

# Linear Algebra

沈威宇

July 1, 2025

# Contents

Chapter 1	Linear Algebra (線性代數)	1
1	Definitions	1
I	General	1
i	Vector space (向量空間)	1
ii	Affine space (仿射空間)	2
iii	Linear independence (線性獨立)	2
iv	Linear map (線性映射)	2
v	Linear transformation (線性變換)	2
vi	Invertible map (可逆映射)	2
vii	Basis (基)	3
viii	Dimension (維度)	3
ix	Bilinear form (雙線性形式)	3
x	Symmetric bilinear form (對稱雙線性形式)	3
xi	Topological vector space (TVS) (拓撲向量空間)	3
xii	Functional (泛函)	3
xiii	Seminorm (半範數)	3
xiv	Norm (範數/模長)	4
xv	Sublinear function (亞線性函數) or quasi-seminorm (準半範數)	4
xvi	Absolutely continuous function (絕對連續函數)	4
xvii	Bounded linear operator (有界線性運算子)	5
xviii	Hahn–Banach theorem (哈恩-巴拿赫定理)	5
xix	p-norm (p-範數)	5
xx	Locally convex topological vector space (LCTVS) or locally convex space	5
xxi	Dot product (點積)	6
xxii	Inner product space (內積空間)	6

xxiii Orthonormal basis (正交基)	6
xxiv Euclidean vector space (歐幾里德向量空間)	6
xxv Manifold (流形)	6
xxvi Coordinate system (座標系統)	7
xxvii Cartesian coordinate system (笛卡爾座標系統)	7
xxviii Balanced set (平衡集), circled set, or disk set	7
xxix Well-defined (定義良好)/Unambiguous (不模糊的)	7
xxx Affine map (仿射映射)	7
xxxi Affine transformation (仿射變換)	7
xxxii Tagent space (切空間)	7
xxxiii Tagent bundle (切叢)	8
xxxiv Von Neumann bounded or bounded	8
xxxv Convex set (凸集)	8
xxxvi (Algebraic) dual (vector) space (對偶空間)	8
xxxvii Orthogonal complement (正交補)	8
xxxviii Normal space (法空間)	8
xxxix Euclidean affine space (歐幾里德仿射空間) or Euclidean space (歐幾里得空間)	8
xl Affine hull (仿射包)	9
xli Convex hull (凸包)	9
xliv Curvilinear coordinate	9
xlvi Projection (投影)	9
xliv Orthographic projection (正射影)	9
xlvi Hilbert space (希爾伯特空間)	9
xlvi Principal angles (主夾角) between subspaces	9
xlvi Lattice points (格子點)	10
xlvi Flat (平面) or affine subspace (仿射子空間)	10
xlvi Direction space	10
I Parallel (平行)	10
II Two-Dimensional Vector Space	10
i Quadrant (象限)	10

III	Three-Dimensional Vector Space . . . . .	10
i	Right-hand rule (右手定則) . . . . .	10
ii	Octant (卦限). . . . .	10
iii	Cross product (叉積), external product (外積), or vector product (向量積). . . . .	11
2	Matrix (矩陣) . . . . .	11
I	Matrix . . . . .	11
i	Size . . . . .	11
ii	Notation . . . . .	12
II	Matrix equality . . . . .	13
III	Basic operations . . . . .	13
i	Matrix addition . . . . .	13
ii	Scalar multiplication. . . . .	13
iii	Subtraction . . . . .	13
iv	Matrix multiplication (矩陣乘法) . . . . .	13
v	Positive integer exponent . . . . .	13
vi	Zero matrix (零矩陣) . . . . .	13
vii	Identity matrix (單位矩陣) . . . . .	13
viii	Kronecker product (克羅內克積). . . . .	14
ix	Hadamard product (哈達瑪乘積), Schur product (舒爾乘積), elementwise product, or entrywise product (逐項乘積) . . . . .	14
x	Property. . . . .	14
xi	Zero exponent . . . . .	15
IV	Matrix row and column operations . . . . .	15
i	Row operations (zh-hant: 列/zh-hans: 行運算) . . . . .	15
ii	(General) row echelon form ( (廣義) zh-hant: 列/zh-hans: 行階梯形式) . . . . .	15
iii	Column operations (zh-hant: 行/zh-hans: 列運算). . . . .	15
iv	Transpose (轉置). . . . .	15
v	Symmetric matrix (對稱矩陣). . . . .	15
vi	Skew-symmetric matrix/antisymmetric matrix (反對稱矩陣). . . . .	16
vii	Transpose identities . . . . .	16
viii	Conjugate transpose (共軛轉置)/Hermitian conjugate (埃爾米特共軛)/Hermitian transpose (埃爾米特轉置). . . . .	16

V	Cofactor matrix, adjugate matrix, determinant, and inverse . . . . .	16
i	Determinant (行列式或判別式) . . . . .	16
ii	Determinant identities . . . . .	16
iii	Cofactor matrix (餘因子矩陣). . . . .	17
iv	Adjugate matrix (伴隨矩陣)/classical adjoint (經典伴隨矩陣) . . . . .	17
v	Inverse (反方陣、逆方陣、反矩陣或逆矩陣), multiplicative inverse (乘法反方陣), or reciprocal . . . . .	17
vi	Negative integer exponent. . . . .	17
vii	Invertible matrix (可逆矩陣), orthogonal matrix (正交矩陣), orthonormal matrix (正則矩陣), non-singular matrix (非奇異矩陣), or non-degenerate matrix (非退化矩陣) . . . . .	17
viii	Non-invertible matrix (不可逆矩陣), non-orthogonal matrix (非正交矩陣), non-orthonormal matrix (非正則矩陣), singular matrix (奇異矩陣), or degenerate matrix (退化矩陣). . . . .	17
ix	Inverse of a two-by-two matrix . . . . .	17
x	Computing inverse with row operations . . . . .	18
xi	Unitary matrix (么正矩陣或酉矩陣) . . . . .	18
xii	Determinant property. . . . .	18
xiii	Inverse identities . . . . .	18
xiv	Vandermonde matrix (范德蒙矩陣). . . . .	18
VI	Diagonal matrix (對角矩陣) . . . . .	18
i	Diagonal matrix . . . . .	18
ii	Diagonal matrix identities . . . . .	19
VII	System of Linear Equations (線性方程組) . . . . .	19
i	System of equations (方程組) . . . . .	19
ii	Linear Equation (線性方程). . . . .	19
iii	System of Linear Equations (線性方程組) . . . . .	19
iv	Coefficient matrix (係數矩陣). . . . .	20
v	Augmented matrix (增廣矩陣) . . . . .	20
vi	Gaussian Elimination (高斯消去法) or row reduction (zh-hant: 列/zh-hans: 行簡化) . . . . .	20
vii	Cramer Rule or Cramer formula (克拉瑪公式) . . . . .	20

3	Linear Subspace (線性子空間)	21
I	Linear subspace.	21
II	Four fundamental subspaces of a matrix (矩陣的四大子空間)	21
III	Column space (zh-hant: 行/zh-hans: 列空間)	21
IV	Row space (zh-hant: 列/zh-hans: 行空間)	21
V	Null space (零空間), kernel (核), or kernel space (核空間)	21
VI	Left null space (左零空間), left kernel (左核), or left kernel space (左核空間)	21
VII	Image space (像空間) or range (值域)	21
VIII	Cokernel (餘核)	21
IX	Constant rank theorem (常秩定理)	22
X	Rank (秩)	22
XI	Rank-nullity theorem (秩—零化度定理)	22
XII	Full rank (滿秩) and rank deficiency (秩虧)	24
XIII	Property of rank.	24
4	Matrix Analysis (矩陣分析)	24
I	Eigenvector or characteristic vector (特徵向量) and characteristic polynomial (特徵多項式)	24
i	Eigenvector or characteristic vector (特徵向量) and eigenvalue (本徵值、固有值、特徵值或特徵根)	24
ii	Characteristic polynomial	24
iii	Property	25
iv	Cayley–Hamilton theorem (凱萊–哈密頓定理)	25
v	Diagonalization (對角化)	25
II	Generalized Definition of Inverse	25
i	One-sided inverse (單邊逆)	25
ii	Generalized inverse or g-inverse (廣義逆)	26
iii	Reflexive generalized inverse (自反廣義逆)	26
iv	Moore–Penrose inverse (摩爾—彭若斯廣義逆) or pseudoinverse (偽逆或擬反)	26
5	Common Transformation Matrices.	27
I	Transformation matrices in two-dimensional space	27
i	Degenerate matrices	27
ii	Stretching (伸縮/縮放/拉伸).	27

iii Shearing (推移)	27
iv Rotation (旋轉)	27
v Reflection (鏡射/反射)	28
vi Orthogona projection (正射影)	28
vii Exmaple 1	28
viii Exmaple 2	28
ix Example 3.	29
x Example 4.	29
II Transformation matrices in three-dimensional space	29
i Stretching (伸縮/縮放/拉伸)	29
ii Shearing (推移)	30
iii Rotation (旋轉)	31
iv Reflection (鏡射/反射)	31
v Orthogona projection (正射影)	31

# Chapter 1 Linear Algebra (線性代數)

## 1 Definitions

### I General

#### i Vector space (向量空間)

A vector space over a field  $F$  is a non-empty set  $V$  together with a binary operation and a binary function that satisfy the eight axioms listed below. In this context, the elements of  $V$  are commonly called vectors, and the elements of  $F$  are called scalars.

The binary operation, called vector addition or simply addition assigns to any two vectors  $v$  and  $w$  in  $V$  a third vector in  $V$  which is commonly written as  $v + w$ , and called the sum of these two vectors.

The binary function, called scalar multiplication, assigns to any scalar  $a$  in  $F$  and any vector  $v$  in  $V$  another vector in  $V$ , which is denoted  $av$ .

To have a vector space, the eight following axioms must be satisfied for every  $u, v$  and  $w$  in  $V$ , and  $a$  and  $b$  in  $F$ .

