高等线性代数及其应用

Advanced Linear Algebra and Its Applications

學院 数据科学学院（SDS）

2025 年 11 月 2 日

Abstract

这是线性代数关于**向量空间**与**线性变换**的部分，我们将介绍一些重要的代数概念。

这是线性代数中重要的前置概念。

我们参考了：以下内容 [1][2][3][4][5][6][7][8][9]

其中，我们着重参考了美国伊利诺伊大学 ***Friedberg*** 教授的 LA（4th）经典教材 [1] 和复旦大学谢启鸿教授主编的《高等代数学》[2]

并且在章节设计上，参考了香港中文大学（深圳）倪維明教授和罗才华教授的课程资料。

对此我们深表感谢！

目录

向量空间与线性映射

Chapter 1

向量空间

在正式介绍线性代数与高等代数的具体知识前，我们将先介绍以下前备知识和前垫概念：

1.1 线性代数的前备知识

在介绍线性代数之前，让我们先从“封闭”这个概念讲起：

【定义】：对于数集 S，若对任意 $a, b \in S$，都有 $a \pm b, ab, \frac{a}{b} \in S$，则称 S 是一个“**封闭**”的数集。

Example

实数集（域）$\mathbb{R}$、有理数集（域）$\mathbb{Q}$ 都是封闭的。

Exercise

无理数集封闭吗？
整数集封闭吗？

【定义】：**未知数/变量（variable）**：表达式中待确定值的符号或可以改变的量，通常用 x, y, z 字母等表示。

Example

表达式 $3x + 5$ 中，x 是未知数，也是变量。选择不同的 x 会给出不同的值（result value）。

【定义】：**多项式（Polynomial）**：由变量（如 x）和系数通过有限次加法、减法、乘法和正整数次幂运算得到的表达式。

Example

$3x^2+2x+1$ 就是关于变量 x 的多项式。

其中变量的最高次数，我们称之为多项式的“次数”。

在刚刚的例子中，次数为 2。

【定义】：**n 元一次方程（n-ary Linear Equation）：** 含有 n 个未知数，且每个未知数次数均为 1 的方程。

其中只含一个未知数且未知数（变量）的最高次数为 1 的方程，称为一次方程，也称线性方程（Linear Equation）。

Example

例如：$2x+3=7$ 是一元一次方程。

例如：$2x+3y=5$ 是二元一次方程。

在实践中，我们常常将多个 n 元一次线性方程组合起来就可以得到线性方程组：

【定义】：**n 元一次线性方程组（n-ary Linear Equations System with m equations）**：由 m 个 n 元一次方程组成的方程组。

Example

例如：$\begin{cases} x+y=1 \\ 2x-y=3 \end{cases}$

就是二元一次线性方程组（$n=2$，$m=2$）。

更一般的，我们来考察 $m\times n$ 的情节：

即由 m 行含有 n 个未知数的一次方程联立而成的方程组：

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1 \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2 \\ \vdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_m \end{cases}$$

其中：a_{ij} 是第 i 个方程中第 j 个未知数的系数；$b_1,b_2,...,b_m$ 是常数项。

由此，我们将这些未知数前面的系数提取并且排列出来，就可以得到“系数矩阵”和“增广矩阵”的概念。

【定义】：**系数矩阵（Coefficient Matrix）：** 由方程组的系数 a_{ij} 构成的 $m\times n$ 矩阵：

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

简记为 $A=(a_{ij})_{m\times n}$。

【定义】：增广矩阵（Augmented Matrix）：在系数矩阵 A 的右侧增加一列常数项（即原方程等式右边的所有元素）$b_1, b_2, \ldots, b_m$ 构成的 $m\times(n+1)$ 矩阵：

$$\tilde{A}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix}$$

简记为 $\tilde{A}=[A\mid b]$，其中 $b=(b_1,b_2,\ldots,b_m)^T$ 是常数项列向量。

矩阵形式表示上述方程组可简洁地表示为矩阵方程：$Ax=b$

其中：A 是系数矩阵（$m\times n$）；$x=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ 是未知数向量（$n\times 1$）；$b=\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$ 是常数项向量（$m\times 1$）。

在刚刚的例子：$\begin{cases} x+y=1 \\ 2x-y=3 \end{cases}$ 中

上述方程组的系数矩阵为 $\begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix}$。

上述方程组的增广矩阵为 $\begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & 3 \end{pmatrix}$。

更一般的，我们通过这种 m×n 的形式，进一步地抽象出一般矩阵的概念：

【定义】：矩阵（Matrix）：由 $m\times n$ 个数 a_{ij} 排成的 m 行 n 列的数表，记为 $A=(a_{ij})_{m\times n}$。

接着我们将讨论一种作用在矩阵上的常规操作：**转置**（Transpose）

【定义】矩阵的转置：将矩阵的行与列按顺序互换，即把原矩阵的第 i 行变为转置后矩阵的第 i 列，第 j 列变为转置后的第 j 行。

用数学符号表示：若原矩阵为 A，其转置矩阵记为 A^T 或者 ${}^tA$

设 A 是一个 $m\times n$ 的矩阵（m 行 n 列），则 A^T 是一个 $n\times m$ 的矩阵，满足：$(A^T)_{i,j}=A_{j,i}$ 其中 $(A^T)_{i,j}$ 表示转置矩阵中第 i 行第 j 列的元素，$A_{j,i}$ 表示原矩阵中第 j 行第 i 列的元素。

Example

例如：设矩阵 A 为：$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$

这是一个 2×3 的矩阵（2 行 3 列）。

其转置矩阵 A^T（或 ${}^tA$ ）为：$A^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$

接下来，我们将考察一些特殊的矩阵：

【定义】：**方阵（Square Matrix）**：行数 m 等于列数 n 的矩阵，即 $m \times m$ 矩阵。

例子：$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$。就是一个三阶方阵。

【定义】：**对角阵（Diagonal Matrix）**：主对角线以外元素全为 0 的方阵，记为 $\text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$。

例子：$\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$。

【定义】：**单位矩阵（Identity Matrix）**：主对角线元素全为 1，其余元素全为 0 的 n 阶方阵，记为 I_n 或 E_n。

例如二阶单位矩阵：主对角线元素全为 1，其余元素全为 0 的 2×2 方阵 $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$。

又如三阶单位矩阵：主对角线元素全为 1，其余元素全为 0 的 3×3 方阵，记为 I_3 或 E_3。$I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

我们现在不加证明的给出：

Fact or Background

单位阵类似于“1”，对于任意的 n 阶方阵 A，我们有：$IA = AI = A$

我们后续将在矩阵乘法中详细解释。

【定义】：**数量矩阵（Scalar Matrix）**：主对角线元素相同，其余元素全为 0 的方阵。（即 kI，k 属于 F）

例子：$\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$。

除了对矩阵元素的研究，我们也可以对矩阵元素的排布与结构进行进一步研究：

- 定义：$W_+ := \{A \in M_n(\mathbb{F}) : {}^tA = A\} \subset M_n(\mathbb{F})$（对称矩阵集合）：
  - 解释：W_+ 是域 $\mathbb{F}$ 上所有 n 阶对称矩阵（symmetric matrix）构成的集合，其中 ${}^tA$ 表示矩阵 A 的转置（这种方式也可以表达转置），对称矩阵满足转置后与自身相等。
  - 例如（取 $\mathbb{F} = \mathbb{R}$，$n = 2$）：矩阵 $A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$，其转置 ${}^tA = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} = A$，故 $A \in W_+$。
- 定义：$W_- := \{A \in M_n(\mathbb{F}) : {}^tA = -A\} \subset M_n(\mathbb{F})$（反对称矩阵集合）：
  - 解释：W_- 是域 $\mathbb{F}$ 上所有 n 阶反对称矩阵构成的集合，反对称矩阵满足转置后与自身的相反数相等，且**主对角线元素全为** 0（因 $(A)_{ii} = -(A)_{ii}$，故 $(A)_{ii} = 0$）。
  - 例如（取 $\mathbb{F} = \mathbb{R}$，$n = 2$）：矩阵 $B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$，其转置 ${}^tB = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = -B$，故 $B \in W_-$。

在介绍完矩阵后，我们可以进入对行列式的讨论：

行列式的概念最早源于对线性方程组的研究，但是在线性代数的发展历程中又起到了重要的作用。接下来，我们将研究一类特殊的线性方程组（n 个未知数，n 个方程）与其构成的方阵，从而引入行列式的概念。

我们将在后续的讲义中解释行列式的性质及由来（其内禀定义）。

1.2 向量与向量空间

这个世界上有许多不同的实体（如向量、函数等），而他们往往具有**两种运算（“加法” 和 “标量乘法”）的相同规则**，由此通过将它们抽象为向量空间，我们只需要研究这个通用框架即可了解他们的**共性**。

1.2.1 向量空间与向量的定义

向量空间（vector space）是线性代数中的核心概念。设 V 是一个非空集合，F 是一个数域。在 V 上定义以下两种运算：

- “加法 addition”：对于任意 $\alpha, \beta \in V$，存在唯一的 $\alpha + \beta \in V$。

- "数乘 scalar multiplication"：对于任意 $k \in F$ 和 $\alpha \in V$，存在唯一的 $k\alpha \in V$。

则称定义在数域 F 上的空间 V 封闭。

【Remark】：

该加法与数乘不一定是传统意义上的加法与数乘。

Definition

如果 V 在满足封闭的前提下，同时满足以下八条公理：

1. 加法交换律：$\alpha + \beta = \beta + \alpha$。
2. 加法结合律：$(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$。
3. 零元存在：存在 $0 \in V$，使得对于任意 $\alpha \in V$，$\alpha + 0 = \alpha$。
4. 负元存在：对于任意 $\alpha \in V$，存在 $\beta \in V$，使得 $\alpha + \beta = 0$。
5. 数乘单位元：$1 \cdot \alpha = \alpha$。
6. 数乘结合律：$k(l\alpha) = (kl)\alpha$。
7. 数乘分配律 1：$k(\alpha + \beta) = k\alpha + k\beta$。
8. 数乘分配律 2：$(k + l)\alpha = k\alpha + l\alpha$。

满足上述条件时，称 V 是数域 F 上的向量空间，记作 $(V, F, +, \cdot)$。

向量空间的定义可简化记忆为"**存在零元，两种运算，八条公理**"，其中**加法和数乘需满足封闭性**。

接下来我们会举几个典型的向量空间的例子来促进理解：

- 实数域上的 n 维向量空间 $\mathbb{R}^n$ 定义：由 n 个实数组成的有序数组 $(x_1, x_2, \ldots, x_n)$ 的集合，其中加法和数乘按分量进行。

  Example

  - $\mathbb{R}^2$：二维平面上的所有向量（如坐标点 (x, y)）；
  - $\mathbb{R}^3$：三维空间中的所有向量（如坐标点 (x, y, z)）。

- 同样的，可以得到复数域上的 n 维向量空间 $\mathbb{C}^n$
- 零空间 $\{0\}$，定义：仅包含零向量的空间，维数为 0。
- $M_{m\times n}(\mathbb{F})$：

- 解释：域 $\mathbb{F}$ 上所有 m 行 n 列矩阵的集合。
- 验证：
  * 加法封闭：两 $m \times n$ 矩阵相加，结果仍为 $m \times n$ 矩阵。
  * 数乘封闭：域中数乘 $m \times n$ 矩阵，结果仍为 $m \times n$ 矩阵。
  * 其他公理（交换律、结合律等）由矩阵运算规则保证。

- $\mathcal{P}_n(\mathbb{F})$：
  - 解释：域 $\mathbb{F}$ 上次数不超过 n 的一元多项式集合，形如 $\sum_{i=0}^{n} a_i x^i$（$a_i \in \mathbb{F}$）。
  - 验证：
    * 加法封闭：两次数不超过 n 的多项式相加，次数仍不超过 n。
    * 数乘封闭：域中数乘次数不超过 n 的多项式，次数仍不超过 n。
    * 其他公理（交换律、结合律等）由多项式运算规则保证。

Definition

向量（vector）： 给定一个向量空间，其空间中的元素，我们称之为向量。

向量空间的元素（即向量）的具体形式（如数组、函数、矩阵等）由具体的研究场景决定。

1.2.2 向量空间中的线性运算

什么是线性运算?

满足上述向量空间八条公理的加法和数乘称为线性运算。

线性运算包含加法和数乘运算：

- **加法：** 对于向量空间 V 中的向量 α, β，它们的和 $\alpha + \beta$ 是按照向量空间所定义的加法规则得到的 V 中的唯一向量。

- **数乘：** 对于数域 F 中的数 k 和向量空间 V 中的向量 α，数乘 $k\alpha$ 是按照向量空间所定义的数乘规则得到的 V 中的唯一向量。

凡是符合上述向量空间中加法和数乘这两种运算性质的运算，我们称之为线性运算。

所以，我们可以给出论断：

Claim

向量空间的本质上是由集合与满足特定公理的线性运算共同构成的一种代数结构，二者缺一不可。

若集合 V 上定义的加法或数乘运算不满足线性运算的 8 条公理，则构成的结构不再是向量空间；

若集合 V_1 和集合 V_2 定义的运算满足公理但具体的定义方式不同（如调整数乘的域或加法规则），则形成不同的向量空间。

Example

例如：集合 $\mathbb{R}^n$ 按标准加法与实数乘构成实向量空间 $\mathbb{R}^n$；
若将数乘的域改为复数域 $\mathbb{C}$，则构成复向量空间 $\mathbb{C}^n$，二者为不同代数结构。

1.2.3 向量空间的五条重要性质

在介绍完向量空间的基本定义，窥探其本质后，我们可以对其数学性质做进一步挖掘：得到其 5 条重要的性质与推论。

Property

在向量空间中成立 $(-1)\cdot\boldsymbol{v}=-\boldsymbol{v}$ 和 $0\cdot\boldsymbol{v}=\boldsymbol{0}$
其中，v 为向量空间中的任意向量。

1. 证明 $0\cdot\boldsymbol{v}=\boldsymbol{0}$:

$$\begin{aligned}\boldsymbol{v}+0\cdot\boldsymbol{v}&=1\cdot\boldsymbol{v}+0\cdot\boldsymbol{v}\\&=(1+0)\cdot\boldsymbol{v}\\&=1\cdot\boldsymbol{v}=\boldsymbol{v}.\end{aligned}$$

   Add $-\boldsymbol{v}$ to both sides:

$$(\boldsymbol{v}+0\cdot\boldsymbol{v})+(-\boldsymbol{v})=\boldsymbol{v}+(-\boldsymbol{v})\implies 0\cdot\boldsymbol{v}=\boldsymbol{0}.$$

2. 证明 $(-1)\cdot\boldsymbol{v}=-\boldsymbol{v}$:

$$\begin{aligned}\boldsymbol{v}+(-1)\cdot\boldsymbol{v}&=1\cdot\boldsymbol{v}+(-1)\cdot\boldsymbol{v}\\&=(1+(-1))\cdot\boldsymbol{v}\\&=0\cdot\boldsymbol{v}=\boldsymbol{0}.\end{aligned}$$

   因此，$(-1)\cdot\boldsymbol{v}$ 是 $\boldsymbol{v}$ 的加法逆元，所以 $(-1)\cdot\boldsymbol{v}=-\boldsymbol{v}$。

接着我们给出其它三条重要性质：

Property

1. **零向量唯一**: 对于任意 $\alpha \in V$，若 $\alpha + 0_1 = \alpha$ 且 $\alpha + 0_2 = \alpha$，令 $\alpha = 0_1$ 代入 $\alpha + 0_2 = \alpha$ 可得 $0_1 + 0_2 = 0_1$，再令 $\alpha = 0_2$ 代入 $\alpha + 0_1 = \alpha$ 可得 $0_2 + 0_1 = 0_2$，所以 $0_1 = 0_2$，即零向量是唯一的。

2. **任一向量其负向量唯一**: 若 $\alpha+\beta = 0$ 且 $\alpha+\gamma = 0$, 则 $\beta = \beta+0 = \beta+(\alpha+\gamma) = (\beta+\alpha)+\gamma = 0+\gamma = \gamma$，所以任一向量的负向量是唯一的。

3. **若 $\lambda\alpha = 0$，则 $\lambda = 0$ 或 $\alpha = 0$**: 假设 $\lambda \neq 0$，因为数域 F 中数的运算性质，λ 存在逆元 λ^{-1}，在 $\lambda\alpha = 0$ 两边同时左乘 λ^{-1}，可得 $\lambda^{-1}(\lambda\alpha) = \lambda^{-1}\cdot 0$，根据数乘结合律和 $0\cdot\alpha = 0$，得到 $(\lambda^{-1}\lambda)\alpha = 0$，即 $1\cdot\alpha = \alpha = 0$，所以若 $\lambda\alpha = 0$，则 $\lambda = 0$ 或 $\alpha = 0$。

1.2.4 向量空间的判断

判断给定集合 V 是否为向量空间，一般按以下步骤：

1. 考察加法和数乘的封闭性，即检查运算结果是否仍在集合 V 内。
2. 考察零元是否存在于集合 V 中。
3. 逐一考察验证八条公理是否成立。

非向量空间的例子)：

Example

- 集合 $V = \{(x,y) \in \mathbb{R}^2 | x+y = 1\}$ 不是向量空间, 因为若 (x_1,y_1) 和 (x_2,y_2) 满足 $x_1+y_1 = 1$ 且 $x_2+y_2 = 1$, (x_1+x_2, y_1+y_2) 不一定满足 $(x_1+x_2)+(y_1+y_2) = 1$，即对加法不封闭，并且该集合不包含零向量。
- 次数严格等于 n 的多项式集合不是向量空间, 因为对加法不封闭。例如: $(x^n + x) + (-x^n + 1) = x + 1$，结果不再是次数严格等于 n 的多项式。

【Remark】:

注意到，多项式空间 $P_n[x]$（次数不超过 n 的实系数多项式全体）是向量空间。

1.2.5 课后习题

Exercise

1. 设 V 是向量空间，$\alpha, \beta \in V$。证明：方程 $\alpha + X = \beta$ 在 V 中有唯一解。
2. 在 $\mathbb{R}^2$ 中，定义 $(a_1, a_2) + (b_1, b_2) = (a_1 + b_1, a_2 b_2)$ 和 $k(a_1, a_2) = (ka_1, a_2)$，证明其不构成向量空间。
3. 设集合 $V = \{(x, y)|x \geq 0, y \in \mathbb{R}\}$，定义普通的向量加法和实数域上的数乘运算。证明 V 不是线性空间。

1.3 子空间

1.3.1 子空间的定义与判定

Definition

子空间（subspace）： 设 V 是数域 F 上的向量空间，若 V 的子集 W 对 V 中定义的加法和数乘运算也构成数域 F 上的向量空间，则称 W 是 V 的子空间。

在这里，沿用上节向量空间的知识，我们不加证明的的给出子空间的判定条件：

Method

【子空间的判定条件】：若定义在数域 F 上的向量空间 V 的子集 W 是子空间，当且仅当以下 2 个判定条件成立：

1. 对任意 $w_1, w_2 \in W$ 以及任意 $a, b \in \mathbb{F}$，都有 $aw_1 + bw_2 \in W$;
2. 包含零向量：W 中存在零向量 $\mathbf{0}$。

我们来考察以下几个例子：

- 对于 $\mathbb{R}^2$ 中 $W = \{(x, y) \mid x + y = 0\}$：
  - 取 $w_1 = (a, -a)$，$w_2 = (b, -b) \in W$，任取 $a, b \in \mathbb{R}$，计算 $aw_1 + bw_2 = (a^2 + b^2, -a^2 - b^2)$，满足 $x + y = 0$，故 $aw_1 + bw_2 \in W$。
  - W 包含零向量 $(0, 0)$（当 $x = y = 0$ 时，$x + y = 0$ 成立）。

  因此 W 是 $\mathbb{R}^2$ 的子空间。
- 对于 $\mathbb{R}^2$ 中 $W_1 = \{(x, 0) \mid x \in \mathbb{R}\}$：
  - 取 $w_1 = (x_1, 0)$, $w_2 = (x_2, 0) \in W_1$, 任取 $a, b \in \mathbb{R}$, $aw_1 + bw_2 = (ax_1 + bx_2, 0) \in W_1$。

  - W_1 包含零向量 $(0,0)$。

    因此 W_1 是 $\mathbb{R}^2$ 的子空间。

- 对于 $\mathbb{R}^2$ 中 $W_2=\{(x,x)\mid x\in\mathbb{R}\}$：

  - 取 $w_1=(x_1,x_1)$，$w_2=(x_2,x_2)\in W_2$，任取 $a,b\in\mathbb{R}$，$aw_1+bw_2=(ax_1+bx_2,ax_1+bx_2)\in W_2$。

  - W_2 包含零向量 $(0,0)$。

    因此 W_2 是 $\mathbb{R}^2$ 的子空间。

- 对于多项式空间 $P_n(x)$ 中 $W_3=\{a_nx^n+\cdots+a_1x\mid a_i\in\mathbb{R},i=1,\ldots,n\}$：

  - 取 $p(x)=a_nx^n+\cdots+a_1x$，$q(x)=b_nx^n+\cdots+b_1x\in W_3$，任取 $a,b\in\mathbb{R}$，$ap(x)+bq(x)=(aa_n+bb_n)x^n+\cdots+(aa_1+bb_1)x\in W_3$。

  - W_3 包含零多项式（所有系数为 0 的多项式）。

    因此 W_3 是 $P_n(x)$ 的子空间。

- 对于多项式空间 $P_n(x)$ 中 $W_4=\{a_nx^n\mid a_n\in\mathbb{R}\}$：

  - 取 $f(x)=a_nx^n$，$g(x)=b_nx^n\in W_4$，任取 $a,b\in\mathbb{R}$，$af(x)+bg(x)=(aa_n+bb_n)x^n\in W_4$。

  - W_4 包含零多项式（当 $a_n=0$ 时）。

    因此 W_4 是 $P_n(x)$ 的子空间。

- 对于矩阵空间 $M_{m\times n}(\mathbb{R})$ 中 W_5（第一行任意、其余行全为 0 的矩阵集合）：

  - 取 $A=\begin{pmatrix}a_{11}&\ldots&a_{1n}\\0&\ldots&0\\\vdots&\ddots&\vdots\\0&\ldots&0\end{pmatrix}$，$B=\begin{pmatrix}b_{11}&\ldots&b_{1n}\\0&\ldots&0\\\vdots&\ddots&\vdots\\0&\ldots&0\end{pmatrix}\in W_5$，任取 $a,b\in\mathbb{R}$，$aA+bB=\begin{pmatrix}aa_{11}+bb_{11}&\ldots&aa_{1n}+bb_{1n}\\0&\ldots&0\\\vdots&\ddots&\vdots\\0&\ldots&0\end{pmatrix}\in W_5$。

  - W_5 包含零矩阵（所有元素为 0 的矩阵）。

    因此 W_5 是矩阵空间 $M_{m\times n}(\mathbb{R})$ 的子空间。

接下来是一个经典的例子：

Example

$\mathbb{R}_+ \subset \mathbb{R}$ 不是 $\mathbb{R}$ 的子空间。
其中 $R_+ = \{x \in R\ x > 0\}$

$\mathbb{R}$ 作为定义在实数域上的向量空间，运算为普通实数加法和实数对向量的数乘。但是 $\mathbb{R}_+$ 不满足：

零向量的存在性：$\mathbb{R}$ 的零向量是 0，但 $\mathbb{R}_+ = \{x > 0\}$ 不包含 0。

数乘封闭性：取 $x = 1 \in \mathbb{R}_+$，数域中取 $a = -1 \in \mathbb{R}$，则数乘结果为 $a \cdot x = -1 \times 1 = -1 \notin \mathbb{R}_+$，说明数乘不封闭。

重新定义 $\mathbb{R}_+$ 上的加法和数乘（脱离 $\mathbb{R}$ 的普通运算），使它满足向量空间公理：定义新运算：

加法：对任意 $x, y \in \mathbb{R}_+$，定义：$x \oplus y := x \cdot y$（用实数乘法代替加法）。

数乘：对任意 $x \in \mathbb{R}_+$，$a \in \mathbb{R}$（数域），定义 $a \odot x := x^a$（用实数的幂运算代替数乘）。

即可满足。

1.3.2　课后练习

Exercise

设 $V = M_{n\times n}(\mathbb{C})$ 是复 $n \times n$ 矩阵构成的 $\mathbb{C}$-向量空间，我们定义

$$W_{*,+} := \{A \in V : {}^t\overline{A} = A\} \quad 和 \quad W_{*,-} := \{A \in V : {}^t\overline{A} = -A\}$$

这里 ${}^t\overline{A} = {}^t\overline{(a_{ij})}_{i,j} := (\overline{a}_{ji})_{i,j}$，其中 $\overline{a}_{ij} = \overline{x + \sqrt{-1}y} := x - \sqrt{-1}y \in \mathbb{C}$。
证明 $W_{*,+}$ 是 V 的 $\mathbb{R}$-向量子空间；
证否 $W_{*,-}$ 是 V 的 $\mathbb{C}$-向量子空间。

1.4 线性组合与张成

1.4.1 线性组合与线性表出

Definition

线性组合（linear combination）： 在向量空间 V 中，对于非空子集 S，若向量 $v \in V$ 能够表示为 S 中有限个向量 $u_1, u_2, \cdots, u_n$ 与数域 F 中系数 $a_1, a_2, \cdots, a_n$ 的组合形式，即

$$v = a_1 u_1 + a_2 u_2 + \cdots + a_n u_n,$$

则称 v 是 S 中向量的线性组合，
其中 $a_1, a_2, \cdots, a_n$ 被称为组合系数。

同样的，我们称 v 可以有 S 中向量的线性表出（Linearly represented）。

根据向量空间的运算性质，对于任意 $v \in V$，都有 $0v = 0$，**这表明零向量是任何非空子集的线性组合**。

【Remark】:

- 线性组合体现了向量空间中向量之间通过数乘与加法构建的联系。这种组合方式基于向量空间所定义的数乘和加法运算规则。
- 组合系数来自于数域 F，数域的性质（如有理数域、实数域、复数域等不同数域）会影响向量线性组合的表现形式和性质。

接下来我们考察两个经典的例子：

在二维向量空间 $\mathbb{R}^2$ 中，对于子集 $S = \{(1,0), (0,1)\}$，向量 $\vec{v} = (3,4)$ 可表示为 $\vec{v} = 3 \times (1,0) + 4 \times (0,1)$，这里 3 和 4 作为组合系数，清晰地展示了 $\vec{v}$ 是 S 中向量的线性组合。

在多项式向量空间 $P_2(\mathbb{R})$（由次数不超过 2 的实系数多项式构成）中，对于子集 $S = \{1, x, x^2\}$，多项式 $p(x) = 2 - 3x + 5x^2$ 可表示为 $p(x) = 2 \times 1 + (-3) \times x + 5 \times x^2$。这表明在多项式向量空间中，通过基向量（这里的 $1, x, x^2$ 类似于向量空间的基）的线性组合可以得到该空间内的具体多项式。

1.4.2 张成（Spanning）

Definition

张成（spanning）：

- （定义 1）：设 S 是向量空间 V 的非空子集，S 的张成 $\mathrm{span}(S)$ 是由 S 中所有向量的线性组合所构成的集合。即：若 $\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n \in V$，则：

$$\mathrm{Span}(\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n) = \{\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_n\vec{v}_n \mid \alpha_1, \alpha_2, \cdots, \alpha_n \in F\}$$

  特别规定 $\mathrm{span}(\varnothing) = \{0\}$，即空集的张成是仅包含零向量的集合。

- （定义 2）：$\mathrm{span}(S)$ 也可被视作 V 中包含子集 S 的最小子空间。

  这意味着一方面 $\mathrm{span}(S)$ 是 V 的子空间，满足子空间的判定条件；另一方面，任何其他包含 S 的 V 的子空间必然也包含 $\mathrm{span}(S)$，体现了 $\mathrm{span}(S)$ 在包含 S 的子空间中的最小性。

容易观察到，向量空间的向量集合张成的结果依然是一个向量空间，满足运算封闭性和 8 条公理。

接下来，我们考察几个典型的例子：

- 在 $P_2(\mathbb{R})$ 中，取 $S = \{1, x, x^2\}$，$\mathrm{span}(S)$ 包含所有形如 $a + bx + cx^2$（$a, b, c \in \mathbb{R}$）的多项式，即 $\mathrm{span}(S) = P_2(\mathbb{R})$，且 span（S）的结果二次多项式空间的全空间。

- 在 $\mathbb{R}^3$ 里，设 $\vec{v} = (1, 2, 3)$，$\mathrm{span}(\{\vec{v}\}) = \{k(1, 2, 3) \mid k \in \mathbb{R}\}$，几何上是过原点且与 $\vec{v}$ 共线的直线，该直线是 $\mathbb{R}^3$ 的一个子空间（见前节）。

