

# Linear Algebra

Vinsong

November 24, 2025

## **Abstract**

The lecture note of 2025 Fall Linear Algebra by professor 李明穗 (Amy Lee) .

# Contents

|          |  |           |
|----------|--|-----------|
| <b>0</b> | <b>Introduction</b>  | <b>2</b>  |
| 0.1      | Geometry . . . . .   | 2         |
| 0.2      | Abstract Algebra . . . . .   | 2         |
| 0.3      | Applied Science . . . . .  | 2         |
| <b>1</b> | <b>Matrices and Gaussian Elimination</b>   | <b>3</b>  |
| 1.1      | Introduction . . . . .   | 3         |
| 1.2      | Geometry of Linear Equation . . . . .  | 4         |
| 1.3      | An Example of Gaussian Elimination . . . . .                                       | 8         |
| 1.4      | Matrix Notation and Matrix Multiplication . . . . .                                | 10        |
| 1.5      | Triangular Factors and Row Exchanges . . . . .                                     | 14        |
| 1.6      | Inverse and Transpose . . . . .  | 18        |
| 1.7      | Transpose $A^T$ . . . . .  | 21        |
| <b>2</b> | <b>Vector Spaces and Linear Equation</b>   | <b>23</b> |
| 2.1      | Vector Spaces and Subspace . . . . .   | 23        |
| 2.2      | The Solution of $m$ Equations in $n$ Unknowns . . . . .                            | 28        |
| 2.3      | Linear Independence, Basis and Dimension . . . . .                                 | 32        |
| 2.4      | The Four Fundamental Subspaces . . . . .   | 37        |
| 2.5      | Graph and Network . . . . .  | 42        |
| 2.6      | Linear Transformation . . . . .  | 42        |
| <b>3</b> | <b>Orthogonality</b>   | <b>49</b> |
| 3.1      | Perpendicular Vectors and Orthogonal Subspaces . . . . .                           | 49        |
| 3.2      | Inner Product and Projections onto Lines . . . . .                                 | 55        |
| 3.3      | Projections and Least Squares Applications . . . . .                               | 57        |
| 3.4      | Orthogonal Bases, Orthogonal Matrices and Gram-Schmidt Orthogonalization . . . . . | 62        |

# Chapter 0

## Introduction

### Lecture 1

#### 0.1 Geometry

2 Sep. 13:20

- linear
- To study geometry with linearity
- In a different dimension:
  - In 2D: **lines**
  - In 3D: **planes**
  - In  $n$ D: **hyperplanes**

#### 0.2 Abstract Algebra

**Definition 0.2.1 (Linear Algebra).** Here is the definition of Linear algebra.

- Algebra is the study of basic "mathematical structure."  
e.g. **Group**, **Ring**, **Field**, ...etc.
- Linear Algebra studies one of the structures called **vector space**.

**Note.** Followed by logical deduction from the basic definition, we can derive some theorems.

#### 0.3 Applied Science

- **Mathematic:** ODE, PDE.
- **Linear Programming:** developing during World War II
- **Image Processing, Computer Vision, Computer Graphic,** etc.

# Chapter 1

## Matrices and Gaussian Elimination

### 1.1 Introduction

The central problem of Linear Algebra is the solution of Linear Equations. The most important and simplest case is when the # of unknowns equals to the # of equations.

**Note.** There are two ways to solve linear equations:

- The method of elimination (**Gaussian Elimination**)
- Determinants (**Crammer's Rule**)

#### 1.1.1 Four aspects to follow

(1) The geometry of linear equations.

**Note.**  $n = 2, n = 3 \rightarrow$  higher dimensional space.

(2) The interpretation of elimination is a factorization of the coefficient matrix.

**Definition.** Some notation to define:

**Definition 1.1.1 (Scalar, Matrix, Vector).**

$$Ax = b \quad \begin{cases} \alpha, \beta, \gamma : & \text{scalar} \\ A, B, C : & \text{matrix} \\ a, b, c : & \text{vector} \end{cases}$$

**Definition 1.1.2 (Lower/Upper triangular matrix).**

$$A = LU \quad \begin{cases} L : & \text{lower triangular matrix} \\ U : & \text{upper triangular matrix} \end{cases}$$

**Definition 1.1.3 (Transpose/Inverse).**

$$A^T / A^{-1} : \quad \begin{cases} A^T : & \text{Transpose of matrix A} \\ A^{-1} : & \text{Inverse of matrix A} \end{cases}$$

(3) Irregular case and Singular case (**no unique solution**):

**Note.** no solution or infinitely many solutions

(4) The # of operations to solve the system by elimination

## 1.2 Geometry of Linear Equation

**Example.** Consider the linear equation below:

$$\begin{cases} 2x - y = 1 \\ x + y = 5 \end{cases}$$

- **approach 1: row picture** → two lines in plane

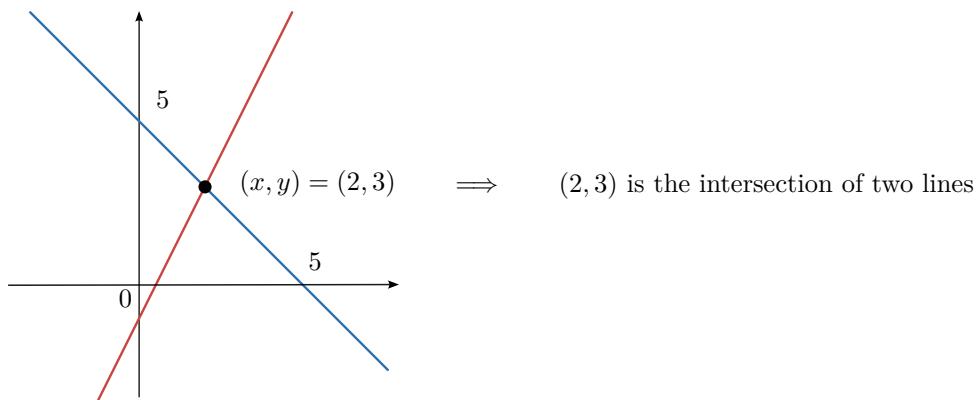


Figure 1.1: Row Picture

- **approach 2: column picture**

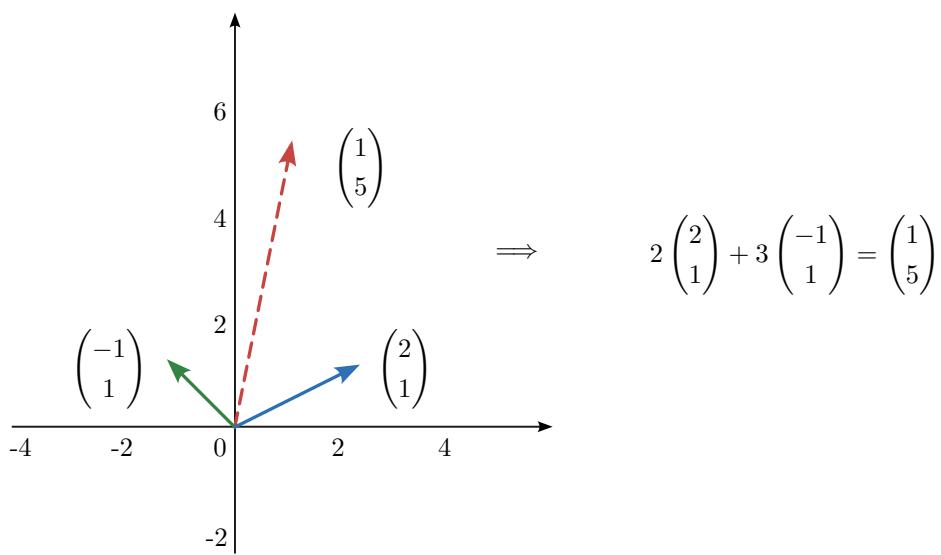


Figure 1.2: Column Picture

**Lemma 1.2.1** (Linear Combination).

$$x \begin{pmatrix} 2 \\ 1 \end{pmatrix} + y \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 5 \end{pmatrix}$$

To find the **Linear Combination** of  $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$  and  $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$  to reach  $\begin{pmatrix} 1 \\ 5 \end{pmatrix}$

**Note.** A vector is a  $n \times 1$  array with  $n$  real numbers,  $c_n$  is

$$\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$$

But in the text, we use

$$(c_1, \dots, c_n)$$

to represent.

**Definition.** Here are some operations on matrix:

**Definition 1.2.1.**

$$\alpha \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}_{n \times 1} = \begin{pmatrix} \alpha \cdot c_1 \\ \vdots \\ \alpha \cdot c_n \end{pmatrix}_{n \times 1}, \quad \alpha \in \mathbb{R}$$

**Definition 1.2.2.**

$$\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} + \begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix} = \begin{pmatrix} c_1 + d_1 \\ \vdots \\ c_n + d_n \end{pmatrix}_{n \times 1}$$

**Definition 1.2.3.**

$$y \in \mathbb{R}$$

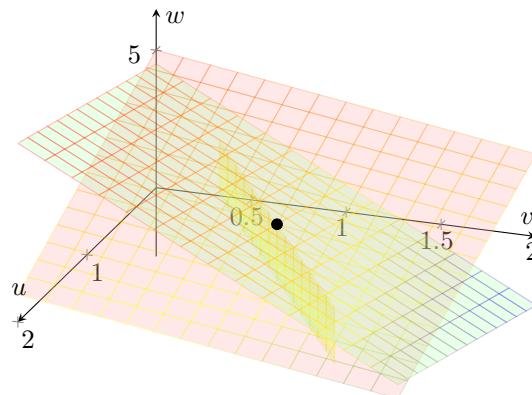
$$y \in \mathbb{R}^2 \implies y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}_{2 \times 1} \quad y_1, y_2 \in \mathbb{R}$$

$$y \in \mathbb{R}^3 \implies y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}_{3 \times 1} \quad y_1, y_2, y_3 \in \mathbb{R}$$

**Example.** Consider the linear equation below:

$$\begin{cases} 2u + v + w = 5 \\ 4u - 6v = -2 \\ -2u + 7u + 2w = 9 \end{cases}$$

- Row picture



$$(u, v, w) = (1, 1, 2)$$

**Lemma 1.2.2.** in  $n$ -dimension, a line require  $(n - 1)$  equation.

**Question.** How to extend into  $n$ -dimensions?

**Answer.** Consider the following steps:

- Each equation represents a plane or hyperplane.
- The first equation produces a  $(n - 1)$ -dimension plane in  $\mathbb{R}^n$
- The second equation produces another  $(n - 1)$ -dimension plane in  $\mathbb{R}^n$
- Their intersection in smaller set of  $(n - 2)$ -dimension
- $(n - 3) \rightarrow (n - 4) \rightarrow \dots \rightarrow 3 \rightarrow 2 \rightarrow 1 \rightarrow$  point

Then we can find the final intersection. (\*)

- Column picture

$$u \begin{pmatrix} 2 \\ 4 \\ -2 \end{pmatrix} + v \begin{pmatrix} 1 \\ -6 \\ 7 \end{pmatrix} + w \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 5 \\ -2 \\ 9 \end{pmatrix} \iff \begin{cases} 2u + v + w = 5 \\ 4u - 6v = -2 \\ -2u + 7u + 2w = 9 \end{cases}$$

RHS is a linear combination of 3 column vectors.

**Theorem 1.2.1.** Solution to a linear equation:

$$(\text{intersection of } 2 \text{ points})_{\text{row pic.}} = (\text{coefficient of linear combination})_{\text{column pic.}}$$

### 1.2.1 Singular Case

(1) Row Picture: In 3D case, they didn't intersect at a point.

- **Case 1:** two parallel

$$\begin{cases} 2u + v + w = 5 \\ 4u + 2v + 2w = 9 \end{cases}$$

- **Case 2:** three plane perpendicular ( $\perp$ )

$$\begin{cases} u + v + w = 2 \cdots (1) \\ 2u + 3w = 5 \cdots (2) \\ 3u + v + 4w = 6 \cdots (3) \end{cases}$$

$$\text{RHS} \Rightarrow (1) + (2) = (3) ; \quad \text{LHS} \Rightarrow (1) + (2) \neq (3)$$

- **Case 2:** three plane have a whole line in common.

$$\begin{cases} u + v + w = 2 \cdots (1) \\ 2u + 3w = 5 \cdots (2) \\ 3u + v + 4w = 7 \cdots (3) \end{cases}$$

$$\text{RHS} \Rightarrow (1) + (2) = (3) ; \quad \text{LHS} \Rightarrow (1) + (2) = (3)$$

- **Case 4:** three parallel

(2) Column Picture:

$$u \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + v \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + w \begin{pmatrix} 1 \\ 3 \\ 4 \end{pmatrix} = b$$

In the case above, three vectors are linear combination to each other, i.e. three vectors share the same plane.

**Lemma 1.2.3 (Singular case).** If the three vectors are linear combination to each other (three vector share a common plane), it must be **singular case**.

- If  $b = \begin{pmatrix} 2 \\ 5 \\ 7 \end{pmatrix}$ , which is on the plane  $\Rightarrow$  too many solution to produce  $b$ .
- If  $b = \begin{pmatrix} 2 \\ 5 \\ 6 \end{pmatrix}$ , which is not on the plane  $\Rightarrow$  no solution.

### 1.2.2 Fundamental Linear Algebra Theorem (Geometry form)

**Theorem 1.2.2 (Fundamental LA Theorem).** Consider a linear system

$$Ax = b, \quad A \in \mathbb{R}^{m \times n}, \quad x \in \mathbb{R}^n, \quad b \in \mathbb{R}^m.$$

If the  $n$  hyperplanes have no only one intersection or infinitely many points, then the  $n$  columns lie in the same plane. (consistency of *row picture* and *column picture*)

**Notation.** Logic notation:

- If ..., then :  $\Rightarrow$
- If and only if :  $\Leftrightarrow$

## Lecture 2

### 1.3 An Example of Gaussian Elimination

9 Sep. 13:20

**Example.** Here is a linear equation.

$$\begin{cases} 2u + v + w &= 5 \\ 4u - 6v &= -2 \\ -2u + 7v + 2w &= 9 \end{cases}$$

$$\left( \begin{array}{cccc} 2 & 1 & 1 & 5 \\ 4 & -6 & 0 & -2 \\ -2 & 7 & 2 & 9 \end{array} \right) \Rightarrow \left( \begin{array}{cccc} 2 & 1 & 1 & 5 \\ 0 & -8 & -2 & -12 \\ 0 & 8 & 3 & 14 \end{array} \right) \Rightarrow \left( \begin{array}{cccc} 2 & 1 & 1 & 5 \\ 0 & -8 & -2 & -12 \\ 0 & 0 & 1 & 2 \end{array} \right) \text{ "pivot"}$$

Then we get  $w = 2$ , we can plug in the equation i.e.

$$\begin{cases} 2u + v + 1w = 5 \\ -8v - 2w = -12 \\ w = 2 \end{cases} \Rightarrow \text{Forward Elimination}$$

Then we substitute into 2nd, 1st equation to get  $v = 1$  and  $u = 1 \Rightarrow \text{Backend Elimination}$

**Note.** By definition, **pivots cannot be zero!**

**Question.** Under what circumstances could the elimination process break down?

**Answer.** Here are some situations.

- Something **must** go wrong in the singular case.
- Something **might** go wrong in the nonsingular case.

A zero appears in a pivot position!

If in the process, there are nonzero pivots, then there's only one solution.

⊗

**Example.**

$$\begin{pmatrix} 2 & 1 & 1 & 5 \\ 4 & -6 & 0 & -2 \\ -2 & 7 & 2 & 9 \end{pmatrix}$$

- (1) If  $a_{11} = 0 \implies$  nonsingular
- (2) If  $a_{22} = 0 \implies$  nonsingular
- (3) If  $a_{33} = 1 \implies$  singular

**Question.** How many separate arithmetical operations does elimination require for  $n$  equations in  $n$  unknowns?

**Answer.** For a single operation.

a single operation = each division & each multiplication-subtraction

⊗

- **FE:**

$$\begin{array}{ccccccccc} x & x & \cdots & x & = & x \\ \vdots & \vdots & & & & \vdots \\ x & x & \cdots & x & = & x \\ & & & \underbrace{x \quad x \quad \cdots \quad x}_{n} & & & \end{array}$$

$$n(n-1) + (n-1)(n-2) + \cdots + (1^2 - 1) = \frac{n^3 - n}{3} \sim \frac{n^3}{3} \text{ steps}$$

- **RHS:**

$$(n-1) + (n-2) + \cdots + 1 = \frac{n(n-1)}{2} \sim \frac{n^2}{2} \text{ steps}$$

- **BF:**

$$1 + 2 + \cdots + n = \frac{n(n+1)}{2} \sim \frac{n^2}{2} \text{ steps}$$

## 1.4 Matrix Notation and Matrix Multiplication

$$\begin{cases} 2u + 4v - 2w = 2 \\ 4u + 9v - 3w = 8 \\ -2u - 3v + 7w = 10 \end{cases} \implies u \begin{pmatrix} 2 \\ 4 \\ -2 \end{pmatrix} + v \begin{pmatrix} 4 \\ 9 \\ -3 \end{pmatrix} + w \begin{pmatrix} -2 \\ 7 \\ 10 \end{pmatrix} = \begin{pmatrix} 2 \\ 8 \\ 10 \end{pmatrix}$$

We can rewrite it in the below form.

$$A = \begin{pmatrix} 2 & 4 & -2 \\ 4 & 9 & -3 \\ -2 & -3 & 7 \end{pmatrix}_{3 \times 3}, \quad x = \begin{pmatrix} u \\ v \\ w \end{pmatrix}_{3 \times 1}, \quad b = \begin{pmatrix} 2 \\ 8 \\ 10 \end{pmatrix}_{3 \times 1} \implies x = \begin{pmatrix} -1 \\ 2 \\ 2 \end{pmatrix}_{3 \times 1}$$

coefficient matrix      unknowns      RHS      solution

$$Ax = b$$

**Definition 1.4.1.** An  $m \times n$  matrix,  $A_{m \times n}$  over  $\mathbb{R}$ , is an array with  $m$  rows and  $n$  columns of real numbers, which can be written as

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \text{ where } a_{ij} \in \mathbb{R}, \quad \begin{cases} i : \text{index of row} \\ j : \text{index of column} \end{cases}$$

