 \\ b_2 a_2 & b_2 a_3 \end{bmatrix} = ab'
$$

Revisiting differention

Some useful results (more can be found in Petersen and Pedersen (2012)) :

Let a, b, v and s be $k \times 1$ vectors, and M be $k \times k$ matrix

- $f(v) = a'v$ leads to $f' = a$
- $f(M) = a'Mb$ leads to $f' = ab'$
- $f(v) = v'Mv$ leads to $f' = (M + M')\,v$ and $f'' = M + M'$
- $f(v) = (v - s)'M(v - s)$ leads to $f' = (M + M')\,(v - s)$ and $f'' = M + M'$

If M is symmetric, then $M + M' = 2M$

Exercise 11 : prove results (1) (3) and (4)