- 在 $\mathbb{R}^3$ 中，集合 $\{(1, 0, 0), (0, 1, 0)\}$ 的张成 $\mathrm{span}\{(1, 0, 0), (0, 1, 0)\}$ 是 $\mathbb{R}^3$ 中形如 $a(1, 0, 0) + b(0, 1, 0) = (a, b, 0)$（$a, b \in \mathbb{R}$）的所有向量构成的集合，从几何角度看，它张成的是一个平面，对应于 xy-平面，是 $\mathbb{R}^3$ 的一个子空间。这一示例直观地展示了张成集合在三维空间中的几何形态以及作为子空间的性质。

- 在矩阵向量空间 $M_{2\times2}(\mathbb{R})$（所有 2×2 实矩阵构成的向量空间）中，对于子集 $S = \left\{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\right\}$，其张成 $\mathrm{span}(S)$ 是所有形如 $a\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix}$（$a, b \in \mathbb{R}$）的矩阵集合，它是 $M_{2\times2}(\mathbb{R})$ 的一个子空间。此例将张成的概念拓展到矩阵向量空间，展示了在不同类型向量空间中张成集合的构建方式和性质。

由此，我们还可以证明一条有用的结论：

Claim

当域 $\mathbb{F}$ 的特征 $\mathrm{char}(\mathbb{F}) \neq 2$ 时，n 阶全矩阵空间 $M_n(\mathbb{F})$ 等于 $\mathbb{F}$ 上由对称矩阵空间 W_+ 和反对称矩阵空间 W_- 张成的线性空间

证明. • 对任意矩阵 $A \in M_n(\mathbb{F})$，构造两个矩阵：

$$A_+ = \frac{1}{2}(A + A^T), \quad A_- = \frac{1}{2}(A - A^T)$$

当 $\mathrm{char}(\mathbb{F}) \neq 2$ 时，2 在 $\mathbb{F}$ 中可逆（即 $\frac{1}{2}$ 存在），因此 A_+ 和 A_- 都是 $\mathbb{F}$ 上的合法矩阵。

- 验证 A_+ 是对称矩阵，A_- 是反对称矩阵：
  - 对 A_+，有：

$$(A_+)^T = \left(\frac{1}{2}(A + A^T)\right)^T = \frac{1}{2}(A^T + (A^T)^T) = \frac{1}{2}(A^T + A) = A_+$$

  - 对 A_-，有：

$$(A_-)^T = \left(\frac{1}{2}(A - A^T)\right)^T = \frac{1}{2}(A^T - (A^T)^T) = \frac{1}{2}(A^T - A) = -A_-$$

- **证明 $M_n(\mathbb{F}) = \mathbf{Span}_{\mathbb{F}}(W_+ \cup W_-)$：**
  - 由分解式 $A = A_+ + A_-$ 可知，任意矩阵都能表示为对称矩阵与反对称矩阵的和，因此 $M_n(\mathbb{F}) \subseteq \mathrm{Span}_{\mathbb{F}}(W_+ \cup W_-)$。
  - 另一方面，对称矩阵和反对称矩阵都是全矩阵空间的元素，它们的张成空间自然包含于全矩阵空间，即 $\mathrm{Span}_{\mathbb{F}}(W_+ \cup W_-) \subseteq M_n(\mathbb{F})$。
  - 综上，两者相等。

- **关于特征不能等于 2 的说明：**
  - 当 $\mathrm{char}(\mathbb{F}) = 2$ 时，域中元素满足 $1 + 1 = 0$，即 $2 = 0$，此时 $\frac{1}{2}$ 无意义，上述分解式 $A_+ = \frac{1}{2}(A + A^T)$ 和 $A_- = \frac{1}{2}(A - A^T)$ 无法定义。
  - 进一步，在特征为 2 的域中，$-1 = 1$，因此反对称矩阵与对称矩阵的定义变得一致（$A^T = -A$ 等价于 $A^T = A$），导致 $W_+ = W_-$，它们的张成空间无法覆盖整个全矩阵空间。
  - 因此，只有当 $\mathrm{char}(\mathbb{F}) \neq 2$ 时，上述分解和结论才成立。

□

1.4.3 课后练习

Exercise

证明 $\mathbb{F}[x]=\mathrm{Span}_{\mathbb{F}}(W_e\cup W_o)$，其中 $W_e:=\{f(x)\in\mathbb{F}[x]\mid f(x)=f(-x)\}$（偶函数多项式集合），$W_o:=\{f(x)\in\mathbb{F}[x]\mid f(x)=-f(-x)\}$（奇函数多项式集合）。（请特别注意 $\mathrm{char}(\mathbb{F})=2$ 的情形）

1.5 线性相关性

线性相关

Definition

线性相关（Linearly dependent）： 若向量空间 V 的子集 S 中存在**有限个**不同向量 $u_1,u_2,\cdots,u_n$ 以及**不全为零**的数域 F 中的系数 $a_1,a_2,\cdots,a_n$，使得

$$a_1u_1+a_2u_2+\cdots+a_nu_n=0,$$

则称子集 S 是**线性相关的**，也说 S 中的**向量线性相关**。

【判定方法】：对于给定向量集合，可通过建立线性方程组，看是否存在非零解来判断线性相关性。

接下来，我们将展示两个例子，以加深理解：

Example

- 在 $\mathbb{R}^2$ 中，向量集合 $S=\{(1,2),(2,4)\}$ 是线性相关的。设 $a(1,2)+b(2,4)=(0,0)$，即 $(a+2b,2a+4b)=(0,0)$，得到方程组 $\begin{cases}a+2b=0\\2a+4b=0\end{cases}$，取 $a=-2$，$b=1$ 就是一组非零解，说明存在零向量的非平凡表示，所以向量集合 S 线性相关。
- 在多项式向量空间 $P_2(\mathbb{R})$ 中，集合 $S=\{1+x,2+2x,x^2\}$ 是线性相关的。设 $a(1+x)+b(2+2x)+cx^2=0$（这里 0 是零多项式），整理得 $(a+2b)+(a+2b)x+cx^2=0$，取 $a=2$，$b=-1$，$c=0$ 就是一组非零解，表明该集合线性相关。

线性无关

Definition

线性无关（Linearly independent）： 若向量空间 V 的子集 S 不是线性相关的，则称 S 是线性无关的，也说 S 中的向量线性无关。

即：若 $a_1 = a_2 = \cdots = a_n = 0$ 时才有 $a_1u_1 + a_2u_2 + \cdots + a_nu_n = 0$，称这种情况为线性无关。

接下来我们对非零向量与空集进行辨析：

Property

- 空集是线性无关的，因为线性相关集合必须非空。（空集不等于 0 集）
- 由单个非零向量构成的集合是线性无关的。因为若 $\{u\}$ 线性相关，则存在非零标量 a 使 $au = 0$，那么 $u = a^{-1}(au) = a^{-1}0 = 0$，与 u 非零矛盾。

让我们来考察接下来两个例子，加深对线性无关的理解：

Example

- 在 $\mathbb{R}^3$ 中，向量集合 $S = \{(1,0,0),(0,1,0),(0,0,1)\}$ 是线性无关的。设 $a(1,0,0)+b(0,1,0)+c(0,0,1) = (0,0,0)$，即 $(a,b,c) = (0,0,0)$，只能得到 $a = b = c = 0$，所以该集合线性无关。
- 在定义在 $[0,1]$ 上的连续函数构成的向量空间 $C[0,1]$）里，集合 $S = \{1,x,x^2\}$ 是线性无关的。假设 $a \times 1 + b \times x + c \times x^2 = 0$（这里 0 是零函数）对所有 $x \in [0,1]$ 都成立，根据多项式函数的性质，只有 $a = b = c = 0$ 时等式才恒成立，所以集合 S 线性无关。

线性无关性在线性代数中具有相当重要的地位，我们接下来将讨论本节的两个主定理：

Theorem

定理 1： 设 V 是一个向量空间，且 $S_1 \subseteq S_2 \subseteq V$。若 S_1 线性相关，则 S_2 线性相关；且若 S_2 线性无关，则 S_1 线性无关。

Theorem

定理 2： 设 S 是向量空间 V 的一个线性无关子集，v 是 V 中一个不属于 S 的向量。那么 $S \cup \{v\}$ 线性相关当且仅当 $v \in \mathrm{span}(S)$。

1.6 基与维度

1.6.1 基的定义

Definition

基/基底（basis)： 向量空间 V 的一个基底 β 是 V 的一个**线性无关子集**，并且它**生成**（generates or spans）V。
基向量（Base vectors)： 如果 β 是 V 的一个基底，我们也说 β 里面的向量是 V 的基向量。

【Remark】:

“基”和“基底”是同一概念的不同中文表述，二者的英文原文均为“basis”（复数形式为“bases”）

因为涉及佐恩引理，极大元和塔的概念, 我们这里不加证明地给出：

Claim

每个向量空间都有一个基，这一论断。

因此，我们对基底应该有如下的认识：
基不仅仅是可以张成全空间的向量组合，同时这个组合线性无关，即判断条件为：
1. 张成全空间；
2. 线性无关性
用下图生动反映了基底与生成集与线性无关集的关系：

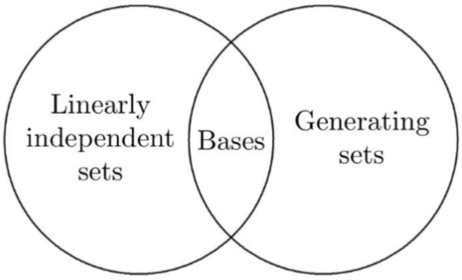

图 1.1: 基和生成集与线性无关集的关系

Example

在 $\mathbb{F}^n$ 中，令 $e_1 = (1,0,0,\ldots,0)$，$e_2 = (0,1,0,\ldots,0)$，$\ldots$，$e_n = (0,0,\ldots,0,1)$；$\{e_1,e_2,\ldots,e_n\}$ 就是 $\mathbb{F}^n$ 的一个基底，被称为 $\mathbb{F}^n$ 的标准基底（standard basis）。
在 $P_n(F)$ 中，集合 $\{1,x,x^2,\ldots,x^n\}$ 是一个基底。我们也称这个基底为 $P_n(F)$ 的标准基底。

【Remark】:

一个常见的误区是：假设 $S=\{s_1,s_2,\cdots,s_n\}$ 是一组基，但是不考虑 V 是否为有限维

接下来，我们给出另一个易错的论断：

Claim

子空间与全空间基的交集不一定是子空间的基。

Example

设 $V=\mathbb{R}^2$（二维实向量空间），子空间 $W=\{(x,-x)\mid x\in\mathbb{R}\}$。可验证 W 对线性组合封闭，因此是 V 的子空间。
定义全空间的基取 V 的一组基 $S=\{(1,0),(0,1)\}$
验证交集不是子空间的基计算 $W\cap S$，即找既属于 W 又属于 S 的向量。但 S 中的向量是 $(1,0)$（纵坐标为 0）和 $(0,1)$（横坐标为 0），都不满足 W 中“横坐标与纵坐标相反”的条件，因此 $W\cap S=\emptyset$（空集）。
而子空间的基必须是非空的线性无关组且能张成子空间，空集显然不满足，因此 $W\cap S$ 不是 W 的基。

1.6.2 基底的重要性质

接下来，我们将展示本节的 2 条主定理：

Theorem

设 V 是一个向量空间，且 $\beta = \{u_1, u_2, \ldots, u_n\}$ 是 V 的一个子集。那么 β 是 V 的一个基底**当且仅当**每个 $v \in V$ 都可以唯一地表示为 β 中向量的线性组合，即可以表示为

$$v = a_1u_1 + a_2u_2 + \cdots + a_nu_n$$

其中 $a_1, a_2, \ldots, a_n$ 是唯一的标量且均属于数域 $\mathbb{F}$。

证明. **正向证明：**

设 β 是 V 的一个基底。如果 $v \in V$，那么 $v \in \text{span}(\beta)$，因为 $\text{span}(\beta) = V$。因此 v 是 β 中向量的线性组合。假设

$$v = a_1u_1 + a_2u_2 + \cdots + a_nu_n \quad \text{和} \quad v = b_1u_1 + b_2u_2 + \cdots + b_nu_n$$

是 v 的两个这样的表示。用第一个方程减去第二个方程得到

$$0 = (a_1 - b_1)u_1 + (a_2 - b_2)u_2 + \cdots + (a_n - b_n)u_n$$

由于 β 是线性无关的，所以 $a_1 - b_1 = a_2 - b_2 = \cdots = a_n - b_n = 0$。因此 $a_1 = b_1$，$a_2 = b_2$，…，$a_n = b_n$，所以 v 可以唯一地表示为 β 中向量的线性组合。

反向证明：

已知每个 $v \in V$ 都能唯一地表示为 $\beta = \{u_1, u_2, \ldots, u_n\}$ 中向量的线性组合，所以 $\text{span}(\beta) = V$，即 β 生成 V。

线性无关：设 $a_1u_1 + a_2u_2 + \cdots + a_nu_n = 0$，又 $0 = 0u_1 + 0u_2 + \cdots + 0u_n$，根据表示唯一性，可得 $a_1 = a_2 = \cdots = a_n = 0$，故 β 线性无关。

综上，β 线性无关且生成 V，所以 β 是 V 的基底。 □

接着，我们将讨论有限生成向量空间的基底：

Theorem

如果一个向量空间 V 由一个有限集 S 生成，那么 S 一定存在某个子集是 V 的一个基底。

因此 V 有一个有限基底。

证明. 如果 $S = \varnothing$ 或 $S = \{0\}$，那么 $V = \{0\}$ 且 $\varnothing$ 是 S 的一个子集，它是 V 的一个基底。

否则 S 包含一个非零向量 u_1。$\{u_1\}$ 是一个线性无关集。如果可能的话，继续在 S 中选择向量 $u_2, \ldots, u_k$，使得 $\{u_1, u_2, \ldots, u_k\}$ 是线性无关的。

由于 S 是一个有限集，我们最终一定会到达一个阶段，此时 $\beta = \{u_1, u_2, \ldots, u_k\}$ 是 S 的一个线性无关子集，但是将 S 中任何不在 β 中的向量添加到 β 中都会产生一个线性相关集。我们声称 β 是 V 的一个基底。

因为根据构造 β 是线性无关的，所以只需证明 β 生成 V。我们需要证明 $S \subseteq \text{span}(\beta)$。令 $v \in S$。如果 $v \in \beta$，那么显然 $v \in \text{span}(\beta)$。否则，如果 $v \notin \beta$，那么前面的构造表明 $\beta \cup \{v\}$ 是线性相关的。

因此 $v \in \text{span}(\beta)$。

因此 $S \subseteq \text{span}(\beta)$。 □

1.6.3 n 维向量及其坐标表示

思考：结合前言知识里面的向量加减法，我们已经有了向量坐标化的思想。现在又在向量空间规定了基的定义，我们就可以提出有序基，标准基底与坐标表示的概念了。

在向量空间中，基是一组线性无关且能生成整个空间的向量集合。而有序基则是为基中的向量赋予了特定顺序的基，即基中向量的排列顺序是固定且有意义的。

Definition

有序基（ordered basis）： 给定的基底中向量的顺序 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 是固定且被明确指定的。则称 $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ 为向量空间 V 的一个有序基。

标准基底（standard base）： F^n 的标准基底是 $\{e_1, e_2, \ldots, e_n\}$ ，e_i 第 i 个分量为 1 ，其余为 0 ，它线性无关且生成 F^n 。

坐标表示 (Coordinate representation)： 对于 F^n 的有序基底 $\beta = \{v_1, v_2, \ldots, v_n\}$ ，向量 $\mathbf{u} \in F^n$ 可唯一表示为 $\mathbf{u} = \sum_{i=1}^{n} c_i v_i$ ，$(c_1, c_2, \ldots, c_n)$ 是 $\mathbf{u}$ 关于 β 的坐标，记为 $[\mathbf{u}]_\beta$ 。

1.6.4 有限维与维度的定义

Definition

维度 (dimension)： V 的每个基底中向量的唯一数量被称为 V 的维度 (dimension)，记为 $\dim(V)$。

一个向量空间如果有一个由有限个向量组成的基底，就被称为有限维的。**一个不是有限维的向量空间被称为无限维的**。

一般来说，本讲义的讨论都局限于有限维中。

从该定义我们可以观察到，若向量空间 V 存在有限基，则 V 的任意基所包含的向量个数必然相同。这一性质表明，基的向量数量是向量空间 V 的固有属性（见后半部分的 lemma），不依赖于基的具体选取。

联系前言部分，我们再通过几个例子，加深对维度的理解

Example

向量空间 $\{0\}$ 的维度为零。
向量空间 $\mathbb{F}^n$ 的维度为 n。
向量空间 $M_{m\times n}(F)$ 的维度为 mn。
向量空间 $P_n(F)$ 的维度为 $n+1$。

接下来，我们来证明：

Lemma

若向量空间 V 存在有限基，则 V 的任意基所包含的**向量个数必然相同**。

证明. **有限维情形：**

设 V 是有限维向量空间，$S_1=\{u_1,u_2,\dots,u_m\}$ 与 $S_2=\{v_1,v_2,\dots,v_n\}$ 是 V 的两个基。

因 S_1 张成 V 且 S_2 线性无关，我们现在需要证明：线性无关集的基数不超过张成集的基数，即 $n\le m$。

因为 $\operatorname{span}(S)=V$，所以 L 中第一个向量 v_1 可由 S 线性表出：$v_1=a_1u_1+a_2u_2+\cdots+a_mu_m$.

由于 $v_1\neq 0$（线性无关集中向量非零），至少有一个系数 $a_i\neq 0$。不妨设 $a_1\neq 0$，则 u_1 可由 $\{v_1,u_2,\dots,u_m\}$ 线性表出，因此：

$\operatorname{span}(\{v_1,u_2,\dots,u_m\})=\operatorname{span}(S)=V.$

重复该过程，假设 $n>m$（即 S_2 的向量个数比 S_1 多）：

S_1 只有 m 个向量，因此最多能替换 m 次。替换 m 次后，S_1 的所有向量 $u_1,u_2,\dots,u_m$ 会被完全替换成 $v_1,v_2,\dots,v_m$，此时张成组为 $\{v_1,v_2,\dots,v_m\}$，且 $\operatorname{span}(\{v_1,v_2,\dots,v_m\})=V$。

但 S_2 还有第 $m+1$ 个向量 v_{m+1}，由于 $v_{m+1}\in V$，它必须能由 $\{v_1,v_2,\dots,v_m\}$ 线性表出——这与“$S_2=\{v_1,v_2,\dots,v_n\}$ 线性无关”矛盾（线性无关组中任意向量都不能由其余向量线性表出）。

同理，交换 S_1 与 S_2 的角色，可得 $m\le n$。

综上，$m=n$，即有限维空间的基有相同元素个数。

无限维情形：

设 V 是无限维向量空间，S_1,S_2 为其两个基。

对任意 $u\in S_1$，因 S_2 张成 V，u 可表为 S_2 中有限个向量的组合。记所有这些有限子集的并为 $F\subseteq S_2$，则 $\operatorname{span}(F)=V$。

由 S_1 的线性无关性，得 $|S_1|\le|F|\le|S_2|$。

同理可得 $|S_2|\le|S_1|$。

由基数反对称性，$|S_1| = |S_2|$，即无限维空间的基有相同基数。

故，向量空间的任意两个基具有相同基数，故维数 $\dim(V)$ 是良定义的。 □

接下来，我们来讨论一个著名的公式：

Theorem

若 W_1 和 W_2 是向量空间 V 的有限维子空间，则子空间 $W_1 + W_2$ 是有限维的，且

$$\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2).$$

证明. 设 W_1, W_2 是向量空间 V 的有限维子空间，令 $\dim(W_1 \cap W_2) = r$。

取 $W_1 \cap W_2$ 的一组基：$\alpha_1, \alpha_2, \ldots, \alpha_r$。

由于 $W_1 \cap W_2 \subseteq W_1$，将这组基扩展为 W_1 的基：$\alpha_1, \alpha_2, \ldots, \alpha_r, \beta_1, \beta_2, \ldots, \beta_s$，此时 $\dim(W_1) = r + s$。

同理，将 $W_1 \cap W_2$ 的基扩展为 W_2 的基：$\alpha_1, \alpha_2, \ldots, \alpha_r, \gamma_1, \gamma_2, \ldots, \gamma_t$，此时 $\dim(W_2) = r + t$。

需证明向量组 $\alpha_1, \ldots, \alpha_r, \beta_1, \ldots, \beta_s, \gamma_1, \ldots, \gamma_t$ 是 $W_1 + W_2$ 的基。

首先证明生成性：对任意 $\xi \in W_1 + W_2$，存在 $\xi_1 \in W_1$，$\xi_2 \in W_2$，使得 $\xi = \xi_1 + \xi_2$。

因 $\xi_1 \in W_1$，故 $\xi_1 = \sum_{i=1}^{r} a_i \alpha_i + \sum_{j=1}^{s} b_j \beta_j$。

因 $\xi_2 \in W_2$，故 $\xi_2 = \sum_{i=1}^{r} c_i \alpha_i + \sum_{k=1}^{t} d_k \gamma_k$。

因此：

$$\xi = \xi_1 + \xi_2 = \sum_{i=1}^{r} (a_i + c_i)\alpha_i + \sum_{j=1}^{s} b_j \beta_j + \sum_{k=1}^{t} d_k \gamma_k$$

即 ξ 可由该向量组线性表示。

再证明线性无关性：假设存在数 $k_1, \ldots, k_r, l_1, \ldots, l_s, m_1, \ldots, m_t$，使得：

$$\sum_{i=1}^{r} k_i \alpha_i + \sum_{j=1}^{s} l_j \beta_j + \sum_{k=1}^{t} m_k \gamma_k = 0 \tag{1}$$

令 $\eta = \sum_{i=1}^{r} k_i \alpha_i + \sum_{j=1}^{s} l_j \beta_j$，则由式 (1) 得 $\eta = -\sum_{k=1}^{t} m_k \gamma_k$。

一方面，$\eta \in W_1$；另一方面，$-\sum_{k=1}^{t} m_k \gamma_k \in W_2$，故 $\eta \in W_1 \cap W_2$。

因此 η 可由 $\alpha_1, \ldots, \alpha_r$ 线性表示，即存在 $n_1, \ldots, n_r$，使得 $\eta = \sum_{i=1}^{r} n_i \alpha_i$。

结合 $\eta = \sum_{i=1}^{r} k_i \alpha_i + \sum_{j=1}^{s} l_j \beta_j$，且 $\alpha_1, \ldots, \alpha_r, \beta_1, \ldots, \beta_s$ 线性无关，可得 $l_j = 0$ $(j = 1, \ldots, s)$，且 $k_i = n_i$ $(i = 1, \ldots, r)$。

此时 $\eta = \sum_{i=1}^{r} n_i \alpha_i = -\sum_{k=1}^{t} m_k \gamma_k$，即：

$$\sum_{i=1}^{r} n_i \alpha_i + \sum_{k=1}^{t} m_k \gamma_k = 0$$

因 $\alpha_1, \ldots, \alpha_r, \gamma_1, \ldots, \gamma_t$ 线性无关，故 $n_i = 0$ $(i = 1, \ldots, r)$，$m_k = 0$ $(k = 1, \ldots, t)$，进而 $k_i = 0$ $(i = 1, \ldots, r)$。

因此该向量组线性无关，确为 W_1+W_2 的基，含 $r+s+t$ 个向量，故 $\dim(W_1+W_2)=r+s+t$。

又 $\dim(W_1)=r+s$，$\dim(W_2)=r+t$，$\dim(W_1\cap W_2)=r$，因此：

$$\dim(W_1+W_2)=\dim(W_1)+\dim(W_2)-\dim(W_1\cap W_2)$$

综上，定理得证。

□

1.6.5 课后练习

Exercise

1. 对于固定的 $a\in\mathbb{R}$，确定由 $\{f\in P_n(\mathbb{R}):f(a)=0\}$ 定义的 $P_n(\mathbb{R})$ 的子空间的维度。
2. 设 W 是有限维向量空间 V 的一个子空间。那么 W 是有限维的，且 $\dim(W)\leq\dim(V)$。此外，如果 $\dim(W)=\dim(V)$，那么 $V=W$。
3. 考虑由所有迹为零的 $n\times n$ 矩阵构成的集合 V（若 $A=(a_{ij})$ 是一个 $n\times n$ 矩阵，A 的迹是和 $\sum_{i=1}^n a_{ii}$）。

   1. 证明它是一个向量空间。
   2. 对于所有迹为零的 2×2 矩阵，求该向量空间的一个基。
   3. 求 V 的一个基及其维数。

4. 设 $V=M_{n\times n}(\mathbb{C})$ 是复 $n\times n$ 矩阵构成的 $\mathbb{C}$-向量空间，我们定义

$$W_{*,+}:=\{A\in V:{}^t\overline{A}=A\}\quad 和\quad W_{*,-}:=\{A\in V:{}^t\overline{A}=-A\}$$

这里 ${}^t\overline{A}={}^t\overline{(a_{ij})}_{i,j}:=(\overline{a}_{ji})_{i,j}$，其中 $\overline{a}_{ij}=\overline{x+\sqrt{-1}y}:=x-\sqrt{-1}y\in\mathbb{C}$。
证明: 集合 $\{E_{ij}+E_{ji}:i\geq j\}\cup\{\sqrt{-1}(E_{ij}-E_{ji}):i>j\}$ 是作为 $\mathbb{R}$-向量空间的 $W_{*,+}$ 的一组基。这里 $E_{ij}:=(a_{kl})_{k,l}$ 表示第 i 行第 j 列元素为 1，其余元素为 0 的矩阵。

E_{ij} 的例子:

例子 1：E_{11}（以 2 阶矩阵为例）$E_{11}=\begin{pmatrix}1&0\\0&0\end{pmatrix}$，即第 1 行第 1 列元素为 1，其余元素为 0。

例子 2：E_{12}（以 2 阶矩阵为例）$E_{12}=\begin{pmatrix}0&1\\0&0\end{pmatrix}$，即第 1 行第 2 列元素为 1，其余元素为 0。

1.7 替换定理

接下来，我们将利用数学归纳法这一强大的归纳工具来证明线性代数中的一个重要定理——替换定理。该证明过程和归纳思想都相当重要，是后文许多思想所在的源泉。

Theorem

替换定理 (Replacement Theorem)：设 V 是一个由恰好含 n 个向量的集合 G 生成的向量空间，L 是 V 的一个恰好含 m 个向量的线性无关子集。那么 $m \leq n$ ，且存在 G 的一个恰好含 $n-m$ 个向量的子集 H ，使得 $L \cup H$ 生成 V 。

Example

例如：设向量空间 $V = \mathbb{R}^3$（三维实数向量空间），而线性无关子集 $L = \{\mathbf{v}_1, \mathbf{v}_2\}$，其中 $\mathbf{v}_1 = (1, 1, 0)$，$\mathbf{v}_2 = (1, -1, 1)$。我们可以找到 3-2=1 个向量，如 $\mathbf{e}_3 = (0, 0, 1)$，使得这三个向量可以生成（span）V。

该定理还有相当丰富的推论：

Corollary & Secondary Conclusion

推论 1：设 V 是一个具有有限基底的向量空间。那么 V 的每一个基底都包含相同数量的向量。(即上一节的重要定理)

Corollary & Secondary Conclusion

推论 2：设 V 是一个维度为 n 的向量空间。

(a)V 的任何有限生成集至少包含 n 个向量，且恰好包含 n 个向量的 V 的生成集是 V 的一个基底，故线性无关。

(b)V 的任何恰好包含 n 个向量的线性无关子集是 V 的一个基底。

(c)V 的每一个线性无关子集都可以扩展为 V 的一个基底。

1.8 子空间上的运算 1

1.8.1 子空间上的加法和减法

设 $W_1, W_2 \subset V$ 是两个子空间，它们的**和（sum）**定义为 $W_1 + W_2 := \{w_1 + w_2 : w_i \in W_i,\ i = 1, 2\}$。

可以证明：它仍然是一个子空间。

Example

例如：若 $S = \{(1,0)\}$，$W = \{(0,1)\}$，则 $S + W = \{(1,0) + (0,1)\} = \{(1,1)\}$，即 $S + W$ 为点 $(1,1)$；
若 S 为 x 轴（向量形如 $(a,0), a \in \mathbb{R}$），W 为 y 轴（向量形如 $(0,b), b \in \mathbb{R}$），则 $S + W = \{(a,0) + (0,b) \mid a, b \in \mathbb{R}\} = \{(a,b) \mid a, b \in \mathbb{R}\}$，即 $S + W$ 为全平面。

【Remark】：

$W_1, W_2 \subset V$ 其各自的基是 $S_1 \cup S_2$，那么这两组基是否为 $W_1 + W_2$ 的基？若 $S_1 \cap S_2 = \emptyset$ 呢？

子空间的和 $W_1 + W_2$ 的基需满足线性无关且张成 $W_1 + W_2$。

一般情况：$S_1 \cup S_2$ 不是 $W_1 + W_2$ 的基。即使 $S_1 \cap S_2 = \emptyset$：仍不一定是基。

原因：若 W_1, W_2 有公共非零向量，其基 S_1, S_2 中可能包含重复向量（或线性相关的向量），导致 $S_1 \cup S_2$ 线性相关。

Example

例如：设 $V = \mathbb{R}^2$，$W_1 = \mathrm{span}\{(1,0)\}$（基 $S_1 = \{(1,0)\}$），$W_2 = \mathrm{span}\{(2,0)\}$（基 $S_2 = \{(2,0)\}$）。
则 $W_1 + W_2 = W_1$，其基为 $\{(1,0)\}$，但 $S_1 \cup S_2 = \{(1,0),(2,0)\}$ 线性相关，不是基。

设 $W_1, W_2 \subset V$ 是两个子空间，它们的**差集（difference set）**定义为 $W_1 - W_2 := \{w_1 - w_2 : w_i \in W_i, i = 1, 2\}$。

Example

例子 1：不同子空间的差集为全空间设 $V=\mathbb{R}^2$，取子空间 $W_1=\{(a,0)\mid a\in\mathbb{R}\}$（x 轴），$W_2=\{(0,b)\mid b\in\mathbb{R}\}$（y 轴）。则 $W_1-W_2=\{(a-0,0-b)\mid a,b\in\mathbb{R}\}=\{(a,-b)\mid a,b\in\mathbb{R}\}=\mathbb{R}^2$（因为任意 $(x,y)\in\mathbb{R}^2$ 可表示为 $(x,0)-(0,-y)$，其中 $(x,0)\in W_1$，$(0,-y)\in W_2$）。
例子 2：相同子空间的差集为自身设 $V=\mathbb{R}^2$，取子空间 $W_1=W_2=\{(a,a)\mid a\in\mathbb{R}\}$（直线 $y=x$）。则 $W_1-W_2=\{(a-b,a-b)\mid a,b\in\mathbb{R}\}=\{(c,c)\mid c\in\mathbb{R}\}=W_1$（因为任意 $(c,c)\in W_1$ 可表示为 $(c,c)-(0,0)$，其中 $(0,0)\in W_2$）。

1.8.2 子空间上的交与并

两个子空间 $W_1,W_2\subset V$ 的交：

Definition

交 (Intersection)： $W_1\cap W_2:=\{w\in V: w\in W_1 \text{ 且} w\in W_2\}$，**它仍然是一个子空间**。

Property

交的性质：若 W_1 和 W_2 是 V 的子空间，则 $W_1\cap W_2$ 一定是 V 的子空间。

- 例如：在 $\mathbb{R}^3$ 中，设 $W_1=\{(x,y,z)\mid x+y=0\}$，$W_2=\{(x,y,z)\mid x-z=0\}$，则 $W_1\cap W_2=\{(x,-x,x)\mid x\in\mathbb{R}\}$，满足子空间判定条件。

【Remark】：

$S_1\cap S_2$ 是否为 $W_1\cap W_2$ 的基?