- $[m \times n]$  is called the **dimensions (size)** of  $A \implies$  dimension of a  $(3 \times 5)$  is  $3 \times 5$
- $[a_{ij}]$  is called the **elements/entry/coefficient** of  $A$
- **Addition:**  $A = (a_{ij})_{m \times n}, \quad B = (b_{ij})_{m \times n}$

$$A + B = (a_{ij} + b_{ij})_{m \times n}$$

- **Multiplication:**  $A = (a_{ij})_{m \times n}, \quad B = (b_{ij})_{n \times r}$

$$AB = (c_{ij})_{m \times r}, \quad \text{where } c_{ij} = \sum_{k=1}^n a_{ik} b_{kj}$$

- **Scalar Multiplication:**

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

- 

$$A_{m \times n} x_{n \times 1} = b_{m \times 1}$$

In particular, if

$$A_{1 \times n} B_{n \times 1} = \mathbf{v} \cdot \mathbf{w} = ()_{1 \times 1}.$$

Then it's the **inner product** of vector  $\mathbf{v}$  and vector  $\mathbf{w}$

**Example.**

$$Ax = \begin{pmatrix} 2 & 4 & -2 \\ 4 & 9 & -3 \\ -2 & 3 & -7 \end{pmatrix} \begin{pmatrix} -1 \\ 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \cdot (-1) & 4 \cdot (2) & -2 \cdot (2) \\ 4 \cdot (-1) & 9 \cdot (2) & -3 \cdot (2) \\ -2 \cdot (-1) & 3 \cdot (2) & -7 \cdot (2) \end{pmatrix} = \begin{pmatrix} 2 \\ 8 \\ 22 \end{pmatrix}$$

$$(-1) \begin{pmatrix} 2 \\ 4 \\ -2 \end{pmatrix} + 2 \begin{pmatrix} 4 \\ 9 \\ 3 \end{pmatrix} + 2 \begin{pmatrix} -2 \\ 3 \\ 7 \end{pmatrix}$$

- (1) by row: 3 inner product
- (2) by column: a linear combination of 3 columns of  $A$

**Example (1A).**  $Ax$  is a combination of columns of  $A$

$$\begin{aligned} A_{m \times n} x_{n \times 1} &= (A_1 | A_2 | \cdots | A_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \\ &= x_1(A_1) + x_2(A_2) + \cdots + x_n(A_n) = \left( \sum_{j=1}^n a_{ij} x_j \right)_{m \times 1} \end{aligned}$$

### 1.4.1 The Matrix Form of One Elimination Step

**Definition (1B).** Matrix form

**Definition 1.4.2.** zero matrix:

$$O = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$$

**Definition 1.4.3.** identity matrix:

$$I = \begin{pmatrix} 1 & \cdots & 0 \\ & \ddots & \\ 0 & \cdots & 1 \end{pmatrix} = I_n = I_{n \times n}; \quad \begin{cases} A_{m \times n} I_n = A_{m \times n} \\ A_{m \times n} = A_{m \times n} I_n \end{cases}$$

**Definition 1.4.4.** elementary matrix (elimination matrix):

$$E_{ij} = \begin{pmatrix} 1 & \cdots & 0 & \cdots & 0 \\ 0 & \ddots & & & \vdots \\ \vdots & & \ddots & & \vdots \\ \vdots & & -\ell & \ddots & 0 \\ 0 & \cdots & 0 & \cdots & 1 \end{pmatrix} \quad \begin{array}{l} \ell : \text{multiplier} \\ \text{i-th row} \\ \text{jth column} \end{array}$$

$$E_{ij} \cdot A = \begin{pmatrix} \cdots & -\ell & \cdots & 1 \end{pmatrix} \begin{pmatrix} & & & \end{pmatrix} \leftarrow \begin{array}{l} \text{i-th} \implies (\text{i-th row}) + (-\ell)(\text{j-th column}) \\ \text{j-th} \implies \text{create zero at } (i,j) \text{ position!} \end{array}$$

**Example.**

$$\begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}_{E_{21}} \begin{pmatrix} 2 & 4 & -2 \\ 4 & 9 & -3 \\ -2 & -3 & 7 \end{pmatrix}_A = \begin{pmatrix} 2 & 4 & -2 \\ 0 & 9 & -3 \\ -2 & -3 & 7 \end{pmatrix}_{EA}$$

**Note.** Here are two properties

1.  $Ax = b \implies E_{ij}Ax = E_{ij}b$
2.  $E_{ij}A \neq AE_{ij}$

### 1.4.2 Matrix Multiplication

- (1) The  $(i, j)$ -th entry of  $AB$  is the inner product of the **i-th** of  $A$  and the **j-th** of  $B$ .
- (2) Each column of  $AB$  is the product of a matrix  $A$  and a **column of  $B$**

$$\begin{aligned} \implies \text{column } j \text{ of } AB &= A \text{ times } \mathbf{j-th} \text{ of } B \\ &= \text{linear combination of columns of } A \\ &= b_{1j}A_{\bullet 1} + b_{2j}A_{\bullet 2} + \cdots + b_{nj}A_{\bullet n} \end{aligned}$$

**Example.**

$$\begin{pmatrix} 3 & 1 & 1 \\ 2 & 0 & -1 \end{pmatrix}_{A_{2 \times 3}} \begin{pmatrix} 5 & 0 & 1 \\ -1 & 0 & 1 \\ 2 & 1 & 2 \end{pmatrix}_{B_{3 \times 3}} = \begin{pmatrix} 16 & 1 & 1 \\ 8 & 0 & -1 \end{pmatrix}_{C_{2 \times 3}}$$

$$\text{1st column of } AB = \begin{pmatrix} 16 \\ 8 \end{pmatrix} = 5 \cdot \begin{pmatrix} 3 \\ 2 \end{pmatrix} + (-1) \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 2 \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

(3) Each row of  $AB$  is a product of a row of  $A$  and a matrix  $B$ .

$$\begin{aligned} \implies & \text{i-th row of } AB = \text{ of } A \text{ times } B. \\ & = \text{linear combination of rows of } B \\ & = a_{i1}B_{1\bullet} + a_{i2}B_{2\bullet} + \cdots + a_{in}B_{n\bullet}. \end{aligned}$$

**Theorem 1.4.1.** Let  $A, B$  and  $C$  be matrices (possibly rectangular). Assume that their dimension permit them to be added and multiplied in the following theorem

(1) The matrix multiplication is associative

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

(2) Matrix operations are distributive

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

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

(3) Matrix multiplication is noncommutative

$$AB \neq BA \quad \text{in general}$$

(4) Identity Matrix

$$A_{n \times n} I_n = I_n A_{n \times n} = A_{n \times n}$$

**Example.**

$$E = \begin{pmatrix} 1 & 0 & 0 \\ \boxed{-2} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad F = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \boxed{1} & 0 & 1 \end{pmatrix}, \quad G = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \boxed{-1} & 1 \end{pmatrix}$$

(1)

$$\underset{\substack{21 \\ 31}}{E} \underset{\substack{21 \\ 31}}{F} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \quad \boxed{=} \quad \underset{\substack{31 \\ 21}}{F} \underset{\substack{31 \\ 21}}{E} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

(2)

$$\underset{\substack{21 \\ 32}}{E} \underset{\substack{21 \\ 32}}{G} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix} \quad \boxed{\neq} \quad \underset{\substack{32 \\ 21}}{G} \underset{\substack{32 \\ 21}}{E}$$

(3)

$$\underset{\substack{32 \\ 31 \\ 21}}{G} \underset{\substack{21 \\ 31}}{F} \underset{\substack{21 \\ 31}}{E} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 3 & -1 & 1 \end{pmatrix} \quad \boxed{\neq} \quad \underset{\substack{21 \\ 31 \\ 32}}{E} \underset{\substack{21 \\ 31}}{F} \underset{\substack{32 \\ 21}}{G} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix}$$

"right order"

**Note.** The product of lower triangular matrices is a lower triangular matrix.

## Lecture 3

### 1.5 Triangular Factors and Row Exchanges

16 Sep. 13:20

$$\boxed{Ax = b}$$

$$\implies \boxed{LUx = b} \implies \begin{cases} Lc = b \\ Ux = c \end{cases}$$

**Example.**

$$Ax = \begin{pmatrix} 2 & 4 & -2 \\ 4 & 9 & -3 \\ -2 & -3 & 7 \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} 2 \\ 8 \\ 10 \end{pmatrix} = b$$

**Remark.**  $\ell$ : multipliers

$$E_{ij}(\ell) : (\text{i-th row}) + (-\ell)(\text{j-th column})$$

$$\begin{pmatrix} 2 & 4 & -2 & 2 \\ 4 & 9 & -3 & 8 \\ -2 & -3 & 7 & 10 \end{pmatrix} \xrightarrow[\substack{R_3+(1)R_1 \\ R_2+(-2)R_1}]{} \begin{pmatrix} 2 & 4 & -2 & 2 \\ 0_{21} & 1 & 1 & 4 \\ 0 & 1 & 5 & 12 \end{pmatrix} \xrightarrow[R_3+(-1)R_2]{} \begin{pmatrix} 2 & 4 & -2 & 2 \\ 0 & 1 & 1 & 4 \\ 0 & 0_{32} & 4 & 8 \end{pmatrix} \quad \text{pivot}$$

$$E_{21}(2) = E = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad E_{31}(-1) = F = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \quad E_{32}(1) = G = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}$$

i.e.

$$E_{21}E_{31}E_{32}Ax = \begin{pmatrix} 2 & 4 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix} = Ux = c = \begin{pmatrix} 2 \\ 4 \\ 8 \end{pmatrix} = E_{21}E_{31}E_{32}b$$

**Question.** How can we undo the steps of Gaussian Elimination?

$$\boxed{E^{-1}F^{-1}G^{-1}GFEA = A = \underbrace{E^{-1}F^{-1}G^{-1}}_{\text{factors of } A} \boxed{U} = LU} \quad \text{i.e. } A = \underbrace{LU}_{\text{factors of } A}$$

$$E^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -(-2) & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad F^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -(1) & 0 & 1 \end{pmatrix}, \quad G^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -(-1) & 1 \end{pmatrix}$$

$$E^{-1}F^{-1}G^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 1 & 1 \end{pmatrix} \implies \text{records everything that has been done so far}$$

### 1.5.1 Triangular Factorization

**Theorem 1.5.1.** If no exchanges are required, the original matrix  $A$  can be written as

$$A = LU$$

- The matrix  $L$  is lower triangular with 1's on the diagonal and the multipliers  $\ell_{ij}$  (taken from elimination) below the diagonal.
- The matrix  $U$  is the upper triangular matrix which appears after forward elimination and before back-substitution; its diagonal entries are the pivots.

**Example.**

$$\begin{pmatrix} 2 & -1 & -1 \\ 0 & -4 & 2 \\ 6 & -3 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & -1 & -1 \\ 0 & -4 & 2 \\ 0 & 0 & 4 \end{pmatrix} \Rightarrow \text{提出2}$$

$$= \begin{pmatrix} 2 & 0 & 0 \\ 0 & -4 & 0 \\ 6 & 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & -1/2 & -1/2 \\ 0 & 1 & -1/2 \\ 0 & 0 & 1 \end{pmatrix}$$

**Question.**

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

"triangular matrix" 有三條對角線

**Answer.**

⊗

### 1.5.2 One Linear System = Two Triangular Systems

$$Ax = b \implies Ux = c \text{ & } Lc = b \implies A = LU$$

**Remark.** The  $LU$  form is unsystematic in one aspect.  $U$  has pivots along its diagonal where  $L$  always has 1's.

We can rewrite  $U$  as

$$U = \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{22} & & u_{2n} \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & u_{nn} \end{pmatrix} = \begin{pmatrix} \color{red}{u_{11}} & 0 & \cdots & 0 \\ 0 & \color{teal}{u_{22}} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & u_{nn} \end{pmatrix} \begin{pmatrix} 1 & \color{red}{u_{12}/u_{11}} & \cdots & \color{red}{u_{1n}/u_{11}} \\ 0 & 1 & & \color{teal}{u_{2n}/u_{22}} \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

**Example.**

$$\begin{aligned} A &= \begin{pmatrix} 3 & 4 \\ 1 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 1/3 & 1 \end{pmatrix} \begin{pmatrix} \color{red}{3} & 4 \\ 0 & \color{red}{2/3} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1/3 & 1 \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & \color{red}{2/3} \end{pmatrix} \begin{pmatrix} 1 & 4/3 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

**Theorem 1.5.2.** If

$$A = L_1 D_1 U_1 \text{ and } A = L_2 D_2 U_2$$

then

$$L_1 = L_2, D_1 = D_2, U_1 = U_2$$

i.e. if  $A$  has  $LDU$  decomposition, then it **is** unique.

### 1.5.3 Row Exchange and Permutation Matrices

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad \text{(Permutation matrix } P_{ij}\text{)}$$

**Note.** Permutation matrix is also an elementary matrix.

**Example.** Here are some of the example:

1°

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} = \begin{pmatrix} 1 \\ 5 \\ 3 \end{pmatrix} \quad [R_2 \leftrightarrow R_3]$$

2°

$$PA = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 4 & 1 \\ 0 & 0 & 3 \\ 0 & 6 & 5 \end{pmatrix} = \begin{pmatrix} 2 & 4 & 1 \\ 0 & 6 & 5 \\ 0 & 0 & 3 \end{pmatrix} \quad [R_2 \leftrightarrow R_3]$$

3°

$$AP = \begin{pmatrix} 2 & 4 & 1 \\ 0 & 0 & 3 \\ 0 & 6 & 5 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 4 \\ 0 & 3 & 0 \\ 0 & 5 & 6 \end{pmatrix} \quad [C_2 \leftrightarrow C_3]$$

**Note.** For the permutation matrix:

1°  $PA$ : Performing row exchange of  $A$

2°  $AP$ : Performing column exchange of  $A$

3°  $\mathcal{P}Ax = Pb$ ; Should we permute the component of  $x = \begin{pmatrix} u \\ v \\ w \end{pmatrix}$  as well? **NONONONONO!!!**

**Example.**

$$A = \begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ d & e & f \end{pmatrix} \quad Ax = b$$

(1) if  $d = 0$ , the problem is incurable. The matrix is singular.

(2) if  $d \neq 0$ ,  $P_{13}A = \begin{pmatrix} d & e & f \\ 0 & 0 & c \\ 0 & a & b \end{pmatrix}$ ; if  $a \neq 0$ ,  $P_{23}P_{13}A = \begin{pmatrix} d & e & f \\ 0 & a & b \\ 0 & 0 & c \end{pmatrix}$

$$\left| \begin{array}{ccc|ccccc} & & & P_{23}P_{13} & & & P_{13}P_{23} & \\ & & & \neq & & & & \\ \text{row} & 1 & 3 & 3 & 1 & 1 & 2 & \\ & 2 \rightarrow 2 \rightarrow 1 & & & 2 \rightarrow 3 \rightarrow 3 & & & \\ & 3 & 1 & 2 & 3 & 2 & 1 & \end{array} \right|$$

**Theorem 1.5.3.** We separate into two cases:

- In the non singular case, there's a permutation matrix  $P$  that reorders the rows of  $A$  to avoid zeros in the pivot positions. In this case,
  - (1)  $Ax = b$  has a unique solution.
  - (2) It is found by elimination with row exchange
  - (3) With the rows reordered in advance,  $PA$  can be factored into  $LU$  ( $PA = LU$ )
- In singular case, no reordering can produce a full set of pivots.

**Example.**

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 3 \\ 2 & 5 & 8 \end{pmatrix} \xrightarrow[\text{row 3}]{\cancel{\ell_{31}} \cancel{\ell_{21}} \ell_{31}=1} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 2 \\ 0 & 3 & 6 \end{pmatrix} \xrightarrow{P_{23}} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 3 & 6 \\ 0 & 0 & 2 \end{pmatrix} = U$$

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} \langle \text{This is WRONG} \rangle = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

**To summarize:** A good code for Gaussian Elimination keeps a record of  $L, U$  and  $P$ . They allow the solution ( $Ax = b$ ) from two triangular systems. If the system  $Ax = b$  has a unique solution, they we say:

1° The system is nonsingular or

2° The matrix is nonsingular

Otherwise, it is singular.

## 1.6 Inverse and Transpose

**Definition 1.6.1.** An  $n \times n$  matrix  $A$  is invertible if  $\exists$  an  $n \times n$  matrix  $B \ni BA = I = AB$

**Theorem 1.6.1.** If  $A$  is invertible, then the matrix  $B$  satisfying  $AB = BA = I$  is unique!

**Proof.** Suppose  $\exists c \neq B \ni AC = CA = I$

$$B = BI = B(AC) = (BA)C = IC = C \text{ i.e } B = C$$

we call this matrix  $B$ , the inverse of  $A$ , and denoted as  $A^{-1}$  ■

**Note.** Not all  $n \times n$  matrices have inverse.

e.g.