1. Associativity of vector addition:  $u + (v + w) = (u + v) + w$ .
2. Commutativity of vector addition:  $u + v = v + u$ .
3. Identity element of vector addition: There exists an element  $0 \in V$ , called the zero vector, such that  $\forall v \in V : v + 0 = v$ .
4. Inverse elements of vector addition:  $\forall v \in V : \exists$  an element  $-v \in V$ , called the additive inverse of  $v$ , such that  $v + (-v) = 0$ .
5. Compatibility of scalar multiplication with field multiplication:  $a(bv) = (ab)v$ .
6. Identity element of scalar multiplication:  $1v = v$ , where  $1$  denotes the multiplicative identity in  $F$ .
7. Distributivity of scalar multiplication with respect to vector addition:  $a(u + v) = au + av$ .
8. Distributivity of scalar multiplication with respect to field addition:  $(a + b)v = av + bv$ .

When the scalar field is the real numbers, the vector space is called a real vector space, and when the scalar field is the complex numbers, the vector space is called a complex vector space. These two cases are the most common ones, but vector spaces with scalars in an arbitrary field  $F$  are also commonly considered. Such a vector space is called an  $F$ -vector space or a vector space over  $F$ .



## ii Affine space (仿射空間)

An affine space is a point set  $A$  together with a vector space  $\vec{A}$ , and a transitive and free action of the additive group of  $\vec{A}$  on the set  $A$ . The vector space  $\vec{A}$  is said to be associated to the affine space, and its elements are called vectors, translations, or sometimes free vectors.

Explicitly, the definition above means that the action is a mapping, generally denoted as an addition,

$$\begin{aligned} A \times \vec{A} &\rightarrow A \\ (a, v) &\mapsto a + v \end{aligned}$$

that has the following properties.

1. Right identity:

$$\forall a \in A : a + 0 = a, \quad \text{where } 0 \text{ is the zero vector in } \vec{A}$$

2. Associativity:

$$\forall v, w \in \vec{A}, \forall a \in A, (a + v) + w = a + (v + w)$$

3. Free and transitive action:

$$\forall a \in A : \text{the mapping } \vec{A} \rightarrow A : v \mapsto a + v \text{ is a bijection.}$$

4. Existence of one-to-one translations

$$\forall v \in \vec{A} : \text{the mapping } \vec{A} \rightarrow A : v \mapsto a + v \text{ is a bijection.}$$

## iii Linear independence (線性獨立)

For a set of vectors  $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ , if the equation:

$$\sum_{i=1}^n c_i \mathbf{v}_i = 0$$

is true only if all coefficients  $c_1, c_2, \dots, c_n$  are zero, then the set of vectors is linearly independent.

## iv Linear map (線性映射)

A map  $T : V \rightarrow W$  between vector spaces  $V$  and  $W$  over field  $F$  is linear if

$$\forall a, b \in F, v, w \in V : T(av + bw) = aT(v) + bT(w).$$

The zero vector always maps to the zero vector under a linear transformation.

## v Linear transformation (線性變換)

A linear transformation is a bijection from an vector space onto itself that is an linear map. I

## vi Invertible map (可逆映射)

A map  $T : V \rightarrow W$  is Invertible if and only if it is bijective. The unique inverse of  $T$ , called  $T^{-1}$ , is defined to be:

$$T^{-1} : W \rightarrow V; \forall v \in V : T(v) \mapsto v = T^{-1}(T(v)).$$

## vii Basis (基)

The basis of a vector space  $X$  over a field  $F$  is a set of linearly independent vectors  $\mathbf{v} = \{v_1, v_2, \dots\} \subseteq X$  such that for any vector  $\mathbf{P} \in X$ , there exists a unique coefficient vector  $(c_1, c_2, \dots) \subseteq F$  such that:

$$\sum_{i=1} c_i v_i = \mathbf{P}.$$

## viii Dimension (維度)

The dimension of a vector space is defined as the cardinality of a basis of it.

The  $n$ -dimensional vector space over field  $F$  is usually written as  $F^n$ .

## ix Bilinear form (雙線性形式)

Let  $V$  be a vector space of dimension  $n$  over a field  $\mathbb{K}$ . A map  $B : V \times V \rightarrow \mathbb{K}$  is a bilinear form on the space if:

1.  $\forall u, v, w \in V : B(u + v, w) = B(u, w) + B(v, w).$
2.  $\forall u, v \in V, \lambda \in \mathbb{K} : B(\lambda u, v) = \lambda B(u, v).$
3.  $\forall u, v, w \in V : B(u, v + w) = B(u, v) + B(u, w).$
4.  $\forall u, v \in V, \lambda \in \mathbb{K} : B(u, \lambda v) = \lambda B(u, v).$

## x Symmetric bilinear form (對稱雙線性形式)

Let  $V$  be a vector space of dimension  $n$  over a field  $\mathbb{K}$ . A map  $B : V \times V \rightarrow \mathbb{K}$  is a symmetric bilinear form on the space if:

1.  $\forall u, v \in V : B(u, v) = B(v, u).$
2.  $\forall u, v, w \in V : B(u + v, w) = B(u, w) + B(v, w).$
3.  $\forall u, v \in V, \lambda \in \mathbb{K} : B(\lambda u, v) = \lambda B(u, v).$

## xi Topological vector space (TVS) (拓撲向量空間)

A topological vector space  $X$  is a vector space over a topological field  $\mathbb{K}$ , such that vector addition  $X \times X \rightarrow X$  and scalar multiplication  $\mathbb{K} \times X \rightarrow X$  are continuous functions.

## xii Functional (泛函)

A functional is a function from a vector space into the field of real or complex numbers.

## xiii Seminorm (半範數)

Given a vector space  $X$  over a ordered field  $F$ , a seminorm on  $X$  is a real-valued function  $p : X \rightarrow \mathbb{R}$  with the following conditions, where  $|s|$  denotes the usual absolute value of a scalar  $s$ :

1. Subadditivity/Triangle inequality:

$$\forall x, y \in X : p(x + y) \leq p(x) + p(y).$$

2. Absolute homogeneity:

$$\forall s \in \mathbb{R}, x \in X : p(sx) = |s|p(x).$$

These conditions implies that:

1. Non-negativity:  $\forall x \in X : p(x) \geq 0$ .

2.  $p(0) = 0$ .

#### xiv Norm (範數/模長)

A norm on  $X$  is a seminorm  $p : X \rightarrow \mathbb{R}$  with the following properties:

- Positive definiteness/Positiveness/Point-separating:

$$\forall x \in X : p(x) = 0 \implies x = 0$$

A vector  $x$  such that  $p(x) = 1$  is called an unit vector.

#### xv Sublinear function (亞線性函數) or quasi-seminorm (準半範數)

Let  $p : X \rightarrow \mathbb{R}$  be a function on a vector space  $X$  over the field  $\mathbb{K}$ , which is either  $\mathbb{R}$  or  $\mathbb{C}$ .  $p$  is called a sublinear function or quasi-seminorm if it satisfies the following conditions:

1. Subadditivity/Triangle inequality:

$$\forall x, y \in X : p(x + y) \leq p(x) + p(y).$$

2. Nonnegative homogeneity:

$$\forall x \in X, r \geq 0 : p(rx) = rp(x).$$

#### xvi Absolutely continuous function (絕對連續函數)

Function  $f : I \subseteq \mathbb{R} \rightarrow \mathbb{R}$  is absolutely continuous on  $I$  if for every positive number  $\varepsilon$ , there is a positive number  $\delta$  such that, for every finite sequence of pairwise disjoint sub-intervals  $(x_k, y_k)$  of  $I$  with  $x_k < y_k \in I$  and cardinality  $N$ ,

$$\sum_{k=1}^N (y_k - x_k) < \delta \implies \sum_{k=1}^N |f(y_k) - f(x_k)| < \varepsilon.$$

The collection of all absolutely continuous functions on  $I$  is denoted  $AC(I)$ .

### xvii Bounded linear operator (有界線性運算子)

A bounded linear operator is a linear transformation  $L : X \rightarrow Y$  between topological vector spaces (TVSs)  $X$  and  $Y$  that maps bounded subsets of  $X$  to bounded subsets of  $Y$ . If  $X$  and  $Y$  are normed vector spaces, then  $L$  is bounded if and only if there exists some  $M > 0$  such that:

$$\forall x \in X : \|Lx\|_Y \leq M\|x\|_X.$$

The smallest such  $M$  is called the operator norm of  $L$  and denoted by  $\|L\|$ . A bounded operator between normed spaces is continuous and vice versa.

### xviii Hahn-Banach theorem (哈恩-巴拿赫定理)

Let  $p : X \rightarrow \mathbb{R}$  be a sublinear functional on a vector space  $X$  over the field  $\mathbb{K}$ , which is either  $\mathbb{R}$  or  $\mathbb{C}$ . If  $f : M \rightarrow \mathbb{K}$  is a linear functional on a vector subspace  $M$  such that

$$\forall m \in M : f(m) \leq p(m),$$

then there exists a linear functional  $F : X \rightarrow \mathbb{K}$  such that

$$\forall m \in M : F(m) = f(m),$$

$$\forall x \in X : F(x) \leq p(x).$$

### xix p-norm (p-範數)

The  $p$ -norm of a  $n$ -dimensional vector  $x = ((x_i)_{i=1}^n)$  is defined to be

$$\|x\|_p = \left( \sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}}, \quad p \in \mathbb{N}_{\infty},$$

where the  $\infty$ -norm is

$$\|x\|_{\infty} = \max_{i=1}^n |x_i|.$$

The 1-norm is also called Manhattan norm (曼哈頓範數); the 2-norm is also called Euclidean norm (歐幾里德範數).

### xx Locally convex topological vector space (LCTVS) or locally convex space

A locally convex topological vector space (LCTVS)  $(X, \mathcal{T})$  is a TVS whose topology is generated by a family of seminorms  $P$  on it such that:

$$\bigcap_{p \in P} \{x : p(x) = 0\} = \{0\}$$

That is,  $\mathcal{T}$  is generated by a basis  $\{U_{\epsilon, x}\}_{\epsilon \in \mathbb{R}_{\geq 0}, x \in X}$  of neighborhoods defined to be:

$$U_{\epsilon, 0} = \{y \in X : \forall p \in P : p(y) < \epsilon\}.$$

$$U_{\epsilon, x} = \{y \in X : y - x \in U_{\epsilon, 0}\}.$$

This definition implies that a LCTVS is necessarily a Hausdorff space (T2 space).