一般情况：$S_1\cap S_2$ 不是 $W_1\cap W_2$ 的基。虽这些向量线性无关，但可能"数量不足"，无法张成整个 $W_1\cap W_2$。

Example

设 $V=\mathbb{R}^2$，$W_1=\operatorname{span}\{(1,0),(1,1)\}=\mathbb{R}^2$（基 $S_1=\{(1,0),(1,1)\}$），$W_2=\operatorname{span}\{(1,0),(2,1)\}=\mathbb{R}^2$（基 $S_2=\{(1,0),(2,1)\}$）。则 $W_1\cap W_2=\mathbb{R}^2$，其基需 2 个线性无关向量，但 $S_1\cap S_2=\{(1,0)\}$，无法张成 $\mathbb{R}^2$。

两个子空间 $W_1,W_2\subset V$ 的并:

Definition

并 (Union)： $W_1\cup W_2:=\{w\in V: w\in W_1 \text{ 或} w\in W_2\}$。

【Remark】:

Union 一般来说并不是一个子空间。

Example

例如：在 $\mathbb{R}^2$ 中，W_1 为 x 轴上的向量集，W_2 为 y 轴上的向量集。$(1,0) \in W_1$，$(0,1) \in W_2$，但 $(1,0)+(0,1)=(1,1) \notin W_1 \cup W_2$，不满足加法封闭性，故 $W_1 \cup W_2$ 不是子空间。

Claim

$W_1 \cup W_2$ 是子空间当且仅当 $W_1 \subseteq W_2$ 或 $W_2 \subseteq W_1$。

证明. 充分性容易证明；

下面证明**必要性**：

假设 $W_1 \cup W_2$ 是子空间，但 $W_1 \nsubseteq W_2$ 且 $W_2 \nsubseteq W_1$

此时：因为 $W_1 \nsubseteq W_2$，所以存在 $w_1 \in W_1 \setminus W_2$（$w_1$ 在 W_1 中，但不在 W_2 中）；

因为 $W_2 \nsubseteq W_1$，所以存在 $w_2 \in W_2 \setminus W_1$（$w_2$ 在 W_2 中，但不在 W_1 中）。

考虑这两个向量的和 $w_1 + w_2$：由于 $W_1 \cup W_2$ 是子空间，它对加法封闭，因此 $w_1 + w_2$ 必须属于 $W_1 \cup W_2$，即要么 $w_1 + w_2 \in W_1$，要么 $w_1 + w_2 \in W_2$。

若 $w_1 + w_2 \in W_1$：

因为 W_1 是子空间，对减法封闭（子空间对数乘和加法封闭，自然对减法封闭：$a - b = a + (-1)b$），所以 $w_2 = (w_1 + w_2) - w_1 \in W_1$。但这与 $w_2 \in W_2 \setminus W_1$（$w_2$ 不在 W_1 中）矛盾。

若 $w_1 + w_2 \in W_2$：

同理，W_2 对减法封闭，所以 $w_1 = (w_1 + w_2) - w_2 \in W_2$。但这与 $w_1 \in W_1 \setminus W_2$（$w_1$ 不在 W_2 中）矛盾。 □

接下来我们来介绍一种特殊的和：

1.8.3 子空间上的直和

Definition

直和/直接和(direct sum)：若 W_1 和 W_2 是向量空间 V 的子空间，且满足 $W_1 \cap W_2 = \{0\}$ 以及 $W_1 + W_2 = V$，则称向量空间 V 是 W_1 和 W_2 的直和，记为 $V = W_1 \oplus W_2$。

接下来我们将介绍多个子空间的直和:

Definition

设 $V_1, V_2, \cdots, V_m$ 是线性空间 V 的子空间并且其和是全空间，且对任意 $2 \le i \le m$，$V_i \cap (V_1 + V_2 + \cdots + V_{i-1}) = \{0\}$，那么：$V_0 = V_1 \oplus V_2 \oplus \cdots \oplus V_m$ 是直和

【Remark】:

注意到：若 $V = \sum_{i=1}^{k} W_i$ 且对于 $i \ne j$ 有 $W_i \cap W_j = \{0\}$，但是不一定有 $V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$。

我们可以找出反例：例如，取 $W_1 = \text{span}\{(1,0)\}$，$W_2 = \text{span}\{(0,1)\}$，以及 $W_3 = \text{span}\{(1,1)\}$，说明不满足条件时不能得到直和结论。

在介绍本节的主定理以前，我们将先介绍 2 个重要的引理：

Lemma

lemma1: 若 S_i 是 W_i 的基（$i = 1, 2$），则 $S_1 \cup S_2$ 就是 $W_1 \oplus W_2$ 的基；
lemma2: 若 $S_1 \cup S_2$ 是 $W_1 \oplus W_2$ 的基，且 S_1 是 W_1 的基，那么 S_2 就是 W_2 的基；反之亦然。

证明. 对于引理 1：

张成性:

对任意 $v \in W_1 \oplus W_2$，由直和定义，$v = w_1 + w_2$（其中 $w_1 \in W_1$，$w_2 \in W_2$）。

因 S_1 是 W_1 的基，故 w_1 可由 S_1 线性表出；同理，w_2 可由 S_2 线性表出。

因此，$v = w_1 + w_2$ 可由 $S_1 \cup S_2$ 线性表出，即 $S_1 \cup S_2$ 张成 $W_1 \oplus W_2$。

线性无关性:

假设 $\sum_{s \in S_1} a_s s + \sum_{s \in S_2} b_s s = 0$（$a_s, b_s$ 为系数）。令 $w_1 = \sum_{s \in S_1} a_s s \in W_1$，$w_2 = \sum_{s \in S_2} b_s s \in W_2$，则 $w_1 + w_2 = 0$。

将等式变形为：$w_1 = -w_2$。

此时：$w_1 \in W_1$（由 w_1 的定义）；$-w_2 \in W_2$（因为 W_2 是子空间，对“数乘”封闭，$w_2 \in W_2$ 则 $-w_2 \in W_2$）。

因此，w_1 同时属于 W_1 和 W_2，即 $w_1 \in W_1 \cap W_2$。

由直和的“交为零空间”（$W_1 \cap W_2 = \{0\}$），得 $w_1 = 0$ 且 $w_2 = 0$。

又因 S_1 是 W_1 的基（线性无关），故 $w_1 = 0$ 蕴含所有 $a_s = 0$；同理，S_2 线性无关，故 $w_2 = 0$ 蕴含所有 $b_s = 0$。

因此，$S_1 \cup S_2$ 线性无关。

对于引理 2：需验证 S_2 满足线性无关且张成 W_2。

张成性：

对任意 $w_2 \in W_2$，因 $S_1 \cup S_2$ 张成 $W_1 \oplus W_2$，故 w_2 可表为：$w_2 = \sum_{s\in S_1} a_s s + \sum_{s\in S_2} b_s s$

令 $w_1 = \sum_{s\in S_1} a_s s \in W_1$，则 $w_2 = w_1 + \sum_{s\in S_2} b_s s$。

由直和的“交为零空间”（$W_1 \cap W_2 = \{0\}$），若 $w_1 \neq 0$，则 $w_1 = w_2 - \sum_{s\in S_2} b_s s \in W_1 \cap W_2$ 且非零 (右边：$w_2 \in W_2$，且 $\sum_{s\in S_2} b_s s \in W_2$（因为 S_2 张成 W_2，子空间对线性组合封闭）。因此，$w_2 - \sum_{s\in S_2} b_s s \in W_2$（子空间对减法封闭）。)，与 $W_1 \cap W_2 = \{0\}$ 矛盾。

因此 $w_1 = 0$，即：$w_2 = \sum_{s\in S_2} b_s s$ 故 S_2 张成 W_2。

线性无关性：

假设 $\sum_{s\in S_2} b_s s = 0$。

因 S_1 是 W_1 的基（线性无关），且 $S_1 \cup S_2$ 线性无关，故组合 $\sum_{s\in S_1} 0 \cdot s + \sum_{s\in S_2} b_s s = 0$ 蕴含所有 $b_s = 0$。因此，S_2 线性无关。 □

接下来我们将讲解本节的主定理：

Theorem

设 $V_1, V_2, \cdots, V_m$ 是线性空间 V 的子空间，$V_0 = V_1 + V_2 + \cdots + V_m$ ，则以下命题**等价**：

1. $V_0 = V_1 \oplus V_2 \oplus \cdots \oplus V_m$ 是直和；
2. 对任意 $2 \le i \le m$ ，$V_i \cap (V_1 + V_2 + \cdots + V_{i-1}) = \{0\}$ ；
3. $\dim(V_1 + V_2 + \cdots + V_m) = \dim V_1 + \dim V_2 + \cdots + \dim V_m$ ；
4. $V_1, V_2, \cdots, V_m$ 的一组基可拼成 V_0 的一组基；
5. V_0 中向量表示为 $V_1, V_2, \cdots, V_m$ 中向量之和时表示唯一，即若 $\alpha \in V_0$ 且 $\alpha = v_1 + v_2 + \cdots + v_m = u_1 + u_2 + \cdots + u_m$ ，其中 $v_i, u_i \in V_i$ ，则 $u_i = v_i$ ($i = 1, 2, \cdots, m$)。

证明. □

1.9 等价关系、等价类与商集

1.9.1 等价关系 (Equivalence Relation)

Definition

设 A 是非空集合，R 是 A 上的二元关系。若 R 满足以下三条性质，则称 R 为 A 上的**等价关系**（equivalence relation）：

1. **自反性** (reflexivity)：对任意 $a \in A$，有 $(a, a) \in R$；（或 $a\ a$）
2. **对称性** (symmetry)：对任意 $a, b \in A$，若 $(a, b) \in R$，则 $(b, a) \in R$；
3. **传递性** (transitivity)：对任意 $a, b, c \in A$，若 $(a, b) \in R$ 且 $(b, c) \in R$，则 $(a, c) \in R$。

若 $(a, b) \in R$，则称 a 与 b 等价，记为 $a \sim b$。

Example

我们有以下几种常见的等价关系：三角形集合上的“相似（is similar to）”关系；整数集 $\mathbb{Z}$ 上的“模 n 同余（is congruent to modulo n）”关系; 函数 $f : X \to Y$ 下有相同的像的关系

相似三角形的定义是“对应角相等，对应边成比例”。
因此：

- 任意三角形与自身相似（角相等、边比例为 $1 : 1$），故 $a \sim a$。
- 对称性：若三角形 A 相似于 B，则 B 的角与 A 相等、边与 A 成比例，因此 B 也相似于 A，故 $a \sim b \iff b \sim a$。
- 若 A 相似于 B，B 相似于 C，则 A 与 C 的对应角都等于 B 的角（相等），对应边比例为“A 对 B 的比例”乘以“B 对 C 的比例”（仍为常数），因此 A 相似于 C，故 $a \sim b \wedge b \sim c \implies a \sim c$。

若整数 a, b 满足“n 整除 $a - b$”，则称 $a \equiv b \pmod n$。

在整数集 $\mathbb{Z}$ 上，定义关系 $R = \{(a, b) \mid a \equiv b \pmod 2\}$（即 a 与 b 奇偶性相同）。验证 R 是等价关系：

1. 自反性：$\forall a \in \mathbb{Z}$，$a - a = 0$ 是 2 的倍数，故 $a \sim a$；
2. 对称性：若 $a \sim b$，则 $a - b$ 是 2 的倍数，显然 $b - a$ 也是 2 的倍数，故 $b \sim a$；即：若 $a \equiv b \pmod n$，则 $n \mid (a - b)$，即存在整数 k 使得 $a - b = nk$。因此 $b - a = -nk = n(-k)$，故 $n \mid (b - a)$，即 $b \equiv a \pmod n$。

3. 传递性：若 $a \sim b$ 且 $b \sim c$，则 $a-b=2k$，$b-c=2m\ (k,m\in\mathbb{Z})$，故 $a-c=2(k+m)$，即 $a\sim c$。

因此，R 是 $\mathbb{Z}$ 上的等价关系。

至于函数：

一般情况：对 $x_1, x_2 \in X$，定义 $x_1 \sim x_2 \iff f(x_1)=f(x_2)$。

- 自反性：对任意 $x\in X$，$f(x)=f(x)$，故 $x\sim x$。
- 对称性：若 $x_1\sim x_2$（即 $f(x_1)=f(x_2)$），则 $f(x_2)=f(x_1)$，故 $x_2\sim x_1$。
- 传递性：若 $x_1\sim x_2\ (f(x_1)=f(x_2))$ 且 $x_2\sim x_3\ (f(x_2)=f(x_3))$，则 $f(x_1)=f(x_3)$，故 $x_1\sim x_3$。

我们在这里还有 1 个例子：

Example

集合 $X=\{a,b,c\}$ 上的关系 $\sim$（由序对 $R=\{(a,a),(b,b),(c,c),(a,b),(b,a)\}$ 定义）——该关系 $\sim$ 的本质是：a 与 a、b 与 b、c 与 c 相关；a 与 b、b 与 a 相关；c 仅与自身相关。

其中：序对 $(x,y)\in R$ 就表示 "x 和 y 满足关系 $\sim$"（即 $x\sim y$）

容易验证其自反与对称性，只需验证传递性即可。

问题：找出那些因恰好违背了定义中的三个条件（自反性、对称性、传递性）之一而不是等价关系的二元关系的例子。

- 违反自反性：设集合 $X=\{a,b\}$，定义关系 $R=\{(a,a),(a,b),(b,a)\}$。此时 $b\not\sim b$，不满足自反性，而对称性、传递性成立。
- 违反对称性：设集合 $X=\{1,2\}$，定义关系 $R=\{(1,1),(2,2),(1,2)\}$。此时 $1\sim 2$ 但 $2\not\sim 1$，不满足对称性，而自反性、传递性成立。
- 违反传递性：设集合 $X=\{1,2,3\}$，定义关系 $R=\{(1,1),(2,2),(3,3),(1,2),(2,1),(2,3),(3,2)\}$。此时 $1\sim 2$ 且 $2\sim 3$，但 $1\not\sim 3$，不满足传递性，而自反性、对称性成立。

1.9.2 等价类

等价类是基于等价关系对集合元素的分类，其定义依赖于等价关系：

Definition

设 R 是集合 A 上的等价关系，对任意 $a \in A$，称集合

$$[a]_R = \{x \in A \mid x \sim a\}$$

为 a 关于 R 的**等价类**（equivalence class of a under ）或称为**陪集**（Coset），简记为 $[a]$, 或：$\bar{a}$。其中 a 称为该等价类的**代表元**。

用数学语言来说：

若定义 $v_1 \sim v_2 \iff v_1 - v_2 \in W$。

那么陪集（即等价类）记为：$\bar{v} = [v] = \{v + w \mid w \in W\}$（即 “$v$ 所在的陪集”，可理解为 “v 被 W 平移后的集合”）。

【Remark】:

每个陪集的 “内部容量”（元素个数）可能无限

Example

对上例中的等价关系 R（模 2 同余）:
- 取 $a = 0$，则 $[0] = \{x \in \mathbb{Z} \mid x \equiv 0 \pmod 2\} = \{\dots, -4, -2, 0, 2, 4, \dots\}$（所有偶数）;
- 取 $a = 1$，则 $[1] = \{x \in \mathbb{Z} \mid x \equiv 1 \pmod 2\} = \{\dots, -3, -1, 1, 3, 5, \dots\}$（所有奇数）。

可见，等价类是按 “等价关系” 对集合元素的**归类**。

等价类具有以下性质：

1. 对任意 $a \in A$，$a \in [a]$（非空性）;
2. 若 $a \sim b$，则 $[a] = [b]$（代表元等价则类相等）;
3. 若 $a \not\sim b$，则 $[a] \cap [b] = \emptyset$（不同等价类互斥）。

1.9.3 划分的定义

集合的**划分**（**Partition**）是将一个集合分解为若干非空且互不相交的子集的集合，使得原集合中的每个元素恰好属于其中一个子集。更形式化地:

Definition

对于集合 A，其划分 $\mathcal{P} = \{A_i\}_{i\in I}$ 需满足以下条件：

- 对所有 $i \in I$，子集 A_i 是 A 的非空子集，即 $A_i \subset A$ 且 $A_i \neq \emptyset$；
- 所有子集的并集等于原集合 A，即 $\bigcup_{i\in I} A_i = A$；
- 任意两个不同的子集互不相交，即对所有 $i \neq j$，有 $A_i \cap A_j = \emptyset$。

划分中的每个子集 A_i 称为该划分的块（block）或部分（part）。

- 假设原集合 $A = \{1,2,3,4\}$。
- 例子 1：$\mathcal{P} = \{\{1,2\},\{3,4\}\}$ 是一个划分。
  - 验证：每个子集非空且是 A 的子集；并集为 $\{1,2,3,4\} = A$；两个子集交集为空。
- 例子 2：$\mathcal{P} = \{\{1\},\{2\},\{3,4\}\}$ 也是一个划分。
- 反例 1：$\mathcal{P} = \{\{1,2\},\{2,3\}\}$ 不是划分（因为元素 2 同时属于两个子集，违反“恰好一个”）。
- 反例 2：$\mathcal{P} = \{\{1,2\},\emptyset,\{3,4\}\}$ 不是划分（因为包含空集 $\emptyset$，违反“非空”要求）。

1.9.4 商集

商集是所有等价类构成的集合，是对原集合的“整体划分”：

Definition

设 R 是集合 A 上的等价关系，称所有等价类组成的集合

$$A/R = \{[a] \mid a \in A\}$$

为 A 关于 R 的**商集**。

Example

对整数域中的等价关系 R（模 2 同余），商集为：

$$\mathbb{Z}/R = \{[0], [1]\}$$

即整数集被划分为“偶数类”和“奇数类”两个部分。

再如，设 $A = \{1, 2, 3, 4\}$，R 是“模 3 同余”关系，则：
- $[1] = \{1, 4\}$，$[2] = \{2\}$，$[3] = \{3\}$，
- 商集 $A/R = \{\{1, 4\}, \{2\}, \{3\}\}$。

重要应用：有理数的构造

整数组成的分数 $\frac{m}{n}$（$n \neq 0$）有“不同形式但值相同”的情况（比如 $\frac{1}{2}$ 和 $\frac{2}{4}$ 应该表示同一个有理数）。

等价关系：定义 $\frac{a}{b} \sim \frac{c}{d} \iff ad = bc$（交叉相乘相等）。这个关系把“数值相等的分数”归为同一等价类。商集构造：有理数 $\mathbb{Q}$ 是“所有分数”模这个等价关系的商集，即 $\mathbb{Q} := \left\{ \frac{m}{n} \mid m \in \mathbb{Z}, n \in \mathbb{Z} \setminus \{0\} \right\} / \sim$。

核心思想：用等价关系“忽略分数的不同表示形式，只关注数值是否相等”，从而严格构造出有理数域。

重要应用：p 有限域的构造

Definition

$\mathbb{F}_p$（也记为 $\mathbb{Z}/p\mathbb{Z}$，其中 p 为素数）是模 p 的剩余类集合，元素为 $\{0, 1, 2, \ldots, p-1\}$。在其上定义模 p 加法和模 p 乘法，满足域的所有公理，因此是一个有限域（也称为素域，因为它无法由更小的域扩张得到）

例子:

Example

$\mathbb{F}_2$：元素为 $\{0, 1\}$
加法：$0+0=0$，$0+1=1$，$1+0=1$，$1+1=0$；
乘法：$0\times0=0$，$0\times1=0$，$1\times0=0$，$1\times1=1$。
$\mathbb{F}_3$：元素为 $\{0, 1, 2\}$
加法：$1+2 \equiv 0 \mod 3$，$2+2 \equiv 1 \mod 3$ 等；
乘法：$2\times2 \equiv 1 \mod 3$（即 2 的乘法逆元是自身）

代数结构性质:

Property

元素个数：恰好有 p 个元素，是素数幂阶有限域（幂次为 1）。

由此等价关系、等价类与商集的逻辑递进关系如下：

1. **基础**：等价关系是集合上满足自反、对称、传递性的二元关系，为元素分类提供“规则”；

2. **分类**：基于等价关系，将集合中相互等价的元素归为一类，形成等价类；

3. **整体划分**：所有等价类的集合构成商集，它是原集合的一个“划分”（各等价类非空、互斥且并集为原集合）。

三者的核心联系：等价关系 $\iff$ 集合的划分（商集本质是划分），而等价类是划分的基本单元。

1.9.5 等价关系的两条重要定理

Theorem

等价关系基本定理 (Fundamental Theorem of Equivalence Relations)：**集合上的等价关系**与该集合的**划分**之间存在自然的**一一对应**关系。

proof.

- **从等价关系到划分**：假设集合 X 上有等价关系 $\sim$。我们要证明所有等价类构成的集合 $\mathcal{P} = X/\sim$ 是 X 的划分。

  1. **等价类是 X 的非空子集**：对任意 $x \in X$，等价类 $[x] = \{y \in X \mid y \sim x\}$。由自反性，$x \sim x$，所以 $x \in [x]$，故 $[x] \neq \emptyset$；且等价类中元素都属于 X，故 $[x] \subset X$。

  2. **所有等价类的并集是 X**：对任意 $x \in X$，$x \in [x]$，而 $[x]$ 是 $\mathcal{P}$ 中的元素，因此 $X = \bigcup_{x \in X} [x]$。

  3. **不同等价类互不相交**：若等价类 $[a]$ 和 $[b]$ 有公共元素 c（即 $c \in [a] \cap [b]$），则 $a \sim c$ 且 $b \sim c$。由对称性，$c \sim b$，再由传递性，$a \sim b$。进而可证 $[a] \subset [b]$ 且 $[b] \subset [a]$，故 $[a] = [b]$。

- **从划分到等价关系**：假设 $\mathcal{P} = \{A_i\}_{i \in I}$ 是 X 的划分。定义二元关系 $\sim$：$a \sim b \iff a, b$ 属于同一个块 A_i。验证其为等价关系：

  1. **自反性**：对任意 $x \in X$，x 属于某个 A_i，故 $x \sim x$。

  2. **对称性**：若 $x \sim y$，则 x, y 属于 A_i，显然 y, x 也属于 A_i，故 $y \sim x$。

  3. **传递性**：若 $x \sim y$ 且 $y \sim z$，则 x, y 属于 A_i，y, z 属于 A_j。因 y 只能属于一个块，故 $A_i = A_j$，所以 x, z 属于 A_i，即 $x \sim z$。

Theorem

投影的泛性质定理（(Universal Property of the Projection)）设 $f : X \to Y$ 是一个函数，且 $\sim$ 是 X 上的一个等价关系。
如果 f 在等价类上是常值的（即若 $a \sim b$ 蕴含 $f(a) = f(b)$），那么存在唯一的函数 $g : X/{\sim} \to Y$ 使得 $f = g \circ p$，其中 $p : X \to X/{\sim}$ 是投影映射。
且如果 f 是满射的，且等价关系由 $a \sim b \iff f(a) = f(b)$ 定义，那么诱导映射 g 是一个双射。

这通常用下述交换图来直观表示：

$$X[r,"f"][d,"p"'] Y X/{\sim}[ru, dashed,"g"'][ru,"\exists!"]$$

这个性质是基本的；它意味着与等价关系相容的函数能唯一地通过商映射分解。
proof.