1°

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

2° if  $Ax = \vec{0}$  has a nonzero solution, then  $A$  has no inverse!

$$x = A^{-1}(Ax) = A^{-1}\vec{0} = \vec{0} \quad (\rightarrow \leftarrow)$$

**Note.** The inverse of  $A^{-1}$  is  $A$  itself. i.e.  $(A^{-1})^{-1} = A$ .

**Note.** If  $A = (a)_{1 \times 1}$  and  $a \neq 0$ , then  $A^{-1} = (\frac{1}{a})$ . The inverse of  $\begin{pmatrix} a & b \\ c & d \end{pmatrix}_{2 \times 2}$  is

$$\frac{1}{\det(A)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \text{ if } \det(A) \neq 0$$

**Note.**

$$A = \begin{pmatrix} d_1 & \cdots & 0 \\ & \ddots & \\ 0 & \cdots & d_n \end{pmatrix} \quad d_i \neq 0, \forall i \quad \Rightarrow A^{-1} = \begin{pmatrix} 1/d_1 & \cdots & 0 \\ & \ddots & \\ 0 & \cdots & 1/d_n \end{pmatrix}$$

**Proposition 1.6.1.** If  $A$  and  $B$  are invertible, then

- $(AB)^{-1} = B^{-1}A^{-1}$
- $(A_1 A_2 \cdots A_n)^{-1} = A_n^{-1} \cdots A_2^{-1} A_1^{-1}$

### 1.6.1 The Calculation of $A^{-1}$ : Gaussian-Jordan Method

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

$$A_{n \times n} B_{n \times n} = I_n$$

$$\begin{aligned} &\Rightarrow A_{n \times n} (B_1 | B_2 | \cdots | B_n)_{n \times n} = (e_1 | e_2 | \cdots | e_n)_{n \times n} \\ &\Rightarrow (\textcolor{red}{AB}_1 | \textcolor{teal}{AB}_2 | \cdots | \textcolor{brown}{AB}_n)_{n \times n} = (\textcolor{red}{e}_1 | \textcolor{teal}{e}_2 | \cdots | \textcolor{brown}{e}_n)_{n \times n} \\ &\Rightarrow AB_1 = e_1; AB_2 = e_2; \cdots; AB_n = e_n \quad \rightarrow \quad n \text{ linear systems: } Ax = b \end{aligned}$$

**Definition 1.6.2 (Gaussian-Jordan Method).** Instead of stopping at  $U$  and switching to back substitution, it continues by subtracting multipliers of a row from the rows above till it reaches a diagonal matrix. Then we divide each row by corresponding pivot.

$$(A | I) \xrightarrow[\textcolor{red}{LU}]{\times L^{-1}} (U | \textcolor{red}{L}^{-1}) \xrightarrow{\times U^{-1}} (I | \textcolor{red}{A}^{-1})$$

$$\left( \begin{array}{ccc|ccc} 2 & -1 & 0 & 1 & 0 & 0 \\ -1 & 2 & -1 & 0 & 1 & 0 \\ 0 & -1 & 2 & 0 & 0 & 1 \end{array} \right) \rightarrow \left( \begin{array}{ccc|ccc} \boxed{2} & -1 & 0 & 1 & 0 & 0 \\ 0 & \boxed{3/2} & -1 & 1/2 & 1 & 0 \\ 0 & 0 & \boxed{4/3} & 1/3 & 2/3 & 1 \end{array} \right)$$

$$\rightarrow \left( \begin{array}{ccc|ccc} \boxed{2} & -1 & 0 & 1 & 0 & 0 \\ 0 & \boxed{3/2} & -1 & 1/2 & 1 & 0 \\ 0 & 0 & \boxed{4/3} & 1/3 & 2/3 & 1 \end{array} \right) \rightarrow \left( \begin{array}{ccc|ccc} 1 & 0 & 0 & 3/4 & 1/2 & 1/4 \\ 0 & 1 & 0 & 1/2 & 1 & 1/2 \\ 0 & 0 & 1 & 1/4 & 1/2 & 3/4 \end{array} \right)$$

$$A^{-1} = \begin{pmatrix} 3/4 & 1/2 & 1/4 \\ 1/2 & 1 & 1/2 \\ 1/4 & 1/2 & 3/4 \end{pmatrix}$$

### 1.6.2 Invertible = Nonsingular

**Question.** What kind of matrices are invertible?

**Answer.** Here are the example:

1° nonzero pivot [Ch1 Ch4](#)

2° nonzero determinants [Ch4](#)

3° independent columns (rows) [Ch2](#)

4° nonzero eigenvalues [Ch5](#)

which will in the whole course

⊗

Suppose a matrix  $A$  has full set of nonzero pivots. By definition,  $A$  is nonsingular and the  $n$  systems

$$Ax_1 = e_1, Ax_2 = e_2, \dots, Ax_n = e_n$$

can be solved by elimination or Gaussian-Jordan Method.

Row exchanges maybe necessary, but the columns of  $A^{-1}$  are uniquely determined.

$$Ax = b \quad PAx = Pb$$

$$PAx_i = Pe_i$$

$$\{Pe_1, Pe_2, \dots, Pe_n\} = \{e_1, e_2, \dots, e_n\}$$

**Note.** Compute  $A^{-1}$ :

1°  $A(x_1 | \dots | x_n) = I = (e_1 | \dots | e_n) \iff Ax_i = e_i, i = 1 \dots n$

2° Gauss-Jordan Method:  $(A | I) \rightarrow (I | A^{-1})$

**Question.** We have found a matrix  $A^{-1} \ni AA^{-1} = I$ . But is  $A^{-1}A = I$

**Answer.** We can do this by recall.

**As previously seen.** Recall that every Gauss-Jordan step is a multiplication of matrices on the left. There are three types of elementary matrices:

1°  $\boxed{E_{ij}(\ell)}$  : to subtract a multiple  $\ell$  of  $j$  row from  $i$  row.

2°  $\boxed{P_{ij}}$  : to exchange row  $i$  and  $j$

3°  $\boxed{D_i(d)}$  : to multiply row  $i$  by  $d$  i.e.  $D_i(d) = \begin{pmatrix} 1 & & & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & d & \vdots \\ \vdots & & \ddots & 0 \\ 0 & & & 1 \end{pmatrix}$  →  $i$ th row

$$\begin{pmatrix} d_1 & & 0 \\ 1 & & \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \begin{pmatrix} 1 & & 0 \\ & d_2 & \\ & & \ddots \\ 0 & & 1 \end{pmatrix} \cdots \begin{pmatrix} 1 & & 0 \\ & 1 & \\ & & \ddots \\ 0 & & d_n \end{pmatrix} = \begin{pmatrix} d_1 & & 0 \\ d_2 & & \\ & \ddots & \\ 0 & & d_n \end{pmatrix}$$

$$\implies \text{DEEPEE}A = I \implies A^{-1}A = I \therefore \text{we have a left inverse!}$$

These are the operation of  $A^{-1}$

(\*)

**Theorem 1.6.2.** For nonsingular and invertible:

- Every nonsingular matrix is invertible.
- Every invertible matrix is nonsingular.

**Theorem 1.6.3.** A square matrix is invertible  $\iff$  it is nonsingular

## Lecture 4

### 1.7 Transpose $A^T$

23 Sep. 13:20

**Proposition 1.7.1.** Here are the proposition of transpose

- $(A + B)^T = A^T + B^T$
- $(A^T)^T = A$
- $(AB)^T = B^T A^T$
- $(A^{-1})^T = (A^T)^{-1}$

**Proof.** Here is the proof

$$1^\circ ((A + B)^T)_{ij} = (A + B)_{ji} = A_{ji} + B_{ji} = (A^T + B^T)_{ij}$$

$$2^\circ ((AB)^T)_{ij} = (AB)_{ji} = \boxed{\sum_{k=1}^n a_{j\ell} b_{ki}} \quad (B^T A^T)_{ij} = \sum_{\ell=1}^n b_{i\ell}^T a_{\ell j}^T = \sum_{\ell=1}^n b_{\ell i} a_{j\ell} = \boxed{\sum_{\ell=1}^n a_{j\ell} b_{\ell i}}$$

3°

■

**Definition 1.7.1.** A symmetric matrix is a matrix which equals its own transpose. i.e.  $A = A^T$

**Example.**

$$\begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix} \text{ YES } \begin{pmatrix} 5 & 4 \\ 1 & 5 \end{pmatrix} \text{ NO } \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \text{ YES}$$

**Note.** A symmetric matrix is **not necessarily** invertible. If it is invertible, then its inverse is symmetric.

**Theorem 1.7.1.** If  $A$  is symmetric and if  $A$  can be factored as  $LDU$ , then  $A = LDU^T$

**Proof.** Here is the proof.

$$1^\circ \quad A = A^T, A = LDU \Rightarrow A^T = (LDU)^T = U^T D^T L^T = A = LDU$$

2° By theorem 1.5.2, the theorem is correct.

$LDU$  is unique if they exist. ■

# Chapter 2

## Vector Spaces and Linear Equation

### 2.1 Vector Spaces and Subspace

To answer the basic questions about the  $\underset{1^{\circ}}{\text{existence}}$  and  $\underset{2^{\circ}}{\text{uniqueness}}$  of the solution of  $Ax = b$ , we need the concept of vector space.

$$\text{Field} \implies \text{Vector Space} \implies \text{Solution of } Ax = b$$

**Definition 2.1.1 (Field).** Let  $F$  be a set with two operations "+" and " $\cdot$ " i.e.

$$+ : F \times F \longrightarrow F$$

$$\cdot : F \times F \longrightarrow F$$

and  $+$ ,  $\cdot$  are well-defined functions. If the system  $(F, +, \cdot)$  satisfies the following conditions, the  $F$  is called a **Field**.

For  $a, b, c \in F$

$$(1) (a + b) + c = a + (b + c)$$

$$(2) a + b = b + a$$

$$(3) \exists 0 \in F \ni a + 0 = 0 + a = a \quad \text{單位元素 (1st operation)}$$

$$(4) \forall a \in F, \exists (-a) \in F \ni a + (-a) = 0 \quad \text{反元素 (1st operation)}$$

$$(5) (a \cdot b) \cdot c = a \cdot (b \cdot c)$$

$$(6) a \cdot b = b \cdot a$$

$$(7) \exists 1 \in F \ni a \cdot 1 = 1 \cdot a = a \quad \text{單位元素 (2nd operation)}$$

$$(8) \forall a \neq 0 \in F, \exists a^{-1} \in F \ni a \cdot a^{-1} = a^{-1} \cdot a = 1 \quad \text{反元素 (2nd operation)}$$

$$(9) a \cdot (b + c) = ab + ac \quad \text{Distribution Law}$$

**Example.**

$$\begin{array}{lll} \mathbb{R}_{(\text{real})} & (\text{YES}) & \mathbb{Q}_{(\text{rational})} & (\text{YES}) & \mathbb{Z}_{(\text{integer})} & (\text{NO}) & \mathbb{C}_{(\text{complex})} & (\text{YES}) & \mathbb{N} & (\text{NO}) \end{array}$$

**Definition 2.1.2** (vector space). Let  $V$  be a set and  $F$  be a field.  $V$  is a vector space over  $F$  if  $\underset{1^{\circ}}{\text{addition}}$  and  $\underset{2^{\circ}}{\text{multiplication by scalar}}$  are defined on  $V$  and they satisfy.

$$+: V \times V \longrightarrow V$$

$$\cdot : F \times V \longrightarrow V$$

- (A1) addition is associated
- (A2) addition is commutative
- (A3)  $\exists$  zero vector  $\in V \ni 0 + v = v + 0, \forall v \in V$
- (A4)  $\forall v \in V, \exists (-v) \in V \ni (-v) + v = 0$
- (M1)  $1 \cdot v = v, v \in V, 1 \in F$
- (M2)  $(\lambda\mu) \cdot v = \lambda(\mu v) v \in V, \lambda, \mu \in F$
- (M3)  $\lambda(v_1 + v_2) = \lambda v_1 + \lambda v_2 v_1, v_2 \in V, \lambda \in F$
- (M4)  $(\lambda + \mu)v = \lambda v + \mu v v \in V, \lambda, \mu \in F$

### 2.1.1 Algebraic Rules of Vector Algebra

**Question.**  $n \in \mathbb{N}, \mathbb{R}^n / \mathbb{R}$  ( $\mathbb{R}^n$  over  $\mathbb{R}$ ) is a vector space?

**Answer.** YES \*

**Example.**

$$\mathbb{C}^n / \mathbb{C}, \mathbb{C}^n / \mathbb{R}, \mathbb{R} / \mathbb{R}$$

**Question.**  $M_{2 \times 2}(\mathbb{R}) / \mathbb{R}$  is a vector space?

$$M_{2 \times 2}(\mathbb{R}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a, b, c, d \in \mathbb{R} \right\}$$

**Answer.** YES \*

**Question.**  $V$  is a vector space?

$$V = \{\text{all } 3 \times 3 \text{ symmetric matrices over } \mathbb{R}\}$$

**Answer.** YES \*

**Question.**  $\mathbb{R}^\infty / \mathbb{R}, \mathbb{R}^\infty = \{(a_1, a_2, \dots) \mid a_i \in \mathbb{R}\}$

**Answer.** YES \*

**Question.** Let  $V = \{f \mid f \text{ is a real-valued function defined on } [0, 1]\}$  define  $(rf)(x) = r \cdot f(x)$ ,  $r \in \mathbb{R}$

**Answer. YES**

(zero vector) = (zero function)

i.e.  $f(x) = 0, \forall x \in [0, 1]$

⊗

**Question.**  $V = \{\text{all positive } \mathbb{R}\}$

$$\begin{cases} x+y &= xy \\ c \cdot x &= x^c \end{cases}, \quad \text{is } V \text{ a v.s. over } \mathbb{R}$$

**Answer. YES**

1° (A1)  $(x + y) + z = x + (y + z)$

2° (A2)  $(x + y) = xy = yx = (y + x)$

3° (A3) zero vector:  $x + 1 = x$

4° (A4)  $x + \frac{1}{x} = \text{zero vector} = 1$

5° (M3)  $\lambda(x + y) = (x + y)^\lambda = (xy)^\lambda = x^\lambda y^\lambda = (\lambda x)(\lambda y) = \lambda x + \lambda y$

6° (M4)  $(\lambda + \mu) \cdot x = x^{(\lambda+\mu)} = x^\lambda \cdot x^\mu = \lambda x \cdot \mu x = \lambda x + \mu x$

All conditions apply.

⊗

## 2.1.2 subspace

**Definition 2.1.3 (subspace).** A **subspace  $W$**  of a vector space  $(V, +, \cdot)$  over  $F$  is a nonempty subset of  $V \ni (W, +, \cdot)$  itself is a vector space over  $F$ .  $W$  is a subspace of  $V$  over  $F$  if and only if  $W$  is closed under addition and scalar multiplication.

**Question.** Does the zero vector belong to subspace?

**Answer. YES**

$W = \{\text{zero vector}\}$  is the smallest possible vector space. (\*)

**Remark.** If  $W_1$  and  $W_2$  are subspaces of  $V$  over  $F$ . Then  $W_1 \cap W_2 \neq \emptyset$

**Note.** If  $W$  is a subspace of  $V/F$ , then we use notation  **$W \leq V$** .

**Question.**  $V = \mathbb{R}^2/\mathbb{R}$  ( $xy$ -plane), What are the subspace of  $V$ ?

**Answer.** Here are all subspace of  $V$

- (i) origin (one point)
- (ii)  $\mathbb{R}^2/\mathbb{R} \leq V$
- (iii) all lines through origin
- (iv) 2nd-quadrant (no zero)

There are much more example. (\*)

**Question.**  $V = M_{n \times n}(\mathbb{R})/\mathbb{R}$

$$S = \{n \times n\text{symmetric matrix}\}$$

$$U = \{n \times n\text{upper triangular matrix}\}$$

$$L = \{n \times n\text{lower triangular matrix}\}$$

**Answer. YES, YES, YES** (\*)

**Theorem 2.1.1 ()**. Let  $V$  be a vector space over  $F$ . A nonempty subset  $W$  of  $V$  is a subspace of  $V$ , if and only if for each pair  $x, y \in W$  and  $\alpha \in F$ :

1° The zero vector  $\in W$ .

2°  $\alpha x + y \in W$

### 2.1.3 Column Space of $A$

**Example.**

$$A_{m \times n} x_{n \times 1} = b_{m \times 1}$$

The first concern is to find all attainable r.h.s. vector  $b$ . For example:

$$\begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \underline{u} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \underline{v} \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$$

**Theorem 2.1.2.** The system is solvable if and only if the vector  $b$  can be expressed as a combination of columns of  $A$

**Note.** The columns of  $A_{m \times n}$  are vectors in  $\mathbb{R}^m$ , the rows of  $A_{m \times n}$  are vectors in  $\mathbb{R}^n$ .

**Example.** Let  $\mathcal{C}(A) = \{\text{all combinations of columns of } A\}$ . Then,  $\mathcal{C}(A)$  is a subspace of  $\mathbb{R}^m/\mathbb{R}$ .

**Proof.** If  $b$  and  $b' \in \mathcal{C}(A)$ ,  $\exists x, x' \ni Ax = b \& Ax' = b'$

$$\forall \alpha \in \mathbb{R}, \quad A(\alpha x + x') = A(\alpha x) + A(x') = A\alpha x + Ax' = \alpha b + b' \in \mathcal{C}(A)$$

$$\implies \mathcal{C}(A) \leq \mathbb{R}^m/\mathbb{R}$$

**Definition 2.1.4.**  $\mathcal{C}(A)$  is called the **column space** of  $A$ . Thus if  $b \in \mathcal{C}(A)$ , then  $Ax = b$  is solvable.

- $A_{m \times n} = 0 \implies \mathcal{C}(A) = 0_{m \times 1}$
- $A_{m \times n} = I_m \implies \mathcal{C}(A) = \mathbb{R}^m$

## Lecture 5

### 2.1.4 Nullspace of $A$

30 Sep. 13:20

**Definition 2.1.5.** Let  $\mathcal{N}(A) = \{x \in \mathbb{R}^n \mid Ax = 0\}$ , then  $\mathcal{N}(A) \leq \mathbb{R}^n/\mathbb{R}$ . Then  $\mathcal{N}(A)$  is called the **null space** of  $A$ .

**Proof.** We proof it with the Theorem 2.1.1

- zero vector is in the  $\mathcal{N}(A)$
- $x, x' \in \mathcal{N}(A) \Rightarrow Ax = 0, Ax' = 0$

$$A(x + x') = Ax + Ax' = 0 + 0 = 0 \Rightarrow x + x' \in \mathcal{N}(A)$$

$$A(\alpha x) = \alpha Ax = \alpha \cdot 0 = 0 \Rightarrow \alpha x \in \mathcal{N}(A), \forall \alpha \in \mathbb{R} \quad \therefore \mathcal{N}(A) \leq \mathbb{R}^n/\mathbb{R}$$

**Note.** The system  $Ax = 0$  is called a homogeneous equation. (齊次)

**Remark.** The solution set of  $Ax = b$  is NOT a subspace of  $\mathbb{R}^n / \mathbb{R}$

$$x, x' \rightarrow Ax = b, Ax' = b$$

$$A(x + x') = Ax + Ax' = 2b \neq b$$

**Example.**

$$\begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \mathcal{N}(A) = \left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

**Example.**

$$\begin{pmatrix} 1 & 4 & 5 \\ 2 & 5 & 7 \\ 3 & 6 & 9 \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \mathcal{N}(A) = \left\{ \begin{pmatrix} t \\ t \\ -t \end{pmatrix}, t \in (-\infty, \infty) \right\}$$

$$\begin{aligned} \mathcal{C}(A) &= \{\text{all combinations of columns of } A\} \\ &= \text{column space of } A \leq \mathbb{R}^m / \mathbb{R} \end{aligned}$$

$$\begin{aligned} \mathcal{N}(A) &= \{x \in \mathbb{R}^n \mid Ax = 0\} \\ &= \text{null space of } A \leq \mathbb{R}^n / \mathbb{R} \end{aligned}$$

## 2.2 The Solution of $m$ Equations in $n$ Unknowns

For  $ax = b$ ,  $a, b, x \in \mathbb{R}$

- (i) if  $a \neq 0 \Rightarrow x = \frac{b}{a}$ , unique
- (ii) if  $a = 0, b = 0 \Rightarrow$  infinitely many solutions.
- (iii) if  $a = 0, b \neq 0 \Rightarrow$  no  $x$  exists.