## xxi Dot product (點積)

The dot product of two vectors  $\mathbf{a} = (a_1, a_2, \dots, a_n)$  and  $\mathbf{b} = (b_1, b_2, \dots, b_n)$  is defined as:

$$\mathbf{a} \cdot \mathbf{b} = \sum_{i=1}^n a_i b_i.$$

## xxii Inner product space (內積空間)

A bar over an expression representing a scalar denotes the complex conjugate of this scalar.

An inner product space is a vector space  $V$  over the field  $F$  together with an inner product, that is, a map

$$\langle \cdot, \cdot \rangle : V \times V \rightarrow F$$

that satisfies the following three properties for all vectors  $x, y, z \in V$  and all scalars  $a, b \in F$ .

- Conjugate symmetry:

$$\langle x, y \rangle = \overline{\langle y, x \rangle}.$$

This implies that  $\langle x, x \rangle$  is a real number.

- Linearity in the first argument:

$$\langle ax + by, z \rangle = a\langle x, z \rangle + b\langle y, z \rangle.$$

- Positive-definiteness: if  $x$  is not zero, then  $\langle x, x \rangle > 0$ .

## xxiii Orthonormal basis (正交基)

An orthonormal basis for an inner product space  $V$  with finite dimension is a basis for  $V$  such that for  $e, f$  in it:

$$\langle e, f \rangle = \begin{cases} 1, & \text{if } e = f, \\ 0, & \text{if } e \neq f. \end{cases}$$

## xxiv Euclidean vector space (歐幾里德向量空間)

A Euclidean vector space  $E$  is a finite-dimensional vector space over the real numbers equipped with an inner product

$$\langle \cdot, \cdot \rangle : E \times E \rightarrow \mathbb{R},$$

and a norm  $\| \cdot \|$  defined as:

$$\|v\| = \sqrt{\langle v, v \rangle},$$

and a metric, called the Euclidean distance  $d(\cdot, \cdot) : E \times E \rightarrow \mathbb{R}$ , defined as:

$$d(v, w) = \|v - w\|.$$

## xxv Manifold (流形)

Let  $M$  be a Hausdorff space. If for any  $x \in M$ , there exists a neighborhood  $U_x$  of  $x$  that is homeomorphic to some open set of the  $m$ -dimensional Euclidean space, then  $M$  is called an  $m$ -dimensional manifold.

## xxvi Coordinate system (座標系統)

A coordinate system is a system that uses one or more numbers, or coordinates, to uniquely determine the points on a manifold.

## xxvii Cartesian coordinate system (笛卡爾座標系統)

The Cartesian coordinate system is a coordinate system in a Euclidean vector space  $E$  determined by an orthogonal basis of unit vectors such that the inner product of two vectors in  $E$  equals to the dot product of them, and that the norm of  $E$  is 2-norm.

## xxviii Balanced set (平衡集), circled set, or disk set

Let  $X$  be a vector space over a field  $\mathbb{K}$  with an absolute value function  $|\cdot|$ . A subset  $S$  of  $X$  is called a balanced set or balanced if:

$$\forall a \in \mathbb{K} \text{ s.t. } |a| \leq 1 : aS \subseteq S.$$

## xxix Well-defined (定義良好)/Unambiguous (不模糊的)

A well-defined expression or unambiguous expression is an expression whose definition assigns it a unique interpretation or value. Otherwise, the expression is said to be not well defined, ill defined or ambiguous.

## xxx Affine map (仿射映射)

Given two affine spaces  $A$  and  $B$  whose associated vector spaces are  $\mathcal{A}$  and  $\mathcal{B}$ , an affine map from  $A$  to  $B$  is a map  $f : A \rightarrow B$  such that

$$\begin{aligned} f : \mathcal{A} &\rightarrow \mathcal{B} \\ b - a &\mapsto f(b) - f(a) \end{aligned}$$

is a well-defined linear map.

## xxxi Affine transformation (仿射變換)

An affine transformation is a bijection from an affine space onto itself that is an affine map.

## xxxii Tangent space (切空間)

Suppose that  $M$  is a  $C^k$  differentiable  $n$ -dimensional manifold (with smoothness  $k \geq 1$ ) and that  $x \in M$ . Pick a coordinate chart  $\varphi : U \rightarrow \mathbb{R}^n$ , where  $U$  is an open subset of  $M$  containing  $x$ . Suppose further that two curves  $\gamma_1, \gamma_2 : (-1, 1) \rightarrow M$  with  $\gamma_1(0) = \gamma_2(0) = x$  are given such that both  $\varphi \circ \gamma_1, \varphi \circ \gamma_2 : (-1, 1) \rightarrow \mathbb{R}^n$  are differentiable. We call these differentiable curves initialized at  $x$ . Then  $\gamma_1$  and  $\gamma_2$  are said to be equivalent at 0 if and only if the derivatives of  $\varphi \circ \gamma_1$  and  $\varphi \circ \gamma_2$  at 0 coincide. This defines an equivalence relation on the set of all differentiable curves initialized at  $x$ , and equivalence classes of such curves are called tangent vectors (切向量) of  $M$  at  $x$ . The tangent space of  $M$  at  $x$ , denoted by  $T_x M$ , is then defined as the set of all tangent vectors at  $x$ , which does not depend on the choice of coordinate chart  $\varphi$ .

### xxxiii Tangent bundle (切叢)

A tangent bundle  $TM$  of a smooth manifold  $M$  is defined to be

$$TM = \{(x, v) : x \in M \wedge v \in T_x M\},$$

where  $T_x M$  is the tangent space at  $x$ .

### xxxiv Von Neumann bounded or bounded

Suppose  $X$  is a topological vector space (TVS) over a field  $\mathbb{K}$ . A subset  $B$  of  $X$  is called von Neumann bounded (or bounded) in  $X$  if for every neighborhood  $V$  of the origin, there exists a real  $r > 0$  such that  $B \subseteq sV$  for all scalars  $s$  satisfying  $|s| \geq r$ .

### xxxv Convex set (凸集)

Let  $S$  be a vector space or an affine space over some ordered field. A subset  $C$  of  $S$  is convex if, for all  $x$  and  $y$  in  $C$ , the line segment connecting  $x$  and  $y$  is included in  $C$ .

### xxxvi (Algebraic) dual (vector) space (對偶空間)

Given any vector space  $V$  over a field  $F$ , the dual space  $V^*$  is defined as the set of all linear maps  $\varphi : V \rightarrow F$ . The dual space  $V^*$  itself thus becomes a vector space over  $F$  when equipped with an addition and scalar multiplication satisfying:

$$\forall \varphi, \psi \in V^*, x \in V, a \in F :$$

$$(\varphi + \psi)(x) = \varphi(x) + \psi(x)$$

$$(a\varphi)(x) = a(\varphi(x))$$

### xxxvii Orthogonal complement (正交補)

An orthogonal complement of a subspace  $W$  of a vector space  $V$  equipped with a bilinear form  $B$  is the set  $W^\perp$  of all vectors in  $V$  that are orthogonal to every vector in  $W$ , of which  $v \in V$  is orthogonal to  $w \in V$  if and only if  $B(v, w) = 0$ .

### xxxviii Normal space (法空間)

Suppose that  $M$  is a  $C^k$  differentiable  $n$ -dimensional manifold (with smoothness  $k \geq 1$ ) in a vector space  $V$  equipped with a bilinear form  $B$  and that  $x \in M$ . The normal space of  $M$  at  $x$ , denoted by  $N_x M$ , is defined as the orthogonal complement of the tangent space of  $M$  at  $x$ . A vector  $v \in N_x M$  such that  $B(v, v) \neq 0$  is called a normal vector (法向量) of  $M$  at  $x$ .

### xxxix Euclidean affine space (歐幾里德仿射空間) or Euclidean space (歐幾里得空間)

A Euclidean space, also known as Euclidean affine space, is an affine space over  $\mathbb{R}$  such that the associated vector space is a Euclidean vector space.

A Euclidean space is equipped with a metric, called the Euclidean distance  $d(\cdot, \cdot) : E \times E \rightarrow \mathbb{R}$ , defined as:

$$d(v, w) = \|v - w\|.$$

## **xl Affine hull (仿射包)**

An affine combination is a linear combination of points where all coefficients sum up to 1.

The affine hull  $\text{aff}(S)$  of a set of points  $S$  in a Euclidean space is the smallest affine space that contains all the points in  $S$ , namely, the set of all affine combinations of the points in  $S$ , i.e.:

$$\text{aff}(S) = \left\{ \sum_{i=1}^k \alpha_i x_i : k > 0, x_i \in S, \alpha_i \in \mathbb{R}, \sum_{i=1}^k \alpha_i = 1 \right\}$$

## **xli Convex hull (凸包)**

A convex combination is a linear combination of points where all coefficients are non-negative and sum up to 1.

The convex hull  $\text{conv}(S)$  of a set of points  $S$  in a Euclidean space is the smallest convex set that contains all the points in  $S$ , namely, the set of all convex combinations of points in  $S$ , i.e.:

$$\text{conv}(S) = \left\{ \sum_{i=1}^k \alpha_i x_i : k > 0, x_i \in S, \alpha_i \in [0, 1], \sum_{i=1}^k \alpha_i = 1 \right\}$$

## **xliv Curvilinear coordinate**

A curvilinear coordinate is a coordinate system for Euclidean space such that there exists an invertible map from a Cartesian coordinate system to it.

## **xlvi Projection (投影)**

A projection on a subset  $U$  of a vector space  $V$  is a linear operator  $P : V \rightarrow U$  such that  $P^2 = P$ . The property  $P^2 = P$  is called idempotency (幂等性).

## **xlvii Orthographic projection (正射影)**

An orthographic on a subset  $U$  of a vector space  $V$  is a projection  $P : V \rightarrow U$  such that

$$\forall v \in V : v - P(v) \in U^*$$

where  $U^*$  is the orthogonal complement of  $U$ .

## **xlviii Hilbert space (希爾伯特空間)**

A Hilbert space is a real or complex inner product space that is also a complete metric space with respect to the distance function induced by the inner product.

## **xlvix Principal angles (主夾角) between subspaces**

Let  $V$  be an inner product space. Given two subspaces  $U, W$  with  $\dim(U) = k \leq \dim(W) = m$ , there exists then a sequence of  $k$  angles  $0 \leq \theta_1 \leq \theta_2 \leq \dots \leq \theta_k \leq \frac{\pi}{2}$  called the principal angles between  $U$  and  $W$ , the first one defined as

$$\theta_1 := \min \left\{ \arccos \left( \frac{|\langle u, w \rangle|}{\|u\| \|w\|} \right) \mid u \in U \wedge w \in W \right\} = \angle(u_1 \in U, w_1 \in W),$$



The vectors  $u_1$  and  $w_1$  are the corresponding principal vectors.