- 证明存在唯一的 $g : X/\sim \to Y$ 使得 $f = g \circ p$

  1. 定义良定义的函数 g：

     因为 f 在每个等价类上是常值的（即若 $a \sim b$，则 $f(a) = f(b)$），对于商集 $X/\sim$ 中的任意等价类 $[x]$，它包含所有与 x 等价的元素 x'（满足 $x' \sim x$）。由于 f 在等价类上常值，所以这些 x' 的函数值 $f(x')$ 都等于 $f(x)$。

     因此，我们可以定义对于其中一个代表元 x，定义 $g([x]) = f(x)$，这个定义不依赖于从等价类 $[x]$ 中选择哪个具体的元素，所以 g 是从商集 $X/\sim$ 到 Y 的“良定义”函数。

  2. 证明 g 的唯一性：

     假设存在另一个函数 $h : X/\sim \to Y$，也满足 $f = h \circ p$。需证：$h = g$。

     对于商集 $X/\sim$ 中的任意等价类 $[x]$，取 $x \in X$（此时 $p(x) = [x]$）。由 $f = h \circ p$ 可得，$f(x) = (h \circ p)(x) = h(p(x)) = h([x])$。而根据 g 的定义，$g([x]) = f(x)$，所以 $h([x]) = g([x])$ 对所有 $[x] \in X/\sim$ 都成立，因此 $h = g$，即 g 是唯一的。

- 证明 f 是满射时，g 是双射

  双射需要同时满足单射（Injective）和满射（Surjective）。

  1. 证明 g 是单射：

     单射的定义是：若 $g([x]) = g([y])$，则 $[x] = [y]$。

由 g 的定义可知，$g([x]) = f(x)$，$g([y]) = f(y)$，所以如果 $g([x]) = g([y])$，那么 $f(x) = f(y)$。

又因为等价关系 $\sim$ 定义为 $a \sim b \iff f(a) = f(b)$，所以 $f(x) = f(y)$ 意味着 $x \sim y$。而等价类相等的充要条件是其代表元等价，即 $x \sim y \implies [x] = [y]$，所以 g 是单射。

2. 证明 g 是满射

满射的定义是：对于任意 $y \in Y$，存在 $[x] \in X/\sim$ 使得 $g([x]) = y$。

因为 f 是满射，所以对于任意 $y \in Y$，都存在 $x \in X$ 使得 $f(x) = y$。取这个 x 所在的等价类 $[x] \in X/\sim$，根据 g 的定义，$g([x]) = f(x) = y$。因此，对于任意 $y \in Y$，都能找到对应的 $[x] \in X/\sim$，所以 g 也是满射。

接下来，我们会对代数学中三个重要对象进行解读：

1.10 群，环，域

1.10.1 群与阿贝尔群

设 G 为非空集合，$*$（往往是 $+$ 或者 $\times$）是 G 上的二元运算。若满足以下公理，则 $(G, *)$ 称为**群**：

1. **封闭性**：对任意 $a, b \in G$，有 $a * b \in G$；
2. **结合律**：对任意 $a, b, c \in G$，有 $(a * b) * c = a * (b * c)$；
3. **单位元存在性**：存在 $e \in G$，使得对任意 $a \in G$，$e * a = a * e = a$；
4. **逆元存在性**：对任意 $a \in G$，存在 $b \in G$，使得 $a * b = b * a = e$。

若群 $(G, *)$ 还满足**交换律**（对任意 $a, b \in G$，$a * b = b * a$），则称 $(G, *)$ 为**阿贝尔群**（或交换群）。

Example

$(\mathbb{Z}, +)$：整数集合对加法构成**阿贝尔群**。加法满足封闭性、结合律、交换律；单位元为 0；对任意 $n \in \mathbb{Z}$，其加法逆元为 $-n$，满足所有阿贝尔群公理。

$(n\mathbb{Z}, +)$ $(n \in \mathbb{N}^+)$：n 的整数倍集合对加法是**阿贝尔群**。加法封闭 $(nk+nl = n(k+l) \in n\mathbb{Z})$、结合、交换；单位元为 0；逆元为 $-nk$（对应 $nk \in n\mathbb{Z}$）。

1.10.2 环与交换环

设 R 为非空集合，配备加法 $+$ 和乘法 $\times$ 两种二元运算。若满足：

1. **加法结构**：$(R, +)$ 是阿贝尔群；
2. **乘法结构**：$(R, \times)$ 是半群（满足封闭性与结合律）；
3. **分配律**：乘法对加法满足左、右分配律 ($a \times (b + c) = (a \times b) + (a \times c)$ 且 $(b + c) \times a = (b \times a) + (c \times a)$)；

则 $(R, +, \times)$ 称为**环**。若环的乘法满足交换律（$a \times b = b \times a$ 对任意 $a, b \in R$），则称为**交换环**。

Example

$(\mathbb{Z}, +, \times)$：整数对加法和乘法构成**交换环**。加法是阿贝尔群；乘法封闭、结合且交换；分配律成立。但 $(\mathbb{Z}, \times)$ 非群（除 ± 1 外无乘法逆元），故为环而非域。
$\mathbb{F}[x]$ (系数在域 $\mathbb{F}$ 上的一元多项式集合)：对加法和乘法是**交换环**。加法是阿贝尔群；乘法封闭、结合且交换；分配律成立。但多项式无普遍乘法逆元，故非域。

1.10.3 域

设 F 为非空集合，配备加法 $+$ 和乘法 $\times$ 两种二元运算。若满足：

1. **加法结构**：$(F, +)$ 是阿贝尔群；
2. **乘法结构**：$(F^{\times}, \times)$ 是阿贝尔群（$F^{\times} = F \setminus \{0\}$，$0$ 为加法单位元）；
3. **分配律**：乘法对加法满足左、右分配律；

则 $(F, +, \times)$ 称为**域**。

可见，域是最完备的。

Example

$\mathbb{F}_p := \mathbb{Z}/p\mathbb{Z}$（p 为素数）
$\mathbb{Z}/p\mathbb{Z}$：表示“整数模 p 的剩余类集合”（p 是素数）。通俗说，是把整数按“除以 p 的余数”分类，每个余数对应一个“剩余类”。例如 $p=5$ 时，余数有 $0,1,2,3,4$，因此 $\mathbb{Z}/5\mathbb{Z}$ 的元素是 $\{\overline{0},\overline{1},\overline{2},\overline{3},\overline{4}\}$，其中 $\overline{k}$ 代表“所有除以 5 余 k 的整数”（比如 $\overline{1}$ 包含 $1,6,11,-4,-9,\cdots$）。
加法结构：$(\mathbb{F}_p,+)$ 是阿贝尔群。单位元：$\overline{0}$（即模 p 余 0 的等价类），因为对于任意 $\overline{k}\in\mathbb{F}_p$，都有 $\overline{k}+\overline{0}=\overline{k}$。逆元：对于任意 $\overline{k}\in\mathbb{F}_p$，其逆元是 $\overline{p-k}$，因为 $\overline{k}+\overline{p-k}=\overline{p}=\overline{0}$。交换性与结合性：由于整数加法满足交换律和结合律，模 p 后的加法也自然满足交换律和结合律。乘法结构：$(\mathbb{F}_p^{\times},\times)$ 是阿贝尔群（$\mathbb{F}_p^{\times}=\mathbb{F}_p\setminus\{\overline{0}\}$）。封闭性：若 $\overline{a},\overline{b}\in\mathbb{F}_p^{\times}$，则 ab 不被 p 整除（因为 a,b 都不被 p 整除，且 p 是素数），所以 $\overline{ab}\in\mathbb{F}_p^{\times}$。逆元存在性：对于 $\overline{a}\in\mathbb{F}_p^{\times}$，因为 $\gcd(a,p)=1$，根据贝祖定理，存在整数 x,y 使得 $ax+py=1$，模 p 后得到 $\overline{a}\cdot\overline{x}=\overline{1}$，所以 $\overline{x}$ 是 $\overline{a}$ 的乘法逆元。交换性与结合性：整数乘法的交换律和结合律在模 p 乘法下也成立。分配律：乘法对加法满足左、右分配律，这是由整数乘法对加法的分配律模 p 后保持的。
$\mathbb{Q}[\sqrt{-1}] := \{x+y\sqrt{-1}\mid x,y\in\mathbb{Q}\}$：是**域**。
集合描述：所有形如“有理数 x 加上有理数 y 乘以 i”的数，即 $x+yi$（$x,y\in\mathbb{Q}$）加法是阿贝尔群；非零元乘法封闭且逆元为 $\frac{x-y\sqrt{-1}}{x^2+y^2}$，故 $(\mathbb{Q}[\sqrt{-1}]^{\times},\times)$ 是阿贝尔群；分配律成立。

而在这么多域中，我们可以提炼出其核心特质：

(i) $x+y=y+x$，$x\times y=y\times x$（交换律）。

(ii) $x+(y+z)=(x+y)+z$，$x\times(y\times z)=(x\times y)\times z$（结合律）。

(iii) 存在 $0,1\in\mathbb{F}$，使得对任意 x，有 $x+0=x$，$1\times x=x$（单位元）。

(iv) 对任意 x，存在 $(-x)\in\mathbb{F}$，使得 $x+(-x)=0$（加法逆元）。

(v) 对任意 $x\neq 0$，存在 $x^{-1}\in\mathbb{F}$，使得 $x\times x^{-1}=1$（乘法逆元）。

(vi) $x\times(y+z)=(x\times y)+(x\times z)$（分配律）。

那么我们称这个封闭的集合为域（Field）

下表可以体现群、环、域的联系与异同

结构	运算数量	加法要求	乘法要求	核心特点	例子对
群	1 种	基本运算	无（单运算）	单运算下的“对称完备性”	$(\mathbb{Z},+)$ 是阿
环	2 种	阿贝尔群	半群（封闭、结合）	“加法完备乘法不完备”的过渡	$(\mathbb{Z},+,\times)$ 是交
域	2 种	阿贝尔群	非零元构成阿贝尔群	两运算都“完备”支持除法	$\mathbb{F}_p$、$\mathbb{Q}[\sqrt{-}$

表 1.1: 群、环、域的联系与异同

接下来我们考察域的一些重要特性，同时定义域的特征 (characteristic of a field F)
:

Fact or Background

域 $\mathbb{F}$ 中的单位元和逆元是唯一的（identities and inverses are unique）

- 加法单位元的唯一性：假设存在两个加法单位元 0_1 和 0_2，根据加法单位元定义，对任意 $x \in \mathbb{F}$ 有 $x + 0_1 = x$ 且 $x + 0_2 = x$。取 $x = 0_2$，得 $0_2 + 0_1 = 0_2$；取 $x = 0_1$，得 $0_1 + 0_2 = 0_1$。由加法交换律，$0_2 + 0_1 = 0_1 + 0_2$，故 $0_1 = 0_2$，加法单位元唯一。
- 乘法单位元的唯一性：假设存在两个乘法单位元 1_1 和 1_2，根据乘法单位元定义，对任意 $x \in \mathbb{F}$ 有 $x \times 1_1 = x$ 且 $x \times 1_2 = x$。取 $x = 1_2$，得 $1_2 \times 1_1 = 1_2$；取 $x = 1_1$，得 $1_1 \times 1_2 = 1_1$。由乘法交换律，$1_2 \times 1_1 = 1_1 \times 1_2$，故 $1_1 = 1_2$，乘法单位元唯一。
- 加法逆元的唯一性：设 y_1 和 y_2 都是 $x \in \mathbb{F}$ 的加法逆元，即 $x+y_1 = 0$ 且 $x+y_2 = 0$。则 $y_1 = y_1 + 0 = y_1 + (x + y_2) = (y_1 + x) + y_2 = 0 + y_2 = y_2$，故加法逆元唯一。
- 乘法逆元的唯一性：设 y_1 和 y_2 都是非零元 $x \in \mathbb{F}$ 的乘法逆元，即 $x \times y_1 = 1$ 且 $x \times y_2 = 1$。则 $y_1 = y_1 \times 1 = y_1 \times (x \times y_2) = (y_1 \times x) \times y_2 = 1 \times y_2 = y_2$，故乘法逆元唯一。

Definition

考虑域 $\mathbb{F}$ 的乘法单位元”1”，定义集合：

$$\mathcal{P} := \left\{ n \in \mathbb{N}_+ \mid \underbrace{1 + 1 + \cdots + 1}_{n\text{个}} = 0 \in \mathbb{F} \right\},$$

若 $\mathcal{P} = \varnothing$（没有正整数能让 n 个 1 相加为 0），则称 $\mathbb{F}$ 是特征 0 的域;
若 $\mathcal{P} \neq \varnothing$，取 $\mathcal{P}$ 中最小的正整数 $p = \min(\mathcal{P})$，则称 $\mathbb{F}$ 是特征 p 的域，记为 $\operatorname{char}(\mathbb{F}) = p$.

其中“n 个 1 相加”是域的加法运算，结果若等于加法单位元”0”，则 n 属于 $\mathcal{P}$。

而我们常见的“1+1=2”的世界中，是 $\mathcal{P} = \varnothing$，即称 $\mathbb{F}$ 是特征 0 的域（如 $\mathbb{Q}, \mathbb{R}, \mathbb{C}$，这些域中“n 个 1 相加”的结果是“整数 n 在域中的嵌入”，永远不为”0”）。

n 个 1 相加为”0”仅在有限域（也叫伽罗瓦域，元素个数有限的域）中发生。

Example

以最典型的有限域 $F_p = \mathbb{Z}/p\mathbb{Z}$（$p$ 为素数）为例：
元素：$\{0, 1, 2, \ldots, p-1\}$（模 p 的剩余类）；
加法：模 p 加法（如 $p = 2$ 时，$1+1 = 2 \equiv 0 \mod 2$；$p = 3$ 时，$1+1+1 = 3 \equiv 0 \mod 3$）。

在这类域中，“p 个 1 相加”的结果是 $p \mod p = 0$，因此 $p \in \mathcal{P}$，**域的特征为 p**。

在利用这个性质之前，我们还要证明一条引理：

Lemma

p 为素数。

证明：假设 p 不是素数，则 p 可分解为两个大于 1 的正整数的乘积，即 $p = m \cdot n$（其中 $1 < m, n < p$）。

因为 $p = \min(\mathcal{P})$，所以 $m \notin \mathcal{P}$ 且 $n \notin \mathcal{P}$（否则 m 或 n 会比 p 更小且属于 $\mathcal{P}$，与“最小”矛盾）。

因此：

$\underbrace{1 + \cdots + 1}_{m} \neq 0, \quad \underbrace{1 + \cdots + 1}_{n} \neq 0.$

域满足乘法对加法的分配律：对任意 $a, b, c \in \mathbb{F}$，有 $a \cdot (b + c) = a \cdot b + a \cdot c$。

这个规律可重复推广到“多个加法项的乘积”：若把 p 个 1 分成 m 组，每组 n 个 1（即 $p = m \cdot n$），则“p 个 1 相加”可表示为：

$$\underbrace{1 + 1 + \cdots + 1}_{p} = \underbrace{\left(\underbrace{1 + \cdots + 1}_{n}\right) + \left(\underbrace{1 + \cdots + 1}_{n}\right) + \cdots + \left(\underbrace{1 + \cdots + 1}_{n}\right)}_{m}.$$

根据分配律，“m 个相同的和相加”可提取“组数 m”作为乘数，因此上式可写成：$\underbrace{1 + \cdots + 1}_{p} = \left(\underbrace{1 + \cdots + 1}_{m}\right) \cdot \left(\underbrace{1 + \cdots + 1}_{n}\right).$

左边等于 0（因为 $p \in \mathcal{P}$），但是右边是两个非零元素的乘积却都不等于 0。但域的定义要求“无零因子”（非零元素的乘积必非零），这与域的公理矛盾。

因此，“p 不是素数”的假设不成立，故 p 必须是素数。

而在这么多的域里面，本讲义研究的重点落在：**“数域”**：

数域可定义为：数域是一个域 $(\mathbb{F}, +, \times)$，其中 $\mathbb{F}$ 是数集（由实数、有理数等数构成的集合），且满足域的所有公理（交换律、结合律、单位元存在、逆元存在、分配律等），同时对加减乘除（除数不为 0）封闭。

1.11 子空间上的运算 2

1.11.1 子空间上的商与商空间

Definition

商空间 (Quotient space)： 所有陪集构成的集合 $V/W := \{\bar{v} \mid v \in V\}$，并在其上定义线性运算：

$\bar{v}_1 + \bar{v}_2 = \overline{v_1 + v_2}, \quad c\bar{v} = \overline{cv} \quad (\forall c \in \mathbb{F})$

可验证：这些运算良定义且使 V/W 成为域 $\mathbb{F}$ 上的向量空间，称为 V 对 W 的商空间。

用更加可以理解的语言说：

Fact or Background

(等价定义 1)：若 V 是数域 $\mathbb{F}$ 上的向量空间，W 是 V 的子空间，定义等价关系：$u \sim v \iff u - v \in W$。所有等价类的集合称为 V 关于 W 的**商空间**，记为 V/W。
(等价定义 2)：设 V 是数域 $\mathbb{F}$ 上的向量空间，W 是 V 的子空间，对任意 $u \in V$，其等价类定义为 $\overline{u} = u + W = \{u + w \mid w \in W\}$。而商空间 V/W 的元素就是等价类。

【Remark】：

商集： 仅为集合，无额外代数结构（如加法、数乘），仅由等价类的“归属”关系构成。
商空间： 是向量空间，继承了原空间的线性结构。对等价类定义加法和数乘：

在向量空间中，0 元是加法的单位元，即满足 $\overline{0} + \overline{u} = \overline{u}$ 的等价类。

根据等价类的加法定义：$\overline{0} + \overline{u} = \overline{0 + u} = \overline{u}$ 而 $\overline{0} = 0 + W = W$（因为 $0 \in V$，其等价类是所有 $0 + w = w$，即子空间 W 本身）。

所以我们得到了：

Property

商空间 V/W 的零元就是子空间 W 本身。

对于商空间的基，我们有以下的方法：

Method

基的构造步骤第一步：取 W 的一组基 $\{w_1, w_2, \ldots, w_k\}$（其中 $k = \dim W$）。
第二步：将其扩充为 V 的一组基：$\{w_1, \ldots, w_k, v_{k+1}, \ldots, v_n\}$（其中 $n = \dim V$）。
第三步：考虑这些基向量在商空间中的等价类 $\{\overline{v}_{k+1}, \overline{v}_{k+2}, \ldots, \overline{v}_n\}$，这组向量就是 V/W 的一组基。

证明. **线性无关：** 假设存在系数 $a_{k+1}, \ldots, a_n \in \mathbb{F}$，使得 $a_{k+1}\overline{v}_{k+1} + \cdots + a_n\overline{v}_n = \overline{0}$ 根据等价类的数乘定义，这等价于 $a_{k+1}v_{k+1} + \cdots + a_nv_n \in W$ 但 W 的基是 $\{w_1, \ldots, w_k\}$，故可设 $a_{k+1}v_{k+1} + \cdots + a_nv_n = b_1w_1 + \cdots + b_kw_k$。

由于 $\{w_1, \ldots, w_k, v_{k+1}, \ldots, v_n\}$ 是 V 的基（线性无关），因此所有系数必须为 0，即 $a_{k+1} = \cdots = a_n = 0$，$b_1 = \cdots = b_k = 0$。

故 $\{\overline{v}_{k+1}, \ldots, \overline{v}_n\}$ 线性无关。

张成性： 对任意 $\overline{u} \in V/W$，$u \in V$ 可表示为 $u = b_1w_1 + \cdots + b_kw_k + a_{k+1}v_{k+1} + \cdots + a_nv_n$ 则其等价类为 $\overline{u} = \overline{b_1w_1 + \cdots + b_kw_k + a_{k+1}v_{k+1} + \cdots + a_nv_n} = \overline{b_1w_1} + \cdots + \overline{b_kw_k} + a_{k+1}\overline{v}_{k+1} + \cdots + a_n\overline{v}_n$ 由于 $b_iw_i \in W$，故 $\overline{b_iw_i} = \overline{0}$，因此 $\overline{u} = a_{k+1}\overline{v}_{k+1} + \cdots + a_n\overline{v}_n$ 即 $\{\overline{v}_{k+1}, \ldots, \overline{v}_n\}$ 张成 V/W。 □

【Remark】:

基的核心要求是两个：线性无关 + 张成整个空间。
这两个要求只和"基元素（陪集）的个数"以及"它们的运算性质"有关，和"每个基元素（陪集）内部的元素个数"无关。
满足这两个条件，就是基——哪怕基里只有 1 个陪集，哪怕每个陪集里有无限个点（内部元素无限）。

1.11.2 子空间上的直积

Definition

两个 $\mathbb{F}$ - 向量空间 W_1 和 W_2 的**直积** (product) 定义为:
$W_1 \times W_2 := \{(w_1, w_2) : w_i \in W_i,\ i = 1, 2\}$。
在通常意义下，它是一个 $\mathbb{F}$-向量空间。

Example

欧氏空间的直积设 $\mathbb{F}=\mathbb{R}$，$W_1=\mathbb{R}^2$（二维实向量空间，元素形如 (a,b)），$W_2=\mathbb{R}^3$（三维实向量空间，元素形如 (c,d,e)）。

则它们的直积为：$W_1\times W_2=\{((a,b),(c,d,e))\mid a,b,c,d,e\in\mathbb{R}\}$

其向量空间运算为分量：

加法：$((a,b),(c,d,e))+((a',b'),(c',d',e'))=((a+a',b+b'),(c+c',d+d',e+e'))$

数乘：对 $\lambda\in\mathbb{R}$，$\lambda\cdot((a,b),(c,d,e))=((\lambda a,\lambda b),(\lambda c,\lambda d,\lambda e))$

1.11.3 双线性与张量积

Definition

双线性映射的定义：设 V,W,U 是域 $\mathbb{F}$ 上的向量空间，映射 $f:V\times W\to U$ 若满足：

- 对第一个变量线性：$\forall v_1,v_2\in V,w\in W,a\in\mathbb{F}$，有 $f(v_1+v_2,w)=f(v_1,w)+f(v_2,w)$ 且 $f(av,w)=af(v,w)$；
- 对第二个变量线性：$\forall v\in V,w_1,w_2\in W,a\in\mathbb{F}$，有 $f(v,w_1+w_2)=f(v,w_1)+f(v,w_2)$ 且 $f(v,aw)=af(v,w)$，

则称 f 为**双线性映射**。

Example

实数乘法：取 $V=\mathbb{R}$，$W=\mathbb{R}$，$U=\mathbb{R}$（均为实向量空间），定义映射 $f:\mathbb{R}\times\mathbb{R}\to\mathbb{R}$ 为 $f(a,b)=a\times b$（普通实数乘法）。

验证双线性：

- 对第一个变量：$f(a_1+a_2,b)=(a_1+a_2)b=a_1b+a_2b=f(a_1,b)+f(a_2,b)$，且 $f(ka,b)=(ka)b=k(ab)=kf(a,b)$；
- 对第二个变量：$f(a,b_1+b_2)=a(b_1+b_2)=ab_1+ab_2=f(a,b_1)+f(a,b_2)$，且 $f(a,kb)=a(kb)=k(ab)=kf(a,b)$。

故实数乘法是双线性映射。

为什么需要张量积：线性代数的核心工具（如矩阵、线性变换、特征值等）直接适用于**线性映射**（定义域为单个向量空间，对“整个输入”线性）。

双线性映射的定义域是直积空间 $V\times W$（其本身是向量空间，是有序对 (v,w) 的集合），但双线性映射不满足 $V\times W\to U$ 的线性映射定义（线性映射要求对‘有序对整体’的数乘齐次性，而双线性映射仅对单个变量满足线性），因此无法直接用线性代数工具（如基、矩阵）研究。

为了用线性代数工具研究双线性关联，需构造新向量空间 (张量积) $V \otimes_{\mathbb{F}} W$，使得所有 $V \times W$ 上的双线性映射，都能唯一对应到 $V \otimes_{\mathbb{F}} W$ 上的线性映射（即把“双线性问题”转化为“线性问题”）。

【Remark】:

张量积空间 $V \otimes_{\mathbb{F}} W$ 不是线性映射的空间（同态空间），它是一个独立的向量空间。它的元素是“张量”。

由此，给出**张量积的构造和定义**:

Method

设 V, W 是域 $\mathbb{F}$ 上的向量空间，构造步骤如下:

1. 定义基础集合 $S := \{(v, w) \mid v \in V, w \in W\}$;
2. 构造 $\mathbb{F}$ 上由 S 张成的向量空间 $\mathrm{Span}_{\mathbb{F}}(S)$（元素为 S 中元素的有限线性组合，形如 $\sum_{i=1}^{n} a_i(v_i, w_i)$，$a_i \in \mathbb{F}, v_i \in V, w_i \in W$）;
3. 定义 $\mathrm{Span}_{\mathbb{F}}(S)$ 的子集 S_0，包含四类元素以刻画双线性规则（如加法、数乘的线性性）:

$$
\begin{aligned}
S_0 := &\{(v_1 + v_2, w) - (v_1, w) - (v_2, w) \mid v_1, v_2 \in V, w \in W\} \\
&\cup \{(v, w_1 + w_2) - (v, w_1) - (v, w_2) \mid v \in V, w_1, w_2 \in W\} \\
&\cup \{a(v, w) - (v, aw) \mid v \in V, w \in W, a \in \mathbb{F}\} \\
&\cup \{a(v, w) - (av, w) \mid v \in V, w \in W, a \in \mathbb{F}\} \subset \mathrm{Span}_{\mathbb{F}}(S),
\end{aligned}
$$

于是，给出张量积（张量积空间）的定义:

Definition

定义: **张量积空间为商空间** $\mathrm{Span}_{\mathbb{F}}(S)/\mathrm{Span}_{\mathbb{F}}(S_0)$，that is $V \otimes_{\mathbb{F}} W := \mathrm{Span}_{\mathbb{F}}(S)/\mathrm{Span}_{\mathbb{F}}(S_0)$;

Definition

秩 1 张量: 对 $v \in V, w \in W$，记 $v \otimes w$ 为张量积空间 $V \otimes W$ 中的元素（称为‘秩 1 张量’）。

张量（tensor）: 是张量积空间 $U \otimes V$ 中的**任意**元素（形如 $\sum_i a_i u_i \otimes v_i$，$a_i \in \mathbb{F}, u_i \in U, v_i \in V$）

同时 $v \otimes w$ 也为 (v, w) 在商空间中的等价类 $\overline{(v, w)}$。

张量积空间的基底:

当 V, W 有限维时，若 $\{e_i\}$ 是 V 的基、$\{f_j\}$ 是 W 的基，则 $\{e_i \otimes f_j\}$ 是 $V \otimes W$ 的基。

【Remark】:

$\{e_i \otimes f_j\}$ 不是一个数值（如 1、2、3），它是张量积空间的基向量，是线性空间中用于构建其他元素的“基本单元”，类似于 $\mathbb{R}^n$ 中标准基 $\mathbf{e}_i$（如 $\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$）的角色，本身没有数值，仅作为线性空间的基元素存在。

注意到：在构造向量空间的时候，我们并不用“直接等式 $=0$”，即 $(v_1+v_2, w)-(v_1, w)-(v_2, w)=0$，而是通过“将差放入 S_0 + 商空间”的方式，是为了先建立无约束的基础空间 $\mathrm{Span}_{\mathbb{F}}(S)$，再通过商空间严格施加双线性约束，确保张量积空间的构造在逻辑上是严谨且自洽的。

如果直接写 $(v_1+v_2, w)-(v_1, w)-(v_2, w)=0$，相当于“强行规定等式成立”，但向量空间的构造需要逻辑上的严谨性：我们需要先承认 $\mathrm{Span}_{\mathbb{F}}(S)$ 中这些元素的**独立性**，再通过“商空间”的方式严格推导出“满足双线性规则的空间”。

下面将给出这个定义的完整证明：

证明. (v_1, w) 和 (v_2, w) 是 S 中两个不同的生成元（因为 $v_1 \neq v_2$），在“自由张成”下，它们是独立的（即无法用对方的线性组合表示）。

(v_1+v_2, w) 是 S 中另一个生成元（因为 v_1+v_2 是 V 中元素，w 是 W 中元素，所以 $(v_1+v_2, w) \in S$）。

在“自由张成”的规则下，(v_1+v_2, w) 并不天然等于 $(v_1, w)+(v_2, w)$——它只是一个独立的基元，和 (v_1, w)、(v_2, w) 之间没有预设的线性关系。

将 $(v_1+v_2, w)-(v_1, w)-(v_2, w)$ 放入 S_0，再构造 $\mathrm{Span}_{\mathbb{F}}(S_0)$（$S_0$ 中元素的线性组合），最后通过商空间来“消去”这些元素：在商空间中，若两个元素的差属于 $\mathrm{Span}_{\mathbb{F}}(S_0)$，则它们被视为“等价”（即“差为 0”）。

因此：(v_1+v_2, w) 与 $(v_1, w)+(v_2, w)$ 的**差**是 $(v_1+v_2, w)-(v_1, w)-(v_2, w)$（属于 S_0，进而属于 $\mathrm{Span}_{\mathbb{F}}(S_0)$）。

因此在商空间 $V \otimes_{\mathbb{F}} W$ 中，(v_1+v_2, w) 的等价类 等于 $(v_1, w)+(v_2, w)$ 的等价类，即 $(v_1+v_2) \otimes w = v_1 \otimes w + v_2 \otimes w$。

类似的，我们可以给出剩下的三个表达式：

第二类：$(v, w_1+w_2)-(v, w_1)-(v, w_2)$——要求“对第二个变量 w 加法线性”，同理。

第三类：$a(v, w)-(v, aw)$——要求“对第二个变量 w 数乘线性”，即数乘 a 可以从 w 外移到 w 内。

第四类：$a(v,w)-(av,w)$——要求“对第一个变量 v 数乘线性”，即数乘 a 可以从 v 外移到 v 内。

□

我们可以用三个例子来演示：

1. **实直线与实直线的张量积**：取 $V=\mathbb{R}$，$W=\mathbb{R}$，则 $V\otimes_{\mathbb{R}}W\cong\mathbb{R}$。因为实数乘法是双线性的，张量积在此场景下可“还原”为实数自身，且 $a\otimes b$ 可对应到 ab（符合商空间中等价类的线性性）。
2. **二维与三维实空间的张量积**：设 $V=\mathbb{R}^2$（基为 $e_1=(1,0),e_2=(0,1)$），$W=\mathbb{R}^3$（基为 $f_1=(1,0,0),f_2=(0,1,0),f_3=(0,0,1)$），则 $V\otimes_{\mathbb{R}}W$ 是 6 维实向量空间，基为 $e_1\otimes f_1,e_1\otimes f_2,e_1\otimes f_3,e_2\otimes f_1,e_2\otimes f_2,e_2\otimes f_3$。
3. **多项式空间的张量积**：取 $V=\mathbb{R}[x]$（一元实多项式空间），$W=\mathbb{R}[y]$（另一元实多项式空间），则 $V\otimes_{\mathbb{R}}W$ 同构于二元实多项式空间 $\mathbb{R}[x,y]$。例如，$(a+bx)\otimes(c+dy)$ 对应二元多项式 $ac+adx+bcy+bdxy$，体现了“变量独立组合”的双线性性质。