Now, consider  $Ax = b$ , if  $A$  is a square, then (i), (ii), (iii) may occur.

- (i)  $A^{-1}$  exists  $\rightarrow x = A^{-1}b$ , unique
- (ii)  $A$  is singular (undetermined case)
- (iii) inconsistent case.

With a rectangular matrix  $A$ ,  $x = A^{-1}b$  will never happen!

**Definition.** Here is the definition of two similar jargon.

**Definition 2.2.1 (row echelon matrix).** An  $m \times n$  matrix  $R$  is called a **row echelon matrix** if

- (i) the nonzero rows come first and the pivots are the first nonzero entries in those rows.
- (ii) below each pivot is a column of zeros
- (iii) each pivot lies to the right of the pivot in the row above.

e.g.

$$\begin{pmatrix} (*) & (*) & (*) & (*) & (*) \\ 0 & (*) & (*) & (*) & 0 \\ 0 & 0 & 0 & (*) & (*) \end{pmatrix}$$

**Definition 2.2.2 (row-reduced echelon matrix).** An  $m \times n$  matrix  $R$  is called a **row-reduced echelon matrix** if

- (i) the nonzero rows come first and the pivots are the first nonzero entries in those rows; **pivots are normalized to be 1.**
- (ii) **Above & Below** each pivot is a column of zeros
- (iii) each pivot lies to the right of the pivot in the row above.

e.g.

$$\begin{pmatrix} [1] & 0 & (*) & 0 & (*) \\ 0 & [1] & (*) & 0 & (*) \\ 0 & 0 & 0 & [1] & (*) \end{pmatrix}$$

**Theorem 2.2.1.** To any  $m \times n$  matrix  $A$ , there exists a permutation matrix  $P$ , a lower triangular matrix  $L$  with unit diagonal and an  $m \times n$  echelon matrix  $U \ni PA = LU$

**OR**

Every  $m \times n$  matrix  $A$  is **row equivalent to** a row echelon matrix.

- Case 1. Homogeneous Case.  $b_{m \times 1} = 0$

$$Ax = 0$$

We call the component of  $x$ , which correspond to columns with pivots the **basic variables**; and these correspond to columns with pivots the **free variables**.

$$\begin{pmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} u \\ v \\ w \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{cases} \text{basic variables: } u, w \\ \text{free variables: } v, y \end{cases}$$

The basic variables are then expressed in terms of free variables.

$$\begin{cases} 3w + y = 0 \\ u + 3v + 3w + 2y = 0 \end{cases} \implies \begin{cases} w = -\frac{1}{3}y \\ u = -3v - y \end{cases}$$

$$x = \begin{pmatrix} u \\ v \\ w \\ y \end{pmatrix} = \begin{pmatrix} -3v - y \\ v \\ -\frac{1}{3}y \\ y \end{pmatrix} = v \begin{pmatrix} -3 \\ 1 \\ 0 \\ 0 \end{pmatrix} + y \begin{pmatrix} -1 \\ 0 \\ -\frac{1}{3} \\ 1 \end{pmatrix}$$

$\begin{pmatrix} -3 \\ 1 \\ 0 \\ 0 \end{pmatrix}$  is obtain from  $x$  by setting  $\begin{cases} v = 1 \\ y = 0 \end{cases}$

$\begin{pmatrix} 1 \\ 0 \\ -\frac{1}{3} \\ 0 \end{pmatrix}$  is obtain from  $x$  by setting  $\begin{cases} v = 0 \\ y = 1 \end{cases}$

**Theorem 2.2.2.** If a homogeneous system  $A_{m \times n}x = 0$  has more unknowns than equations ( $m < n$ ), it has a nontrivial solution.

$$(A_{m \times n}) \longrightarrow (A_{m \times n})^n$$

at most  $m$  pivot, at most  $m$  basic variables, at least  $(n - m)$  free variables.

**Note.** The nullspace is a subspace of the same dimension as the number of free variables.

- Case 2. Inhomogeneous Case:  $b \neq 0$

$$Ax = b \rightarrow Ux = c \text{ where } c = L^{-1}b$$

$$\begin{aligned} & \begin{pmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 5 \\ -1 & -3 & 3 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \\ \implies & \begin{pmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} u \\ v \\ w \\ y \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 - 2b_1 \\ b_3 - 2b_2 + 5b_1 \end{pmatrix} \quad \rightarrow \quad b_3 - 2b_2 + 5b_1 = 0 \end{aligned}$$

We know that  $Ax = b$  is solvable  $\Rightarrow b \in \mathcal{C}(A)$

– 1 & 3: basic variables

–  $\mathcal{C}(A) = \text{the set of combinations of } \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} \text{ & } \begin{pmatrix} 3 \\ 9 \\ 3 \end{pmatrix}$   
, which is also  $\left\{ \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \mid b_3 - 2b_2 + 5b_1 = 0 \right\} \perp \begin{pmatrix} 5 \\ -2 \\ 1 \end{pmatrix}$

### Example.

$$b = \begin{pmatrix} 1 \\ 5 \\ 5 \end{pmatrix}$$

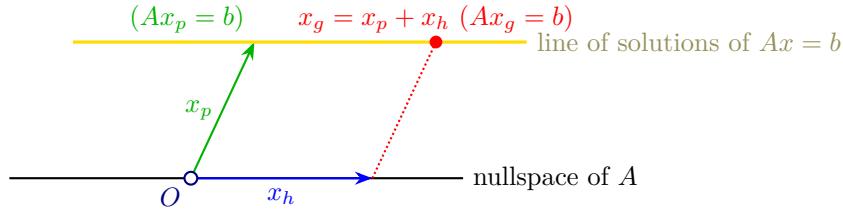
$$\begin{aligned} & \begin{pmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} u \\ v \\ w \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix} \quad \Rightarrow \quad \begin{cases} w = 1 - \frac{1}{3}y \\ u = -2 - 3v - y \end{cases} \\ & x = \begin{pmatrix} u \\ v \\ w \\ y \end{pmatrix} = \begin{pmatrix} -2 - 3v - y \\ v \\ 1 - \frac{1}{3}y \\ y \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \\ 1 \\ 0 \end{pmatrix} + v \begin{pmatrix} -3 \\ 1 \\ 0 \\ 0 \end{pmatrix} + y \begin{pmatrix} -1 \\ 0 \\ -\frac{1}{3} \\ 1 \end{pmatrix} \quad \text{shift} \quad \text{solution to } Ax=0 \text{ (nullspace)} \end{aligned}$$

Shift: particular solution to  $Ax = b$  (set all free variables to be zero)

$$x_{\text{general}} = x_{\text{particular}} + x_{\text{homogeneous}}; \quad x_g = x_p = x_h$$

Generally, the general solution is fills a two-dimensiona; surface (but NOT a subspace since it doesn't contain the zero vector (origin))

It is paralled to the Nullspace of  $A$



### 2.2.1 Steps to obtain the solution to $Ax = b$

- Reduce  $Ax = b$  to  $Ux = c$  to determine basic/free variables.
- Set all free variables to zero to find particular solution,  $x_p$
- set RHS = 0. Give each free variables 1 others 0, in terms, find the homogeneous sloution,  $x_h$

$$\implies x_g = x_p + x_h$$

**Definition 2.2.3 (rank).**  $A_{m \times n}$  if there are  $r$  pivots, there are  $r$  basic variables and  $n - r$  free variables. The number of pivots,  $r$ , is called the **rank** of the matrix.

**Theorem 2.2.3.** Suppose elimination reduce  $A_{m \times n}x = b$  to  $Ux = c$  and there are  $r$  pivots and the last  $(m - r)$  rows of  $U$  are zero. Then there is a solution only if last  $(m - r)$  elements of  $c$  are zeros.

- If  $r = m$ , there's always a solution. The general solution is the sum of particular solution and a homogeneous solution.
- If  $r = n$ , there are No free variables and the null space contains  $x = 0$  only. The number  $r$  is called the rank of  $A$ .

Two extreme case:  $A_{m \times n}x = b$

- If  $r = n \rightarrow$  No free variables  $\rightarrow \mathcal{N}(A) = \{x \in \mathbb{R}^n \mid Ax = 0\} = \{0\}$
- If  $r = m \rightarrow$  No zero rows in  $U \rightarrow \mathcal{C}(A) = \mathbb{R}^m \Rightarrow \exists$  solution for all  $b$

## 2.3 Linear Independence, Basis and Dimension

In the elimination process, we refer to the number,  $r$ , of pivots as the rank of  $A$ . This definition is purely computational rather than mathematical. We shall give a formal definition later.

Now we shall discuss the following four ideas:

- linear independence or dependence
- spanning** a subspace
- basis** for a subspace
- dimension** of a subspace

**Definition 2.3.1.** Let  $V$  be a vector space over  $F$ . A nonempty subset  $S$  of  $V$  is said to be linearly dependent if there exist distinct vectors  $v_1, v_2, \dots, v_n$  in  $S$  and scalar  $\alpha_1, \alpha_2, \dots, \alpha_n$  in  $F$ , not all of which are zero such that

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n = 0$$

A set which is not linearly dependent is called linearly independent. If  $S = \{v_1, v_2, \dots, v_n\}$  then we say that  $v_1, v_2, \dots, v_n$  are linearly dependent/independent.

## Lecture 6

**Remark (1).** To show that  $v_1, \dots, v_n$  are linearly independent. We verify if

14 Oct. 13:20

$$c_1 v_1 + c_2 v_2 + \dots + c_n v_n = 0 \text{ for some } c_i \in F$$

then  $c_i$  must be zero for all  $i$ .

**Example.** In  $\mathbb{R}^2$ , if  $v_1, v_2$  are not colinear(共線) then they are linearly independent.

$v_1 (\neq 0)$  and  $v_2 (\neq 0)$  are linearly dependent  $\iff v_1, v_2$  are on the same line

Any 3 vectors in  $\mathbb{R}^2$  are linearly dependent.

**Remark (2).** If  $v_1 = v_2$ , then the set  $\{v_1, \dots, v_n\}$  is linearly dependent.

$$\alpha v_1 + (-\alpha)v_2 = 0$$

**Remark (3).** Any set which contains a linearly dependent subset is linearly dependent.

**Remark (4).** Any subset of a linearly independent set is linearly independent.

**Remark (5).** Any set which contains 0 vector is linearly dependent.

**Example.**

$$A = \begin{pmatrix} 1 & 2 & 1 & -1 \\ 3 & 2 & -3 & 0 \\ -4 & -4 & 2 & 1 \\ -2 & 0 & -4 & 0 \end{pmatrix}$$

$$\begin{matrix} v_1 & v_2 & v_3 & v_4 \end{matrix}$$

The columns of  $A$  are linearly dependent.

$$c_1 v_1 + c_2 v_2 + c_3 v_3 + c_4 v_4 = 0$$

$$(v_1 \ v_2 \ v_3 \ v_4) \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} = 0 \implies 4v_1 + (-3)v_2 + 2v_3 + 0v_4 = 0$$

**Example.**

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & -1 & -1 & 1 \end{pmatrix}$$

The columns of  $A$  are linearly **independent**

**Note.** We showed that the nullspace of  $A$  is  $\{0\}$  only. That is exactly the same as saying the columns of  $A$  are linearly independent.

**Example.**

$$U = \begin{pmatrix} \boxed{1} & 3 & 3 & 2 \\ 0 & 0 & \boxed{3} & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

**Proposition 2.3.1 (2F).** The  $r$  nonzero rows of echelon matrix  $U$  are linearly independent, and so are  $r$  columns that contain pivots.

**Example.** In  $\mathbb{R}^n$ ,  $e_1, e_2, \dots, e_n$  are linearly **independent**.

**To summarize:** To check any set of vectors  $v_1, v_2, \dots, v_n (\in \mathbb{R}^n)$  are linearly independent.

Let  $A = (v_1 | v_2 | \dots | v_n)_{m \times n}$ , then solve  $Ax_{n \times 1} = 0$ .

1° if  $\exists$  solution  $x \neq 0$ , then  $v_i$ 's are linearly **dependent**.

2° if there are no free variables (i.e.  $\text{rank}(A) = n$ ), **nullspace = {0}** then  $v_i$ 's are linearly **independent**.

3° if  $\text{rank}(A) < n$ , then  $v_i$ 's are linearly **dependent**.

4° special case: if  $v_i \in \mathbb{R}^>$  and  $n > m$ , then  $v_i$ 's are linearly **dependent**.

**Proposition 2.3.2.** A set of  $n$  vectors in  $\mathbb{R}^m$  must be linearly dependent if  $n > m$ .

### 2.3.1 Spanning a Subspace

**Definition 2.3.2 (2H).** Let  $S$  be a subset vectors in  $V/F$ .

The subspace spanned by  $S$  is defined to be the intersection  $W$  of all subspaces of  $V$  which contain  $S$ .

When  $S$  is finite,  $S = \{v_1, \dots, v_n\}$ , we call  $W$  the subspace spanned by  $v_1, \dots, v_n$  and denoted as  $W = \langle v_1, \dots, v_n \rangle$  or  $W = \text{span}(S) = \langle S \rangle$ .

**Theorem 2.3.1.** [The subspace spanned by a nonempty subset  $S$ ] of a vector space  $V$  is [the set  $T$  of all linear combinations of vectors in  $S$ ].

**Proof.** We need to show  $W = T$ .

**Claim.**  $W = T$  if and only if  $W \subseteq T$  and  $T \subseteq W$ .

- Let  $W$  be the subspace spanned by  $S$ ,  $S \subseteq W$  ( $S$  不一定有包含 0 vector 所以不能用  $\leq$ ).  
So every linear combination of vectors in  $S$  is in  $W$ .  $\Rightarrow T \subseteq W$ .  
 $(\because W \text{ is a subspace which is a vector space})$
- on the other hand,  $T$  is a subspace containing  $S$ .  
 $(\because x, y \in T, \alpha \in F \Rightarrow \alpha x + y \in T)$

So,  $W \subseteq T$  by definition  $\Rightarrow W = T$ .

(Intersection of all subspaces containin  $S$ ) ■

**Example.**  $\mathcal{C}(A) =$  space spanned by columns of  $A$ .

**Example.**  $w_1 = (1, 0, 0)$ ,  $w_2 = (0, 1, 0)$ ,  $w_3 = (0, 0, 1)$ , span a space  $\mathbb{R}^3$ .

$w_1 = (1, 0, 0)$ ,  $w_2 = (0, 1, 0)$ ,  $w_3 = (-3, 0, 0)$ , span a plane  $\mathbb{R}^2$ .

**Note.** Spanning involves the **columns space**, independence involves the **null space**.

### 2.3.2 Basis

**Definition 2.3.3 (2I).** A basis for a vector space is a set of vectors that satisfies

- it is linearly independent AND
- it span the vector space

If the basis of  $V$  is finite, then  $V$  is finite-dimensional (f-dim).

**Remark (1).** There's one and **only one** way to write every  $v \in V$  as a linear combination of the basis elements.

**Remark (2).** In  $\mathbb{R}^n$ ,

$$e_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}_{n \times 1} \quad \begin{array}{l} \uparrow \\ i^{th} \\ \downarrow \end{array}$$

then  $\{e_1, \dots, e_n\}$  is a basis for  $\mathbb{R}^n$ . The basis is called the **standard basis**.

$$\forall x = (x_1, \dots, x_n) \in \mathbb{R}^n, \quad x = \sum_{i=1}^n x_i e_i$$

The standard basis is not the only basis for  $\mathbb{R}^n$ . In fact, there are infinitely many bases for  $\mathbb{R}^n$ . For any nonsingular matrix  $A_{n \times n}$ , the columns of  $A$  are the basis for  $\mathbb{R}^n$ .

**Example.**

$$A = \begin{pmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 5 \\ -1 & -3 & 3 & 0 \end{pmatrix}_{3 \times 4} \quad \rightarrow \quad U = \begin{pmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}_{3 \times 4}$$

The columns of  $U$  that contain pivots are a basis for  $\mathcal{C}(U)$ .

Note that  $\mathcal{C}(U)$  is generated by  $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$  and  $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ , which is a  $xy$ -plane within  $\mathbb{R}^3$ .

**Remark.**  $\mathcal{C}(U)$  is NOT same as  $\mathcal{C}(A)$ .

**Theorem 2.3.2 (2J).** Any two bases for  $V$  contain the same number of vectors. This number is called the *dimension* of  $V$ .

**Proof.** Suppose  $v_1, \dots, v_m$  and  $w_1, \dots, w_n$  are bases for  $V$ , and suppose  $m < n$ .

For  $j = 1, \dots, n$ ,

$$w_j = a_{1j}v_1 + \dots + a_{mj}v_m \quad \text{for some } a_{ij} \in F.$$

Let

$$w = [w_1, \dots, w_n] = VA = [v_1, \dots, v_m] \begin{bmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}.$$

The matrix  $A$  is  $m \times n$  with  $m < n$ . By Theorem 2C,  $\exists$  nontrivial  $C$  such that  $AC = 0$ .

$$VAC = WC = 0.$$

Hence the columns of  $W$  are linearly dependent. But the columns of  $W$  are basis elements, contradiction  $\Rightarrow m \not< n$ .

Similarly, we can show that  $n \not< m$ , so we conclude  $m = n$ . ■

**Theorem 2.3.3 (2L).** Any linearly independent set in a finite-dimensional vector space  $V$  can be extended to a basis. Any spanning set of  $V$  can be reduced to a basis.