The other principal angles and vectors are then defined recursively via

$$\theta_i := \min \left\{ \arccos \left( \frac{|\langle u, w \rangle|}{\|u\| \|w\|} \right) \mid u \in U \wedge w \in W \wedge \forall j \in \mathbb{N} \wedge j < i : \langle u_i, u_j \rangle = \langle w_i, w_j \rangle = 0 \right\} = \angle(u_i \in U, w_i \in W)$$

#### **xlvi Lattice points (格子點)**

In a Euclidean vector space, vectors whose components are all integers are called lattice points.

#### **xlvii Flat (平面) or affine subspace (仿射子空間)**

A flat or an affine subspace of an affine space is a subspace of it that is an affine space.

#### **xlviii Direction space**

The direction space of an affine subspace is the tangent space of it at any point on it. (Note that the tangent spaces of it at any two points on it are the same.)

#### **I Parallel (平行)**

Two vectors  $u, v$  in a vector space over field  $F$  are called to be parallel if there exists a scalar  $x \in F$  such that  $xu = v$ .

Two affine subspaces are called to be parallel if the direction spaces of them are the same.

### **II Two-Dimensional Vector Space**

#### **i Quadrant (象限)**

I(+, +), II(-, +), III(-, -), IV(+, -).

### **III Three-Dimensional Vector Space**

#### **i Right-hand rule (右手定則)**

In a right-handed Cartesian coordinate system, if you point the thumb of your right hand in the positive  $x$ -axis direction and your index finger in the positive  $y$ -axis direction, then your middle finger (extended perpendicularly from the palm) will point in the positive  $z$ -axis direction. If not specified otherwise, right-handed Cartesian coordinate system is usually used.

#### **ii Octant (卦限)**

I(+, +, +), II(+, +, -), III(+, -, +), IV(+, -, -), V(-, +, +), VI(-, +, -), VII(-, -, +), VIII(-, -, -).

### iii Cross product (叉積), external product (外積), or vector product (向量積)

The cross product of vector  $\mathbf{a} = (a_1, a_2, a_3)$  and  $\mathbf{b} = (b_1, b_2, b_3)$  is defined as:

$$\mathbf{a} \times \mathbf{b} = \det \begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix},$$

where  $\mathbf{i} = (1, 0, 0)$ ,  $\mathbf{j} = (0, 1, 0)$ ,  $\mathbf{k} = (0, 0, 1)$ .

Properties:

- Self cross product is the zero vector:

$$\mathbf{a} \times \mathbf{a} = \mathbf{0}.$$

- Anticommutative law (反交換律):

$$\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}.$$

- Distributive over addition and subtraction:

$$\mathbf{a} \times (\mathbf{b} \pm \mathbf{c}) = \mathbf{a} \times \mathbf{b} \pm \mathbf{a} \times \mathbf{c}.$$

- Commutative and associative over scalar multiplication:

$$(r\mathbf{a}) \times \mathbf{b} = a \times (r\mathbf{b}) = r(\mathbf{a} \times \mathbf{b}).$$

- Triple product identity (三重積恆等式)/Jacobi identity (雅可比恆等式):

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) + \mathbf{B} \times (\mathbf{C} \times \mathbf{A}) + \mathbf{C} \times (\mathbf{A} \times \mathbf{B}) = \mathbf{0}.$$

$$\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = \mathbf{B} \cdot (\mathbf{C} \times \mathbf{A}) = \mathbf{C} \cdot (\mathbf{A} \times \mathbf{B}) = \det \begin{pmatrix} \mathbf{A} \\ \mathbf{B} \\ \mathbf{C} \end{pmatrix}$$

$$(\mathbf{A} \times \mathbf{B}) \times \mathbf{C} = (\mathbf{A} \cdot \mathbf{C})\mathbf{B} - (\mathbf{B} \cdot \mathbf{C})\mathbf{A}.$$

## 2 Matrix (矩陣)

### I Matrix

A matrix is a rectangular array of numbers (or other mathematical objects), called the entries (元或元素) of the matrix. Most commonly, a matrix over a field  $F$  is a rectangular array of elements of  $F$ .

#### i Size

The size of a matrix is defined by the number of rows (zh-hant: 列/zh-hans: 行) and columns (zh-hant: 行/zh-hans: 列) it contains. A matrix with  $m$  rows and  $n$  columns is called an  $m \times n$  matrix, or  $m$ -by- $n$  matrix, where  $m$  and  $n$  are called its dimensions (階). A matrix with the same number of rows and columns  $n$  is called a  $n$  square matrix (方陣). A matrix with an infinite number of rows or/and columns is called an infinite matrix.

The set of all  $m \times n$  matrices over  $\mathbb{F}$  is commonly denoted as  $\mathbb{F}^{m \times n}$ .

## ii Notation

The entry in the  $i$ -th row and  $j$ -th column, called the  $i, j$ th entry (第  $i, j$  元) or  $ij$ th entry (第  $ij$  元), is commonly denoted as  $a_{i,j}$  or  $a_{ij}$ .

An  $m \times n$  matrix  $\mathbf{A}$  is commonly denoted as:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix},$$

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},$$

$$\mathbf{A} = [a_{ij}]_{1 \leq i \leq m, 1 \leq j \leq n},$$

$$\mathbf{A} = (a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n},$$

$$\mathbf{A} = [a_{ij}]_{m \times n},$$

or

$$\mathbf{A} = (a_{ij})_{m \times n}.$$

The  $i, i$ th entries are called diagonal entries (對角線元).

A matrix with only one row is called a row matrix or row vector, and a matrix with only one column is called a column matrix or column vector.

The  $i$ th row of the matrix  $\mathbf{A}$  is commonly denoted as  $\mathbf{A}_{i,:}$ , and the  $j$ th column is commonly denoted as  $\mathbf{A}_{:,j}$ .

The matrix formed by merging two matrices  $\mathbf{A}$  and  $\mathbf{B}$  with the same number of rows left and right is denoted as:

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \end{bmatrix}$$

or

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \end{pmatrix}$$

The matrix formed by merging two matrices  $\mathbf{A}$  and  $\mathbf{B}$  with the same number of columns up and down is denoted as:

$$\begin{bmatrix} \mathbf{A} \\ \mathbf{B} \end{bmatrix}$$

or

$$\begin{pmatrix} \mathbf{A} \\ \mathbf{B} \end{pmatrix}$$

## II Matrix equality

We say that two matrices  $\mathbf{A} = [a_{ij}]_{m \times n}$  and  $\mathbf{B} = [b_{ij}]_{m \times n}$  are equal if

$$\forall 1 \leq i \leq m, 1 \leq j \leq n : a_{ij} = b_{ij}.$$

## III Basic operations

### i Matrix addition

The sum of two  $m \times n$  matrices  $\mathbf{A}$  and  $\mathbf{B}$  is calculated entrywise.

### ii Scalar multiplication

The product  $c\mathbf{A}$  of a scalar  $c$  is calculated by multiplying every entry of  $\mathbf{A}$  by  $c$ .

### iii Substraction

The subtraction of two  $m \times n$  matrices  $\mathbf{A}$  and  $\mathbf{B}$  is defined as  $\mathbf{A} - \mathbf{B} = \mathbf{A} + (-1)\mathbf{B}$ .

### iv Matrix multiplication (矩陣乘法)

Multiplication of two matrices is defined only if the number of columns of the left matrix is the same as the number of rows of the right matrix. The matrix product  $\mathbf{AB}$  of an  $m \times n$  matrix  $\mathbf{A}$  and  $n \times p$  matrix  $\mathbf{B}$  is the  $m \times p$  matrix whose  $ij$ th entry is the dot product of the  $i$ th row of  $\mathbf{A}$  and the  $j$ th column of  $\mathbf{B}$ .

### v Positive integer exponent

Given a square matrix  $\mathbf{A}$  and a positive integer  $n$ ,  $\mathbf{A}^n$  is defined as  $\prod_{i=1}^n \mathbf{A}$ .

### vi Zero matrix (零矩陣)

A zero matrix  $\mathbf{O}$  over a field  $F$  is a matrix of which all entries are the additive identity of  $F$ , that is, let  $m \times n$  zero matrix over  $F$  be  $\mathbf{O}_{m \times n}$ , for any matrices  $\mathbf{A}_{m \times n}$  and  $\mathbf{B}_{n \times p}$  over  $F$ , it satisfies:

$$\begin{aligned}\mathbf{A}_{m \times n} + \mathbf{O}_{m \times n} &= \mathbf{A}_{m \times n} \\ \mathbf{A}_{m \times n} - \mathbf{A}_{m \times n} &= \mathbf{O}_{m \times n} \\ \mathbf{O}_{m \times n} \mathbf{B}_{n \times p} &= \mathbf{O}_{m \times p} \\ \mathbf{A}_{m \times n} \mathbf{O}_{n \times p} &= \mathbf{O}_{m \times p}\end{aligned}$$

### vii Identity matrix (單位矩陣)

An identity matrix  $\mathbf{I}$  over a field  $F$  is a square matrix of which all diagonal entries are the multiplicative identity of  $F$  and all off-diagonal entries are the multiplicative absorbing element  $0_F$  of  $F$  (i.e.  $\forall x \in F : x \times 0_F = 0_F$ ), that is, let  $n \times n$  identity matrix over  $F$  be  $\mathbf{I}_n$ , for any  $m \times n$  matrices  $\mathbf{A}$  over  $F$ , it satisfies:

$$\mathbf{A} \mathbf{I}_n = \mathbf{I}_m \mathbf{A} = \mathbf{A}$$

### viii Kronecker product (克羅內克積)

If  $\mathbf{A}$  is an  $m \times n$  matrix and  $\mathbf{B}$  is a  $p \times q$  matrix, then the Kronecker product  $\mathbf{A} \otimes \mathbf{B}$  is the  $pm \times qn$  block matrix:

$$(\mathbf{A} \otimes \mathbf{B})_{i,j} = a_{\lceil i/p \rceil, \lceil j/q \rceil} b_{(i-1) \bmod p+1, (j-1) \bmod q+1}$$