接下来我们考察张量积的**维度**：

Theorem

设 U,V 是 $\mathbb{F}$-向量空间，分别以 S_1,S_2 为基。则 $U\otimes V$ 的基为 $S=\{\varepsilon\otimes\eta:\varepsilon\in S_1,\eta\in S_2\}$。
因此，$\dim_{\mathbb{F}}(U\otimes V)=\dim_{\mathbb{F}}(U)\cdot\dim_{\mathbb{F}}(V)$。

证明：我们证明集合 $S=\{\varepsilon\otimes\eta:\varepsilon\in S_1,\eta\in S_2\}$ 是 $U\otimes V$ 的基。
(i) $\mathrm{Span}_{\mathbb{F}}(S)=U\otimes V$。

$U\otimes V$ 中的任意元素都可写成有限和 $\sum_{i=1}^n u_i\otimes v_i$，其中 $u_i\in U$，$v_i\in V$。由于 S_1 是 U 的基，每个 u_i 可表示为 $u_i=\sum_{k=1}^s a_{ik}\varepsilon_k$，其中 $a_{ik}\in\mathbb{F}$。类似地，每个 v_i 可表示为 $v_i=\sum_{j=1}^t b_{ij}\eta_j$。由张量积的双线性性，我们有：

$$u_i\otimes v_i=\left(\sum_{k=1}^s a_{ik}\varepsilon_k\right)\otimes\left(\sum_{j=1}^t b_{ij}\eta_j\right)=\sum_{\substack{1\le k\le s\\1\le j\le t}}a_{ik}b_{ij}(\varepsilon_k\otimes\eta_j).$$

于是和式 $\sum_{i=1}^n u_i\otimes v_i$ 是 $\varepsilon_k\otimes\eta_j$ 型元素的线性组合。因此，S 张成 $U\otimes V$。
(ii) S 线性无关。

假设

$$\sum_{k,j}c_{kj}\varepsilon_k\otimes\eta_j=0,$$

其中仅有有限个 c_{kj} 非零。对每一对 (p,q)（其中 $\varepsilon_p \in S_1$，$\eta_q \in S_2$），考虑由对偶基定义的线性泛函 $\phi_p : U \to \mathbb{F}$ 和 $\psi_q : V \to \mathbb{F}$，即 $\phi_p(\varepsilon_k) = \delta_{pk}$，$\psi_q(\eta_j) = \delta_{qj}$（克罗内克 delta δ_{ij} 定义为 $\delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$）。

定义双线性映射 $f_{pq} : U \times V \to \mathbb{F}$ 为 $f_{pq}(u,v) = \phi_p(u)\psi_q(v)$。由泛性质，这诱导出线性映射 $\widetilde{f}_{pq} : U \otimes V \to \mathbb{F}$，使得 $\widetilde{f}_{pq}(u \otimes v) = f_{pq}(u,v)$。

于是有

$$\widetilde{f}_{pq}\left(\sum_{k,j} c_{kj}\varepsilon_k \otimes \eta_j\right) = \sum_{k,j} c_{kj}\widetilde{f}_{pq}(\varepsilon_k \otimes \eta_j) = \sum_{k,j} c_{kj}\phi_p(\varepsilon_k)\psi_q(\eta_j) = \sum_{k,j} c_{kj}\delta_{pk}\delta_{qj} = c_{pq}.$$

由于该和为零，故 $c_{pq} = 0$。这对所有 p,q 成立，因此所有系数均为零。故 S 线性无关。

综上，S 是 $U \otimes V$ 的基，且 $\dim_{\mathbb{F}}(U \otimes V) = \#(S_1) \cdot \#(S_2) = \dim_{\mathbb{F}}(U) \cdot \dim_{\mathbb{F}}(V)$。□

【Remark】:

张量是张量积空间 $U\otimes V$ 中的任意元素(形如 $\sum_i a_i u_i \otimes v_i$, $a_i \in \mathbb{F}, u_i \in U, v_i \in V$),其中 $u_i \otimes v_i$ 称为‘秩 1 张量’;
而**张量积**包含两层含义:
一是‘张量积操作’ $\otimes : U \times V \to U \otimes V$（构造秩 1 张量的运算）;
二是‘张量积空间’ $U \otimes V$（所有张量构成的向量空间）;张量是张量积操作与空间的共同产物。

1.11.4 COMPOSITION

设 $\{W_i\}_{i\in I}$ 是向量空间 V 的子空间族，I 为指标集（可有限或无限）。

1. 和（Sum）： $\sum_{i\in I} W_i$

定义：由**有限个** W_i 中元素的和构成，即

$$\sum_{i\in I} W_i := \left\{ \sum_{i\in J} w_i \,\middle|\, w_i \in W_i,\ J \subset I,\ \#(J) < \infty \right\} \subset V.$$

（无限和无定义，故限制为有限和，元素可表示为有限个 W_i 中元素的和。）

例子：取 $V = \mathbb{R}^3$，$I = \{1,2,3\}$，$W_1 = \mathrm{span}\{(1,0,0)\}$（x 轴），$W_2 = \mathrm{span}\{(0,1,0)\}$（y 轴），$W_3 = \mathrm{span}\{(0,0,1)\}$（z 轴）。则 $\sum_{i=1}^3 W_i = \mathbb{R}^3$，因任意 $(a,b,c) = a(1,0,0) + b(0,1,0) + c(0,0,1)$（有限和）。

2. 交（Intersection）： $\bigcap_{i\in I} W_i$

定义：所有 W_i 的**公共元素**组成的集合，即

$$\bigcap_{i\in I} W_i = \{v \in V \mid v \in W_i,\ \forall i \in I\},$$

是 V 的子空间。

例子：取 $V=\mathbb{R}^2$，$I=\{1,2\}$，$W_1=\operatorname{span}\{(1,0)\}$（x 轴），$W_2=\operatorname{span}\{(1,1)\}$（直线 $y=x$）。

则 $\bigcap_{i=1}^2 W_i=\{0\}$（仅零向量是公共元素）。

3. 并（Union）： $\bigcup_{i\in I} W_i$

定义：所有 W_i 的元素“合并”而成的集合，即

$$\bigcup_{i\in I} W_i=\{v\in V\mid \exists i\in I,\ v\in W_i\}.$$

例子：取 $V=\mathbb{R}^2$，$I=\{1,2\}$，$W_1=\operatorname{span}\{(1,0)\}$（x 轴），$W_2=\operatorname{span}\{(0,1)\}$（y 轴）。则 $\bigcup_{i=1}^2 W_i$ 是 x 轴和 y 轴上所有点的集合。

【Remark】：

但是该“并”**不是向量空间**（如 $(1,0)+(0,1)=(1,1)\notin\bigcup_{i=1}^2 W_i$，不满足加法封闭性）。

4. 内直和（Direct Sum）：$\bigoplus_{i\in I} W_i$（作为 V 的子集）

定义：由**有限个** W_i 的直和“扩展”到无限指标，即

$$\bigoplus_{i\in I} W_i:=\bigcup_{\substack{J\subset I\\ \#(J)<\infty}}\left(\bigoplus_{i\in J} W_i\right)\subset V.$$

有限直和 $\bigoplus_{i\in J} W_i$ 满足“分解唯一性”，内直和要求这种唯一性推广到“无限和中仅有限项非零的分解”。

我们可以考察几个经典的例子：

Example

1. 取 $V=\mathbb{R}^\infty$，无限维实向量空间，元素为**有限非零**的实数列 $(a_1,a_2,\dots)$。$I=\mathbb{N}$，$W_i=\operatorname{span}\{e_i\}$（$e_i$ 是第 i 个分量为 1、其余为 0 的向量）。

则 $\bigoplus_{i\in\mathbb{N}} W_i=V$，因任意 $v\in V$ 是有限个 e_i 的线性组合，且分解唯一（每个分量对应唯一的 W_i 中元素）。

2. 又如 $\bigoplus_{i\in I}\mathbb{F}$ 是“有限支撑”的元组空间：给指标集 I 中每个元素 i 分配一个 $\mathbb{F}$ 的拷贝，空间中的元素是“只有有限个位置非零的元组”。

2.1：例子 1（有限指标集）：取 $I=\{1,2,3\}$，则 $\bigoplus_{i=1}^3\mathbb{F}=\mathbb{F}\oplus\mathbb{F}\oplus\mathbb{F}=\mathbb{F}^3$（就是普通的 3 维向量空间，因为有限个分量自然“有限支撑”）。

2.2：例子 2（无限指标集）：取 $I=\mathbb{N}$（自然数集），则 $\bigoplus_{n\in\mathbb{N}}\mathbb{F}$ 中的元素是“只有有限个自然数位置非零的序列”。比如 $(a_0,a_1,0,0,\dots)$（前两个分量非零，其余为 0）属于这个空间；但无限非零的序列（如 $(1,1,1,\dots)$）不属于，因为“非零分量的位置集合（支撑）”不是有限的。

5. 直积（Product）：$\prod_{i\in I} W_i$

定义：由所有"序列"$(w_i)_{i\in I}$（其中 $w_i \in W_i$）组成的向量空间，运算为**逐分量**进行（和、数乘对每个分量单独操作）。

例子：取 $I = \mathbb{N}$，$W_i = \mathbb{R}$（对所有 $i \in \mathbb{N}$）。则 $\prod_{i\in\mathbb{N}} \mathbb{R} = \mathbb{R}^{\mathbb{N}}$（所有实数列的集合，包括**无限非零项**的序列，如 $(1, 1, 1, \ldots)$）。

6. 张量积（Tensor Product）：$\bigotimes_{i\in I} W_i$

定义：多重张量积的推广，先通过"有限张量积"（如 $W_1 \otimes W_2$）的**泛性质**（将多线性映射转化为线性映射），再扩展到无限指标（核心是"捕捉多重线性关系"）。

例子 1：有限维向量空间的张量积（双线性性的直观体现）

取指标集 $I = \{1, 2\}$，向量空间 $W_1 = \mathbb{R}^2$（其标准基为 $\mathbf{e}_1 = (1, 0)$、$\mathbf{e}_2 = (0, 1)$），$W_2 = \mathbb{R}^3$（其标准基为 $\mathbf{f}_1 = (1, 0, 0)$、$\mathbf{f}_2 = (0, 1, 0)$、$\mathbf{f}_3 = (0, 0, 1)$）。

张量积 $W_1 \otimes W_2$ 是 **6 维实向量空间**，它的一组基为 $\{\mathbf{e}_i \otimes \mathbf{f}_j \mid i = 1, 2;\ j = 1, 2, 3\}$（共 $2 \times 3 = 6$ 个基元素）。

若取 W_1 中的元素 $\mathbf{v} = 2\mathbf{e}_1 + 3\mathbf{e}_2$（即向量 $(2, 3)$），W_2 中的元素 $\mathbf{w} = 4\mathbf{f}_1 + 5\mathbf{f}_2 + 6\mathbf{f}_3$（即向量 $(4, 5, 6)$），则它们的张量积为：

$$\mathbf{v}\otimes\mathbf{w} = (2\mathbf{e}_1+3\mathbf{e}_2)\otimes(4\mathbf{f}_1+5\mathbf{f}_2+6\mathbf{f}_3) = 8(\mathbf{e}_1\otimes\mathbf{f}_1)+10(\mathbf{e}_1\otimes\mathbf{f}_2)+12(\mathbf{e}_1\otimes\mathbf{f}_3)+12(\mathbf{e}_2\otimes\mathbf{f}_1)+15(\mathbf{e}_2\otimes\mathbf{f}_2)+18(\mathbf{e}_2\otimes\mathbf{f}_3).$$

这个例子直观体现了张量积的**双线性性**：对 W_1 和 W_2 中的元素分别保持线性，并且通过"基向量的张量积"生成了新空间的基。

例子 2：多项式空间的张量积（多重线性与变量组合）

取 $W_1 = \mathbb{R}[x]$（一元实系数多项式空间，其基为 $\{1, x, x^2, \ldots\}$），$W_2 = \mathbb{R}[y]$（另一元实系数多项式空间，其基为 $\{1, y, y^2, \ldots\}$）。

张量积 $W_1 \otimes W_2$ 自然同构于**二元实系数多项式空间** $\mathbb{R}[x, y]$。

若 $p(x) = 1 + 2x \in \mathbb{R}[x]$，$q(y) = 3 + 4y + 5y^2 \in \mathbb{R}[y]$，则 $p(x) \otimes q(y)$ 对应二元多项式：

$$(1+2x)\otimes(3+4y+5y^2) = 3(1\otimes 1)+4(1\otimes y)+5(1\otimes y^2)+6(x\otimes 1)+8(x\otimes y)+10(x\otimes y^2),$$

也就是二元多项式 $3 + 4y + 5y^2 + 6x + 8xy + 10xy^2$。

该例子体现了张量积"捕捉变量间独立组合的多重线性关系"的特点——将单变量多项式的"线性组合"扩展为双变量多项式的"双线性组合"。

例子 3：线性变换空间的张量积（算子的张量表示）

设 $W_1 = \mathcal{L}(V, V')$（从向量空间 V 到向量空间 V' 的线性变换空间），$W_2 = \mathcal{L}(W, W')$（从向量空间 W 到向量空间 W' 的线性变换空间）。

张量积 $W_1 \otimes W_2$ 可自然对应到 ** 双线性变换空间 ** $\mathcal{L}(V \times W, V' \times W')$（或者更精细地说，对应 $\mathcal{L}(V \otimes W, V' \otimes W')$）。

取 $V = \mathbb{R}^2$、$V' = \mathbb{R}^3$，定义线性变换 $T \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^3)$ 为 $T(\mathbf{e}_1) = (1, 0, 0)$、$T(\mathbf{e}_2) = (0, 1, 0)$；取 $W = \mathbb{R}^3$、$W' = \mathbb{R}^2$，定义线性变换 $S \in \mathcal{L}(\mathbb{R}^3, \mathbb{R}^2)$ 为 $S(\mathbf{f}_1) = (1, 0)$、$S(\mathbf{f}_2) =$

$(0,1)$、$S(\mathbf{f}_3) = (1,1)$；则张量积 $T \otimes S$ 作用于 $\mathbb{R}^2 \otimes \mathbb{R}^3$ 中的元素 $\mathbf{v} \otimes \mathbf{w}$ 时，满足 $(T \otimes S)(\mathbf{v} \otimes \mathbf{w}) = T(\mathbf{v}) \otimes S(\mathbf{w})$。例如，元素 $\mathbf{e}_1 \otimes \mathbf{f}_1$ 在 $T \otimes S$ 下的像为 $T(\mathbf{e}_1) \otimes S(\mathbf{f}_1) = (1,0,0) \otimes (1,0)$。

该例子体现了张量积在 ** 算子理论 ** 中的作用：将“单个空间上的线性变换”扩展为“两个空间上的双线性变换”，是研究多线性算子的复合与表示的核心工具。

总结：运算与直和条件的逻辑

运算	核心特征	直和相
和 $\sum$	有限项和，不要求唯一性	是直和的“基础集合”，需额外条件
交 $\bigcap$	公共元素，刻画子空间的重叠	直和要求“子空间间无交（除
并 $\bigcup$	元素合并，一般不封闭	仅用于直观理解“集合合
内直和 $\bigoplus$	有限直和的无限扩展，分解唯一（有限非零项）	是和 $\sum$ 满足“分解唯一性”后的结
直积 $\prod$	允许无限非零项的序列，分量独立运算	与内直和的区别：直积包含“无限非零
张量积 $\bigotimes$	捕捉多重线性关系，泛性质为核心	与直和无关，是另一类“组合”子空

Chapter 2

线性变换

2.1 线性变换及其核与像空间

2.1.1 映射

Definition

映射（mapping）：设 V 和 W 是向量空间，一个从 V 到 W 的**映射**（或**函数**）是一个规则 T，它为每个 $v \in V$ 指派唯一的元素 $T(v) \in W$。记作 $T : V \to W$。

映射可以看作函数这一概念的推广

然后我们可以自然地定义:**映射的定义域、值域与陪域**(Domain, Range and Codomain):

对于映射 $T : V \to W$,

- V 称为 T 的**定义域**;
- W 称为 T 的**陪域**;
- 集合 $\{T(v) : v \in V\}$ 称为 T 的**值域**，记作 $\mathrm{R}(T)$。

2.1.2 线性变换的定义及相关术语

Definition

线性变换（linear transformation）：设 V 和 W 是数域 F 上的向量空间。一个映射 $T : V \to W$ 称为**线性变换**，如果对所有 $u, v \in V$ 和 $c \in F$，满足：

1. $T(u + v) = T(u) + T(v)$;
2. $T(cv) = cT(v)$。

一般来说，线性变换的记号为 T，当然也可记作 Φ（fai）。并且线性变换也称为**线性映射**

当 $V = W$ 时，线性变换也可以称为：**线性算子（Linear operator）**。

我们首先考察其中 2 种特殊的线性映射——**零变换与恒等变换**（Zero transformation and identity transformation）：

Definition

- **零变换** $O : V \to W$ 定义为对所有 $v \in V$，$O(v) = 0_W$。
- **恒等变换** $I_V : V \to V$ 定义为对所有 $v \in V$，$I_V(v) = v$。

【Remark】：

对于任意的线性变换 T，我们有 $T\,0 = 0$. 但是如果 $T(x) = 0$, 我们无法推出 T 为零变换或者 $x = 0$。（见下部分的核空间）

接下来我们来考察一下线性变换的实例——微分与积分的运算：

考虑向量空间 $P_n(F)$（次数不超过 n 的多项式），定义**微分算子** $D : P_n(F) \to P_n(F)$ 为

$$D(p(x)) = \frac{d}{dx}p(x).$$

D 是线性变换，因为：

1. $D(p(x) + q(x)) = D(p(x)) + D(q(x))$；
2. $D(cp(x)) = cD(p(x))$ 对所有 $c \in F$。

定义**积分算子** $T : P_n(F) \to P_{n+1}(F)$ 为

$$T(p(x)) = \int_0^x p(t)dt.$$

T 也是线性变换。

2.1.3 线性变换的性质

设 $T : V \to W$ 是线性变换，有以下几条常见的性质：

Property

- $T(0)=0$，即：$T(0_V)=0_W$；

  proof：$T(0)=T(0\cdot 0)=0\cdot T(0)=0$

- $T(-v)=-T(v)$ 对所有 $v\in V$；

- $T(c_1v_1+c_2v_2+\cdots+c_kv_k)=c_1T(v_1)+c_2T(v_2)+\cdots+c_kT(v_k)$ 对所有 $v_1,v_2,\ldots,v_k\in V$ 和 $c_1,c_2,\ldots,c_k\in F$；

- 如果 T 和 S 都是线性映射，那么：也是 $T+S$ 线性映射;

  proof:$(T+S)(av1+bv2)=T(av1+bv2)+S(av1+bv2)=aT(v1)+bT(v2)+aS(v1)+bS(v2)=a(T+S)(v1)+b(T+S)(v2)$.

- 如果 T 和 S 都是线性映射，那么：也是 $T\circ S$ 线性映射。

  proof:$(T\circ S)(au1+bu2)=T(S(au1+bu2))=T(aS(u1)+bS(u2))=aT(S(u1))+bT(S(u2))=a(T\circ S)(u1)+b(T\circ S)(u2)$.

由最后两条性质，我们可以定义线性变换的叠加与复合：

2.1.4 线性变换的叠加与复合

首先，我们介绍加法与数乘，即线性变换的叠加：

Definition

线性变换的加法与数乘：设 $T,U:V\to W$ 是线性变换，$c\in F$。定义：

- $(T+U)(v)=T(v)+U(v)$ 对所有 $v\in V$；
- $(cT)(v)=cT(v)$ 对所有 $v\in V$。

其中 $T+U$ 和 cT 都是线性变换，我们分别称之为：线性变换的加法与数乘。

Definition

线性变换的复合：设 $T:V\to W$ 和 $U:W\to Z$ 是线性变换。它们的**复合** $U\circ T:V\to Z$ 定义为

$$(U\circ T)(v)=U(T(v))\quad \text{对所有}v\in V.$$

$U\circ T$ 也是线性变换。我们称之为：线性变换的复合。

Property

线性变换的复合满足结合律：若 $T : V \to W$，$U : W \to Z$，$S : Z \to Y$，则

$$S \circ (U \circ T) = (S \circ U) \circ T.$$

证明. proof:

□

2.1.5 线性变换 T 的核与像

Definition

核空间/零空间（Kernel space /null space）：设 $T : V \to W$ 是线性变换。T 的**核空间**（或**零空间**）定义为

$$N_T = N\ T\ = ker(T) = \{v \in V : T(v) = 0_W\}.$$

换言之，核空间的几何意义就是记录**线性变换 T 把定义域空间里面的哪些向量信息搞丢了**（被压缩到陪域空间的 0 向量）。

Definition

值域/像空间（Range/Image space）：T 的**值域**（或**像空间**）定义为

$$\mathrm{R}(T) = \{T(v) : v \in V\}.$$

换言之，值域 $R\ T$ 则是记录了定义域空间经过线性变换 T，在**新的空间输出的全部信息**。但是注意到初学者可能将值域和陪域空间 W 相等。这是错误的，R 不一定就是 W，空间 W 中可能有不属于 R 的元素（即没有被线性变换 T 映射到的元素）

Theorem

设 $T : V \to W$ 是线性变换，则：$ker(T)$ 和 $R(T)$ 均为向量空间，进一步地，$\mathrm{ker}(T)$ 是 V 的子空间；$\mathrm{R}(T)$ 是 W 的子空间。

证明. proof: 要证明 $\mathrm{Ker}(T)$ 和 $\mathrm{Im}(T)$ 是子空间。

- 核（Kernel）：设 $v_1, v_2 \in \mathrm{Ker}(T)$，$a, b \in \mathbb{F}$。则 $T(av_1 + bv_2) = aT(v_1) + bT(v_2) = 0$，因此 $av_1 + bv_2 \in \mathrm{Ker}(T)$。
- 像（Image）：设 $w_1, w_2 \in \mathrm{Im}(T)$，故存在 $v_1, v_2 \in V$ 使得 $T(v_i) = w_i$。对于 $a, b \in \mathbb{F}$，有 $aw_1 + bw_2 = aT(v_1) + bT(v_2) = T(av_1 + bv_2) \in \mathrm{Im}(T)$。

□

接下来我们通过一个例子来更好地理解核与值域的概念：

Example

对于 $T:\mathbb{R}^3\to\mathbb{R}^2$，$T(x,y,z)=(x-y,y+z)$

求 N_T。

令 $T(x,y,z)=(0,0)$，即 $\begin{cases}x-y=0\\ y+z=0\end{cases}$，解得 $x=y,y=-z$。取 $(x,y,z)=(1,1,-1)$，所以 $N_T=\text{Span}\{(1,1,-1)\}$。

求 $R(T)$

对于任意 $(a,b,c)\in\mathbb{R}^3$，$T(a,b,c)=(a-b,b+c)$。因为 $a-b$ 和 $b+c$ 都可以遍历所有实数 $\mathbb{R}$，所以 $R(T)=\mathbb{R}^2$。$R(T)$ 是 T 作用下 $\mathbb{R}^3$ 在 $\mathbb{R}^2$ 中的像集。

有意思的是，R 和 N 不一定是不同的。

Example

如线性变换 $T:R^2\to R^2$ 的例子：$T(a,b)=(b,0)$，我们有 $N(T)=R(T)$。

现在我们来考察一下线性变换对线性相关性的保持性：

Theorem

设 $T:V\to W$ 是线性变换，$B=\{v_1,v_2,\ldots,v_n\}$ 是 V 的基。那么：$\mathrm{R}(T)=\text{span}(\{T(v_1),T(v_2),\ldots,T(v_n)\})$。

证明. proof:

要证明 $R(T)=\text{span}\{T(v_1),T(v_2),\ldots,T(v_n)\}$，只需证明两边互相包含：

1. 任取 $w\in R(T)$，则存在 $v\in V$ 使 $T(v)=w$。因 $\{v_i\}$ 是基，$v=\sum a_iv_i$，故 $w=\sum a_iT(v_i)\in\text{span}\{T(v_i)\}$。

2. 任取 $w\in\text{span}\{T(v_i)\}$，则 $w=\sum b_iT(v_i)=T\left(\sum b_iv_i\right)\in R(T)$。

综上，$R(T)=\text{span}\{T(v_1),\ldots,T(v_n)\}$。

□

【Remark】：

注意到：$\mathrm{R}(T)=\text{span}(\{T(v_1),T(v_2),\ldots,T(v_n)\})$，只需张成，无需是基（线性无关）。我们可以证明：只有当 T 是单射时，才有线性无关这一条件。

Corollary & Secondary Conclusion

C1：对于任意的向量集 $S \in V$，若 S 线性相关，则向量集合 $T(S)$ 也线性相关。

C2：特别的，若 S 线性无关，$T(S)$ 无法保证依然线性无关。

2.1.6 线性算子幂的核与像

对于线性算子的幂，我们有：

Property

设 V 是 $\mathbb{C}$-向量空间，$T: V \to V$ 是 $\mathbb{C}$-线性算子。
那么对任意正整数 $i \in \mathbb{N}_+$，$\ker(T^i)$ 和 $\mathrm{Im}(T^i)$ 都是 V 的 $\mathbb{C}$-向量子空间。

证明. (i) 先证对任意 $i \in \mathbb{N}_+$，T^i 是 $\mathbb{C}$-线性映射。

我们对 i 用数学归纳法。当 $i = 1$ 时，显然 T 是 $\mathbb{C}$-线性映射。假设当 $i = k-1$ 时 T^{k-1} 是 $\mathbb{C}$-线性映射，我们证明 $i = k$ 时结论成立。对任意标量 $a, b \in \mathbb{C}$ 和向量 $u, v \in V$，有

$$\begin{aligned} T^k(au+bv) &= T^{k-1}(T(au+bv)) = T^{k-1}(aT(u)+bT(v)) \\ &= aT^{k-1}(T(u)) + bT^{k-1}(T(v)) = aT^k(u) + bT^k(v). \end{aligned}$$

由数学归纳法，对所有 $i \in \mathbb{N}_+$，T^i 是 $\mathbb{C}$-线性映射。

证明 $\ker(T^i)$ 是 $\mathbb{C}$-向量子空间。对任意标量 $a, b \in \mathbb{C}$ 和向量 $u, v \in \ker(T^i)$，有

$$T^i(au+bv) = aT^i(u) + bT^i(v) = a \cdot 0 + b \cdot 0 = 0.$$

因此 $au + bv \in \ker(T^i)$，故 $\ker(T^i)$ 是 $\mathbb{C}$-向量子空间。

证明 $\mathrm{Im}(T^i)$ 是 $\mathbb{C}$-向量子空间。对任意标量 $a, b \in \mathbb{C}$ 和向量 $u, v \in \mathrm{Im}(T^i)$，即存在 $u', v' \in V$ 使得 $u = T^i(u')$，$v = T^i(v')$。则

$$T^i(au'+bv') = aT^i(u') + bT^i(v') = au + bv.$$

因此 $au + bv \in \mathrm{Im}(T^i)$，故 $\mathrm{Im}(T^i)$ 是 $\mathbb{C}$-向量子空间。 □

并且，我们还可以观察到，随着幂的次数逐渐增大：

Fact or Background

任意正整数 $i > j \in \mathbb{N}_+$，有 $\mathrm{Im}(T^i) \subset \mathrm{Im}(T^j)$：即即复合越多，值域空间越小；
而 $\ker(T^i) \supset \ker(T^j)$ 对任意正整数 $i > j$ 成立：即复合越多，kernal 空间越大；

Lemma

像的稳定性： $\mathrm{Im}(T^i) = \mathrm{Im}(T^j) \implies \mathrm{Im}(T^{j+n}) = \mathrm{Im}(T^j)$

证明. 设 V 是 $\mathbb{C}$-向量空间，$T: V \to V$ 是 $\mathbb{C}$-线性映射。已知存在正整数 $i > j$ 使得 $\mathrm{Im}(T^i) = \mathrm{Im}(T^j)$，需证对任意非负整数 n，有 $\mathrm{Im}(T^{j+n}) = \mathrm{Im}(T^j)$。