**Proof.** Let  $v_1, \dots, v_k$  be linearly independent over  $F$ . Then  $\langle v_1, \dots, v_k \rangle \leq V$ .

If  $\langle v_1, \dots, v_k \rangle = V$ , then  $\langle v_1, \dots, v_k \rangle$  is a basis of  $V$ . Otherwise,  $\exists x \in V$  such that  $x \notin \langle v_1, \dots, v_k \rangle$ . Then  $x, v_1, \dots, v_k$  are linearly independent. If not,  $\exists c \neq 0$ , and  $\exists \alpha_1, \dots, \alpha_k$ , not all zero, such that

$$cx + \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k = 0.$$

$$\Rightarrow x = c^{-1} \alpha_1 v_1 + c^{-1} \alpha_2 v_2 + \cdots + c^{-1} \alpha_k v_k.$$

$$\Rightarrow x \in \langle v_1, \dots, v_k \rangle, \text{ contradiction.}$$

Then repeat the process, i.e. is  $\langle x, v_1, \dots, v_k \rangle = V$ ? Since  $V$  is finite-dimensional, the process will terminate after finite steps.

The 2nd half of the theorem can be proved similarly (exercise). ■

## 2.4 The Four Fundamental Subspaces

Usually there are two ways to describe a subspace

- (i) a set of vectors that span the space.  
(e.g. the column space of  $A_{m \times n}$ ,  $\mathcal{C}(A)$ )
- (ii) a list of constraints that imposed on a subspace.  
(e.g. the null space of  $A_{m \times n}$ ,  $\mathcal{N}(A) = \{x \mid Ax = 0\}$ )

Here we will discuss four fundamental subspaces associated to  $A_{m \times n}$

- (1) the **column space** of  $A$  denoted by  $\mathcal{C}(A)$
- (2) the **null space** of  $A$  denoted by  $\mathcal{N}(A)$
- (3) the **row space** of  $A$  the columns spaces of  $A^T$ , denoted by  $\mathcal{C}(A^T)$
- (4) the **left null space** of  $A$  denoted by  $\mathcal{N}(A^T)$ , i.e.  $\{y \mid A^T y = 0\}$ 
  - If  $A_{m \times n}$ , then  $\mathcal{C}(A), \mathcal{N}(A^T) \leq \mathbb{R}^m$  and  $\mathcal{N}(A), \mathcal{C}(A^T) \leq \mathbb{R}^n$ .

### 2.4.1 Row space $\mathcal{C}(A^T)$

The **row space** of  $A$  (the subspace spanned by the rows of  $A$ ),  $\mathcal{C}(A^T)$ . For an echelon matrix, its  $r$  nonzero rows are independent and its row space is ***r*-dimensional**.

**Proposition 2.4.1 (2M).** The row space of  $A$  has the same dimension  $r$  as the row space of echelon form  $U$  of  $A$ , and they have the same basis.

$$\mathcal{C}(A^T) = \mathcal{C}(U^T)$$

But in general,  $\mathcal{C}(A) \neq \mathcal{C}(U)$ .

## Lecture 7

### 2.4.2 Nullspace $\mathcal{N}(A)$

21 Oct. 13:20

The nullspace of  $A_{m \times n}$ ,  $\{x \mid Ax = 0\} = \{x \mid Ux = 0\}$

$\therefore$  The nullspace of  $A$  is the same as the nullspace of  $U$

**Proposition 2.4.2 (2N).** The nullspace  $\mathcal{N}(A)$  is of dimension  $n - r$

A basis of  $\mathcal{N}(A)$  can be constructed by reducing to  $Ux = 0$  which has  $n - r$  free variables corresponding to the columns of  $U$  that do not contain pivots. Let each free variable  $1$ , in turn, and others  $0$ , and solve  $Ux = 0$ . The  $n - r$  vectors produced in this manner will be a basis of  $\mathcal{N}(A)$ .

$$\dim(\mathcal{N}(A)) = n - r$$

The  $\mathcal{N}(A)$  is also called the **kernel of  $A$** ,  $\ker(A)$ , and its dimension is called the **nullity of  $A$** .

$$\ker(A) = \mathcal{N}(A)$$

### 2.4.3 Column space $\mathcal{C}(A)$

The  $\mathcal{R}$  in  $\mathcal{R}(A)$  stands for “range” which is consistent with the usual idea of range of  $f$

Let  $f(x) = A_{m \times n}x_{n \times 1}$ , the

- the domain of  $f$  is  $\mathbb{R}^n$
- the range of  $f$  is  $\{b \in \mathbb{R}^m \mid Ax = b\} = \mathcal{C}(A) = \mathcal{R}(A)$
- the kernel of  $f$  is  $\{x \in \mathbb{R}^n \mid Ax = f(x) = 0\} = \mathcal{N}(A) = \ker(A)$

If  $U$  is the echelon form of  $A$ ,  $\mathcal{C}(A) \neq \mathcal{C}(U)$ , but they have the same dimension. For  $U$ , the columns with pivots form a basis of  $\mathcal{C}(U)$ . Then, the corresponding columns in  $A$  form a basis of  $\mathcal{C}(A)$ . Since the two systems  $Ax = 0$ ,  $Ux = 0$  are equivalent and have the same solutions. A nontrivial solution  $x$  means a linear combination of columns of  $U$ , hence the same linear combination of columns of  $A$ .

So, if the set of columns of  $U$  is independent, then so are the corresponding **columns** of  $A$  and vice versa.

To find a basis of  $\mathcal{C}(A)$ , we pick those columns of  $A$ , which correspond to the columns of  $U$  with pivots.

**Proposition 2.4.3 (2O).** The dimension of the column space = rank  $r$ , which also equals the dimension of the row space.

$\therefore$  # of independent columns = # of independent rows =  $r$

or more formally,

$$\text{rank}(A) = r = \text{row rank} = \text{column rank}$$

#### 2.4.4 Left nullspace $\mathcal{N}(A^T)$

$$\underset{n \times m}{A^T} \underset{m \times 1}{y} = \underset{n \times 1}{0} = (\underset{1 \times m}{y^T} \underset{m \times n}{A})^T$$

$$(\# \text{ of basic variables}) + (\# \text{ of free variables}) = (\# \text{ of variables}) = n$$

$$\boxed{\dim(\mathcal{C}(A)) + \dim(\mathcal{N}(A)) = \# \text{ of columns of } A}$$

For  $A^T$ , which has  $m$  columns, the column space of  $A^T$  is the row space of  $A$  which has dimension  $\text{rank}(A)$ . So,

$$\dim(\mathcal{N}(A^T)) = m - \text{rank}(A)$$

i.e.

$$\boxed{\dim(\mathcal{C}(A^T)) + \dim(\mathcal{N}(A^T)) = \# \text{ of columns of } A^T}$$

**Proposition 2.4.4 (2P).** The left nullspace  $\mathcal{N}(A^T)$  is of dimension  $m - r$

The left nullspace contain the coefficients that make the rows of  $A$  combined to a zero vector (linear dependent).

To find  $y \ni y^T A = 0$

$$\text{Suppose that } PA = LU \longrightarrow \boxed{\underset{m \times m}{L^{-1}P} \underset{m \times n}{A} = \underset{m \times n}{U}}$$

The last  $m - r$  rows of  $L^{-1}P$  must be a basis for the left nullspace. ( $\therefore$  the last  $m - r$  rows of  $L^{-1}P$  are independent and  $\dim(\mathcal{N}(A^T))$  is  $m - r \rightarrow$  it is a basis of  $\mathcal{N}(A^T)$ )

**Theorem 2.4.1 (Fundamental Theorem of Linear Algebra).** Let  $A$  be an arbitrary  $m \times n$  matrix, then

$$\dim(\mathcal{C}(A)) = \dim(\mathcal{C}(A^T)) = \text{rank}(A)$$

$$\dim(\mathcal{N}(A)) = n - \text{rank}(A); \quad \dim(\mathcal{N}(A^T)) = m - \text{rank}(A)$$

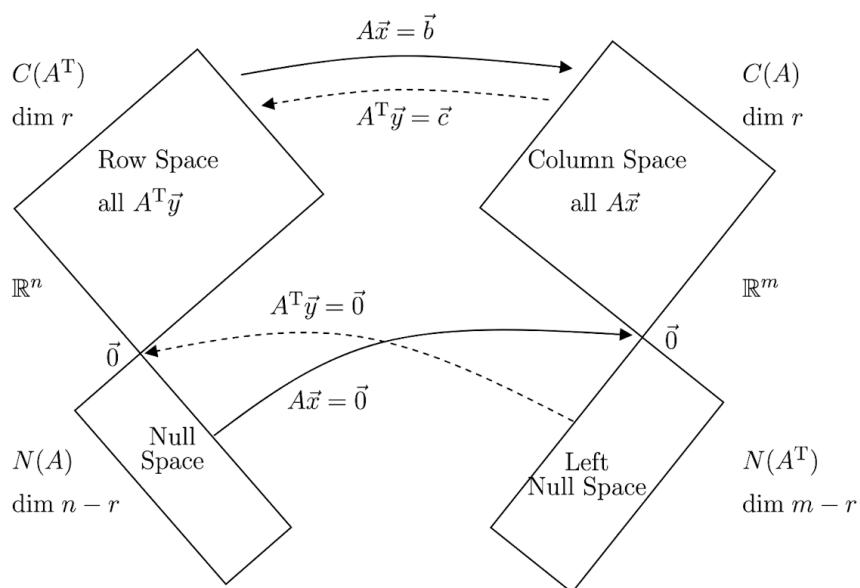


Figure 2.1: Fundamental Theorem of Linear Algebra

**Example.** Find out the basis for the four fundamental subspaces of the matrix

$$A = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 2 & 3 & 4 & 1 \\ 4 & 3 & 6 & 1 \end{pmatrix} \quad \rightarrow \quad U = \begin{pmatrix} \boxed{1} & 0 & 1 & 0 \\ 0 & \boxed{1} & 2/3 & 1/3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad r = 2$$

1°  $\mathcal{C}(A)$

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}, \begin{pmatrix} 0 \\ 3 \\ 3 \end{pmatrix} \right\} \quad \dim(\mathcal{C}(A)) = r = 2$$

2°  $\mathcal{N}(A)$

$$Ax = 0 \longrightarrow Ux = 0 \longrightarrow U \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{cases} x_1 + x_3 = 0 \\ x_2 + \frac{2}{3}x_3 + \frac{1}{3}x_4 = 0 \end{cases}$$

$$(a) \ x_3 = 1, x_4 = 0 \longrightarrow \begin{pmatrix} -1 \\ -2/3 \\ 1 \\ 0 \end{pmatrix} = v_2$$

$$(b) \ x_3 = 0, x_4 = 1 \longrightarrow \begin{pmatrix} 0 \\ -1/3 \\ 0 \\ 1 \end{pmatrix} = v_2$$

Hence,  $\mathcal{B} = \mathcal{N}(A)$  is  $\{v_1, v_2\}$  and

$$\dim(\mathcal{N}(A)) = n - r = 4 - 2 = 2$$

3°  $\mathcal{C}(A^T)$

$$U = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 2/3 & 1/3 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} S_1 \\ S_2 \\ 0 \end{pmatrix} \quad \rightarrow \quad \mathcal{B} = \{S_1^T, S_2^T\}, \quad \dim(\mathcal{C}(A^T)) = r = 2$$

4°  $\mathcal{N}(A^T) \longrightarrow \mathcal{N}(B)$

$$B = \begin{pmatrix} 1 & 2 & 4 \\ 0 & 3 & 3 \\ 1 & 4 & 6 \\ 0 & 1 & 1 \end{pmatrix} = A^T \quad \rightarrow \quad \begin{pmatrix} \boxed{1} & 2 & 4 \\ 0 & \boxed{1} & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \rightarrow \quad \begin{cases} y_1 + 2y_3 = 0 \\ y_2 + y_3 = 0 \end{cases}$$

$$z = 1 \longrightarrow \begin{pmatrix} -2 \\ -1 \\ 1 \end{pmatrix} \quad \therefore \mathcal{B} = \left\{ \begin{pmatrix} -2 \\ -1 \\ 1 \end{pmatrix} \right\}, \quad \dim(\mathcal{N}(A^T)) = m - r = 3 - 2 = 1$$

Check orthogonality

**Proposition 2.4.5 (2Q).** We can find the existence and uniqueness of solution of  $Ax = b$ .

- **Existence** of inverse:

The system  $Ax = b$  has at least one solution  $x$  for each  $b$  iff the columns span  $\mathbb{R}^m$  ( $r = m$ ). In this case,

$$\exists n \times m \text{ "right" inverse } C \ni AC = I$$

This is possible only if  $m \leq n$ .

- **Uniqueness** of inverse:

The system  $Ax = b$  has at most one solution  $x$  for each  $b$  iff the columns are independent ( $r = n$ ). In this case,

$$\exists n \times m \text{ "left" inverse } B \ni BA = I$$

This is possible only if  $m \geq n$ .

**Proof.** We separately prove the two parts.

- **Existence:**

$$Ax = b \text{ has a solution for each } b \Leftrightarrow b \in \mathcal{C}(A), \forall b \in \mathbb{R}^m \Rightarrow \mathcal{C}(A) = \mathbb{R}^m$$

Let  $e_1, e_2, \dots, e_m$  be the standard basis of  $\mathbb{R}^m$ .

Then  $\exists x_1, x_2, \dots, x_m \ni Ax_i = e_i, \forall i = 1, 2, \dots, m$

Let  $C = (x_1 \mid x_2 \mid \dots \mid x_m)$ , then  $AC = A(x_1 \mid x_2 \mid \dots \mid x_m) = (e_1 \mid e_2 \mid \dots \mid e_m) = I_m$ .

- **Uniqueness:**

$$Ax = b \text{ has at most one solution for each } b \in \mathbb{R}^m$$

$\Leftrightarrow \forall b \in \mathbb{R}^m$ , if  $b$  can be represented as linear combination of columns of  $A$ , then it is unique

Hence, proof is complete. ■

**Example.**

$$A = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 5 & 0 \end{pmatrix}_{2 \times 3} \quad m = 2, n = 3, r = 2 \quad \rightarrow \quad \exists \text{ right inverse } C \ni AC = I$$

1°

$$AC = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 5 & 0 \end{pmatrix} \begin{pmatrix} 1/4 & 0 \\ 0 & 1/5 \\ c_{31} & c_{32} \end{pmatrix} = I_2 \quad \Rightarrow \quad C \text{ is not unique}$$

2°

$$\begin{pmatrix} 1/4 & 0 \\ 0 & 1/5 \\ c_{31} & c_{32} \end{pmatrix} \begin{pmatrix} 4 & 0 & 0 \\ 0 & 5 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{impossible since LHS is } 3 \times 2$$

3°

$$A_2 = \begin{pmatrix} 4 & 0 \\ 0 & 5 \\ 0 & 0 \end{pmatrix}_{3 \times 2} \quad m = 3, n = 2, r = 2 \quad \rightarrow Ax = b \quad \begin{pmatrix} 4 & 0 \\ 0 & 5 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

**Note.** The following statements about a square matrix  $A_{n \times n}$  are equivalent:

- (1)  $A$  is nonsingular (invertible)
- (2) The columns of  $A$  span  $\mathbb{R}^n$ , so  $Ax = b$  has **only one** solution  $\forall b \in \mathbb{R}^n$
- (3) The columns of  $A$  are independent, so  $Ax = 0$  has **only one trivial solution**  $x = 0$
- (4) The rows of  $A$  span  $\mathbb{R}^n$
- (5) The rows of  $A$  are independent
- (6) Elimination can be completed:  $PA = LDU$  with all  $d_i \neq 0$
- (7)  $\exists A^{-1} \ni AA^{-1} = A^{-1}A = I_n$
- (8) Determinant of  $A$   $\det(A) \neq 0$
- (9) Zero is NOT an eigenvalue of  $A$
- (10)  $A^T A$  is **positive definite** (正定)

## 2.5 Graph and Network

skip

## 2.6 Linear Transformation

We have seen that a matrix move subspaces around. For example,  $A$  maps  $\mathcal{N}(A)$  to the **zero vector** and move all vectors into its **column space**  $\mathcal{C}(A)$ . Let  $A$  be an  $n \times n$  matrix and  $x \in \mathbb{R}^n$ , so  $A$  transforms  $x$  into  $Ax \in \mathcal{C}(A)$ .

### 2.6.1 Notation of Linear Transformation

**Example.** Here are some examples of linear transformations:

1°

$$A = \begin{pmatrix} c & 0 \\ 0 & c \end{pmatrix} \quad A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} cx_1 \\ cx_2 \end{pmatrix} = c \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad (\text{scaling by } c)$$

2°

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix} \quad (\text{rotation by } 90^\circ)$$

3°

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_1 \end{pmatrix} \quad (\text{reflection about } x_1 = x_2)$$

4°

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ 0 \end{pmatrix} \quad (\text{projection onto } x_1\text{-axis})$$

## Lecture 8

**Definition 2.6.1 (2T).** Let  $V, W$  be vector spaces over a field  $\mathbb{F}$ . A linear transform from  $V$  to  $W$  is a function  $T : V \rightarrow W$  such that preserves the operations on  $V$  and  $W$ , i.e.

$$\begin{cases} T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v}), & \forall \mathbf{u}, \mathbf{v} \in V; \\ T(c\mathbf{u}) = cT(\mathbf{u}), & \forall \mathbf{u} \in V, c \in \mathbb{F}. \end{cases}$$

**Example.**

$$T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$$

$$T : (x_1, x_2, x_3) \mapsto (x_2, x_3, x_1)$$

$T$  is a linear transform.

**Example.**

$$A = \frac{d}{dt} : \mathbb{P}_n(\mathbb{R}) \rightarrow \mathbb{P}_{n-1}(\mathbb{R})$$

$$p(t) \in \mathbb{P}_n(\mathbb{R}), p(t) = a_0 + a_1t + a_2t^2 + \cdots + a_nt^n$$

See the attributes below:

$$AP = \frac{d}{dt}(a_0 + a_1t + a_2t^2 + \cdots + a_nt^n) = a_1 + 2a_2t + \cdots + na_nt^{n-1}$$

The nullspace of  $A$  is all constant polynomials.

$$\mathcal{C}(AP) = \mathbb{P}_{n-1}(\mathbb{R})$$

the basis is  $\{1, t, t^2, \dots, t^{n-1}\}$  and  $\text{rank}(\mathcal{C}(A)) = n$ .

$$\text{nullity}(A) + \text{rank}(A) = 1 + n = \dim(\mathbb{P}_n(\mathbb{R})).$$

**Example.**

$$A = \int_0^t : \mathbb{P}_n(\mathbb{R}) \rightarrow \mathbb{P}_{n+1}(\mathbb{R})$$

See the attributes below:

$$AP = \int_0^t (a_0 + a_1t + a_2t^2 + \cdots + a_nt^n) dt = a_0t + \frac{a_1t^2}{2} + \frac{a_2t^3}{3} + \cdots + \frac{a_nt^{n+1}}{n+1} + C$$

The nullspace of  $A$  is all constant polynomials.

$$\mathcal{N}(AP) = \{0\}$$

The range of  $A$

$$\mathcal{C}(AP) = \mathbb{P}_{n+1}(\mathbb{R}) - \{\text{constant}\}/\{0\}$$

**Example.**

$$T : \mathbb{R}^3 \rightarrow \mathbb{R}$$

$$T : (x_1, x_2, x_3) \mapsto 2x_1 + 3x_2 - x_3, x_i \in \mathbb{R}$$

$T$  is a linear transform.