### ix Hadamard product (哈達瑪乘積), Schur product (舒爾乘積), elementwise product, or entrywise product (逐項乘積)

For two  $m \times n$  matrices  $\mathbf{A}$  and  $\mathbf{B}$ , the Hadamard product  $\mathbf{A} \odot \mathbf{B}$  or  $\mathbf{A} \circ \mathbf{B}$  is an  $m \times n$  matrix defined as:

$$(\mathbf{A} \odot \mathbf{B})_{ij} = (\mathbf{A})_{ij}(\mathbf{B})_{ij}.$$

### x Property

For scalars  $r$  and  $s$ , matrices  $A, B, C, D$ , zero matrix  $O$ , and identity matrix  $I$ , assuming that all of the following operations hold,  $\neq$  means not necessarily equals:

$$A + B = B + A$$

$$A + (B + C) = (A + B) + C$$

$$r(A + B) = rA + rB$$

$$rA + sA = (r + s)A$$

$$(rs)A = r(sA) = s(rA)$$

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

$$r(AB) = (rA)B = A(rB)$$

$$A(B + C) = AB + AC$$

$$(A + B)C = AC + BC$$

$$(A + B)(C + D) = AC + AD + BC + BD$$

$$AB \neq BA$$

$$(AB = AC \wedge A \neq O) \nRightarrow B = C$$

$$A^2 = O \nRightarrow A = O$$

$$AB = O \nRightarrow (A = O \vee B = O)$$

$$A^2 = I \iff A = A^{-1} \nRightarrow A = \pm I$$

$$AB = B \nRightarrow A = I$$

$$ACB = ADB \nRightarrow C = D$$

$$A^2 = I \nRightarrow A = rI$$

$$AB = I \nRightarrow (A = rI \vee B = sI)$$

$$A \odot B = B \odot A$$

$$A \odot (B \odot C) = (A \odot B) \odot C$$

$$A \odot (B + C) = A \odot B + A \odot C$$

## xi Zero exponent

Given an  $n \times n$  matrix  $\mathbf{A}$  and  $n \times n$  identity matrix  $\mathbf{I}$ ,  $\mathbf{A}^0$  is defined as  $\mathbf{I}$ .

## IV Matrix row and column operations

### i Row operations (zh-hant: 列/zh-hans: 行運算)

Including:

- Row addition (zh-hant: 列/zh-hans: 行相加): Adding a row to another.
- Row multiplication (zh-hant: 列/zh-hans: 行加倍): Multiplying all entries of a row by a non-zero constant.
- Row switching (zh-hant: 列/zh-hans: 行交換): Interchanging two rows of a matrix.

If one matrix can be transformed to another through a sequence of row operations, the two matrices are called row-equivalent.

### ii (General) row echelon form ( (廣義) zh-hant: 列/zh-hans: 行階梯形式)

A matrix is in row echelon form if:

- All rows having only zero entries are at the bottom.
- The leading entry, that is, the left-most non-zero entry, of every non-zero row is on the right of the leading entry of every row above.
- Some texts add the condition that the leading coefficient must be 1.

### iii Column operations (zh-hant: 行/zh-hans: 列運算)

Including:

- Column addition (zh-hant: 行/zh-hans: 列相加): Adding a column to another.
- Column multiplication (zh-hant: 行/zh-hans: 列加倍): Multiplying all entries of a column by a non-zero constant.
- Column switching (zh-hant: 行/zh-hans: 列交換): Interchanging two columns of a matrix.

If one matrix can be transformed to another through a sequence of column operations, the two matrices are called column-equivalent.

### iv Transpose (轉置)

The transpose of an  $m \times n$  matrix  $\mathbf{A}$  is the  $n \times m$  matrix  $\mathbf{A}^T$ :

$$(\mathbf{A}^T)_{i,j} = a_{j,i}$$

### v Symmetric matrix (對稱矩陣)

$\mathbf{A}$  is called a symmetric matrix if  $\mathbf{A} = \mathbf{A}^T$ .

## vi Skew-symmetric matrix/antisymmetric matrix (反對稱矩陣)

$\mathbf{A}$  is called a skew-symmetric matrix or an antisymmetric matrix if  $\mathbf{A} = -\mathbf{A}^\top$ .

## vii Transpose identities

$$\left( \sum_{i=1}^n \mathbf{A}_i \right)^\top = \sum_{i=1}^n \mathbf{A}_i^\top$$
$$\left( \prod_{i=1}^n \mathbf{A}_i \right)^\top = \prod_{i=1}^n \mathbf{A}_{n-i+1}^\top$$

## viii Conjugate transpose (共軛轉置)/Hermitian conjugate (埃爾米特共軛)/Hermitian transpose (埃爾米特轉置)

The conjugate transpose of an  $m \times n$  matrix  $\mathbf{A}$  is the  $n \times m$  matrix  $\mathbf{A}^*$ :

$$(\mathbf{A}^*)_{i,j} = \overline{a_{j,i}}$$

## V Cofactor matrix, adjugate matrix, determinant, and inverse

Let  $\mathbf{M}_{ij}$  of  $\mathbf{A}$  be the determinant of the submatrix of  $\mathbf{A}$  without the  $i$ th row and  $j$ th column.

### i Determinant (行列式或判別式)

The determinant of a  $1 \times 1$  matrix is the only element of it; the determinant of an  $n \times n$  matrix  $\mathbf{A}$ , denoted  $\det(\mathbf{A})$  or

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix},$$

is

$$\det(\mathbf{A}) = \sum_{i=1}^n a_{ij}(-1)^{i+j} \mathbf{M}_{ij} \quad j \leq n \in \mathbb{N}$$

$$\det(\mathbf{A}) = \sum_{j=1}^n a_{ij}(-1)^{i+j} \mathbf{M}_{ij} \quad i \leq n \in \mathbb{N}$$

### ii Determinant identities

Let  $\mathbf{A}$  and  $\mathbf{A}_i$  be square matrices.

$$\det\left(\sum_{i=1}^n \mathbf{A}_i\right) = \sum_{i=1}^n \det(\mathbf{A}_i)$$

$$\det\left(\prod_{i=1}^n \mathbf{A}_i\right) = \prod_{i=1}^n \det(\mathbf{A}_i)$$

$$\det(\mathbf{A}) \neq 0 \implies \det(\mathbf{A}^n) \neq 0$$

### iii Cofactor matrix (餘因子矩陣)

The cofactor matrix of an  $n \times n$  matrix  $\mathbf{A}$  is the  $n \times n$  matrix  $\text{cof}(\mathbf{A})$ :

$$[\text{cof}(\mathbf{A})]_{ij} = ((-1)^{i+j} \mathbf{M}_{ij})$$

### iv Adjugate matrix (伴隨矩陣)/classical adjoint (經典伴隨矩陣)

The adjugate matrix or classical adjoint matrix of an  $n \times n$  matrix  $\mathbf{A}$  is the  $n \times n$  matrix  $\text{adj}(\mathbf{A})$ :

$$\text{adj}(\mathbf{A}) = \text{cof}(\mathbf{A})^T$$

### v Inverse (反方陣、逆方陣、反矩陣或逆矩陣), multiplicative inverse (乘法反方陣), or reciprocal

The inverse of an  $n \times n$  matrix  $\mathbf{A}$  is an  $n \times n$  matrix  $\mathbf{A}^{-1}$  such that:

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$$

### vi Negative integer exponent

Given an invertible matrix  $\mathbf{A}$  and a negative integer  $n$ ,  $\mathbf{A}^n$  is defined as  $(\mathbf{A}^{-1})^n$ .

### vii Invertible matrix (可逆矩陣), orthogonal matrix (正交矩陣), orthonormal matrix (正則矩陣), non-singular matrix (非奇異矩陣), or non-degenerate matrix (非退化矩陣)

An invertible matrix, orthogonal matrix, orthonormal matrix, non-singular matrix, or non-degenerate matrix is a matrix such that its inverse exists.

Let  $\mathbf{A}$  be an  $n \times n$  matrix, the following prepositions are equivalent to " $\mathbf{A}$  is an invertible matrix":

$$\exists \mathbf{A}^{-1}$$

$$\exists (\mathbf{A}^T)^{-1}$$

$$\exists (\mathbf{A}^T \mathbf{A})^{-1}$$

$$\text{rank}(\mathbf{A}) = n$$

$\mathbf{A}$ 存在非零的特徵值

### viii Non-invertible matrix (不可逆矩陣), non-orthogonal matrix (非正交矩陣), non-orthonormal matrix (非正則矩陣), singular matrix (奇異矩陣), or degenerate matrix (退化矩陣)

A non-invertible matrix, non-orthogonal matrix, non-orthonormal matrix, singular matrix, or degenerate matrix is a matrix that is not an invertible matrix.

### ix Inverse of a two-by-two matrix

For an invertible  $2 \times 2$  matrix  $\mathbf{A}$ :

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \implies \mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$



## x Computing inverse with row operations

Let  $\mathbf{A}$  be an  $n \times n$  matrix and  $\mathbf{I}$  be the  $n \times n$  identity matrix. If  $(\mathbf{A} \ \mathbf{I})$  can become  $(\mathbf{I} \ \mathbf{B})$  after row operations, then

$$\mathbf{B} = \mathbf{A}^{-1}$$

## xi Unitary matrix (么正矩陣或酉矩陣)

An unitary matrix is a square matrix whose inverse equals to its conjugate transpose.

## xii Determinant property

- When a row (or column) is added to another, the determinant remains unchanged.
- When all entries of a row (or column) is multiplied by a real number  $n$ , the determinant is multiplied by  $n$ .
- When two rows (or columns) are swapped, the determinant is multiplied by  $-1$ .
- Given matrix  $\mathbf{A}$  and its inverse  $\mathbf{A}^{-1}$ ,  $\det(\mathbf{A}^{-1}) = \frac{1}{\det(\mathbf{A})}$ .