采用数学归纳法证明：

基例 $n = 1$：因 $i > j$，故 $i \geq j+1$。由像的包含性结论 $\mathrm{Im}(T^i) \subset \mathrm{Im}(T^{j+1}) \subset \mathrm{Im}(T^j)$，且已知 $\mathrm{Im}(T^i) = \mathrm{Im}(T^j)$，故 $\mathrm{Im}(T^{j+1}) = \mathrm{Im}(T^j)$。

归纳假设：假设对非负整数 $k-1$，有 $\mathrm{Im}(T^{j+k-1}) = \mathrm{Im}(T^j)$。

归纳步骤（$n = k$）：由像的定义，$\mathrm{Im}(T^{j+k}) = T^{j+k}(V) = T(T^{j+k-1}(V))$。

由归纳假设 $\mathrm{Im}(T^{j+k-1}) = \mathrm{Im}(T^j)$，即 $T^{j+k-1}(V) = T^j(V)$，故：

$$\mathrm{Im}(T^{j+k}) = T(T^j(V)) = T^{j+1}(V) = \mathrm{Im}(T^{j+1}) = \mathrm{Im}(T^j)$$

因此，由数学归纳法，对任意非负整数 n，$\mathrm{Im}(T^{j+n}) = \mathrm{Im}(T^j)$。 □

Lemma

核的稳定性： $\ker(T^i) = \ker(T^j) \implies \ker(T^{j+n}) = \ker(T^j)$

证明. 设 V 是 $\mathbb{C}$-向量空间，$T: V \to V$ 是 $\mathbb{C}$-线性映射。已知存在正整数 $i > j$ 使得 $\ker(T^i) = \ker(T^j)$，需证对任意非负整数 n，有 $\ker(T^{j+n}) = \ker(T^j)$。

采用数学归纳法证明：

基例 $n = 1$：因 $i > j$，故 $i \geq j+1$。由核的包含性结论 $\ker(T^j) = \ker(T^i) \supset \ker(T^{j+1}) \supset \ker(T^j)$，故 $\ker(T^{j+1}) = \ker(T^j)$。

归纳假设：假设对非负整数 $k-1$，有 $\ker(T^{j+k-1}) = \ker(T^j)$。

归纳步骤（$n = k$）：任取 $v \in \ker(T^{j+k})$，则 $T^{j+k}(v) = 0$。

分解得 $T^{j+k-1}(T(v)) = 0$，故 $T(v) \in \ker(T^{j+k-1})$。

由归纳假设 $\ker(T^{j+k-1}) = \ker(T^j)$，得 $T(v) \in \ker(T^j)$，即 $T^j(T(v)) = 0$，即 $T^{j+1}(v) = 0$，故 $v \in \ker(T^{j+1})$。由基例 $n = 1$ 知 $\ker(T^{j+1}) = \ker(T^j)$，故 $v \in \ker(T^j)$，即 $\ker(T^{j+k}) \subset \ker(T^j)$。

又由核的包含性结论 $\ker(T^{j+k}) \supset \ker(T^j)$，故 $\ker(T^{j+k}) = \ker(T^j)$。

因此，由数学归纳法，对任意非负整数 n，$\ker(T^{j+n}) = \ker(T^j)$。 □

Lemma

有限维下像与核的等价性: 设 $\dim_{\mathbb{C}}(V) < \infty$，那么对任意正整数 i, j:

1. $\mathrm{Im}(T^i) = \mathrm{Im}(T^j) \implies \ker(T^i) = \ker(T^j)$;
2. $\ker(T^i) = \ker(T^j) \implies \mathrm{Im}(T^i) = \mathrm{Im}(T^j)$。

证明. **证明 $\mathbf{Im}(T^i) = \mathbf{Im}(T^j) \implies \ker(T^i) = \ker(T^j)$**

若 $i = j$，结论显然成立。下设 $i > j$。

由 $\mathrm{Im}(T^i) = \mathrm{Im}(T^j)$，得 $\dim_{\mathbb{C}}(\mathrm{Im}(T^i)) = \dim_{\mathbb{C}}(\mathrm{Im}(T^j))$。根据秩-零化度定理（对任意正整数 k，$\dim_{\mathbb{C}}(V) = \dim_{\mathbb{C}}(\mathrm{Im}(T^k)) + \dim_{\mathbb{C}}(\ker(T^k))$），有:

$$\dim_{\mathbb{C}}(\ker(T^i)) = \dim_{\mathbb{C}}(V) - \dim_{\mathbb{C}}(\mathrm{Im}(T^i)) = \dim_{\mathbb{C}}(V) - \dim_{\mathbb{C}}(\mathrm{Im}(T^j)) = \dim_{\mathbb{C}}(\ker(T^j))$$

又由核的包含性结论 $\ker(T^i) \supset \ker(T^j)$，结合维数相等，得 $\ker(T^i) = \ker(T^j)$。

证明 $\ker(T^i) = \ker(T^j) \implies \mathbf{Im}(T^i) = \mathbf{Im}(T^j)$

若 $i = j$，结论显然成立。下设 $i > j$。

由 $\ker(T^i) = \ker(T^j)$，得 $\dim_{\mathbb{C}}(\ker(T^i)) = \dim_{\mathbb{C}}(\ker(T^j))$。根据秩-零化度定理，有:

$$\dim_{\mathbb{C}}(\mathrm{Im}(T^i)) = \dim_{\mathbb{C}}(V) - \dim_{\mathbb{C}}(\ker(T^i)) = \dim_{\mathbb{C}}(V) - \dim_{\mathbb{C}}(\ker(T^j)) = \dim_{\mathbb{C}}(\mathrm{Im}(T^j))$$

又由像的包含性结论 $\mathrm{Im}(T^i) \subset \mathrm{Im}(T^j)$，结合维数相等，得 $\mathrm{Im}(T^i) = \mathrm{Im}(T^j)$。 □

【Remark】:

在无限维的情况下，该引理不成立:

证明. 令 $V = \mathbb{C}^\infty$（所有复数列 $(a_1, a_2, \ldots, a_n, \ldots)$ 构成的复向量空间），定义两个线性映射:

T_1（左移映射）: $T_1(a_1, a_2, \ldots, a_n, \ldots) = (a_2, a_3, \ldots, a_n, \ldots)$;

T_2（右移 + 零填充映射）: $T_2(a_1, a_2, \ldots, a_n, \ldots) = (0, a_1, a_2, \ldots, a_n, \ldots)$。

分析 T_1: 像（Im）的分析: 对任意 $i, j \in \mathbb{N}_+$，$\mathrm{Im}(T_1^i) = \mathbb{C}^\infty$。

因为任给复数列 $(b_1, b_2, \ldots)$,总能找到原像 $(b_{i+1}, b_{i+2}, \ldots)$,使得 $T_1^i(b_{i+1}, b_{i+2}, \ldots) = (b_1, b_2, \ldots)$。

同理 $\mathrm{Im}(T_1^j) = \mathbb{C}^\infty$，故 $\mathrm{Im}(T_1^i) = \mathrm{Im}(T_1^j) = V$。

核（Ker）的分析: $\mathrm{Ker}(T_1^i)$ 是满足 $T_1^i(a_1, a_2, \ldots) = 0$ 的列，即 $(a_{i+1}, a_{i+2}, \ldots) = (0, 0, \ldots)$，因此: $\mathrm{Ker}(T_1^i) = \{(a_1, \ldots, a_i, 0, 0, \ldots) \mid a_1, \ldots, a_i \in \mathbb{C}\}$,

同理 $\mathrm{Ker}(T_1^j) = \{(a_1, \ldots, a_j, 0, 0, \ldots) \mid a_1, \ldots, a_j \in \mathbb{C}\}$。

当 $i \neq j$ 时（如 $i > j$），$\mathrm{Ker}(T_1^i)$ 与 $\mathrm{Ker}(T_1^j)$ 包含的序列长度不同（前 i 项 vs 前 j 项非零），故 $\mathrm{Ker}(T_1^i) \neq \mathrm{Ker}(T_1^j)$。

分析 T_2: 说明 “$\mathrm{Ker}(T^i) = \mathrm{Ker}(T^j)$ 不蕴含 $\mathrm{Im}(T^i) = \mathrm{Im}(T^j)$”

核（Ker）的分析：$T_2^i(a_1, a_2, \dots) = (0, \dots, 0, a_1, a_2, \dots)$（前 i 个分量为 0），要使其为零向量，需 $a_1 = a_2 = \cdots = 0$，因此对任意 $i \in \mathbb{N}_+$，$\mathrm{Ker}(T_2^i) = \{0\}$，故 $\mathrm{Ker}(T_2^i) = \mathrm{Ker}(T_2^j) = \{0\}$。

像（Im）的分析：$\mathrm{Im}(T_2^i) = \{(0, \dots, 0, a_1, a_2, \dots) \mid a_1, a_2, \dots \in \mathbb{C}\}$（前 i 个分量为 0）。当 $i \neq j$ 时（如 $i > j$），$\mathrm{Im}(T_2^i)$ 中序列的“零前缀长度”为 i，$\mathrm{Im}(T_2^j)$ 为 j，显然不相等（例如 $(0, 1, 0, 0, \dots) \in \mathrm{Im}(T_2^2)$ 但不属于 $\mathrm{Im}(T_2^3)$），故 $\mathrm{Im}(T_2^i) \neq \mathrm{Im}(T_2^j)$。 □

Theorem

定义 $\mathrm{Ker}(T^\infty) := \bigcup_{i\geq 1} \mathrm{Ker}(T^i)$ 且 $\mathrm{Im}(T^\infty) := \bigcap_{i\geq 1} \mathrm{Im}(T^i)$。
那么当 $\dim_{\mathbb{C}}(V) < \infty$ 时，$V = \mathrm{Ker}(T^\infty) \oplus \mathrm{Im}(T^\infty)$。

证明. 有限维情形的证明：$V = \mathrm{Ker}(T^\infty) \oplus \mathrm{Im}(T^\infty)$ 我们分三步证明：

步骤 1：存在正整数 k，使得 $\mathrm{Ker}(T^\infty) = \mathrm{Ker}(T^k)$ 且 $\mathrm{Im}(T^\infty) = \mathrm{Im}(T^k)$ 线性映射的像列是递减的：$\mathrm{Im}(T) \supseteq \mathrm{Im}(T^2) \supseteq \mathrm{Im}(T^3) \supseteq \cdots$。

对应的维数序列 $\dim(\mathrm{Im}(T^i))$ 是非递增的，且下界为 0。由于 V 是有限维的，维数序列必然稳定，即存在正整数 k，使得对所有 $n \in \mathbb{N}_+$，有 $\dim(\mathrm{Im}(T^{k+n})) = \dim(\mathrm{Im}(T^k))$。

有限维子空间“维数相等则子空间相等”，故 $\mathrm{Im}(T^{k+n}) = \mathrm{Im}(T^k)$ 对所有 $n \in \mathbb{N}_+$ 成立。

结合上面的的结论（$\mathrm{Im}(T^i) = \mathrm{Im}(T^j) \implies \mathrm{Ker}(T^i) = \mathrm{Ker}(T^j)$），得 $\mathrm{Ker}(T^{k+n}) = \mathrm{Ker}(T^k)$ 对所有 $n \in \mathbb{N}_+$ 成立。

因此，$\mathrm{Ker}(T^\infty) = \bigcup_{i\geq 1} \mathrm{Ker}(T^i) = \mathrm{Ker}(T^k)$，$\mathrm{Im}(T^\infty) = \bigcap_{i\geq 1} \mathrm{Im}(T^i) = \mathrm{Im}(T^k)$。

步骤 2：$\mathrm{Ker}(T^k) + \mathrm{Im}(T^k)$ 是直和（即 $\mathrm{Ker}(T^k) \cap \mathrm{Im}(T^k) = \{0\}$）任取 $u \in \mathrm{Ker}(T^k) \cap \mathrm{Im}(T^k)$，由 $u \in \mathrm{Im}(T^k)$，存在 $v \in V$ 使得 $u = T^k(v)$。由 $u \in \mathrm{Ker}(T^k)$，得 $T^k(u) = 0$。

代入 $u = T^k(v)$，得 $T^k(T^k(v)) = T^{2k}(v) = 0$，即 $v \in \mathrm{Ker}(T^{2k})$。

由步骤 1，$\mathrm{Ker}(T^{2k}) = \mathrm{Ker}(T^k)$，故 $v \in \mathrm{Ker}(T^k)$，因此 $u = T^k(v) = 0$（因 $v \in \mathrm{Ker}(T^k)$，$T^k(v) = 0$）。故 $\mathrm{Ker}(T^k) \cap \mathrm{Im}(T^k) = \{0\}$，即 $\mathrm{Ker}(T^k) + \mathrm{Im}(T^k) = \mathrm{Ker}(T^k) \oplus \mathrm{Im}(T^k)$。

步骤 3：$V = \mathrm{Ker}(T^k) \oplus \mathrm{Im}(T^k)$

由秩 - 零化度定理，对 T^k 有 $\dim(V) = \dim(\mathrm{Ker}(T^k)) + \dim(\mathrm{Im}(T^k))$。

结合步骤 2，$\dim(\mathrm{Ker}(T^k) + \mathrm{Im}(T^k)) = \dim(\mathrm{Ker}(T^k)) + \dim(\mathrm{Im}(T^k)) - \dim(\mathrm{Ker}(T^k) \cap \mathrm{Im}(T^k)) = \dim(V) + 0 = \dim(V)$。又因 $\mathrm{Ker}(T^k) + \mathrm{Im}(T^k) \subseteq V$，且 V 是有限维的，故 $\mathrm{Ker}(T^k) + \mathrm{Im}(T^k) = V$。

结合步骤 2 的直和，得 $V = \mathrm{Ker}(T^k) \oplus \mathrm{Im}(T^k)$。

再由步骤 1 的 $\mathrm{Ker}(T^\infty)=\mathrm{Ker}(T^k)$、$\mathrm{Im}(T^\infty)=\mathrm{Im}(T^k)$，最终得 $V=\mathrm{Ker}(T^\infty)\oplus\mathrm{Im}(T^\infty)$。

□

【Remark】:

该定理在无限维的时候不成立

证明. 构造无限维空间 $V=\mathbb{C}^\infty$（所有复数列构成的空间），并定义两个线性映射:

T_1（左移映射）：$T_1(a_1,a_2,\dots)=(a_2,a_3,\dots)$；

T_2（右移 + 零填充映射）：$T_2(a_1,a_2,\dots)=(0,a_1,a_2,\dots)$。

分析 T_1:$\mathrm{Ker}(T_1^\infty)=\bigcup_{i\geq 1}\mathrm{Ker}(T_1^i)$。其中 $\mathrm{Ker}(T_1^i)=\{(a_1,\dots,a_i,0,0,\dots)\mid a_1,\dots,a_i\in\mathbb{C}\}$（前 i 项任意，后续为 0）。

当 i 取遍所有正整数时，$\bigcup_{i\geq 1}\mathrm{Ker}(T_1^i)=\mathbb{C}^\infty=V$（任何序列都可被某个 i 包含）。$\mathrm{Im}(T_1^\infty)=\bigcap_{i\geq 1}\mathrm{Im}(T_1^i)$。因 $\mathrm{Im}(T_1^i)=\mathbb{C}^\infty$（任给序列都有原像），故交集为 $\mathbb{C}^\infty=V$。

此时 $\mathrm{Ker}(T_1^\infty)\oplus\mathrm{Im}(T_1^\infty)=V\oplus V\neq V$（直和要求交为 $\{0\}$，但此处交为 V）。

分析 T_2：$\mathrm{Ker}(T_2^\infty)=\bigcup_{i\geq 1}\mathrm{Ker}(T_2^i)$。

其中 $\mathrm{Ker}(T_2^i)=\{0\}$（因 $T_2^i(a_1,a_2,\dots)=(0,\dots,0,a_1,a_2,\dots)$，要为 0 则所有 $a_j=0$），故并集为 $\{0\}$。$\mathrm{Im}(T_2^\infty)=\bigcap_{i\geq 1}\mathrm{Im}(T_2^i)$。其中 $\mathrm{Im}(T_2^i)=\{(0,\dots,0,a_{i+1},a_{i+2},\dots)\mid a_{i+1},\dots\in\mathbb{C}\}$（前 i 项为 0）。当 i 取遍所有正整数时，交集为 $\{0\}$（只有全零序列属于所有 $\mathrm{Im}(T_2^i)$）。

此时 $\mathrm{Ker}(T_2^\infty)\oplus\mathrm{Im}(T_2^\infty)=\{0\}\oplus\{0\}=\{0\}\neq V$。

综上，无限维空间中的结论不成立。 □

Method

在吗无限维空间中，如复数域的无限维空间，我们常常考察这两个线性映射:

T_1（左移映射）：$T_1(a_1,a_2,\dots)=(a_2,a_3,\dots)$；

T_2（右移 + 零填充映射）：$T_2(a_1,a_2,\dots)=(0,a_1,a_2,\dots)$。

Method

若需证明两个子空间（如 A, B）相等，常用路径如下：

1. **互为子集法**（最常用）：若能证明子空间满足 $A \subseteq B$ 且 $B \subseteq A$，则直接由集合相等的定义得 $A = B$

   - 【路径 1】（经典方法）：任取子空间 A 中的元素 v，需证 $v \in B$；再任取子空间 B 中的元素 w，需证 $w \in A$。由双向包含 $A \subseteq B$ 且 $B \subseteq A$，得 $A = B$。
   - 【路径 2】：双向包含夹逼法：若能证明子空间满足 $A \subseteq B \subseteq C \subseteq A$，则直接由集合相等的定义得 $A = B$（本质是"互为子集法"的简洁表述）。
   - 【路径 3】（如涉及值域空间的相等）若存在连接 A, B 的线性映射 $T : A \to B$ 或 $T : B \to A$，可通过映射的"像"（记为 $\mathrm{Im}(T)$）分析：

     用 $T(A)$ 或 $T(B)$ 表示全部的映射在对应子空间上的像（值域子空间），结合像的性质（如线性封闭性）推导 A 与 B 的包含关系。

2. **维度夹逼法：**如果 $A \subseteq B$ 同时：$dim(A) = dim(B)$，那么，我们可以认为 $A = B$。

3. **数学归纳法**：适用于涉及"整数指标的子空间序列"（如 A_n, B_n，$n \in \mathbb{N}$）的相等证明：先验证基例（如 $n = 1$ 时 $A_1 = B_1$），再假设 $n = k$ 时 $A_k = B_k$，递推证明 $n = k + 1$ 时 $A_{k+1} = B_{k+1}$。

2.1.7 课后练习

Exercise

1. 设 $T : V \to W$ 和 $U : W \to Z$ 是线性变换。证明：

   (a) $\mathrm{R}(U \circ T) \subseteq \mathrm{R}(U)$;

   (b) $\ker(T) \subseteq \ker(U \circ T)$;

2. 设 $T : V \to V$ 是线性变换，满足 $T^2 = T$（即 T 是幂等变换）。证明：

   (a) $V = \ker(T) \oplus \mathrm{R}(T)$;

   (b) 若 $v \in \ker(T) \cap \mathrm{R}(T)$，则 $v = 0$。

2.2 维度定理

这一节，我们将证明一个重要的定理.

但在此之前，我们将讨论两个核心概念：秩和零度，尤其是秩（rank），它将贯穿线

性代数的始终。

我们已经知道：核空间记录的是被线性变换 T “搞丢”的信息空间；而值域 $R\,T$ 则是记录了经过线性变换 T，在新的空间输出的全部信息。

我们有必要**考察核空间与值域的维度**，以**量化地**考察线性变换对信息传递的具体情况。

Definition

零度： 核空间 $(N(T))$ 的维度，简记为 *nullity*。
秩： 值域空间 $(R(T))$ 的维度，简记为 *rank*（或 r(T)）

现在我们可以论述**维度定理（秩-零度定理,***Rank-Nullity***）**；

Theorem

维度定理 (Dimension Theorem): 设 $T : V \to W$ 是线性变换，且 V 是有限维的。则

$$\dim(V) = \dim(\ker(T)) + \dim(\mathrm{R}(T)).$$

其中 $\dim(\ker(T))$ 是 T 的**零度**，$\dim(\mathrm{R}(T))$ 是 T 的**秩**。

证明. 设 $n = \dim_{\mathbb{F}}(V)$，$k = T$ 的零度（Nullity）。设 $\{u_1, \dots, u_k\}$ 是 $\mathrm{Ker}(T)$ 的一组基。将其扩充为 V 的一组基 $\mathcal{B} = \{u_1, \dots, u_k, w_1, \dots, w_{n-k}\}$。我们断言 $\{T(w_1), \dots, T(w_{n-k})\}$ 是 $\mathrm{Im}(T)$ 的一组基。

- ** 张成性 **：对任意 $w \in \mathrm{Im}(T)$，存在 $v = \sum_{i=1}^{k} a_i u_i + \sum_{j=1}^{n-k} b_j w_j$ 使得 $w = T(v)$。由于 $T(u_i) = 0$，故 $w = \sum_{j=1}^{n-k} b_j T(w_j)$。
- ** 线性无关性 **：假设 $\sum_{j=1}^{n-k} b_j T(w_j) = 0$。则 $T(\sum_{j=1}^{n-k} b_j w_j) = 0$，故 $\sum_{j=1}^{n-k} b_j w_j \in \mathrm{Ker}(T)$。因此，存在 a_i 使得 $\sum_{j=1}^{n-k} b_j w_j = \sum_{i=1}^{k} a_i u_i$。由 $\mathcal{B}$ 的线性无关性，所有 $b_j = 0$（且所有 $a_i = 0$）。

因此，T 的秩（Rank）满足 $\mathrm{Rank}(T) = n - k = \dim_{\mathbb{F}}(V) - 零度(T)$。 □

2.3 线性变换的矩阵表示

这一节，你将彻底理解矩阵的本质：**线性变换在两个向量空间的两组基底下的线性表示**。

首先我们先从“基决定”性质开始：

2.3.1 线性变换的基决定性质

我们先做出如下断言：线性变换 $T: V \to W$ 的作用可通过其在定义域向量空间 V 的一组基向量与陪域 W 上的一组向量上表现的行为完全确定。

详细的说，若向量空间 V 有一组有序基 $\{v_1, v_2, \ldots, v_n\}$，对于线性变换 $T: V \to W$，只要确定 T 对每个基向量 v_i 的作用（即 $T(v_i)$ 的值），那么 T 对 V 中任意向量 $x = \sum_{i=1}^{n} a_i v_i$ 的作用 $T(x) = \sum_{i=1}^{n} a_i T(v_i)$ 就被唯一确定。已知有序基向量的像，就可以通过线性组合即可得到任意向量的像。

Example

二维空间中，若已知 T 对基向量 $\mathbf{e}_1 = (1,0)$ 和 $\mathbf{e}_2 = (0,1)$ 的变换结果，则任意向量 $\mathbf{v} = a\mathbf{e}_1 + b\mathbf{e}_2$ 的变换结果为 $T(\mathbf{v}) = aT(\mathbf{e}_1) + bT(\mathbf{e}_2)$。

下面我们将为这段表述给出精准的表达与证明：

Theorem

设 V 和 W 是数域 F 上的向量空间，假设 $\{v_1, v_2, \ldots, v_n\}$ 是 V 的一组有序基。对于 W 中的 $w_1, w_2, \ldots, w_n$，存在唯一的线性变换 $T: V \to W$，使得对于 $i = 1, 2, \ldots, n$，有 $T(v_i) = w_i$。

【Remark】：

$w_1, w_2, \ldots, w_n$ 不一定是 W 的基，甚至它们张成的子空间 $\operatorname{span}\{w_1, w_2, \ldots, w_n\}$ 也可以不等于 W。线性变换的唯一性定理对 W 的维数没有要求，$\dim(W)$ 可以等于 n、大于 n 或小于 n，只要 $w_1, \ldots, w_n$ 是 W 中任意给定的向量即可。

证明. 任取 $x \in V$，因为 $\{v_1, v_2, \ldots, v_n\}$ 是 V 的基，所以 $x = \sum_{i=1}^{n} a_i v_i$，其中 $a_1, a_2, \ldots, a_n$ 是唯一的标量。定义 $T: V \to W$，$T(x) = \sum_{i=1}^{n} a_i w_i$。

验证 T 满足条件：当 x 取基向量 v_i 时，其系数 $a_i = 1$，其余 $a_j = 0$（$j \neq i$）。代入 T 的定义式，得 $T(v_i) = 1 \cdot w_i + 0 \cdot w_1 + \cdots + 0 \cdot w_{i-1} + 0 \cdot w_{i+1} + \cdots + 0 \cdot w_n = w_i$，即根据 T 的构造方式，直接得出 $T(v_i) = w_i$。

证明 T 的唯一性：假设存在另一个线性变换 $U: V \to W$ 满足 $U(v_i) = w_i$，则对于任意向量 $x = \sum_{i=1}^{n} a_i v_i$，有 $U(x) = U(\sum_{i=1}^{n} a_i v_i) = \sum_{i=1}^{n} a_i U(v_i) = \sum_{i=1}^{n} a_i w_i = T(x)$，故 $U = T$。 □

于是我们有以下的推论：

Corollary & Secondary Conclusion

设 V 和 W 是向量空间，假设 V 有一组有限基 $\{v_1, v_2, \ldots, v_n\}$。如果 $U, T: V \to W$ 是线性的，且对于 $i = 1, 2, \ldots, n$，$U(v_i) = T(v_i)$，那么 $U = T$。

我们将通过一个例子来理解上述定理与推论：

Example

设 $T:\mathbb{R}^2\to\mathbb{R}^2$ 是由 $T(a_1,a_2)=(2a_2-a_1,3a_1)$ 定义的线性变换，$U:\mathbb{R}^2\to\mathbb{R}^2$ 是线性的，已知 $U(1,2)=(3,3)$，$U(1,1)=(1,3)$，且 $\{(1,2),(1,1)\}$ 是 $\mathbb{R}^2$ 的一组基。

根据上述推论，因为 T 和 U 都是从 $\mathbb{R}^2$ 到 $\mathbb{R}^2$ 的线性变换，且在基 $\{(1,2),(1,1)\}$ 上取值相同（U 在基向量上的值已知，T 作用在基向量上也可计算得到相同结果），所以 $U=T$。

接下来，我们将把**矩阵表示与线性变换**联系起来：

2.3.2 坐标向量与线性变换的矩阵表示

Definition

设 $\beta=\{u_1,u_2,\ldots,u_n\}$ 是有限维向量空间 V 的一个有序基。对 $x\in V$，令 $a_1,a_2,\ldots,a_n$ 是唯一的标量，使得

$$x=\sum_{i=1}^{n}a_iu_i.$$

我们定义 x 相对于 β 的**坐标向量**，记为 $[x]_\beta$，为

$$[x]_\beta=\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}.$$

Example

例如：设 $V=P_2(R)$，且令 $\beta=\{1,x,x^2\}$ 是 V 的标准有序基。若 $f(x)=4+6x-7x^2$，则

$$[f]_\beta=\begin{pmatrix}4\\6\\-7\end{pmatrix}.$$

Definition

用上述记号，我们称由 $A_{ij}=a_{ij}$ 定义的 $m\times n$ 矩阵 A 为从向量空间 V→W 的线性变换 T 在 V 的有序基 β 和 W 的 γ 下的**矩阵表示**，并记为 $A=[T]_\beta^\gamma$。

其中 A 的第 j 列正是 $[T(v_j)]_\gamma$。即 V 的有序基 β 中的每个基向量 v，在作用 T 后

得到的在 W 中的向量 w，再把 w 用 W 中的有序基 γ 线性表示后找出坐标，竖直地排列起来。

即：

$$A = \begin{pmatrix} | & | & & | \\ [T(v_1)]_\gamma & [T(v_2)]_\gamma & \cdots & [T(v_n)]_\gamma \\ | & | & & | \end{pmatrix}.$$

若 $V = W$ 且 $\beta = \gamma$，则记为 $A = [T]_\beta$。

这个概念可能在理解上存在困难，所以我们将提供两个例子：

Example

例 1：设 $T : \mathbb{R}^2 \to \mathbb{R}^3$ 是由

$$T(a_1, a_2) = (a_1 + 3a_2, 0, 2a_1 - 4a_2)$$

定义的线性变换。令 β 和 γ 分别是 $\mathbb{R}^2$ 和 $\mathbb{R}^3$ 的标准有序基。

故：

$$T(1,0) = (1,0,2) = 1e_1 + 0e_2 + 2e_3$$

且

$$T(0,1) = (3,0,-4) = 3e_1 + 0e_2 - 4e_3.$$

因此

$$[T]_\beta^\gamma = \begin{pmatrix} 1 & 3 \\ 0 & 0 \\ 2 & -4 \end{pmatrix}.$$

Example

例 2：设 $T : P_3(R) \to P_2(R)$ 是由 $T(f(x)) = f'(x)$ 定义的线性变换。令 β 和 γ 分别是 $P_3(R)$ 和 $P_2(R)$ 的标准有序基。

则:

$$\begin{aligned} T(1) &= 0 \cdot 1 + 0 \cdot x + 0 \cdot x^2, \\ T(x) &= 1 \cdot 1 + 0 \cdot x + 0 \cdot x^2, \\ T(x^2) &= 0 \cdot 1 + 2 \cdot x + 0 \cdot x^2, \\ T(x^3) &= 0 \cdot 1 + 0 \cdot x + 3 \cdot x^2. \end{aligned}$$

所以:

$$[T]_\beta^\gamma = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

2.3.3 矩阵对向量的作用

在刚刚的小节中，我们已经明白了将 $T:\mathbb{R}^n \to \mathbb{R}^m$ 用矩阵表达的形式，那么 T 对某个向量的作用具体是怎么样呢?