**Example.**

$$T : \mathbb{R}^3 \rightarrow \mathbb{R}$$

$$T : (x_1, x_2, x_3) \mapsto 2x_1^2 + 3x_2 - x_3, \quad x_i \in \mathbb{R}$$

$T$  is NOT a linear transform.

$$\therefore T(x+y) \neq T(x) + T(y)$$

**Theorem 2.6.1.** Let  $T : V \rightarrow W$  be a linear transform, where  $V, W$  are vector spaces over a field  $\mathbb{F}$ .

(i) If  $M$  is a subspace of  $V$ , then

$$T(M) = \{x \in W \mid \exists \mathbf{m} \in M, \text{ such that } T(\mathbf{m}) = x\}$$

is a subspace of  $W$ .

(ii) If  $N$  is a subspace of  $W$ , then

$$T^{-1}(N) = \{\mathbf{v} \in V \mid T(\mathbf{v}) \in N\}$$

is a subspace of  $V$ .

**Proof.** Here is the proof:

(i) Let  $M \leq V$ ,  $y_1, y_2 \in T(M) \subseteq W$ , and  $\alpha \in \mathbb{F}$ .

$$y_1, y_2 \in T(M) \Rightarrow \exists x_1, x_2 \in M \text{ s.t. } T(x_1) = y_1, T(x_2) = y_2$$

Then

$$T(\alpha x_1 + x_2) = \alpha T(x_1) + T(x_2)$$

since  $T$  is a linear transformation.

Also

$$\alpha x_1 + x_2 \in M$$

since  $M$  is a subspace of  $V$ .

Therefore

$$\alpha y_1 + y_2 = \alpha T(x_1) + T(x_2) = T(\alpha x_1 + x_2) \in T(M)$$

so  $T(M)$  is a subspace of  $W$ .

(ii) Let  $x_1, x_2 \in T^{-1}(N)$  and  $\alpha \in \mathbb{F}$ .

$$T(\alpha x_1 + x_2) = \alpha T(x_1) + T(x_2) \in N$$

since  $N \leq W$  and  $T(x_1), T(x_2) \in N$ .

Therefore

$$\alpha x_1 + x_2 \in T^{-1}(N)$$

and  $T^{-1}(N)$  is a subspace of  $V$ .

■

**Definition 2.6.2.**  $T : V \rightarrow W$  over a field  $\mathbb{F}$  is a linear transform. Then  $T^{-1}(\mathbf{0}_W)$  is called the nullspace (kernel) of  $T$ , where  $\mathbf{0}_W$  is the zero vector in  $W$ .  $T(V)$  is called the range (image) of  $T$ .

$$\dim(T^{-1}(\mathbf{0}_W)) = \text{nullity}(T)$$

$$\dim(T(V)) = \text{rank}(T)$$

### 2.6.2 Matrix Representation of Linear Transformations

**Question.** What is the transformation taken  $A : \mathbb{R}^2 \rightarrow \mathbb{R}^3$

$$x_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}_{\in \mathbb{R}^2} \rightarrow \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}_{\in \mathbb{R}^3}, \quad x_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}_{\in \mathbb{R}^2} \rightarrow \begin{pmatrix} 4 \\ 6 \\ 8 \end{pmatrix}_{\in \mathbb{R}^3}$$

**Answer.**

$$A = \begin{pmatrix} 2 & 4 \\ 3 & 6 \\ 4 & 8 \end{pmatrix}_{3 \times 2} = (T)_{\{e_1, e_2\}}^{\{e_1, e_2, e_3\}}$$

(\*)

**Example.**

$$T : \mathbb{P}_3(\mathbb{R}) \rightarrow \mathbb{P}_2(\mathbb{R}), \text{ i.e. } T(f) = \frac{d}{dt}(f)$$

The ordered basis of two vector spaces are

$$\begin{cases} \mathcal{B}_1 = \mathcal{B}(\mathbb{P}_3(\mathbb{R})) : \{1, t, t^2, t^3\} \\ \mathcal{B}_2 = \mathcal{B}(\mathbb{P}_2(\mathbb{R})) : \{1, t, t^2\} \end{cases}$$

Then we have

$$(T)_{\mathcal{B}_2}^{\mathcal{B}_1} = \begin{matrix} 1 \\ t \\ t^2 \end{matrix} \begin{pmatrix} 1 & t & t^2 & t^3 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}_{3 \times 4} \quad \text{e.g. } (T) \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}_{3 \times 1}$$

**Example.**

$$T : \mathbb{P}_3(\mathbb{R}) \rightarrow \mathbb{P}_3(\mathbb{R}), \text{ i.e. } T(f) = \frac{d}{dt}(f)$$

We have to handle the  $t^3$  term, which means

$$(T)_{\mathcal{B}_1}^{\mathcal{B}_1} = (T)_{\mathcal{B}_1} = \begin{matrix} 1 \\ t \\ t^2 \\ t^3 \end{matrix} \begin{pmatrix} 1 & t & t^2 & t^3 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}_{4 \times 4}, \quad \text{this is called a } \langle \text{differentiation matrix} \rangle$$

**Example.**

$$\int_0^t : \mathbb{P}_3(\mathbb{R}) \rightarrow \mathbb{P}_4(\mathbb{R})$$

The ordered basis of two vector spaces are

$$\begin{cases} \mathcal{B}_1 = \mathcal{B}(\mathbb{P}_3(\mathbb{R})) : \{1, t, t^2, t^3\}, \dim(\mathbb{P}_3(\mathbb{R})) = 4 \\ \mathcal{B}_2 = \mathcal{B}(\mathbb{P}_4(\mathbb{R})) : \{1, t, t^2, t^3, t^4\}, \dim(\mathbb{P}_4(\mathbb{R})) = 5 \end{cases}$$

hence we have

$$(T)_{\mathcal{B}_2}^{\mathcal{B}_1} = \begin{pmatrix} 1 & t & t^2 & t^3 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & \frac{1}{3} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{4} \end{pmatrix}_{5 \times 4} \quad \text{which is called an (integration matrix)}$$

and we also have

$$\mathcal{C}(T) = \text{span}\{t, t^2, t^3, t^4\}, \text{rank}(T) = 4$$

$$\mathcal{N}(T) = \{0\}, \text{nullity}(T) = 0$$

**Example.**

$$\mathbb{P}_2(\mathbb{R}) \xrightarrow{\int_t} \mathbb{P}_3(\mathbb{R}) \xrightarrow{\frac{d}{dt}} \mathbb{P}_2(\mathbb{R})$$

$$\left( \frac{d}{dt} \int_0^t \right) = \left( \frac{d}{dt} \right)_{3 \times 4} \left( \int_0^t \right)_{4 \times 3} = \left( I \right)_{3 \times 3} \quad \text{Diff is the left inverse of Int}$$

**2.6.3 Rotation  $Q$ , Projection  $P$ , Reflection  $R$** 

We introduce three important linear transformations in  $\mathbb{R}^2$ :

$$Q = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}, \quad P = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad R = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$1^\circ$  Rotation:  $Q$  rotates vectors by an angle  $\theta$ .

$$(Q) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}_{2 \times 2}$$

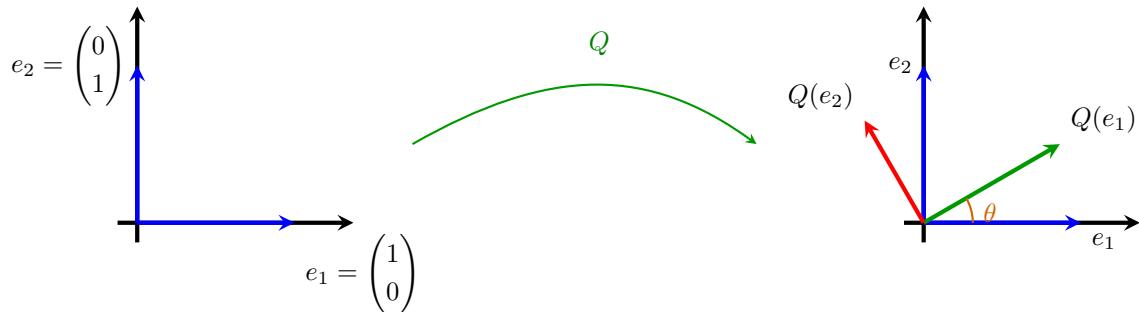


Figure 2.2: Rotation in  $\mathbb{R}^2$

$$Q \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}, \quad Q \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\sin \theta \\ \cos \theta \end{pmatrix}$$

- $Q_{-\theta} \cdot Q_\theta = \mathbf{1}_{\mathbb{R}^2}$
- $Q_\theta \cdot Q_\theta = Q_{2\theta}$
- $Q_\theta \cdot Q_\phi = Q_{\theta+\phi}$

2° Projection:  $P$  projects vectors onto the  $\theta$ -line.

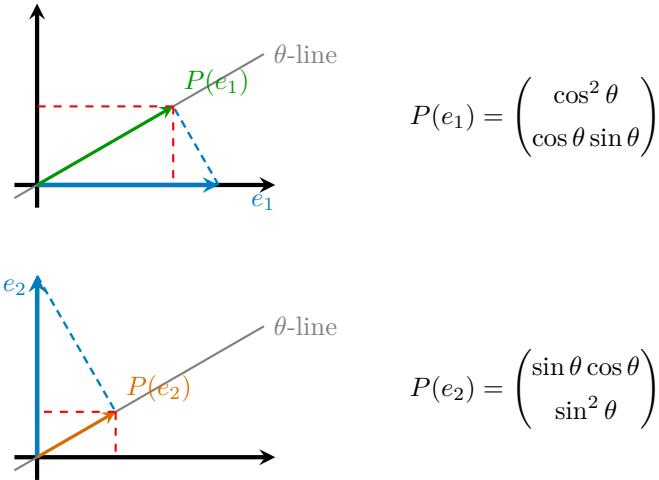


Figure 2.3: Projection onto a line at angle  $\theta$

$$P = \begin{pmatrix} \cos^2 \theta & \sin \theta \cos \theta \\ \sin \theta \cos \theta & \sin^2 \theta \end{pmatrix}$$

Here are some properties of projection:

- $P^2 = P$
- Symmetric:  $P^T = P$
- $P^{-1}$  does not exist.

1°

$$\begin{aligned} P \left( \alpha \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix} \right) &= \alpha P \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix} \\ &= \alpha \begin{pmatrix} \cos^2 \theta & \sin \theta \cos \theta \\ \sin \theta \cos \theta & \sin^2 \theta \end{pmatrix} \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix} = \alpha \begin{pmatrix} \cos^3 \theta + \cos \theta \sin^2 \theta \\ \sin \theta \cos^2 \theta + \sin^3 \theta \end{pmatrix} = \alpha \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix} \end{aligned}$$

2°

$$\begin{aligned} P \left( \alpha \begin{pmatrix} -\sin \theta \\ \cos \theta \end{pmatrix} \right) &= \alpha P \begin{pmatrix} -\sin \theta \\ \cos \theta \end{pmatrix} \\ &= \alpha \begin{pmatrix} \cos^2 \theta & \sin \theta \cos \theta \\ \sin \theta \cos \theta & \sin^2 \theta \end{pmatrix} \begin{pmatrix} -\sin \theta \\ \cos \theta \end{pmatrix} = \alpha \begin{pmatrix} -\sin \theta \cos^2 \theta + \cos \theta \sin \theta \\ -\sin^2 \theta \cos \theta + \sin^3 \theta \end{pmatrix} = \alpha \begin{pmatrix} 0 \\ 0 \end{pmatrix} \end{aligned}$$

Thus,

$\alpha \begin{pmatrix} -\sin \theta \\ \cos \theta \end{pmatrix}$  is in the nullspace of  $P$ .

3° Reflection:  $R$  reflects vectors across the  $\theta$ -line.

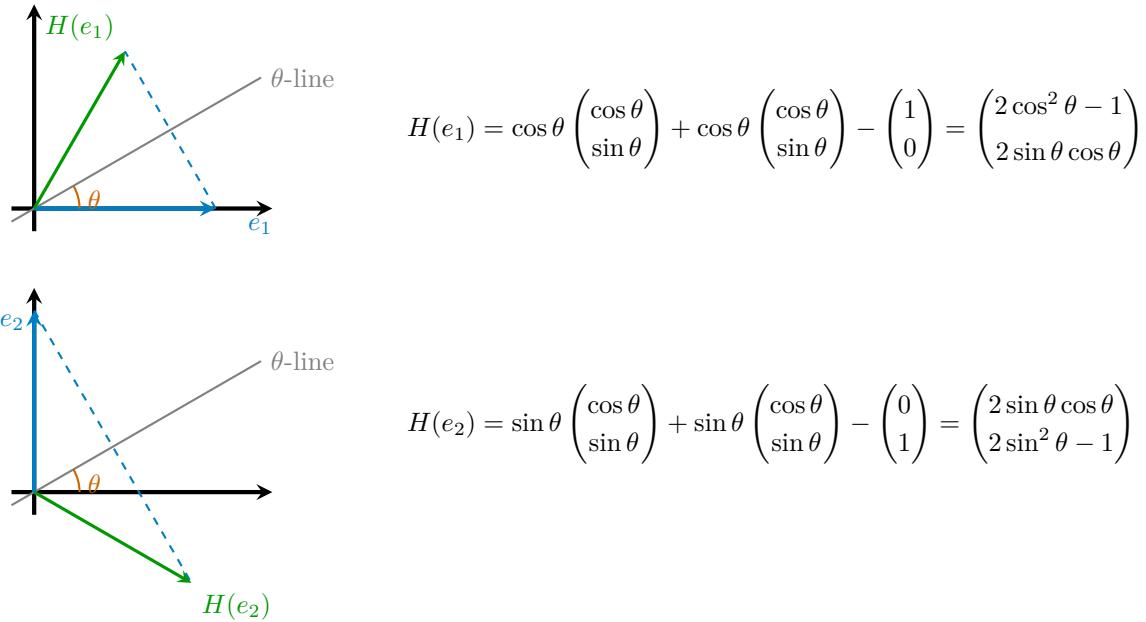


Figure 2.4: Reflection across a line at angle  $\theta$

$$H = \begin{pmatrix} 2 \cos^2 \theta - 1 & 2 \sin \theta \cos \theta \\ 2 \sin \theta \cos \theta & 2 \sin^2 \theta - 1 \end{pmatrix}$$

Here are some properties of reflection:

- $H^2 = I$
- $H^{-1} = H$
- $H = 2P - I$  ( $Hx + x = 2Px$ )

**Note.** If first basis vector is on the  $\theta$ -line, and the second basis vector is perpendicular to the  $\theta$ -line, then

$$P^* = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad H^* = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = 2P^* - I, \quad Q^* = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

# Chapter 3

## Orthogonality

### 3.1 Perpendicular Vectors and Orthogonal Subspaces

There are three important concepts in this section:

- (i) The length of vector
- (ii) The test for perpendicularity
- (iii) How to create perpendicular vectors from linearly independent vectors

Now we start to discuss:

(i) **The length of vector:**

The length (or norm) of a vector, in  $\mathbb{R}^n$ , that satisfies the Pythagorean theorem is defined as:

**Definition 3.1.1.** Let  $\mathbf{x} \in \mathbb{R}^n$  be

$$\mathbf{x} = (x_1, x_2, \dots, x_n) = x_1 e_1 + x_2 e_2 + \dots + x_n e_n \in \mathbb{R}^n$$

then

$$\|\mathbf{x}\|^2 = \sum_{i=1}^n x_i^2 = \mathbf{x}^T \mathbf{x}$$

(ii) **The test for perpendicularity:**

**Definition 3.1.2.** Given  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ , then if  $\mathbf{x} \perp \mathbf{y}$ , then by Pythagorean theorem, we have

$$\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 = \|\mathbf{x} - \mathbf{y}\|^2$$

Then we can deduce that

$$x_1^2 + x_2^2 + \dots + x_n^2 + y_1^2 + y_2^2 + \dots + y_n^2 = (x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2$$

then we have

$$\mathbf{x}^T \mathbf{y} = 0$$

**Definition 3.1.3 (Inner Product).** Let  $V$  be a vector space over a field  $\mathbb{F}$  ( $\mathbb{R}, \mathbb{C}$ ). An inner product on  $V$  is a function that assigns to every ordered pair of vectors  $\mathbf{x}$  and  $\mathbf{y}$  in  $V$ , a scalar in  $\mathbb{F}$ , denoted as

$$\langle \mathbf{x}, \mathbf{y} \rangle$$

$\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in V, c \in \mathbb{F}$ , we have

- (a)  $\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$
- (b)  $\langle c \mathbf{x}, \mathbf{y} \rangle = c \langle \mathbf{x}, \mathbf{y} \rangle$
- (c)  $\langle \mathbf{y}, \mathbf{x} \rangle = \overline{\langle \mathbf{x}, \mathbf{y} \rangle}$  (where  $\overline{a+bi} = a-bi$  complex conjugate)
- (d)  $\langle \mathbf{x}, \mathbf{x} \rangle > 0$ , if  $\mathbf{x} \neq \mathbf{0}$

**Note (1).** If  $\mathbb{F} = \mathbb{R}$ , (c) will reduce to  $\langle \mathbf{y}, \mathbf{x} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle$ .

**Note (2).** Inner product is linear in the first component.

**Definition 3.1.4 (Standard Inner Product).** Let  $V = \mathbb{R}^n / \mathbb{R}$ , defined

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y} = \sum_{i=1}^n x_i y_i$$

This is called the standard inner product on  $\mathbb{R}^n$ .