## xiii Inverse identities

Let  $\mathbf{A}$  and  $\mathbf{A}_i$  be invertible matrices,  $\mathbf{B}$  be a matrix of the same size of  $\mathbf{A}$ .

$$\mathbf{A}^{-1} = \frac{\text{adj}(\mathbf{A})}{\det(\mathbf{A})}$$

$$(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$$

$$\left( \prod_{i=1}^n \mathbf{A}_i \right)^{-1} = \prod_{i=1}^n (\mathbf{A}_{n-i+1})^{-1}$$

$$(\mathbf{A}^{-1} \mathbf{B} \mathbf{A})^n = \mathbf{A}^{-1} \mathbf{B}^n \mathbf{A}, \quad n \in \mathbb{N}$$

## xiv Vandermonde matrix (范德蒙矩陣)

A Vandermonde matrix  $\mathbf{V}$  is:

$$V_{i,j} = \alpha_i^{j-1}$$

An  $n \times n$  Vandermonde matrix  $\mathbf{V}$  satisfies:

$$\det(V) = \prod_{1 \leq i < j \leq n} (\alpha_j - \alpha_i)$$

## VI Diagonal matrix (對角矩陣)

### i Diagonal matrix

A diagonal matrix over a field  $F$  is a square matrix over  $F$  of which all non-diagonal entries are the additive identity of  $F$ .

## ii Diagonal matrix identities

If:

$$\mathbf{D} = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix}$$

where all  $d_i \neq 0$ , then:

$$\mathbf{D}^{-1} = \begin{bmatrix} \frac{1}{d_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{d_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{d_n} \end{bmatrix}$$

Let  $\mathbf{E}, \mathbf{F}$  be two  $n \times n$  diagonal matrices:

$$\mathbf{EF} = [e_{ij}f_{ij}]_{1 \leq i, j \leq n}$$

## VII System of Linear Equations (線性方程組)

### i System of equations (方程組)

- **Consistent system (相容方程組):** A system of equations is considered consistent if it has at least one solution.
- **Inconsistent/contradictory system (不相容方程組):** A system of equations is considered inconsistent or contradictory if it has no solution.
- **Dependent system (相依方程組):** A consistent system of equations is considered dependent if it has infinite solutions.
- **Independent system (獨立方程組):** A consistent system of equations is considered independent if it has finite solutions.

### ii Linear Equation (線性方程)

A linear equation is an equation that may be put in the form:

$$\sum_{k=1}^n a_k x_k = c,$$

where  $a_k$  and  $c$  are called coefficients.

If  $c = 0$ , the linear equation is called a homogeneous linear equation; otherwise, it is called a heterogeneous linear equation.

### iii System of Linear Equations (線性方程組)

A system of linear equations is a collection of two or more linear equations involving the same tuple of variables.

#### iv Coefficient matrix (係數矩陣)

The coefficients of a system of linear equations written as a matrix, with one row for each equation.

#### v Augmented matrix (增廣矩陣)

Coefficient matrix with the constants on the right side of the equal signs of a system of linear equations augmented to the last column of the matrix.

#### vi Gaussian Elimination (高斯消去法) or row reduction (zh-hant: 列/zh-hans: 行簡化)

An algorithm to solve a system of linear equations by using row operations to modify its augmented matrix to row echelon form. Two systems of linear equations have the same solution if and only if their augmented matrices are row-equivalent.

#### vii Cramer Rule or Cramer formula (克拉瑪公式)

Given a system of linear equations that can be written in the form:

$$\mathbf{AX} = \mathbf{C}$$

where  $\mathbf{A}$  is an  $n \times n$  coefficient matrix,  $\mathbf{X}$  is a variable matrix:

$$\mathbf{X} = (x_1, x_2, x_3, \dots, x_n)^T,$$

and  $\mathbf{C}$  is a coefficient matrix:

$$\mathbf{C} = (c_1, c_2, c_3, \dots, c_n)^T.$$

Let  $\mathbf{A}_i$  be  $\mathbf{A}$  with the  $i$ th column replaced by  $\mathbf{C}$ . The solution to the system of linear equations is given by:

$$\begin{cases} x_i = \frac{\det(\mathbf{A}_i)}{\det(\mathbf{A})}, & \det(\mathbf{A}) \neq 0 \\ \text{no solution,} & \det(\mathbf{A}) = 0 \wedge \sum_{i=1}^n \det(\mathbf{A}_i) \neq 0 \\ \text{arbitrary,} & \det(\mathbf{A}) = 0 \wedge \sum_{i=1}^n \det(\mathbf{A}_i) = 0 \end{cases}$$

where  $x_i = \frac{\det(\mathbf{A}_i)}{\det(\mathbf{A})}$  can be geometrically viewed as the volume of the parallelogram formed by the column vectors of  $\mathbf{A}_i$  divided by the volume of the parallelogram formed by the column vectors of  $\mathbf{A}$ .

Particularly, when  $\mathbf{C}$  is a zero matrix, that is,

$$\mathbf{AX} = \mathbf{0},$$

the solution to the system of linear equations is given by:

$$\begin{cases} x_i = 0, & \det(\mathbf{A}) \neq 0 \\ \text{arbitrary,} & \det(\mathbf{A}) = 0 \end{cases}$$

### 3 Linear Subspace (線性子空間)

#### I Linear subspace

A linear subspace of a vector space  $V$  over a field  $\mathbb{K}$  is a nonempty subset  $W$  of  $V$  such that,

$$\forall w_1, w_2 \in W, a, b \in \mathbb{K} : aw_1 + bw_2 \in W.$$

#### II Four fundamental subspaces of a matrix (矩陣的四大子空間)

The four fundamental subspaces of a matrix are the row space, column space, null space, and left null space.

#### III Column space (zh-hant: 行/zh-hans: 列空間)

The column space of  $\mathbf{A}$ , denoted as  $\text{Col}(\mathbf{A})$  or  $\mathbf{C}(\mathbf{A})$ , is defined as the set of all linear combinations of the column vectors of  $\mathbf{A}$ .

#### IV Row space (zh-hant: 列/zh-hans: 行空間)

The row space of  $\mathbf{A}$  is defined as the set of all linear combinations of the row vectors of  $\mathbf{A}$ , that is,  $\text{Col}(\mathbf{A}^T)$ .

#### V Null space (零空間), kernel (核), or kernel space (核空間)

The kernel of a linear map  $L : V \rightarrow W$ , denoted as  $\ker(L)$  or  $\text{Null}(L)$ , is defined as:

$$\ker(L) = \{v \in V : L(v) = 0\}.$$

$\dim(\ker(\mathbf{A}))$  is called nullity (零化度).

#### VI Left null space (左零空間), left kernel (左核), or left kernel space (左核空間)

The left null space of a linear map  $L : V \rightarrow W$  is defined as  $\ker(L^{-1})$ .

#### VII Image space (像空間) or range (值域)

The image space of a linear map  $T : V \rightarrow W$ , denoted as  $\text{im}(T)$

$$\text{im}(T) = \{T(v) : v \in V\}.$$

#### VIII Cokernel (餘核)

The cokernel of a linear map  $T : V \rightarrow W$ , denoted as  $\text{coker}(T)$

$$\text{coker}(T) = W \setminus \text{im}(T).$$

## IX Constant rank theorem (常秩定理)

For any matrix  $\mathbf{A}$ :

$$\dim(\text{Col}(\mathbf{A})) = \dim(\text{Col}(\mathbf{A}^\top))$$

*Proof.*

Let  $\mathbf{A}$  be an  $m \times n$  matrix,  $r = \dim(\text{Col}(\mathbf{A}))$ ,  $c_1, c_2, \dots, c_r$  be a basis of  $\text{Col}(\mathbf{A})$ , matrix  $\mathbf{C} = [c_1, c_2, \dots, c_r]$ , and  $r \times n$  matrix  $\mathbf{R}$  such that  $\mathbf{A} = \mathbf{CR}$ .

Since:

$$\begin{aligned}\text{Col}(\mathbf{A}^\top) &\subseteq \text{Col}(\mathbf{R}^\top), \\ \dim(\text{Col}(\mathbf{A}^\top)) &\leq \dim(\text{Col}(\mathbf{R}^\top))\end{aligned}$$

Because there are only  $r$  rows in  $\mathbf{R}$ ,

$$\dim(\text{Col}(\mathbf{R}^\top)) \leq \dim(\text{Col}(\mathbf{A}))$$

Similarly, we can obtain that

$$\dim(\text{Col}(\mathbf{A})) \leq \dim(\text{Col}(\mathbf{R}^\top)) \leq \dim(\text{Col}(\mathbf{A}^\top)).$$

Therefore,

$$\dim(\text{Col}(\mathbf{A})) = \dim(\text{Col}(\mathbf{R}^\top)) = \dim(\text{Col}(\mathbf{A}^\top)).$$

□

## X Rank (秩)

If  $\mathbf{A}$  is an  $m \times n$  matrix, then the rank of  $\mathbf{A}$ , denoted as  $\text{rank}(\mathbf{A})$  or  $\text{rk}(\mathbf{A})$ , is the maximum number of linearly independent columns (or rows) in  $\mathbf{A}$ , namely:

$$\text{rank}(\mathbf{A}) = \dim(\text{Col}(\mathbf{A}))$$

## XI Rank-nullity theorem (秩—零化度定理)

For any  $n \times n$  matrix  $\mathbf{A}$ :

$$\text{rank}(\mathbf{A}) + \dim(\ker(\mathbf{A})) = n$$

*Proof.*

Let  $\mathbf{A}$  be an  $m \times n$  matrix with  $r$  linearly independent columns (i.e.,  $\text{rank}(\mathbf{A}) = r$ ). We will show that there exists a set of  $n - r$  linearly independent solutions to the homogeneous system  $\mathbf{Ax} = \mathbf{0}$ .

To do this, we will produce an  $n \times (n - r)$  matrix  $\mathbf{X}$  whose columns form a basis of the null space of  $\mathbf{A}$ .

Without loss of generality, assume that the first  $r$  columns of  $\mathbf{A}$  are linearly independent. So, we can write

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_2 \end{pmatrix},$$

where

- $\mathbf{A}_1$  is an  $m \times r$  matrix with  $r$  linearly independent column vectors, and
- $\mathbf{A}_2$  is an  $m \times (n - r)$  matrix such that each of its  $n - r$  columns is a linear combination of the columns of  $\mathbf{A}_1$ .