我们给出如下的公式：

$$[T(v)]_\gamma = [T]_\beta^\gamma \cdot [v]_\beta$$

这一等式将抽象的线性变换 T 转化为具体的矩阵乘法运算，实现了线性变换的数值化表示。

证明. 1. 对于系数矩阵 $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$（共 n 个，$\forall v \in V_n$ 须 n 个系数确定），与 V_n 中的标准有序基 $\{e_1, e_2, \ldots, e_n\}$ 线性组合。可得 V_n 中任意一个向量 v，即 $\forall v \in V_n$，

$$v = \sum_{i=1}^n x_i \cdot e_i$$

称 $x = [v]_\beta = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = (x_1, x_2, \ldots, x_n)^T$（列向量）。

2. 注意到每一个 $e_j \in V_n$，在 W_m 中须 m 个参数确定（$\because \dim W_m = m$，$\therefore$ 每个 $T(e_j) = \sum_{i=1}^m a_{ij} \cdot w_i$，$a_{ij} \in \mathbb{R}$）。

3. 用 T 把 V_n 中的 v 映射到 W_m 并用 w_i 表示。

$$T(v) = T\left(\sum_{i=1}^n x_i e_i\right) = \sum_{i=1}^n x_i T(e_i) = x_1 T(e_1) + x_2 T(e_2) + \cdots + x_n T(e_n).$$

等价于以下式子:

$$x_1\begin{pmatrix}a_{11}\\a_{21}\\\vdots\\a_{m1}\end{pmatrix}+x_2\begin{pmatrix}a_{12}\\a_{22}\\\vdots\\a_{m2}\end{pmatrix}+\cdots+x_n\begin{pmatrix}a_{1n}\\a_{2n}\\\vdots\\a_{mn}\end{pmatrix}=\begin{pmatrix}x_1a_{11}+x_2a_{12}+\cdots+x_na_{1n}\\\vdots\\x_1a_{m1}+\cdots+x_na_{mn}\end{pmatrix}\in\mathbb{R}^m$$

进一步，可写为矩阵形式：

$$\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{pmatrix}\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=\begin{pmatrix}\sum_{j=1}^n a_{1j}x_j\\\sum_{j=1}^n a_{2j}x_j\\\vdots\\\sum_{j=1}^n a_{mj}x_j\end{pmatrix}\in\mathbb{R}^m.$$

其中 $m\times n$ 矩阵 $[T]_\beta^\gamma=([T(e_1)]_\gamma\ [T(e_2)]_\gamma\ \cdots\ [T(e_n)]_\gamma)$。

□

Example

设 $T:\mathbb{R}^2\to\mathbb{R}^3$ 是由

$$T(a_1,a_2)=(a_1+3a_2,0,2a_1-4a_2)$$

定义的线性变换。令

$$[T]_\beta^\gamma=\begin{pmatrix}1&3\\0&0\\2&-4\end{pmatrix}.$$

对于向量 $v=(4,5)\in\mathbb{R}^2$，其坐标向量为 $[v]_\beta=\begin{pmatrix}4\\5\end{pmatrix}$。根据矩阵表示公式：

$$[T(v)]_\gamma=[T]_\beta^\gamma\cdot[v]_\beta=\begin{pmatrix}1&3\\0&0\\2&-4\end{pmatrix}\begin{pmatrix}4\\5\end{pmatrix}=\begin{pmatrix}1\cdot4+3\cdot5\\0\cdot4+0\cdot5\\2\cdot4+(-4)\cdot5\end{pmatrix}=\begin{pmatrix}19\\0\\-12\end{pmatrix}.$$

因此，$T(4,5)=(19,0,-12)$。

对于向量 $u=(1,3)\in\mathbb{R}^2$，其坐标向量为 $[u]_\beta=\begin{pmatrix}1\\3\end{pmatrix}$。根据矩阵表示公式：

$$[T(u)]_\gamma=[T]_\beta^\gamma\cdot[u]_\beta=\begin{pmatrix}1&3\\0&0\\2&-4\end{pmatrix}\begin{pmatrix}1\\3\end{pmatrix}=\begin{pmatrix}1\cdot1+3\cdot3\\0\cdot1+0\cdot3\\2\cdot1+(-4)\cdot3\end{pmatrix}=\begin{pmatrix}10\\0\\-10\end{pmatrix}.$$

因此，$T(1,3) = (10, 0, -10)$。

Example

设 $T : P_3(\mathbb{R}) \to P_2(\mathbb{R})$ 是由 $T(f(x)) = f'(x)$ 定义的线性变换。令 β 和 γ 分别是 $P_3(\mathbb{R})$ 和 $P_2(\mathbb{R})$ 的标准有序基

$$[T]_\beta^\gamma = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

对于多项式 $p(x) = 3x + 7 \in P_3(\mathbb{R})$，其坐标向量为 $[p]_\beta = \begin{pmatrix} 7 \\ 3 \\ 0 \\ 0 \end{pmatrix}$。

根据矩阵表示公式：

$$[T(p)]_\gamma = [T]_\beta^\gamma \cdot [p]_\beta = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 7 \\ 3 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \cdot 7 + 1 \cdot 3 + 0 \cdot 0 + 0 \cdot 0 \\ 0 \cdot 7 + 0 \cdot 3 + 2 \cdot 0 + 0 \cdot 0 \\ 0 \cdot 7 + 0 \cdot 3 + 0 \cdot 0 + 3 \cdot 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ 0 \end{pmatrix}.$$

因此，$T(3x + 7) = 3$。

对于多项式 $q(x) = 2x^2 + 3x + 10 \in P_3(\mathbb{R})$，其坐标向量为 $[q]_\beta = \begin{pmatrix} 10 \\ 3 \\ 2 \\ 0 \end{pmatrix}$。根据矩阵表示公式：

$$[T(q)]_\gamma = [T]_\beta^\gamma \cdot [q]_\beta = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 10 \\ 3 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \cdot 10 + 1 \cdot 3 + 0 \cdot 2 + 0 \cdot 0 \\ 0 \cdot 10 + 0 \cdot 3 + 2 \cdot 2 + 0 \cdot 0 \\ 0 \cdot 10 + 0 \cdot 3 + 0 \cdot 2 + 3 \cdot 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \\ 0 \end{pmatrix}.$$

因此，$T(2x^2 + 3x + 10) = 3 + 4x$。

2.4 线性变换的叠加和复合与矩阵的加法与乘法

2.4.1 矩阵的加法 · 线性变换的叠加

设 V 和 W 是有限维向量空间，β 和 γ 分别是它们的有序基，$T, U : V \to W$ 是线性变换。则：

(a) $[T+U]_\beta^\gamma = [T]_\beta^\gamma + [U]_\beta^\gamma$

(b) 对于任意标量 a，$[aT]_\beta^\gamma = a[T]_\beta^\gamma$

证明：设 $\beta = \{v_1, v_2, \ldots, v_n\}$，$\gamma = \{w_1, w_2, \ldots, w_m\}$。存在唯一标量 a_{ij}，b_{ij}（$1 \le i \le m$，$1 \le j \le n$），使得

$$T(v_j) = \sum_{i=1}^m a_{ij} w_i, \quad U(v_j) = \sum_{i=1}^m b_{ij} w_i.$$

进而

$$(T+U)(v_j) = \sum_{i=1}^m (a_{ij} + b_{ij}) w_i,$$

可得

$$([T+U]_\beta^\gamma)_{ij} = a_{ij} + b_{ij} = ([T]_\beta^\gamma + [U]_\beta^\gamma)_{ij},$$

证明了 (a)，(b) 证明类似。

接下来，我们通过一个例子来理解一下矩阵的加法：$T : \mathbb{R}^2 \to \mathbb{R}^3$，$U : \mathbb{R}^2 \to \mathbb{R}^3$ 为线性变换，定义为

$$T(a_1, a_2) = (a_1 + 3a_2, 0, 2a_1 - 4a_2), \quad U(a_1, a_2) = (a_1 - a_2, 2a_1, 3a_1 + 2a_2).$$

β，γ 为 $\mathbb{R}^2$，$\mathbb{R}^3$ 标准有序基。已知

$$[T]_\beta^\gamma = \begin{pmatrix} 1 & 3 \\ 0 & 0 \\ 2 & -4 \end{pmatrix}, \quad [U]_\beta^\gamma = \begin{pmatrix} 1 & -1 \\ 2 & 0 \\ 3 & 2 \end{pmatrix}.$$

计算得

$$(T+U)(a_1, a_2) = (2a_1 + 2a_2, 2a_1, 5a_1 - 2a_2),$$

$$[T+U]_\beta^\gamma = \begin{pmatrix} 2 & 2 \\ 2 & 0 \\ 5 & -2 \end{pmatrix},$$

即 $[T+U]_\beta^\gamma = [T]_\beta^\gamma + [U]_\beta^\gamma$，示例验证定理 2.8。

2.4.2 矩阵的乘法 · 线性变换的复合

有了上面的讨论与前一节的铺垫，我们自然可以引出矩阵的乘法

设 A 是一个 $m \times n$ 矩阵，B 是一个 $n \times p$ 矩阵。我们定义 A 与 B 的乘积，记为 AB，是一个 $m \times p$ 矩阵，满足

$$(AB)_{ij} = \sum_{k=1}^n A_{ik} B_{kj} \quad \text{对于} 1 \le i \le m,\ 1 \le j \le p.$$

我们将此线性变换的角度来进行解释与证明：

设 $T:V\to W$ 和 $U:W\to Z$ 为线性变换，且令 $A=[U]_{\beta}^{\gamma}$，$B=[T]_{\alpha}^{\beta}$，其中 $\alpha=\{v_1,v_2,\ldots,v_n\}$，$\beta=\{w_1,w_2,\ldots,w_m\}$，$\gamma=\{z_1,z_2,\ldots,z_p\}$ 分别是 V、W 和 Z 的有序基。

而由矩阵乘法的几何意义我们知道：A 的列与 B 的行数量必须一致。

我们想要定义两个矩阵的乘积 AB，使得 $AB=[UT]_{\alpha}^{\gamma}$。考虑矩阵 $[UT]_{\alpha}^{\gamma}$。对于 $1\le j\le n$，我们有：

$$
\begin{aligned}
(UT)(v_j)=U(T(v_j))=U\left(\sum_{k=1}^{m}B_{kj}w_k\right)&=\sum_{k=1}^{m}B_{kj}U(w_k)\\
&=\sum_{k=1}^{m}B_{kj}\left(\sum_{i=1}^{p}A_{ik}z_i\right)=\sum_{i=1}^{p}\left(\sum_{k=1}^{m}A_{ik}B_{kj}\right)z_i\\
&=\sum_{i=1}^{p}C_{ij}z_i
\end{aligned}
$$

其中

$$C_{ij}=\sum_{k=1}^{m}A_{ik}B_{kj}$$

所以，综合上面的讨论，我们总结出：V、W、Z 为有限维向量空间，有序基分别为 α、β、γ ，$T:V\to W$，$U:W\to Z$ 为线性变换，则

$$[UT]_{\alpha}^{\gamma}=[U]_{\beta}^{\gamma}[T]_{\alpha}^{\beta}$$

下面简要介绍以下矩阵乘法的 2 条重要性质：

矩阵乘积的转置性质：

若 A 是 $m\times n$ 矩阵，B 是 $n\times p$ 矩阵，则 $(AB)^t=B^tA^t$ ，推导过程为：

$$(AB)_{ij}^t=(AB)_{ji}=\sum_{k=1}^{n}A_{jk}B_{ki},\quad (B^tA^t)_{ij}=\sum_{k=1}^{n}(B^t)_{ik}(A^t)_{kj}=\sum_{k=1}^{n}B_{ki}A_{jk}$$

同阶对角阵的乘法：

设两个 n 阶对角阵 A 和 B,其主对角线元素分别为 $a_{11},a_{22},\ldots,a_{nn}$ 和 $b_{11},b_{22},\ldots,b_{nn}$，

即：$A=\begin{pmatrix} a_{11} & 0 & \cdots & 0\\ 0 & a_{22} & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & a_{nn}\end{pmatrix}$， $B=\begin{pmatrix} b_{11} & 0 & \cdots & 0\\ 0 & b_{22} & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & b_{nn}\end{pmatrix}$

乘法规则乘积仍为对角阵，其主对角线元素为对应位置元素的乘积：

$$AB = \begin{pmatrix} a_{11}b_{11} & 0 & \cdots & 0 \\ 0 & a_{22}b_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn}b_{nn} \end{pmatrix}$$

但是，我们也要注意到矩阵乘法和一般乘法的**特殊性**：

以下是对矩阵乘法反直觉特点的提取及举例：

1. 没有交换律：矩阵乘法一般不满足交换律，即 $AB \neq BA$，当 $AB = BA$ 时，矩阵 A 和 B 可交换。

例子：设 $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ ，$B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ 。$AB = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ ，$BA = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ ，显然 $AB \neq BA$ 。

【Remark】：

同阶对角阵相乘满足交换性, 交换律成立：$AB = BA$,

2. 没有消去律: 若 $AC = BC$ ，不能得出 $A = B$ 。

例子: 设 $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ ，$B = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ ，$C = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ 。$AC = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ ，$BC = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ ，此时 $AC = BC$ ，但 $A \neq B$ 。

3. 相乘得零但可能无零因子: 若 $AB = 0$ ，不能得出 $A = 0$ 或 $B = 0$ 。

例子: 设 $A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ ，$B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ 。$AB = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ ，这里 $AB = 0$ ，但 $A \neq 0$ 且 $B \neq 0$ 。

2.4.3 矩阵乘法的向量理解

左乘 A：A 左乘列向量，即 $A\cdot$ 列向量 = 列向量。

用符号表示为：

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = x \cdot \mathrm{Col}_1 + y \cdot \mathrm{Col}_2 + z \cdot \mathrm{Col}_3$$

(Col_1、Col_2、Col_3 分别表示矩阵的列向量)。

我们做一点点变换：我们只取左边矩阵的其中的一行：

$$\begin{pmatrix} a_{i1} & a_{i2} & a_{i3} \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} = x \cdot \mathrm{Col}_1 + y \cdot \mathrm{Col}_2 + z \cdot \mathrm{Col}_3$$

在这里：$\mathrm{Col}_1 = a_{i1}$ 同样的：$\mathrm{Col}_2 = a_{i2}$，$\mathrm{Col}_3 = a_{i3}$.

同样的，我们有：

$$\begin{pmatrix} a_{i1} & a_{i2} & a_{i3} \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = x_2 \cdot a_{i1} + y_2 \cdot a_{i2} + z_2 \cdot a_{i2}$$

以此类推到每一个 x_j，当右边 j 个 x 排成一列时：就有了右乘：

右乘 A：A 右乘行向量，即行向量 $\cdot A =$ 行向量。

用符号表示为：

$$\begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} ---\mathrm{row}_1--- \\ ---\mathrm{row}_2--- \\ ---\mathrm{row}_3--- \end{pmatrix} = x_1 \cdot \mathrm{row}_1 + x_2 \cdot \mathrm{row}_2 + x_3 \cdot \mathrm{row}_3$$

我们将通过两个例题来加深理解（例子要修改!!!!!!!）：

- **左乘例题**：求矩阵 A，满足

$$A \cdot \begin{bmatrix} 1 & 2 & 1 \\ 3 & 8 & 1 \\ 0 & 4 & 1 \end{bmatrix} = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 2 & -3 \\ 0 & 4 & 1 \end{pmatrix}$$

  分析过程：因为 A 在左边，考虑 $\begin{bmatrix} 1 & 2 & 1 \\ 3 & 8 & 1 \\ 0 & 4 & 1 \end{bmatrix}$ 的行向量 r_1、r_2、r_3，r_1 不变，$r_1 \times (-3) \rightarrow r_2$ 上，r_3 不变，得出

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

  本质上就是排列起来的三个行向量形成矩阵，然后经过 LA 的变换，变成了另外三个排列起来的行向量。

- **右乘例题**：求矩阵 B，满足

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -3 \\ 0 & 4 & 1 \end{bmatrix} \cdot B = \begin{pmatrix} 1 & 2 & 1 \\ 3 & 8 & 1 \\ 0 & 4 & 1 \end{pmatrix}$$

分析过程：因为 B 在右边，考虑 $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -3 \\ 0 & 4 & 1 \end{bmatrix}$ 的列向量 c_1、c_2、c_3。

- 观察目标矩阵的第一列：c_1 不变。
- 观察目标矩阵的第二列：$2c_2 + c_3$（即原矩阵第二列乘以 2 加上第三列）。
- 观察目标矩阵的第三列：$c_1 - c_3$（即原矩阵第一列减去第三列）。

得出

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

综合以上内容，我们不难得到向量层面的理解：

- 左乘矩阵：影响被乘矩阵的行向量，对应行变换。
- 右乘矩阵：影响被乘矩阵的列向量，对应列变换。
- 通过分析行/列向量的线性组合，可以直接构造出对应的变换矩阵。

2.5 可逆与同构

2.5.1 单射与满射

对于 $T: V \to W$：

(a) **单射（Injection）**，或称一对一（One-to-one）：若 $T(x) = T(y) \implies x = y$。

观察到：定义域里面除了 0，没有其它元素会被映射成 0，即 $N(T) = 0$。

我们可以进一步的论证：

Theorem

若 $T(v) = 0$，则 $v = 0$ 与单射等价。

即：T 是单射当且仅当 $\ker(T) = \{0\}$。

证明. **步骤 1：** 证明"若 $T(v) = 0$ 则 $v = 0 \implies T$ 是单射"设 $u, v \in V$ 满足 $T(u) = T(v)$。由线性映射的性质，$T(u - v) = T(u) - T(v) = 0$。

根据题设"若 $T(w) = 0$ 则 $w = 0$"，可得 $u - v = 0$，即 $u = v$。因此，T 是单射。

步骤 2： 证明"T 是单射 $\implies$ 若 $T(v) = 0$ 则 $v = 0$"设 T 是单射，且 $v \in V$ 满足 $T(v) = 0$。

注意到线性映射的基本性质：$T(0_V) = 0_W$（0_V 是 V 的零向量，0_W 是 W 的零向量）。

因此 $T(v) = T(0_V)$。又因为 T 是单射，故 $v = 0_V$，即"若 $T(v) = 0$ 则 $v = 0$"。综上，"若 $T(v) = 0$ 则 $v = 0$"与"T 是单射"互为充要条件，即二者等价。 □

同时，单射还有如下的性质：

Property

单射与线性无关性： 若 T 为单射，则任一线性无关组 $\{x_1, x_2, \ldots, x_n\} \subset V$ 的像 $\{T(x_1), T(x_2), \ldots, T(x_n)\}$ 仍线性无关。

证明. 证明：假设 $\sum a_i T(x_i) = 0$，则 $T\left(\sum a_i x_i\right) = 0$。由 T 单射，$\sum a_i x_i = 0$，故 $a_i = 0$，矛盾。 □

(b) **满射（Surjection）**，或称映上（Onto）：若对 W 中每个向量 w_i，至少存在 $\alpha \in V$ 使 $T(\alpha) = w_i$。

Example

$T(f(x)) = f'(x)$ 是满射非单射；$T(f(x)) = \int_0^x f(t)dt$ 是单射非满射。

2.5.2 逆与可逆

Definition

对于线性变换 $T : V \to W$，若存在另一个线性变换 $U : W \to V$，使得 $TU = I_W$ 且 $UT = I_V$，则称 U 为 T 的逆变换（或 T 的逆，记为 T^{-1}）。若 T 存在逆，则称 T 可逆（Invertible）。

例如：考虑二维向量空间 $\mathbb{R}^2$，定义线性变换 $T : \mathbb{R}^2 \to \mathbb{R}^2$，对于任意向量 $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$，$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x + y \\ x + y \end{pmatrix}$。

令 $U\begin{pmatrix} m \\ n \end{pmatrix} = \begin{pmatrix} m - n \\ -m + 2n \end{pmatrix}$：容易验证：
$TU = I_{\mathbb{R}^2}$ 且 $UT = I_{\mathbb{R}^2}$。
所以 T 可逆，且逆为 U；
但是，我们又要介绍一对常见的误区。

【Remark】:

设 $T: P_2 \to P_3$，$T(f(x)) = \int_0^x f(t)dt$；$U: P_3 \to P_2$，$U(f(x)) = f'(x)$。
计算得 $(TU)(x) = f(x) - f(0) \neq I_{P_3}(f(x))$，$UT(f(x)) = f(x) = I_{P_2}(f(x))$，故 T、U 并非可逆关系。

Corollary & Secondary Conclusion

如果线性变换 T 可逆，那么**逆唯一**。

证明. □

Theorem

对于线性变换 $T: V \to W$，$\left([T]_\beta^\gamma\right)^{-1} = [T^{-1}]_\gamma^\beta$（反向）。

证明. □

接下来，我们将讨论：**逆的必要条件**

设 $T: V \to W$ 是有限维向量空间之间的线性映射，其矩阵表示为 $m \times n$ 矩阵 A。若 T 可逆，则：

1. V 和 W 的维数相等，即 $\dim(V) = \dim(W) = n$；
2. A 必为 $n \times n$ 方阵，且 $\mathrm{rank}(A) = n$（即 A 满秩）。

2.5.3 同构与同构映射

同构（Isomorphic）与同构映射（Isomorphism）是一种相当重要的映射，其背后的思想既有坐标化的思想，又有转换的思想，通过找到统一性与普遍性，连结特殊性。

Definition

若存在可逆的线性变换 $T: V \to W$，则称 V 和 W 同构，其中 T 称为为同构映射。

其具有以下是性质：

(a) 若 V 和 W 同构，则 $\dim(V) = \dim(W)$；反之，若 $\dim(V) = \dim(W)$，则 V 和 W 同构。

(b) 若 V_0 是 V 的子空间，则 $\dim(V_0) = \dim(T(V_0))$。

并且，我们可以证明以下论断：

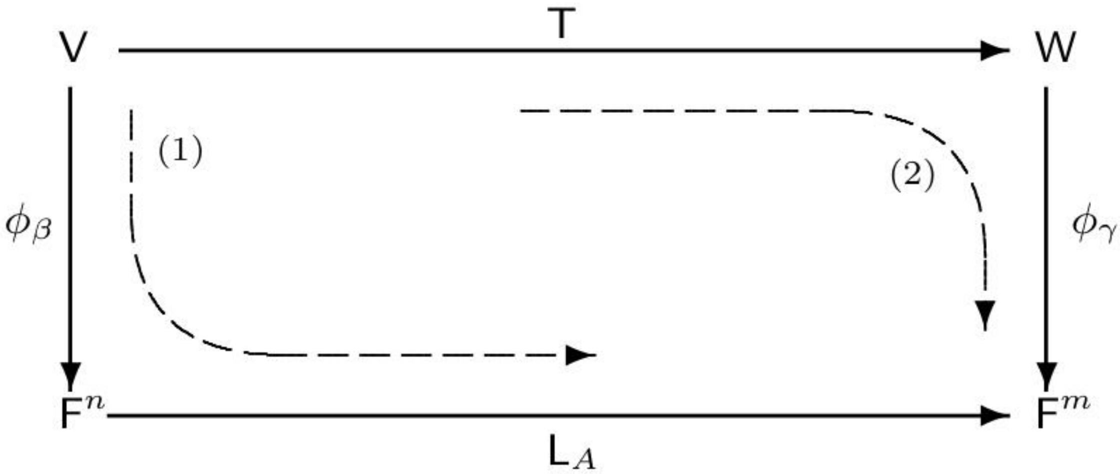

图 2.1: 同构映射概念图

Claim

若 V 和 W 同构（$\dim(V) = \dim(W)$），则以下三条等价：
T 是一对一的；
T 是满射的；
$\dim V = \dim W = \gamma(T)$。

2.5.4 向量空间的标准表示相关内容

设 β 是数域 F 上 n 维向量空间 V 的一个有序基。关于 β 的 V 的表示是函数 $\phi_\beta : V \to F^n$ ，对每个 $x \in V$，定义为 $\phi_\beta(x) = [x]_\beta$ 。

我们现在考察一个具体的例子来理解这个概念：

设 $\beta = \{(1,0),(0,1)\}$ 且 $\gamma = \{(1,2),(3,4)\}$。容易看出 β 和 γ 是 $\mathbb{R}^2$ 的有序基。对于 $x = (1,-2)$，我们有 $\phi_\beta(x) = [x]_\beta = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$ ，$\phi_\gamma(x) = [x]_\gamma = \begin{pmatrix} -5 \\ 2 \end{pmatrix}$ 。

让我们再进一步，设 V 和 W 分别是维数为 n 和 m 的向量空间，且设 $T : V \to W$ 是一个线性变换。定义 $A = [T]_\beta^\gamma$ ，其中 β 和 γ 分别是 V 和 W 的任意有序基。我们现在能够使用 ϕ_β 和 ϕ_γ 来研究线性变换 T 和 $L_A : F^n \to F^m$ 之间的关系。

首先考虑图。注意到有两个将 V 映射到 F^m 的线性变换复合：

1. 用 ϕ_β 将 V 映射到 F^n ，然后用 L_A 进行变换；这产生复合 $L_A\phi_\beta$ 。

2. 用 T 将 V 映射到 W ，然后用 ϕ_γ 进行变换，得到复合 $\phi_\gamma T$ 。

这两个复合由图中的虚线箭头表示。显然，我们可以得出 $L_A\phi_\beta = \phi_\gamma T$ ，即该图是“可交换的”。直观地说，这种关系表明，在通过 ϕ_β 和 ϕ_γ 分别将 V 和 W 与 F^n 和 F^m

等同之后，我们可以将 T 与 L_A “等同”。**此图允许我们将抽象向量空间上的运算转换为可以量化研究的，结构清晰的数域 F^n 和 F^m 上的矩阵运算。**

因此矩阵与线性变换的联系可见下图：

$$
\begin{array}{ccc}
v & \xrightarrow{T} & w \\
\beta \downarrow & & m \downarrow \gamma \\
[v]_\beta & \xrightarrow{[T]_\beta^\gamma} & [T(v)]_\gamma
\end{array}
$$

在这张图中，ϕ_β 直接作用在向量 v 上，将其变成可用矩阵表示的：$[x]_\beta$；ϕ_γ 也是同理

我们再来考察一个例子，加深对上面讨论的理解：

Example

对于线性变换 $T : P_3(R) \to P_2(R)$（$T(f(x)) = f'(x)$）。设 β 和 γ 分别是 $P_3(R)$ 和 $P_2(R)$ 的标准有序基，且设 $\phi_\beta : P_3(R) \to \mathbb{R}^4$ 和 $\phi_\gamma : P_2(R) \to \mathbb{R}^3$ 分别是 $P_3(R)$ 和 $P_2(R)$ 对应的标准表示。若 $A = [T]_\beta^\gamma$.