**Proposition 3.1.1.** If  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ , then

- Let  $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y}$  be standard inner product.
- Let  $\langle \mathbf{x}, \mathbf{y} \rangle = 0$  if and only if  $\mathbf{x} \perp \mathbf{y}$ .

**Example.** If  $\langle \cdot, \cdot \rangle$  is any inner product on  $V$ , and  $r > 0$ , we define

$$\langle \mathbf{x}, \mathbf{y} \rangle' = r \langle \mathbf{x}, \mathbf{y} \rangle$$

$$1^\circ \quad \langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle' = r \langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = r \langle \mathbf{x}, \mathbf{z} \rangle + r \langle \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle' + \langle \mathbf{y}, \mathbf{z} \rangle'$$

$$2^\circ \quad \langle c\mathbf{x}, \mathbf{y} \rangle' = r \langle c\mathbf{x}, \mathbf{y} \rangle = c \cdot r \langle \mathbf{x}, \mathbf{y} \rangle = c \langle \mathbf{x}, \mathbf{y} \rangle'$$

**Example.** Let  $V = \{f \mid f : \text{real-valued continuous functions on } [0, 1]\} = \mathcal{C}([0, 1])$ . For  $f, g \in V$ , define

$$\langle f, g \rangle = \int_0^1 f(t)g(t)dt$$

**Example.** Let  $V = \mathbb{C}^n$ ,  $\mathbb{C}^n$  is a vector space over  $\mathbb{C}$ . For  $\mathbf{x}, \mathbf{y} \in V$ , define

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \bar{\mathbf{y}} = \sum_{i=1}^n x_i \bar{y}_i$$

$$\langle \mathbf{y}, \mathbf{x} \rangle = \mathbf{y}^T \bar{\mathbf{x}} = \overline{\mathbf{x}^T \bar{\mathbf{y}}} = \overline{\langle \mathbf{x}, \mathbf{y} \rangle}.$$

**Example.** Let  $V = \mathbb{C}$ ,  $\mathbb{C}$  is a vector space over  $\mathbb{C}$ . If  $\mathbf{x}, \mathbf{y} \in \mathbb{C}$ ,  $x = a + bi$ ,  $y = c + di$ , define

$$\langle \mathbf{x}, \mathbf{y} \rangle = (a + bi)(c - di)$$

$$1^\circ \quad \langle \mathbf{y}, \mathbf{x} \rangle = (c + di)(a - bi) = \overline{(a + bi)(c - di)} = \overline{\langle \mathbf{x}, \mathbf{y} \rangle}$$

$$2^\circ \quad \langle \mathbf{x}, \mathbf{x} \rangle = (a + bi)(a - bi) = a^2 + b^2 > 0 \text{ if } \mathbf{x} \neq 0$$

## Lecture 9

**Definition 3.1.5 (inner product space).** An **inner product space** is a real or complex vector space (i.e. a vector space over the field  $\mathbb{R}$  or  $\mathbb{C}$ ) together with a specified inner product on that space.

4 Nov. 13:20

**Definition 3.1.6 (orthogonal).** In an inner product space  $V$ ,  $\mathbf{x}$  is **orthogonal** to  $\mathbf{y}$  if  $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ . A set  $S$  of vectors in  $V$  is called **orthogonal set** if all pairs of distinct vectors in  $S$  are orthogonal. An **orthonormal set** is an orthogonal set of unit vectors.

$$\langle \mathbf{v}, \mathbf{v} \rangle = \|\mathbf{v}\|^2 = 1, \quad \forall \mathbf{v} \in S$$

**Proposition 3.1.2.** An orthogonal set of nonzero vectors is linearly independent.

**Proof.** Let  $v_1, \dots, v_n$  be nonzero distinct vectors in  $S$ , and  $c_1, \dots, c_n \in \mathbb{F}$

$$c_1v_1 + c_2v_2 + \dots + c_nv_n = \sum_{i=1}^n c_i v_i = y$$

$$\langle y, v_j \rangle = \left\langle \sum_{i=1}^n c_i v_i, v_j \right\rangle = \sum_{i=1}^n c_i \langle v_i, v_j \rangle = c_j \langle v_j, v_j \rangle = c_j \|v_j\|^2$$

Then we have  $y = 0 \iff c_j = 0, \forall j$

$\therefore \{v_1, v_2, \dots, v_n\}$  is linearly independent. ■

**Example.**  $\{e_1, e_2, \dots, e_n\}$  is an orthonormal set (basis) for  $\mathbb{R}^n$

In  $\mathbb{R}^2$ ,

$$1^\circ \quad \{e_1, e_2\}$$

$$2^\circ \quad \left\{ \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\}$$

$$3^\circ \quad \left\{ \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}, \begin{pmatrix} -\sin \theta \\ \cos \theta \end{pmatrix} \right\}$$

### 3.1.1 Orthogonal Subspaces

**Definition 3.1.7 (3B).** Let  $W_1$  and  $W_2$  be subspaces of an inner product space  $V$ . We say that  $W_1$  is **orthogonal** to  $W_2$  ( $W_1 \perp W_2$ ) if

$$\langle \mathbf{w}_1, \mathbf{w}_2 \rangle = 0, \quad \forall \mathbf{w}_1 \in W_1, \forall \mathbf{w}_2 \in W_2$$

**Note.** In  $\mathbb{R}^3$ , the  $xy$ -plane is **NOT** orthogonal to the  $yz$ -plane. Because vectors along the  $y$ -axis are in both planes, and their inner product is not zero.

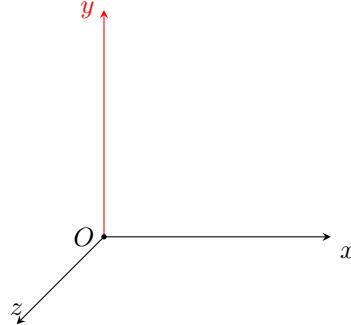


Figure 3.1: The  $xy$ -plane and the  $yz$ -plane in  $\mathbb{R}^3$

**Example.** In  $\mathbb{R}^3$ , the subspace spanned by  $(1, 2, 3)^T$  is orthogonal to the subspace spanned by  $(1, 1, -1)^T$ .

**Example.** In  $\mathbb{R}^3$ , the subspace spanned by  $(1, 2, 3)^T$  is orthogonal to the subspace spanned by  $\{(1, 1, -1)^T, (5, -4, 1)^T\}$ .

**Theorem 3.1.1 (3C).**  $A_{m \times n}$  The row space is orthogonal to the null space (in  $\mathbb{R}^n$ ), and the column space is orthogonal to the left null space (in  $\mathbb{R}^m$ ).

**Proof.** This is the proof.

- 1°     •  $\mathbf{v} \in$  row space of  $A$ , then we have  $\mathbf{b} = A^T y$  for some  $y \in \mathbb{R}^m$ .
- $\mathbf{w} \in$  null space of  $A$ , then we have  $A\mathbf{w} = 0$ .

$$\mathbf{v}^T \mathbf{w} = (A^T y)^T \mathbf{w} = y^T (A\mathbf{w}) = y^T 0 = 0$$

- 2°     •  $b \in \mathcal{C}(A) \Rightarrow Ax = b$  is solvable.
- $y \in \mathcal{N}(A^T) \Rightarrow A^T y = 0$ .

$$b^T y = (Ax)^T y = x^T (A^T y) = x^T 0 = 0$$

■

**Example.**

$$A = \begin{pmatrix} 1 & 3 & 4 \\ 5 & 2 & 7 \end{pmatrix}_{2 \times 3} \rightarrow U = \begin{pmatrix} 1 & 3 & 4 \\ 0 & -13 & -13 \end{pmatrix}_{2 \times 3}$$

- $\mathcal{C}(A) = \text{span} \left\{ \begin{pmatrix} 1 \\ 5 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \end{pmatrix} \right\}$ , rank = 2
- $\mathcal{C}(A^T) = \text{span} \left\{ \begin{pmatrix} 1 \\ 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 5 \\ 2 \\ 7 \end{pmatrix} \right\}$ , rank = 2
- $\mathcal{N}(A) = \text{span} \left\{ \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \right\}$ ,  $n - \text{rank} = 1$
- $\mathcal{N}(A^T) = \text{span} \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}$ ,  $m - \text{rank} = 0$

$$\boxed{\mathcal{C}(A) \perp \mathcal{N}(A^T) \in \mathbb{R}^2} \quad \boxed{\mathcal{C}(A^T) \perp \mathcal{N}(A) \in \mathbb{R}^3}$$

**Note.** The nullspace  $\mathcal{N}(A)$  doesn't contain "some" of vectors orthogonal to the row space. It contains "every" such vector.

**Proposition 3.1.3.** Let  $V$  be an inner product space, and let  $W$  be a subspace of  $V$ . Then the set is defined

$$U = \{ \mathbf{v} \in V \mid \langle \mathbf{v}, \mathbf{w} \rangle = 0, \forall \mathbf{w} \in W \}$$

Then  $U$  is a subspace of  $V$ .

**Definition 3.1.8.** The subspace  $U$  is called the **orthogonal complement** of  $W$  in  $V$ , denoted by  $W^\perp$  ( $W$ -perp). By definition of nullspace  $\mathcal{N}(A)$ , we have

$$\mathcal{N}(A) = (\mathcal{C}(A^T))^\perp, \quad \text{or} \quad \mathcal{C}(A) = (\mathcal{N}(A^T))^\perp$$

**Theorem 3.1.2 ((3D) Fundamental Theorem of Linear Algebra).** The nullspace is the orthogonal complement of the row space in  $\mathbb{R}^n$ , and the left nullspace is the orthogonal complement of the column space in  $\mathbb{R}^m$ .

**Proposition 3.1.4 (3E).** The equation  $Ax = b$  is solvable if and only if

$$b^T y = 0, \quad \forall y \in \mathcal{N}(A^T)$$

**Note.** Solvability of  $Ax = b$ :

- Direct approach:  $b$  must be a combination of the columns of  $A$ .
- Indirect approach:  $b$  must be orthogonal to every vector that is orthogonal to the columns of  $A$ .

### 3.1.2 The Matrix and the Subspaces

$U$  and  $W$  can be orthogonal without being complements when their dimensions are too small. In  $\mathbb{R}^3$

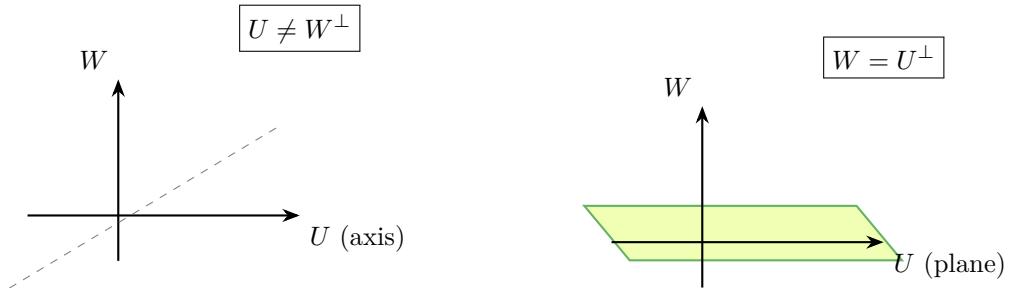


Figure 3.2: Orthogonal but not complements

$$W = U^\perp \Rightarrow U = W^\perp \text{ or } U^{\perp\perp} = U$$

When the space is split into orthogonal parts (i.e.  $V = U+W = U+U^\perp$ ), so every vector ( $x = x_U + x_{U^\perp}$ ).

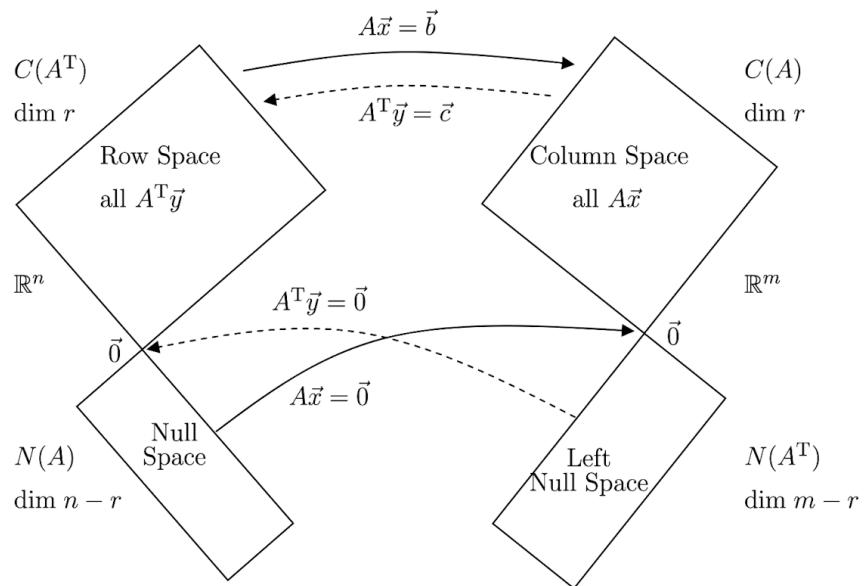


Figure 3.3: Fundamental Theorem of Linear Algebra

**Proposition 3.1.5 (3F).** The mapping from row space to column space is actually invertible. Every matrix  $A_{m \times n}$  transforms its row space to its column space. (On these  $r$ -dimensional subspaces,  $A$  is invertible.)

$$A_{m \times n} : \mathbb{R}_{\substack{x \\ Ax=b}}^n \xrightarrow{A} \mathbb{R}_{\substack{b \\ A^T b=0}}^m \quad Ax = b$$

$$A_{n \times m}^T : \mathbb{R}_{\substack{b \\ A^T b=0}}^m \xrightarrow{A^T} \mathbb{R}_{\substack{x \\ A^T b=0}}^n \quad A^T b \stackrel{?}{=} x \quad x = A^{-1}b$$

- $A^T$  moves the space correctly but NOT the individual vectors.
- When  $A^{-1}$  fails to exist, we can substitute. It's called the **pseudoinverse**, denoted by  $A^+$ .

$$\begin{cases} A^+ A x = x, & \forall x \in \mathcal{C}(A^T) \\ A^+ b = 0, & \forall b \in \mathcal{N}(A^T) \end{cases}$$

## 3.2 Inner Product and Projections onto Lines

Inner product  $x^T y = \begin{cases} = 0, & \text{if } x \perp y \\ \neq 0 \end{cases}$

**Question.** Practical applications?

Least squares solution to an overdetermined system. i.e. given a vector  $b$  not falling in the desired space, we have to project to the subspace. Then we get the approximate solution.

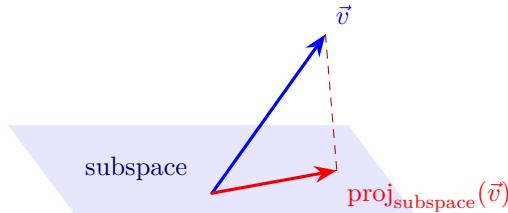


Figure 3.4: Projection onto a subspace (in  $\mathbb{R}^3$ )

**Question.** Practical applications?

A formula for the projection, we need the basis.

### 3.2.1 Inner Product and Schwarz Inequality

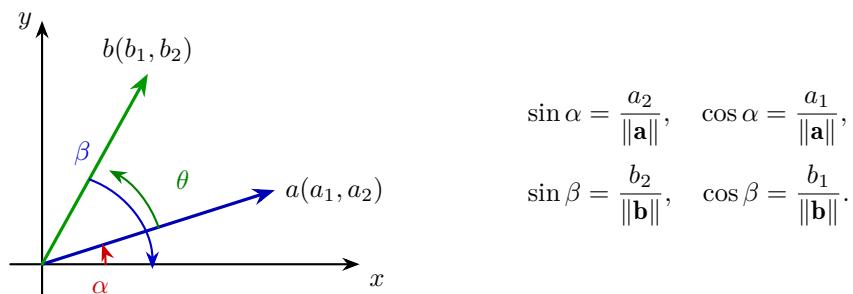


Figure 3.5: Angle between two vectors in  $\mathbb{R}^2$

$$\cos \theta = \cos(\beta - \alpha) = \cos \beta \cos \alpha + \sin \beta \sin \alpha = \frac{a_1 b_1 + a_2 b_2}{\|a\| \|b\|} = \frac{a^T b}{\|a\| \|b\|}$$

**Proposition 3.2.1 (3G).** The cosine of the angle between any two vectors  $a, b \in \mathbb{R}^n$  is

$$\cos \theta = \frac{a^\top b}{\|a\| \|b\|}$$

If we consider the relationship between  $\|a\|, \|b\|$  and  $\|b - a\|$ , then we have

$$\|b - a\|^2 = \|a\|^2 + \|b\|^2 - 2\|a\|\|b\| \cos \theta \quad (\text{Law of Cosines})$$

**Projection onto a Line:**

$$\begin{aligned} (b - p) \perp a &\iff (b - p)^\top a = 0 \iff (b - \alpha a)^\top a = 0 \\ &\iff b^\top a - \alpha a^\top a = 0 \iff \alpha = \frac{a^\top b}{a^\top a} \end{aligned}$$

**Proposition 3.2.2 (3H).** The projection of  $b$  onto the line through  $0 \& a$  is

$$p = \frac{a^\top b}{a^\top a} \cdot a$$

**Theorem 3.2.1 (3I Schwarz Inequality).** For any two vectors in inner product space satisfy the **Cauchy-Schwarz inequality**:

$$|a^\top b| \leq \|a\| \|b\| \quad \text{or } |\langle a, b \rangle| \leq \|a\| \|b\|$$

with equality if and only if  $b = \alpha a$ , for some  $\alpha \in \mathbb{F}$ .