This means that  $\mathbf{A}_2 = \mathbf{A}_1 \mathbf{B}$  for some  $r \times (n - r)$  matrix  $\mathbf{B}$ , and hence,

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_1 \mathbf{B} \end{pmatrix}.$$

Let

$$\mathbf{X} = \begin{pmatrix} -\mathbf{B} \\ \mathbf{I}_{n-r} \end{pmatrix},$$

where  $\mathbf{I}_{n-r}$  is the  $(n - r) \times (n - r)$  identity matrix. So,  $\mathbf{X}$  is an  $n \times (n - r)$  matrix such that

$$\mathbf{A}\mathbf{X} = \begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_1 \mathbf{B} \end{pmatrix} \begin{pmatrix} -\mathbf{B} \\ \mathbf{I}_{n-r} \end{pmatrix} = -\mathbf{A}_1 \mathbf{B} + \mathbf{A}_1 \mathbf{B} = \mathbf{0}_{m \times (n-r)}.$$

Therefore, each of the  $n - r$  columns of  $\mathbf{X}$  are particular solutions of  $\mathbf{A}\mathbf{x} = \mathbf{0}_{\mathbb{F}^m}$ .

Furthermore, the  $n - r$  columns of  $\mathbf{X}$  are linearly independent because  $\mathbf{X}\mathbf{u} = \mathbf{0}_{\mathbb{F}^n}$  will imply  $\mathbf{u} = \mathbf{0}_{\mathbb{F}^{n-r}}$  for  $\mathbf{u} \in \mathbb{F}^{n-r}$ :

$$\mathbf{X}\mathbf{u} = \mathbf{0}_{\mathbb{F}^n} \implies \begin{pmatrix} -\mathbf{B} \\ \mathbf{I}_{n-r} \end{pmatrix} \mathbf{u} = \mathbf{0}_{\mathbb{F}^n} \implies \begin{pmatrix} -\mathbf{B}\mathbf{u} \\ \mathbf{u} \end{pmatrix} = \begin{pmatrix} \mathbf{0}_{\mathbb{F}^r} \\ \mathbf{0}_{\mathbb{F}^{n-r}} \end{pmatrix} \implies \mathbf{u} = \mathbf{0}_{\mathbb{F}^{n-r}}.$$

Therefore, the column vectors of  $\mathbf{X}$  constitute a set of  $n - r$  linearly independent solutions for  $\mathbf{A}\mathbf{x} = \mathbf{0}_{\mathbb{F}^m}$ .

We next prove that any solution of  $\mathbf{A}\mathbf{x} = \mathbf{0}_{\mathbb{F}^m}$  must be a linear combination of the columns of  $\mathbf{X}$ .

Let

$$\mathbf{u} = \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{pmatrix} \in \mathbb{F}^n$$

be any vector such that  $\mathbf{A}\mathbf{u} = \mathbf{0}_{\mathbb{F}^m}$ . Since the columns of  $\mathbf{A}_1$  are linearly independent,  $\mathbf{A}_1 \mathbf{x} = \mathbf{0}_{\mathbb{F}^m}$  implies  $\mathbf{x} = \mathbf{0}_{\mathbb{F}^r}$ .

Therefore,

$$\begin{aligned} \mathbf{A}\mathbf{u} &= \mathbf{0}_{\mathbb{F}^m} \\ \implies \begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_1 \mathbf{B} \end{pmatrix} \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{pmatrix} &= \mathbf{A}_1 \mathbf{u}_1 + \mathbf{A}_1 \mathbf{B} \mathbf{u}_2 = \mathbf{A}_1 (\mathbf{u}_1 + \mathbf{B} \mathbf{u}_2) = \mathbf{0}_{\mathbb{F}^m} \\ \implies \mathbf{u}_1 + \mathbf{B} \mathbf{u}_2 &= \mathbf{0}_{\mathbb{F}^r} \\ \implies \mathbf{u}_1 &= -\mathbf{B} \mathbf{u}_2 \\ \implies \mathbf{u} &= \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{pmatrix} = \begin{pmatrix} -\mathbf{B} \\ \mathbf{I}_{n-r} \end{pmatrix} \mathbf{u}_2 = \mathbf{X} \mathbf{u}_2. \end{aligned}$$

This proves that any vector  $\mathbf{u}$  that is a solution of  $\mathbf{A}\mathbf{x} = \mathbf{0}$  must be a linear combination of the  $n - r$  special solutions given by the columns of  $\mathbf{X}$ . And we have already seen that the columns of  $\mathbf{X}$  are linearly independent. Hence, the columns of  $\mathbf{X}$  constitute a basis for the null space of  $\mathbf{A}$ . Therefore, the nullity of  $\mathbf{A}$  is  $n - r$ . Since  $r$  equals the rank of  $\mathbf{A}$ , it follows that  $\text{rank}(\mathbf{A}) + \text{ker}(\mathbf{A}) = n$ . This concludes our proof.  $\square$

## XII Full rank (滿秩) and rank deficiency (秩虧)

For an  $m \times n$  matrix  $\mathbf{A}$ , if

$$\text{rank}(\mathbf{A}) = \min(m, n)$$

$\mathbf{A}$  is called to have full rank; otherwise it is called to be rank-deficient (欠秩或秩虧的).

The rank deficiency of a  $m \times n$  matrix  $\mathbf{A}$  is defined as

$$\min(m, n) - \text{rank}(\mathbf{A})$$

## XIII Property of rank

- For any  $m \times n$  matrix  $\mathbf{A}$ :

$$\text{rank}(\mathbf{A}) \leq \min(m, n)$$

- A square matrix is invertible if and only if it has full rank.
- A matrix is a zero matrix if and only if its rank is zero.

## 4 Matrix Analysis (矩陣分析)

### I Eigenvector or characteristic vector (特徵向量) and characteristic polynomial (特徵多項式)

#### i Eigenvector or characteristic vector (特徵向量) and eigenvalue (本徵值、固有值、特徵值或特徵根)

Given a linear transformation  $L: V \rightarrow V$ , an eigenvector of it is defined as the vector  $v \in V$  such that:

$$\mathbf{A}v = \lambda v,$$

where  $\lambda$  is a scalar, which is called the eigenvalue of  $\mathbf{A}$  or the eigenvalue of  $v$ .

#### ii Characteristic polynomial

For an  $n \times n$  matrix  $\mathbf{A}$ , the characteristic polynomial of it is defined as

$$p_{\mathbf{A}}(\lambda) := \det(\lambda \mathbf{I}_n - \mathbf{A}),$$

where  $\lambda$  is a scalar variable.

A scalar  $\lambda$  is an eigenvalue of  $\mathbf{A}$  if and only if

$$p_{\mathbf{A}}(\lambda) = 0.$$

### iii Property

- The number of distinct eigenvalues of a square matrix is less than or equal to the number of linearly independent eigenvectors.
- The number of distinct eigenvalues of a  $n \times n$  matrix is less than or equal to  $n$ .
- The algebraic multiplicity of a root  $\lambda$  of the characteristic polynomial of a square matrix is greater than or equal to the number of linearly independent eigenvectors of it whose eigenvalue is  $\lambda$ .

### iv Cayley–Hamilton theorem (凱萊–哈密頓定理)

For a square matrix  $\mathbf{A}$  and its characteristic polynomial  $p_{\mathbf{A}}$ , the Cayley–Hamilton theorem states that:

$$p_{\mathbf{A}}(\mathbf{A}) = 0$$

### v Diagonalization (對角化)

The diagonalization of an  $n \times n$  matrix  $\mathbf{A}$  is the construction of an  $n \times n$  matrix  $\mathbf{P}$  and an  $n \times n$  diagonal matrix  $\mathbf{D}$  such that:

$$\mathbf{A} = \mathbf{P}^{-1} \mathbf{D} \mathbf{P}$$

An  $n \times n$  matrix can be diagonalized if it has  $n$  distinct eigenvalues.

The steps of the diagonalization of an  $n \times n$  matrix  $\mathbf{A}$  is as follows:

1. For each eigenvalue  $\lambda$  of the matrix  $\mathbf{A}$ , solve the eigenvector  $\mathbf{v}$ :

$$(\mathbf{A} - \lambda \mathbf{I})\mathbf{v} = 0.$$

2. Compose matrix  $\mathbf{P}$  with each eigenvector of  $\mathbf{A}$  as a row.

Common application:

$$\mathbf{A}^n = (\mathbf{P}^{-1} \mathbf{D} \mathbf{P})^n = \mathbf{P}^{-1} \mathbf{D}^n \mathbf{P}$$

## II Generalized Definition of Inverse

### i One-sided inverse (單邊逆)

- Right inverse (右逆): For any  $n \times m$  matrix  $\mathbf{A}$  with rank  $\text{rank}(\mathbf{A}) = n$ , there exists an  $m \times n$  matrix  $\mathbf{A}_R^{-1}$ , called the right inverse of  $\mathbf{A}$ , such that

$$\mathbf{A} \mathbf{A}_R^{-1} = \mathbf{I}_n,$$

where  $\mathbf{I}_n$  is the  $n \times n$  identity matrix.

- Left inverse (左逆): For any  $n \times m$  matrix  $\mathbf{A}$  with rank  $\text{rank}(\mathbf{A}) = m$ , there exists an  $m \times n$  matrix  $\mathbf{A}_L^{-1}$ , called the left inverse of  $\mathbf{A}$ , such that

$$\mathbf{A}_L^{-1} \mathbf{A} = \mathbf{I}_m,$$

where  $\mathbf{I}_m$  is the  $m \times m$  identity matrix.



Properties:

- A matrix doesn't necessarily have a right inverse and doesn't necessarily have a left inverse.
- A square matrix has a right inverse if and only if it has left inverse if and only if it is invertible.
- For any invertible matrix, there exists only one right inverse and one left inverse, which are its multiplicative inverse.

## ii Generalized inverse or g-inverse (廣義逆)

A matrix  $A^g \in \mathbb{C}^{n \times m}$  is a generalized inverse of a matrix  $A \in \mathbb{C}^{m \times n}$  if:

$$AA^gA = A.$$

Properties:

- For all matrices, there exists at least one generalized inverse, but it's not necessarily unique.
- For any invertible matrix, there exists only one generalized inverse, which is its multiplicative inverse.