**Proof.**

$$\begin{aligned} \|b - p\|^2 &= \left\|b - \frac{a^\top b}{a^\top a} a\right\|^2 = b^\top b - 2 \cdot \frac{(a^\top b)^2}{a^\top a} + \left(\frac{a^\top b}{a^\top a}\right)^2 a a^\top \\ &= \frac{(b^\top b)(a^\top a) - (a^\top b)^2}{a^\top a} \geq 0 \\ &\Rightarrow |a^\top b| \leq \|a\| \|b\| \end{aligned}$$

and the equality holds  $\iff \|b - p\| = 0 \iff b = p = \alpha a$ . ■

**Example.** Project  $(1, 1, 1) \rightarrow (1, 2, 3)$

$$p = \frac{a^\top b}{a^\top a} a = \frac{6}{14} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 3/7 \\ 6/7 \\ 9/7 \end{pmatrix}$$

### 3.2.2 Projections of Rank One

**Question.** What is the matrix this linear transformation that maps  $b$  to  $p$

$$p = \frac{a^\top b}{a^\top a} \cdot a = \frac{aa^\top}{a^\top a} b$$

The projection matrix is  $P = \frac{aa^\top}{a^\top a}$

**Note.** Here are some properties of  $P$ :

$$1^\circ \ P \text{ is symmetric: } P^T = P$$

$$2^\circ \ P^2 = P \ (\text{idempotent})$$

**Proof.** Here are the proofs:

$$1^\circ \ P^T = \left( \frac{aa^T}{a^T a} \right)^T = \frac{(a^T)^T a^T}{a^T a} = \frac{aa^T}{a^T a} = P$$

$$2^\circ \ P^2 = \left( \frac{aa^T}{a^T a} \right) \left( \frac{aa^T}{a^T a} \right) = \frac{aa^T aa^T}{(a^T a)^2} = \frac{a(a^T a)a^T}{(a^T a)^2} = \frac{aa^T}{a^T a} = P$$

Proof complete. ■

- $\text{rank}(P) = 1$ , nullspace of  $P$  is the space orthogonal to  $a$ . i.e.

$$\mathcal{N}(P) \perp \mathcal{C}(P)$$

which is not general. It is right here because  $\mathcal{C}(P) = \mathcal{C}(P^T) = \text{span}(a)$ .

**Remark (Scaling).** Project  $b$  onto  $a$ , which can be scaled arbitrarily. i.e. project onto  $\alpha a$

$$p = \frac{a'a'^T}{a'^T a'} = \frac{(\alpha a)(\alpha a)^T}{(\alpha a)^T (\alpha a)} = \frac{\alpha^2 a a^T}{\alpha^2 a^T a} = \frac{a a^T}{a^T a} = p \quad (\text{remains the same})$$

## Lecture 10

### 3.3 Projections and Least Squares Applications

18 Nov. 13:20

$$\begin{cases} a_1 x = b_1 \\ a_2 x = b_2 \\ a_3 x = b_3 \end{cases} \quad \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} x = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

$$E^2 = (a_1 x - b_1)^2 + (a_2 x - b_2)^2 + (a_3 x - b_3)^2$$

$$1^\circ \text{ if } b = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} x_0 \text{ then } E^2 = 0$$

2° Or consider

$$\begin{aligned} \frac{dE^2}{dx} &= 2[a_1(a_1 x - b_1) + a_2(a_2 x - b_2) + a_3(a_3 x - b_3)] = 0 \\ \Rightarrow \bar{x} &= \frac{a_1 b_1 + a_2 b_2 + a_3 b_3}{a_1^2 + a_2^2 + a_3^2} = \underbrace{\frac{a^T b}{a^T a}}_{\alpha} \end{aligned}$$

Hence, we call the projection of  $b$  onto  $a$  as  $p_a b = \bar{x}a = \frac{a^T b}{a^T a}a$ , which is also called the least squares solution.

$$a^T(b - \bar{x}a) = 0 = a^T b - \bar{x}a^T a$$

### 3.3.1 Least Squares Problem with Several Variables

$$A_{m \times n} x_{n \times 1} = b_{m \times 1} \quad (m > n)$$

- If  $b \in \text{Col}(A)$ , then the system is solvable.
- If the equations contain errors, then  $b$  might not belong to  $\mathcal{C}(A)$

$$A_{3 \times 2} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}$$

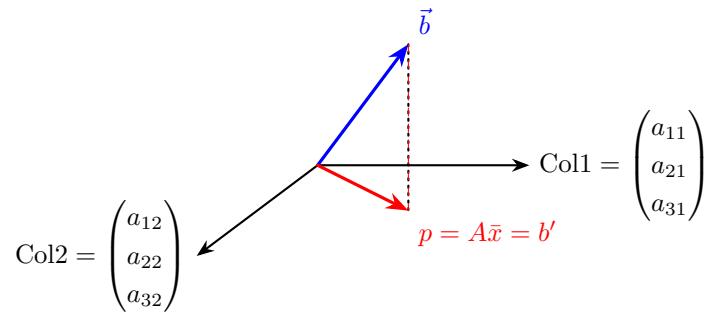


Figure 3.6: Projection of  $\vec{b}$  onto the span of the columns of  $A$

- $Ax = b$  has error,  $b \notin \mathcal{C}(A)$
- $Ax = b'$  is solvable  $\Rightarrow b' \in \mathcal{C}(A) \Leftrightarrow \exists \bar{x}_{n \times 1} \ni A\bar{x} = p = b'$

**Note.** To find  $x$ , we do it in three ways:

1° The vectors perpendicular to  $\mathcal{C}(A)$  are in  $\mathcal{N}(A^T)$

$$A^T(b - A\bar{x}) = 0 \Rightarrow A^T A\bar{x} = A^T b$$

2° The error vector must be perpendicular to each column of  $A$ .

If  $A = [a_1 \ a_2 \ \cdots \ a_n]$

$$\begin{cases} a_1^T(b - A\bar{x}) = 0 \\ a_2^T(b - A\bar{x}) = 0 \\ \vdots \\ a_n^T(b - A\bar{x}) = 0 \end{cases} \Rightarrow A^T(b - A\bar{x}) = 0 \Rightarrow A^T A\bar{x} = A^T b \quad (\textcolor{teal}{A'} \textcolor{teal}{x}' = b')$$

3° The third way is to differentiate the sum of squares.

$$E^2 = \|Ax - b\|^2 = (Ax - b)^T(Ax - b) \Rightarrow A^T Ax - A^T b = 0 \Rightarrow A^T A\bar{x} = A^T b$$

**Proposition 3.3.1 (3L).** The least-squares solution to an inconsistent system  $Ax = b$  of  $m$  equations in  $n$  unknowns satisfies

$$A^T A \bar{x} = A^T b$$

The above equation is referred to as the **Normal Equation**.

**Note.** The properties of  $A^T A$ :

1°  $A^T A$  is symmetric.

**Proof.**  $(A^T A)^T = A^T (A^T)^T = A^T A$  ■

2° The  $(i, j)$ <sup>th</sup> entry of  $A^T A$  is the inner product of the  $i$ <sup>th</sup> and  $j$ <sup>th</sup> columns of  $A$ .

3°  $A^T A$  has the same nullspace of  $A$  (i.e.,  $\mathcal{N}(A^T A) = \mathcal{N}(A)$ ).

**Proof.** We follow the two directions:

- $Ax = 0 \Rightarrow A^T Ax = 0 \quad \therefore \mathcal{N}(A) \subseteq \mathcal{N}(A^T A)$
- if  $A^T Ax = 0$ , then

$$x^T A^T Ax = (Ax)^T (Ax) = \|Ax\|^2 = 0 \Rightarrow Ax = 0 \quad \therefore \mathcal{N}(A^T A) \subseteq \mathcal{N}(A)$$

Proof complete. ■

4°  $A^T A$  is positive definite, i.e., for any non-zero vector  $x$ ,

$$x^T A^T Ax = (Ax)^T (Ax) = \|Ax\|^2 \geq 0$$

with equality if and only if  $Ax = 0$ .

**Proposition 3.3.2 (3L (contd.)).** The least-squares solution to the inconsistent system  $Ax = b$  is the solution of the normal equation

$$A^T A \bar{x} = A^T b$$

**Proposition 3.3.3 (3M).** If the columns of  $A$  are linearly independent (rank =  $n$ ), then  $A^T A$  is invertible and

$$\bar{x} = (A^T A)^{-1} A^T b$$

The projection of  $b$  onto  $\mathcal{C}(A)$  is therefore

$$p_{\mathcal{C}(A)} = A \bar{x} = A(A^T A)^{-1} A^T b$$

**Proof.** We consider  $\text{rank}(A) = r = n \Rightarrow \mathcal{N}(A) = \{0\} \Rightarrow \mathcal{C}(A^T A) = \{0\}$

$\therefore$  rank of  $A^T A = n$  which means  $A^T A$  has full rank.

$\therefore A^T A$  is invertible. ■

**Note.** If  $\text{rank}(A) < n$ , then  $A^T A$  is singular and the linear system has infinitely many solutions.

**Example.**

$$A = \begin{pmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{pmatrix}_{3 \times 2}, \quad b = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}_{3 \times 1}, \quad Ax = b$$

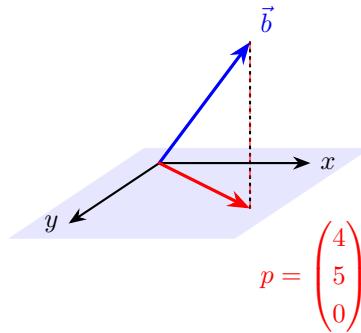


Figure 3.7: Projection of  $\vec{b}$  onto the span of the columns of  $A$

$$1^\circ \bar{x} = (A^T A)^{-1} A^T b \begin{pmatrix} 2 \\ 1 \end{pmatrix} \Rightarrow p = A\bar{x} = \begin{pmatrix} 4 \\ 5 \\ 0 \end{pmatrix}$$

$$2^\circ \mathcal{C}(A) = \text{span} \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix} \right\} = xy - \text{plane} \quad \therefore p = \begin{pmatrix} 4 \\ 5 \\ 0 \end{pmatrix}$$

**Remark.** 1° The normal equation  $A^T A \bar{x} = A^T b$  is indeed consistent.

2° If  $b \in \mathcal{C}(A)$ , then  $p = b$ .

3° Suppose  $b \perp \mathcal{C}(A)$ , then  $p = 0$ .

4° When  $A$  is square and invertible, then  $\mathcal{C}(A) = \mathbb{R}^n$

$$p = A(A^T A)^{-1} A^T b = b$$

5° If  $A_{m \times 1} = a$ , then  $A^T A = a^T a$

$$\bar{x} = (a^T a)^{-1} a^T b = \frac{a^T b}{a^T a} = \alpha$$

### 3.3.2 Projection Matrices

Let  $A$  be an  $m \times n$  matrix over  $\mathbb{R}$ ,  $\mathcal{C}(A) \leq \mathbb{R}^m$ .

Let  $b \notin \mathcal{C}(A)$ , the closest point to  $b$  in  $\mathcal{C}(A)$  is  $p = A(A^T A)^{-1} A^T b$ .

Let  $\mathcal{P} = A(A^T A)^{-1} A^T$ .

i.e. The matrix projects any vector  $b$  onto  $\mathcal{C}(A)$ .

i.e.  $p = \mathcal{P}b$  is the component of  $b$  in  $\mathcal{C}(A)$ .

i.e.  $b - \mathcal{P}b$  (error) is the component of  $b$  in orthogonal complement  $\mathcal{N}(A^T)$ .

**Corollary 3.3.1.**

$$I = \underset{\text{projection onto } \mathcal{C}(A)}{\mathcal{P}} + \underset{\text{projection onto } \mathcal{C}(A)^\perp}{(I - \mathcal{P})}$$

**Theorem 3.3.1 (3N).** Here are some properties of projection matrix

1.  $\mathcal{P}^2 = \mathcal{P}$
2.  $\mathcal{P}^T = \mathcal{P}$

### 3.3.3 Least Square Fitting of Data

$$C + Dt = b$$

Given  $m$  data points

$$\begin{cases} C + Dt_1 = b_1 \\ C + Dt_2 = b_2 \\ \vdots \\ C + Dt_m = b_m \end{cases} \Rightarrow \begin{pmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_m \end{pmatrix} \begin{pmatrix} C \\ D \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

$$\begin{aligned} \min E^2 &= \|b - (C + Dt)\|^2 = \|b - Ax\|^2 \\ &= \sum_{i=1}^m (b_i - C - Dt_i)^2 \end{aligned}$$

**Example.**  $(t_i, b_i) : (-1, 1), (1, 1), (2, 3)$

$$A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \\ 1 & 2 \end{pmatrix}, \quad x = \begin{pmatrix} C \\ D \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} \Rightarrow \bar{x} = \begin{pmatrix} \bar{C} \\ \bar{D} \end{pmatrix} = (A^T A)^{-1} A^T b = \begin{pmatrix} \frac{9}{7} \\ \frac{4}{7} \end{pmatrix}$$

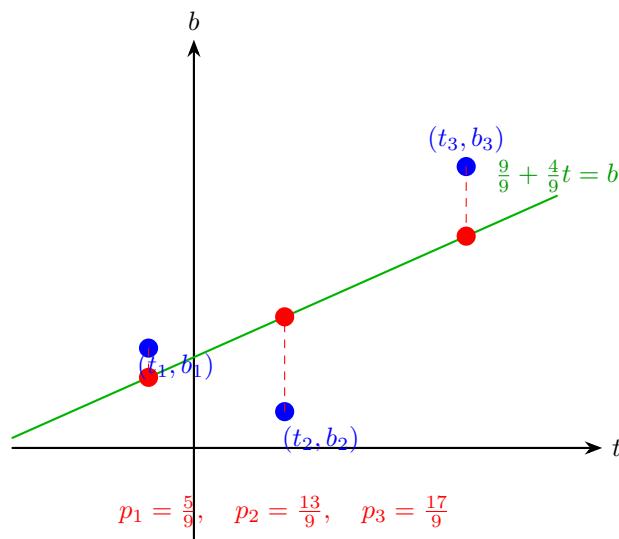


Figure 3.8: Least Squares Line Fitting

$$\mathcal{P} = A(A^T A)^{-1} A^T = \frac{1}{14} \begin{pmatrix} 13 & 3 & -2 \\ 3 & 5 & 6 \\ -2 & 6 & 10 \end{pmatrix}_{3 \times 3} \Rightarrow p = \mathcal{P}b = \frac{1}{7} \begin{pmatrix} 5 \\ 13 \\ 17 \end{pmatrix}$$

$$\therefore \text{error vector} = b - p = \frac{1}{7} \begin{pmatrix} 2 \\ -6 \\ 4 \end{pmatrix}$$

$$P = \frac{1}{14} \begin{pmatrix} 13 & 3 & -2 \\ 3 & 5 & 6 \\ -2 & 6 & 10 \end{pmatrix} \rightarrow U = \frac{1}{14} \begin{pmatrix} 13 & 3 & -2 \\ 0 & \boxed{\frac{56}{13}} & \frac{84}{13} \\ 0 & 0 & 0 \end{pmatrix} \Rightarrow \mathcal{C}(P) = \text{span} \left\{ \begin{pmatrix} 13 \\ 3 \\ -2 \end{pmatrix}, \begin{pmatrix} 3 \\ 5 \\ 6 \end{pmatrix} \right\} \Rightarrow x - 3y + 2z = 0$$

$$\mathcal{C}(A) = \text{span} \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix} \right\} \Rightarrow x - 3y + 2z = 0 \Rightarrow \mathcal{C}(P) = \mathcal{C}(A)$$

### 3.4 Orthogonal Bases, Orthogonal Matrices and Gram-Schmidt Orthogonalization

**Recall.** The vectors  $q_1, q_2, \dots, q_k$  are orthonormal if

$$q_i^T q_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

#### 3.4.1 Orthogonal Matrices

**Definition 3.4.1 (3Q).** An orthognral matrix  $Q$  is a square matrix satisfying  $Q^T Q = I$ . If  $Q = [q_1 \ q_2 \ \dots \ q_n]$ , then

$$Q^T Q = \begin{pmatrix} q_1^T \\ q_2^T \\ \vdots \\ q_n^T \end{pmatrix} (q_1 \ q_2 \ \dots \ q_n) = \begin{pmatrix} q_1^T q_1 & q_1^T q_2 & \cdots & q_1^T q_n \\ q_2^T q_1 & q_2^T q_2 & \cdots & q_2^T q_n \\ \vdots & \vdots & \ddots & \vdots \\ q_n^T q_1 & q_n^T q_2 & \cdots & q_n^T q_n \end{pmatrix}_{n \times n} = I_n$$

i.e. The columns of  $Q$  are orthonormal and  $Q^{-1} = Q^T$ .

**Example.** Here are some examples:

- Rotation matrix  $\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$  YES

- Permutation matrix  $\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$  YES

**Proposition 3.4.1 (3R).** Here are some properties

- $\|Qx\| = \|x\|, \forall x$
- $\langle Qx, Qy \rangle = \langle x, y \rangle, \forall x, y$

The properties preserve

1° length

2° inner product

3° angle (since  $\cos \theta = \frac{\langle x, y \rangle}{\|x\|\|y\|}$ )

**Remark.** Since  $Q^{-1} = Q^T$ , we also have  $QQ^T = I$ . Therefore, the rows of a square matrix are orthonormal whenever the columns are orthonormal.

We've learned that any vector is a combination of basis vectors. The problem becomes how to find the coefficients of the combination.

Let  $\{q_1, q_2, \dots, q_n\}$  be an orthonormal basis, then for any vector  $b$

$$b = x_1 q_1 + x_2 q_2 + \dots + x_n q_n$$

try to compute  $x_i$ 's:

$$q_1^T b = x_1 q_1^T q_1 + x_2 q_1^T q_2 + \dots + x_n q_1^T q_n = x_1$$

Similarly, we have  $x_i = q_i^T b, i = 1, 2, \dots, n$ . i.e.

$$b = (q_1^T b) q_1 + (q_2^T b) q_2 + \dots + (q_n^T b) q_n$$

for the matrix form

$$b = (q_1 \ q_2 \ \dots \ q_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = Qx \quad x = Q^{-1}b = Q^T b = \begin{pmatrix} q_1^T b \\ q_2^T b \\ \vdots \\ q_n^T b \end{pmatrix}$$

**Recall.**  $\mathcal{P} = \frac{a^T b}{a^T a} \cdot a$

Therefore, we can rewrite  $b$  as

$$b = \frac{q_1^T b}{q_1^T q_1} q_1 + \frac{q_2^T b}{q_2^T q_2} q_2 + \dots + \frac{q_n^T b}{q_n^T q_n} q_n = \mathcal{P}_{q_1} b + \mathcal{P}_{q_2} b + \dots + \mathcal{P}_{q_n} b \quad (\text{since } q_i^T q_i = 1)$$

i.e. The sum of the projections of  $b$  onto each basis vector equals to  $b$  itself.