## iii Reflexive generalized inverse (自反廣義逆)

A reflexive generalized inverse of a matrix  $\mathbf{A}$  is defined as a matrix  $\mathbf{A}^g$  such that:

$$\mathbf{A}\mathbf{A}^g\mathbf{A} = \mathbf{A}$$

$$\mathbf{A}^g\mathbf{A}\mathbf{A}^g = \mathbf{A}^g$$

Properties:

- For all matrices, there exists at least one reflexive generalized inverse, but it's not necessarily unique.
- For any invertible matrix, there exists only one reflexive generalized inverse, which is its multiplicative inverse.

## iv Moore–Penrose inverse (摩爾—彭若斯廣義逆) or pseudoinverse (偽逆或擬反)

A Moore–Penrose inverse of a matrix  $\mathbf{A}$  is defined as a matrix  $\mathbf{A}^+$  such that:

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

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

$$(\mathbf{A}\mathbf{A}^+)^* = \mathbf{A}\mathbf{A}^+$$

$$(\mathbf{A}^+\mathbf{A})^* = \mathbf{A}^+\mathbf{A}$$

Properties:

- For all matrices, there exists only one Moore–Penrose inverse.
- For any invertible matrix, there exists only one Moore–Penrose inverse, which is its multiplicative inverse.

## 5 Common Transformation Matrices

### I Transformation matrices in two-dimensional space

#### i Degenerate matrices

Degenerate to 1 dimension  $((a \vee b) \wedge (d \vee e) \neq 0)$ :

$$\begin{bmatrix} a & b \\ ca & cb \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax + by \\ c(ax + by) \end{bmatrix}$$
$$\begin{bmatrix} 0 & 0 \\ d & e \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ dx + ey \end{bmatrix}$$

Degenerate to 0 dimension:

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

#### ii Stretching (伸縮/縮放/拉伸)

A stretch is a linear transformation which enlarges all distances in a particular direction by a constant factor  $k$  but does not affect distances in the perpendicular direction. We only consider stretches along the  $x$ -axis and  $y$ -axis. (Note that if  $k > 1$ , then this really is a "stretch"; if  $k < 1$ , it is technically a "compression", but we still call it a stretch. Also, if  $k = 1$ , then the transformation is an identity, i.e. it has no effect.)

The matrix associated with a stretch by a factor  $k$  along the  $x$ -axis is given by:

$$\begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix}$$

The matrix associated with a stretch by a factor  $k$  along the  $y$ -axis is given by:

$$\begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix}$$

#### iii Shearing (推移)

A shearing in the  $x$ -direction dependent on the  $y$  coordination by a factor  $k$  is given by:

$$\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$$

A shearing in the  $y$ -direction dependent on the  $x$  coordination by a factor  $k$  is given by:

$$\begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}$$

#### iv Rotation (旋轉)

A rotation by an angle  $\theta$  counterclockwise about the origin is given by:

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

## v Reflection (鏡射/反射)

A reflection about a line that passes the original and with  $(a, b) = \sqrt{a^2 + b^2}(\cos \theta, \sin \theta)$  being a vector in the direction of it is given by:

$$\begin{bmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{bmatrix} = \frac{1}{a^2 + b^2} \begin{bmatrix} a^2 - b^2 & 2ab \\ 2ab & b^2 - a^2 \end{bmatrix}$$

## vi Orthogona projection (正射影)

An orthogonal projection about a line that passes the original and with  $(a, b) = \sqrt{a^2 + b^2}(\cos \theta, \sin \theta)$  being a vector in the direction of it is given by:

$$\begin{bmatrix} \cos^2(\theta) & \cos(\theta)\sin(\theta) \\ \cos(\theta)\sin(\theta) & \sin^2(\theta) \end{bmatrix} = \frac{1}{a^2 + b^2} \begin{bmatrix} a^2 & ab \\ ab & b^2 \end{bmatrix}$$

## vii Exmample 1

**Question:** The cricle  $C : x^2 + y^2 = 1$  is transformed by  $\mathbf{A} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$  to obtain a new curve  $C'$ . Find the equation of  $C'$  and its area.

**Answer:**

$$\begin{aligned} \begin{bmatrix} x' \\ y' \end{bmatrix} &= \mathbf{A} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x \\ 3y \end{bmatrix} \\ \begin{cases} x' = 2x \\ y' = 3y \end{cases} &\Rightarrow \begin{cases} x = \frac{x'}{2} \\ y = \frac{y'}{3} \end{cases} \\ \left(\frac{x'}{2}\right)^2 + \left(\frac{y'}{3}\right)^2 = 1 &\Rightarrow \frac{x'^2}{4} + \frac{y'^2}{9} = 1 \\ C' : \frac{x'^2}{4} + \frac{y'^2}{9} &= 1 \end{aligned}$$

The area of  $C'$  is  $\pi \cdot 2 \cdot 3 = 6\pi$ . Since  $\det(\mathbf{A}) = 6$ , and the area of  $C$  is  $\pi$ , the area of  $C'$  equals  $\det(\mathbf{A})$  multiplied by the area of  $C$ .

## viii Exmample 2

**Question:** The line  $\mathbf{L} : 4x - 3y = 5$  is transformed in the  $x$ -direction dependent on the  $y$  coordination by twice and then in the  $y$ -direction dependent on the  $x$  coordination by negative three times to obtain a new line  $\mathbf{L}'$ . Find the equation of  $\mathbf{L}'$ .

**Answer:**

$$\begin{aligned} \mathbf{A} &= \begin{bmatrix} 1 & 2 \\ -3 & 1 \end{bmatrix} \\ \mathbf{A}^{-1} &= \frac{1}{7} \begin{bmatrix} 1 & -2 \\ 3 & 1 \end{bmatrix} \end{aligned}$$

$$\begin{bmatrix} x \\ y \end{bmatrix} = \mathbf{A}^{-1} \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \frac{x' - 2y'}{7} \\ \frac{3x' + y'}{7} \end{bmatrix}$$

$$\mathbf{L}' : 4\frac{x-2y}{7} + 3\frac{3x+y}{7} = 5$$

$$\mathbf{L}' : 5x + 11y + 35 = 0$$

### ix Example 3

**Question:** Reflect the point  $P(5, 2)$  over the line  $\mathbf{M} : x - y + 1 = 0$  to obtain a new point  $Q$ . Find  $Q$ .

**Answer:**

Shift one unit to the right:  $\mathbf{M}$  becomes  $\mathbf{M}' : x - y = 0$  and  $C$  becomes  $C'(6, 2)$ .

The reflection matrix is:

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 1 & 0 \end{bmatrix}$$

$$\mathbf{D}' = \mathbf{AC}' = (2, 6)$$

Shift one unit to the left:  $D(1, 6)$

### x Example 4

**Question:** The line  $x + y - 2 = 0$  is transformed by  $\mathbf{A} = \begin{bmatrix} a & b \\ 1 & 2 \end{bmatrix} \in \mathbb{R}^{2 \times 2}$  to obtain a new line  $2x + 3y - 4 = 0$ . Find  $a$  and  $b$ .

**Answer:**

$$\mathbf{A} \begin{bmatrix} x \\ -x + 2 \end{bmatrix} = \begin{bmatrix} x' \\ \frac{-2x' + 4}{3} \end{bmatrix}$$

$$x - 2x + 4 = \frac{-2x' + 4}{3}$$

$$x' = \frac{3}{2}x - 4$$

$$ax - bx - 2b = x' = \frac{3}{2}x - 4$$

$$(a, b) = \left(-\frac{1}{2}, -2\right)$$

## II Transformation matrices in three-dimensional space

### i Stretching (伸縮/縮放/拉伸)

The matrix associated with a stretch by a factor  $k$  along the  $x$ -axis is given by:

$$\begin{bmatrix} k & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The matrix associated with a stretch by a factor  $k$  along the  $y$ -axis is given by:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The matrix associated with a stretch by a factor  $k$  along the  $z$ -axis is given by:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & k \end{bmatrix}$$

## ii Shearing (推移)

A shearing in the  $x$ -direction dependent on the  $y$  coordination by a factor  $k$  is given by:

$$\begin{bmatrix} 1 & k & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

A shearing in the  $x$ -direction dependent on the  $z$  coordination by a factor  $k$  is given by:

$$\begin{bmatrix} 1 & 0 & k \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

A shearing in the  $y$ -direction dependent on the  $x$  coordination by a factor  $k$  is given by:

$$\begin{bmatrix} 1 & 0 & 0 \\ k & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

A shearing in the  $y$ -direction dependent on the  $z$  coordination by a factor  $k$  is given by:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & k \\ 0 & 0 & 1 \end{bmatrix}$$

A shearing in the  $z$ -direction dependent on the  $x$  coordination by a factor  $k$  is given by:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ z & 0 & 1 \end{bmatrix}$$

A shearing in the  $z$ -direction dependent on the  $y$  coordination by a factor  $k$  is given by:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & k & 1 \end{bmatrix}$$

### iii Rotation (旋轉)

A rotation by an angle  $\theta$  counterclockwise about a line that passes the origin and with  $(a, b, c)$  being a vector in the direction of it is given by:

$$\frac{1}{a^2 + b^2 + c^2} \begin{bmatrix} a^2(1 - \cos \theta) + \cos \theta & ab(1 - \cos \theta) - c \sin \theta & ac(1 - \cos \theta) + b \sin \theta \\ bc(1 - \cos \theta) + c \sin \theta & b^2(1 - \cos \theta) + \cos \theta & bc(1 - \cos \theta) - a \sin \theta \\ ac(1 - \cos \theta) - b \sin \theta & bc(1 - \cos \theta) + a \sin \theta & c^2(1 - \cos \theta) + \cos \theta \end{bmatrix}$$

### iv Reflection (鏡射/反射)

A reflection about a plane  $ax + by + cz = 0$  is given by:

$$\mathbf{I} - \frac{1}{a^2 + b^2 + c^2} \begin{bmatrix} 2a^2 & 2ab & 2ac \\ 2ab & 2b^2 & 2bc \\ 2ac & 2bc & 2c^2 \end{bmatrix},$$

where  $\mathbf{I}$  is the  $3 \times 3$  identity matrix.

### v Orthogonal projection (正射影)

An orthogonal projection about a plane  $ax + by + cz = 0$  is given by:

$$\mathbf{I} - \frac{1}{a^2 + b^2 + c^2} \begin{bmatrix} a^2 & ab & ac \\ ab & b^2 & bc \\ ac & bc & c^2 \end{bmatrix}$$

where  $\mathbf{I}$  is the  $3 \times 3$  identity matrix.