则

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

。

考虑多项式 $p(x) = 2 + x - 3x^2 + 5x^3$ 。我们证明 $L_A\phi_\beta(p(x)) = \phi_\gamma T(p(x))$ 。现在

$$
L_A\phi_\beta(p(x)) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ -3 \\ 5 \end{pmatrix} = \begin{pmatrix} 1 \\ -6 \\ 15 \end{pmatrix}
$$

。

但由于 $T(p(x)) = p'(x) = 1 - 6x + 15x^2$ ，我们有

$$
\phi_\gamma T(p(x)) = \begin{pmatrix} 1 \\ -6 \\ 15 \end{pmatrix}
$$

。

所以 $L_A\phi_\beta(p(x)) = \phi_\gamma T(p(x))$ 。

2.5.5 同构的三大基本定理

Theorem

第一同构定理（first isomorphism）：给定域 $\mathbb{F}$ 上的线性映射 $T:V\to W$，存在 $\mathbb{F}$-线性的自然投影（natural projection）$p:V\to V/\ker(T)$（$p(v)=\bar{v}:=v+\ker(T)$，即将 v 映射到其所在的陪集），使得 $T=\bar{T}\circ p$，其中 $\bar{T}:V/\ker(T)\to \mathrm{Im}(T)$ 是线性同构，定义为 $\bar{T}(\bar{v})=T(v)$。

证明. 1. 映射 $p:V\to V/\ker(T)$，$v\mapsto\bar{v}$ 是线性且满射的。

首先验证线性性：对任意 $v,w\in V$，$p(v+w)=(v+w)+\ker(T)=(v+\ker(T))+(w+\ker(T))=p(v)+p(w)$；对任意 $a\in\mathbb{F}$，$p(av)=av+\ker(T)=a(v+\ker(T))=a\cdot p(v)$。

再验证满射性：对商空间 $V/\ker(T)$ 中任意陪集 $\bar{v}=v+\ker(T)$，存在 $v\in V$ 使得 $p(v)=\bar{v}$，故 p 是满射。

2. 定义 $\bar{T}:V/\ker(T)\to\mathrm{Im}(T)$ 为 $\bar{T}(\bar{v})=T(v)$。下证其良定义：若 $\bar{v}=\bar{w}$，则 $v-w\in\ker(T)$，由核的定义知 $T(v-w)=0$，即 $T(v)=T(w)$，故 $\bar{T}$ 良定义。

3. $\bar{T}$ 是线性的：对任意 $a,b\in\mathbb{F}$，$\bar{v},\bar{w}\in V/\ker(T)$，$\bar{T}(a\bar{v}+b\bar{w})=\bar{T}(\overline{av+bw})=T(av+bw)=aT(v)+bT(w)=a\bar{T}(\bar{v})+b\bar{T}(\bar{w})$。

4. $\bar{T}$ 是单射的：若 $\bar{T}(\bar{v})=0$，则 $T(v)=0$，故 $v\in\ker(T)$，从而 $\bar{v}=\ker(T)$（商空间零元），即 $\bar{T}$ 的核仅含零元，故单射。

5. $\bar{T}$ 是满射的：对任意 $w\in\mathrm{Im}(T)$，存在 $v\in V$ 使 $T(v)=w$，则 $\bar{T}(\bar{v})=T(v)=w$，故满射。

因此，$\bar{T}$ 是同构且 $T=\bar{T}\circ p$。

□

【Remark】：

因此，由该定理可知：线性映射的“像空间”与“核的商空间”代数结构完全一致——所以，维度定理的逻辑自洽。

Theorem

第二同构定理（second isomorphism）：设 V 和 W 是域 $\mathbb{F}$ 上的向量子空间，则自然 $\mathbb{F}$-线性映射 $T:V\to(V+W)/W$ 可通过投影 $p:V\to V/(V\cap W)$ 分解，即 $T=\bar{T}\circ p$，其中 $\bar{T}:V/(V\cap W)\xrightarrow{\cong}(V+W)/W$，并且 $\bar{T}$ 是同构。

即：对两个 $\mathbb{F}$- 子空间 V,W，有同构 $V/(V\cap W)\cong(V+W)/W$。

证明. 1. **定义映射 T 并验证线性性与满射性**：定义 $T:V\to(V+W)/W$ 为 $T(v)=\bar{v}=v+W$（因为 $v\in V\subset V+W$，所以 $v\in V+W$，满足商空间元素的定义）。

- **线性性**：对任意 $v_1, v_2 \in V$，$a, b \in \mathbb{F}$，

$$T(av_1 + bv_2) = (av_1 + bv_2) + W = a(v_1 + W) + b(v_2 + W) = aT(v_1) + bT(v_2),$$

故 T 是线性映射。

- **满射性**：对任意陪集 $\bar{x} = x + W \in (V + W)/W$，由 $V + W$ 的定义，$x = v + w$（其中 $v \in V$，$w \in W$）。此时 $\bar{x} = (v + w) + W = v + W = T(v)$，故存在 $v \in V$ 使得 $T(v) = \bar{x}$，即 T 是满射。

2. **计算 T 的核** $\ker(T)$：核的定义为 $\ker(T) = \{v \in V \mid T(v) = 0_{(V+W)/W}\}$，其中 $0_{(V+W)/W} = W$（商空间的零元是 W 本身）。因此，$T(v) = W \iff v + W = W \iff v \in W$。又因 $v \in V$，故 $\ker(T) = \{v \in V \mid v \in W\} = V \cap W$。

3. **应用第一同构定理**：第一同构定理指出：若 $T : V \to W'$ 是满射线性映射，则 $V/\ker(T) \cong \mathrm{Im}(T)$。此处 T 是满射，且 $\mathrm{Im}(T) = (V + W)/W$，$\ker(T) = V \cap W$，因此：

$$V/(V \cap W) \cong (V + W)/W.$$

记此同构为 $\bar{T} : V/(V \cap W) \to (V + W)/W$，则自然有 $T = \bar{T} \circ p$（其中 $p : V \to V/(V \cap W)$ 是投影映射 $v \mapsto v + (V \cap W)$）。

□

Theorem

第三同构定理 (third isomorphism)：设 $U \subset W \subset V$ 是域 $\mathbb{F}$ 上的向量子空间，则自然 $\mathbb{F}$-线性映射 $T : V \to (V/U)/(W/U)$ 可通过投影 $p : V \to V/W$ 分解，即 $T = \bar{T} \circ p$，其中 $\bar{T} : V/W \xrightarrow{\cong} (V/U)/(W/U)$，并且 $\bar{T}$ 是同构。

即：对嵌套子空间 $U \subset W \subset V$，有同构 $(V/U)/(W/U) \cong V/W$。

证明. 1. **定义辅助映射 ϕ 并验证线性性与满射性**：定义 $\phi : V/U \to (V/U)/(W/U)$ 为 $\phi(\bar{v}_U) = \overline{\bar{v}_U}_{W/U}$（即 $\bar{v}_U = v + U$ 在商空间 $(V/U)/(W/U)$ 中的陪集）。

- **线性性**：对任意 $\bar{v}_U, \bar{w}_U \in V/U$，$a, b \in \mathbb{F}$，

$$\phi(a\bar{v}_U + b\bar{w}_U) = \phi(\overline{av + bw}_U) = \overline{\overline{av + bw}_U}_{W/U} = a\overline{\bar{v}_U}_{W/U} + b\overline{\bar{w}_U}_{W/U} = a\phi(\bar{v}_U) + b\phi(\bar{w}_U),$$

故 ϕ 是线性映射。

- **满射性**：对任意陪集 $\bar{\bar{x}}_{W/U} \in (V/U)/(W/U)$，其中 $\bar{x}_U = x + U \in V/U$，显然 $\phi(\bar{x}_U) = \bar{\bar{x}}_{W/U}$，故 ϕ 是满射。

2. **计算 ϕ 的核** $\ker(\phi)$：核的定义为 $\ker(\phi) = \{\bar{v}_U \in V/U \mid \phi(\bar{v}_U) = 0_{(V/U)/(W/U)}\}$，其中 $0_{(V/U)/(W/U)} = W/U$（商空间的零元是 W/U 本身）。因此，$\phi(\bar{v}_U) = W/U \iff$

$\bar{v}_U + W/U = W/U \iff \bar{v}_U \in W/U$。即 $\ker(\phi) = \{\bar{v}_U \in V/U \mid \bar{v}_U \in W/U\} = W/U$。

3. **应用第一同构定理**：第一同构定理指出：若 $\phi : V/U \to (V/U)/(W/U)$ 是满射线性映射，则 $(V/U)/\ker(\phi) \cong \mathrm{Im}(\phi)$。此处 ϕ 是满射，且 $\mathrm{Im}(\phi) = (V/U)/(W/U)$，$\ker(\phi) = W/U$，因此：

$$(V/U)/(W/U) \cong (V/U)/(W/U) \implies (V/U)/(W/U) \cong V/W$$

（通过自然对应，商空间的"商"可转化为原空间的商 V/W）。记此同构为 $\bar{T} : V/W \to (V/U)/(W/U)$，则自然有 $T = \bar{T} \circ p$（其中 $p : V \to V/W$ 是投影映射 $v \mapsto v + W$）。

□

因此，三大定理可以由下面的逻辑表现清楚：

- 第一定理：用线性映射的"核"分解定义域，建立商空间到像空间的同构，有线性映射的"像空间"与"核的商空间"代数结构完全一致。
- 第二定理：联系"子空间的交"与"子空间的和的商空间"的结构。
- 第三定理：联系"商空间的商空间"与"大商空间"的结构。

2.5.6 同构空间与对偶空间

Definition

同态空间（Homomorphism space）：设 V 和 W 是两个 $\mathbb{F}$-向量空间，从 V 到 W 的同态空间（简称 Hom space）定义为：

$$\mathrm{Hom}_{\mathbb{F}}(V, W) := \{\text{所有 } \mathbb{F}\text{-线性映射 } T : V \to W\}.$$

若 $V = W$，也记作 $\mathrm{End}_{\mathbb{F}}(V)$。

我们知道：从 $\mathbb{F}^n$ 到 $\mathbb{F}^m$ 的线性映射 T，可唯一由一个 $m \times n$ 矩阵 A 表示（满足 $T(\mathbf{x}) = A\mathbf{x}$，$\mathbf{x} \in \mathbb{F}^n$）。所以我们可以考察以下同态空间及其性质：

Theorem

所有从 $\mathbb{F}^n$ 到 $\mathbb{F}^m$ 的线性映射 T 构成的空间 $\mathrm{Hom}_{\mathbb{F}}(\mathbb{F}^n, \mathbb{F}^m)$，与"所有 $m \times n$ 矩阵"构成的空间 $M_{m\times n}(\mathbb{F})$ **同构**。即：$\mathrm{Hom}_{\mathbb{F}}(\mathbb{F}^n, \mathbb{F}^m) \cong M_{m\times n}(\mathbb{F})$；

证明. □

Corollary & Secondary Conclusion

当 V 是有限维 $\mathbb{F}$- 向量空间时，$\dim_{\mathbb{F}} \mathrm{Hom}_{\mathbb{F}}(V,W) = \dim_{\mathbb{F}}(V) \cdot \dim_{\mathbb{F}}(W)$。

证明. 线性映射的“自由度”由基的像唯一确定：

设 V 的基为 $\{v_1, v_2, \dots, v_n\}$ $(n = \dim V)$, W 的基为 $\{w_1, w_2, \dots, w_m\}$ $(m = \dim W)$。

对任意线性映射 $T: V \to W$，它的行为由 $T(v_1), T(v_2), \dots, T(v_n)$ 唯一决定（因为 V 中任意向量 $v = a_1 v_1 + \cdots + a_n v_n$，都有 $T(v) = a_1 T(v_1) + \cdots + a_n T(v_n)$）。每个 $T(v_i)$ 可表示为 W 基的线性组合：$T(v_i) = b_{i1} w_1 + b_{i2} w_2 + \cdots + b_{im} w_m$ $(b_{ij} \in \mathbb{F})$。

这意味着，确定 T 需要选择 n 个“W 中的向量”，每个向量有 m 个独立的标量系数。

因此，所有线性映射的“自由度总数”是 $n \times m$（即 $\dim V \times \dim W$），对应矩阵空间 $M_{m\times n}(\mathbb{F})$ 的维数（$m \times n$ 个独立标量）。 □

Corollary & Secondary Conclusion

$\mathrm{Hom}_{\mathbb{F}}\left(\bigoplus_{i\in I} \mathbb{F}, \mathbb{F}^m\right) \cong \prod_{i\in I} \mathbb{F}^m$。

证明. 从 $\bigoplus_{i\in I} \mathbb{F}$ 到 $\mathbb{F}^m$ 的线性映射 T，由它在“直和的基元素”（每个分量的单位元，如第 i 个分量为 1、其余为 0 的元组）上的作用唯一确定。

对每个 $i \in I$，T 在第 i 个基元素上的像属于 $\mathbb{F}^m$；而所有这样的线性映射，等价于“为每个 $i \in I$ 选取一个 $\mathbb{F}^m$ 中的元素”——这正是直积 $\prod_{i\in I} \mathbb{F}^m$ 的定义（直积中元素是“对每个 $i \in I$ 取一个 $\mathbb{F}^m$ 元素”，无“有限非零”限制）。 □

同样的，我们可以得到：

Corollary & Secondary Conclusion

$\mathrm{Hom}_{\mathbb{F}}\left(\bigoplus_{i\in I} \mathbb{F}, \bigoplus_{j\in J} \mathbb{F}\right) \cong \prod_{i\in I}\left(\bigoplus_{j\in J} \mathbb{F}\right)$。

接下来我们讲介绍：**对偶空间**。

Definition

设 V 是 $\mathbb{F}$- 向量空间，**对偶空间** V^* 定义为“从 V 到数域 $\mathbb{F}$ 的所有线性映射构成的空间”，即：$V^* := \mathrm{Hom}_{\mathbb{F}}(V, \mathbb{F})$

Lemma

若 $\dim_{\mathbb{F}}(V) < \infty$（$V$ 是有限维向量空间），则：$\dim_{\mathbb{F}}(V^*) = \dim_{\mathbb{F}}(V)$
即：$V^* \simeq V$

对偶基（Dual Basis）：为了显式构造 V^* 的基，并建立 V^* 与 V 之间的同构，引入对偶基：

Definition

设 $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ 是 V 的一个基，定义 V^* 中的线性函数组 $\mathcal{B}^* = \{v_1^*, v_2^*, \ldots, v_n^*\}$，其中每个 $v_i^* : V \to \mathbb{F}$ 满足：$v_i^*(v_j) = \delta_{ij}$ （δ_{ij}为克罗内克函数，$i = j$时为 1，否则为 0）此时，$\mathcal{B}^*$ 是 V^* 的一个基（称为 $\mathcal{B}$ 的对偶基）。

通过线性延拓，可诱导出 V 与 V^* 之间的同构（例如 $v_i \mapsto v_i^* \mapsto v_i$）。

Example

$\mathbb{R}^n$ 是实数域上的 n 维向量空间（元素为 n 维实向量 $x = (x_1, x_2, \ldots, x_n)^T$）。其对偶空间 $(\mathbb{R}^n)^*$ 是所有从 $\mathbb{R}^n$ 到 $\mathbb{R}$ 的线性泛函的集合。

原空间 $\mathbb{R}^n$ 的标准基 $\{e_1, e_2, \ldots, e_n\}$，就像 n 维空间里的坐标轴：

e_1 是 “x_1 轴” 的单位向量，e_2 是 “x_2 轴” 的单位向量，⋯，e_n 是 “x_n 轴” 的单位向量。

任意向量 $x = (x_1, x_2, \ldots, x_n)^T$，都能拆成“各坐标轴分量的组合”：$x = x_1e_1 + x_2e_2 + \cdots + x_ne_n$。

而对偶基可以理解为 “刻度读取器”。**对偶空间里的元素是线性泛函**（把向量变成数的 “函数”）。对偶基 $\{e_1^*, e_2^*, \ldots, e_n^*\}$，就像 “读取各坐标轴刻度的工具”：$e_i^*$ 是专门 “读第 i 个坐标轴刻度” 的工具。当你用 e_i^* 去 “读” 向量 x 时，得到的就是 x 在第 i 个坐标轴上的分量 x_i，即 $e_i^*(x) = x_i$。

对偶基对 “精准读取坐标轴” 的验证：$e_i^*(e_j) = \delta_{ij}$，其实是说：用 “读第 i 个坐标轴” 的工具 e_i^*，去读 “第 j 个坐标轴的单位向量 e_j”，结果是什么？

如果 $i = j$：读的是 “第 i 个坐标轴的单位向量”，那它在第 i 个坐标轴上的刻度当然是 1，所以 $e_i^*(e_i) = 1$。

如果 $i \neq j$：读的是 “第 j 个坐标轴的单位向量”，它在第 i 个坐标轴上的刻度是 0，所以 $e_i^*(e_j) = 0$。

这就像 “读 x 轴刻度的工具，读 y 轴的单位向量时，得到的 x 轴刻度是 0；读 x 轴的单位向量时，得到的 x 轴刻度是 1”。

现在任取一个线性泛函 $f \in (\mathbb{R}^n)^*$，根据线性性，对 x 有：$f(x) = f\left(x_1e_1 + x_2e_2 + \cdots + x_ne_n\right)$ 由线性映射的 “可加性” 和 “齐次性”，展开得：$f(x) = x_1 \cdot f(e_1) + x_2 \cdot f(e_2) + \cdots + x_n \cdot f(e_n)$

那么对偶基如何表示这个线性泛函？对偶基 $\{e_1^*, e_2^*, \ldots, e_n^*\}$ 的定义是：对任意 $x = (x_1, \ldots, x_n)^T$，$e_i^*(x) = x_i$（即 e_i^* 是 “提取第 i 个分量” 的线性泛函）。令 $a_i = f(e_i)$（a_i 是实数，因为 $f(e_i) \in \mathbb{R}$），则上式可改写为：$f(x) = a_1 \cdot e_1^*(x) + a_2 \cdot e_2^*(x) + \cdots + a_n \cdot e_n^*(x)$

由于上式对所有 $x \in \mathbb{R}^n$ 成立，因此作为线性泛函（映射），f 可表示为对偶基的线性组合：$f = a_1e_1^* + a_2e_2^* + \cdots + a_ne_n^*$

Example

【有限基】：$\mathbb{R}^2$：取 $n=2$，对偶基为 e_1^*, e_2^*，其中：$e_1^*(x_1,x_2)=x_1$（提取第一个分量），$e_2^*(x_1,x_2)=x_2$（提取第二个分量）。
现在有一个线性泛函 $f:\mathbb{R}^2\to\mathbb{R}$，满足：$f(1,0)=3$（即 $f(e_1)=3$），$f(0,1)=5$（即 $f(e_2)=5$）。根据上述推导，f 可表示为对偶基的线性组合：$f=3e_1^*+5e_2^*$.

【Remark】：

f 和 f（x）不一样！

验证：对任意 $x=(x_1,x_2)$，$f(x)=3e_1^*(x)+5e_2^*(x)=3x_1+5x_2$ 而由线性性，$f(x_1,x_2)=x_1f(1,0)+x_2f(0,1)=3x_1+5x_2$，完全一致。

Example

【无限基】：$\mathbb{F}[x]$ 是数域 $\mathbb{F}$ 上的所有多项式构成的向量空间（比如 $\mathbb{F}=\mathbb{R}$ 时，元素是 x^2+3x+5、$7x^3-2$ 等）。它的基为 $\mathcal{B}_{\prime}=\{1,x,x^2,x^3,\ldots\}$

注意：这个基是**无限基**（因为多项式的次数可以任意高，基中包含无限多个单项式）。

对偶基的构造：“提取系数”的线性泛函对偶空间 $V^*=\mathrm{Hom}_{\mathbb{F}}(V,\mathbb{F})$ 是“所有从 $\mathbb{F}[x]$ 到 $\mathbb{F}$ 的线性映射（线性泛函）”的集合。我们为原基 $\mathcal{B}_{\prime}$ 中的每个单项式 x^i，构造一个线性泛函 $\delta_{x^i}\in V^*$，满足：$\delta_{x^i}(x^j)=\delta_{ij}$　直观来说，每个 δ_{x^i} 是“提取多项式某一项系数”的工具：δ_1（对应 $x^0=1$）：作用在多项式上，提取常数项。

-例如 $\delta_1(5x^3+2x+7)=7$，因为 $\delta_1(x^j)=\delta_{0j}$（$j=0$ 时为 1，$j\geq 1$ 时为 0）。δ_x（对应 x^1）：作用在多项式上，提取一次项系数。

-例如 $\delta_x(5x^3+2x+7)=2$，因为 $\delta_x(x^j)=\delta_{1j}$（$j=1$ 时为 1，其他为 0）。δ_{x^2}：提取二次项系数，δ_{x^3} 提取三次项系数，依此类推。

这样构造出的 $\mathcal{B}^*=\{\delta_1,\delta_x,\delta_{x^2},\delta_{x^3},\ldots\}$ 就是原基 $\mathcal{B}_{\prime}$ 的对偶基。

Example

例如：组合 $3\delta_1+2\delta_x$：作用在多项式上，结果为“$3\times$ 常数项　$+2\times$ 一次项系数”。组合 $c_0\delta_1+c_1\delta_x+\cdots+c_n\delta_{x^n}$：仅涉及前 $n+1$ 个系数的线性组合。

无限维的特殊性：对偶基张成的空间是对偶空间的真子集

对偶空间 V^* 包含所有从 $\mathbb{F}[x]$ 到 $\mathbb{F}$ 的线性泛函，其中存在大量无法用“有限个对偶基元素组合”表示的泛函。

例子：“无限系数和”的线性泛函假设 $\mathbb{F}=\mathbb{R}$，定义线性泛函 $g:\mathbb{F}[x]\to\mathbb{R}$：对任意多项式 $p(x)=a_0+a_1x+a_2x^2+\cdots+a_nx^n$，$g(p(x))=a_0+a_1+a_2+\cdots+a_n$（即“所有系数的和”）。g 是线性的：因为 $g(kp+q)=kg(p)+g(q)$（线性映射的可加性、齐次性）。

但 g 不在 $\mathrm{Span}_{\mathbb{F}}(\mathcal{B}^*)$ 中：因为 $\mathrm{Span}_{\mathbb{F}}(\mathcal{B}^*)$ 中的泛函只能涉及有限个系数的组合，而 g 需要"同时提取无限多个系数并求和"，无法用 $\mathcal{B}^*$ 中元素的有限线性组合表示。

因此，$\mathrm{Span}_{\mathbb{F}}(\mathcal{B}^*)$ 是 V^* 的真子集（即 $\mathrm{Span}_{\mathbb{F}}(\mathcal{B}^*) \subsetneq V^*$）。

有限维与无限维的核心差异: 在有限维空间（如 $\mathbb{R}^n$）中，对偶基的张成空间就是整个对偶空间（因为基是有限的，线性组合能覆盖所有线性泛函）。

但在无限维空间（如 $\mathbb{F}[x]$）中，对偶空间的"规模"远大于"原基对偶基的张成空间"——无限维空间的对偶空间无法被"原基的对偶基的有限组合"完全覆盖。

2.5.7 零化子及其性质

问题：哪些线性泛函 f 会'忽略' W 中的向量（$\forall w \in W, f(w) = 0$)?

其中**忽略**可以理解为：子空间 W 里的向量，在 f 这个"测量工具"下，测不出任何"数值"(结果都是 0)

Example

假设 V 是三维空间，W 是 xy 平面（所有 z 坐标为 0 的向量构成的子空间）。
如果有个泛函 $f(v)$ 是 v 的 z 坐标，那对 W 里的任意向量 w（z 坐标都是 0），都有 $f(w) = 0$。这说明：f 专门"测 z 方向"，而 W 里的向量"没 z 方向的分量"，所以被 f"忽略"了。

原空间 V 有"包含、直和、商空间"这些结构，对偶空间 V^* 的结构会和它们"反向对应"，相当于从"镜子里看原空间的结构"，用泛函的角度重新解释原空间。

于是我们可以导入零化子的概念及重要性质：

零化子的定义：

Definition

设 V 是域 $\mathbb{F}$ 上的向量空间，对子集 $S \subset V$，定义 S 在对偶空间 V^* 中的**零化子**(annihilator) 为：
$\mathrm{Ann}(S) := \{f \in V^* \mid f(s) = 0, \ \forall s \in S\}$ 即对偶空间 V^* 中所有在 S 上恒为 0 的**线性泛函的集合**。

直和分解下的零化子性质:

Theorem

假设:$V = W \oplus W'$ (W, W' 是 V 的子空间，且直和分解成立)，则存在同构映射 $i_{W'} : W^* \oplus W'^* \xrightarrow{\sim} V^*$，其中对任意 $w^* \in W^*$、$w'^* \in W'^*$，以及 $v \in W$、$v' \in W'$，映射规则为：$w^* + w'^* \mapsto (v + v' \mapsto w^*(v) + w'^*(v'))$.

即：该线性同构将 W^* 与 W'^* 的直和 $W^* \oplus W'^*$ 映射到 V^*，将 W^* 和 W'^* 的泛函“拼接”为 V^* 中在 W, W' 上分别作用的泛函。

它通过“分解 - 拼接”的方式，反映了原空间直和结构在对偶空间上的对应关系，是研究对偶空间与原空间子空间关系的核心工具

【Remark】:

这里下标用 W'，是为了强调“以 W' 为参照”的拼接方式——尤其是后续讨论“零化子”时，W' 的角色更直接.

若将下标换为 W（定义 $i_W : W^* \oplus W'^* \to V^*$），纯数学上是可行的，但需重新定义映射规则（比如调整泛函作用的“分解方向”）。

但为了与后续性质（如 $\mathrm{Ann}(W)$ 与 W'^* 的关联）符号一致，通常选择 W' 作为下标——它能更自然地体现“W' 是 W 的补空间，且其对偶空间直接生成 W 的零化子”这一联系。

Example

$V = \mathbb{R}^2$ 的直和分解取 $V = \mathbb{R}^2$（实向量空间）；

并令：$W = \mathrm{span}\{(1,0)\}$（x 轴，一维子空间）；$W' = \mathrm{span}\{(0,1)\}$（y 轴，一维子空间）。

取 $w^* \in W^*$ 满足 $w^*((1,0)) = 2$（即 $w^*((x,0)) = 2x$），再取 $w'^* \in W'^*$ 满足 $w'^*((0,1)) = 3$（即 $w'^*((0,y)) = 3y$）。

W^* 是 W 的对偶空间，即“所有从 W 到数域 $\mathbb{F}$ 的线性泛函（线性映射）”构成的向量空间；

$w^* \in W^*$ 表示：w^* 是定义在 W 上的线性泛函，它满足“线性性”——对任意 $v, u \in W$，$a, b \in \mathbb{F}$，有 $w^*(av + bu) = aw^*(v) + bw^*(u)$，且作用是将 W 中的向量映射到数域 $\mathbb{F}$ 中。

W 的基为 $\{(1,0)\}$，因此 W 上的线性泛函 $w^* \in W^*$ 由“作用于基向量 $(1,0)$ 的值”唯一确定。

若 $w^*((1,0)) = a$ $(a \in \mathbb{R})$，则对任意 $(x,0) \in W$，有 $w^*((x,0)) = ax$。

同理，W' 的基为 $\{(0,1)\}$，若 $w'^*((0,1)) = b$ $(b \in \mathbb{R})$，则对任意 $(0,y) \in W'$，有 $w'^*((0,y)) = by$。

证明. **正向验证：**

根据同构映射 $i_{W'}$，拼接得到 V 上的线性泛函 $f = i_{W'}(w^* + w'^*)$。现在验证 f 对 V 中任意向量 (x,y) 的作用：将 (x,y) 分解为 $(x,0) + (0,y)$ $((x,0) \in W$，$(0,y) \in W')$，则：$f((x,y)) = w^*((x,0)) + w'^*((0,y)) = 2x + 3y$

显然，$f : \mathbb{R}^2 \to \mathbb{R}$ 是线性泛函（满足线性性：$f(kv) = kf(v)$，$f(v_1 + v_2) = f(v_1) + f(v_2)$），因此 $f \in V^*$

反向验证（泛函的“分解”）：

再取 V^* 中任意线性泛函，比如 $f((x,y)) = 5x - 4y$。我们可以将其“分解”为 W 和 W' 上的泛函：

对 W 中的 $(x,0)$，定义 $w^*((x,0)) = 5x$（即 $w^*((1,0)) = 5$，属于 W^*）；对 W' 中的 $(0,y)$，定义 $w'^*((0,y)) = -4y$（即 $w'^*((0,1)) = -4$，属于 W'^*）。

此时 $f = i_{W'}(w^* + w'^*)$，并且我们可以归纳出说明每个 V 上的泛函都能通过这种“拼接”得到，即映射 $i_{W'}$ 是**满射**；

同时，若两个不同的“拼接泛函”作用于 V 上的结果相同，则它们的 w^* 和 w'^* 必须分别相同（由直和分解的唯一性），因此 $i_{W'}$ 是**单射**。综上，$i_{W'}$ 是**双射且保持线性结构，故为同构映射**。

□

基于上述直和与同构，零化子具有以下关键性质：

Property

零化子与对偶空间的像相等：$\text{Ann}(W) = i_{W'}(W'^*)$

特别地，商空间 V/W 的对偶空间与 $\text{Ann}(W)$ 同构：$(V/W)^* \simeq \text{Ann}(W)$

证明. **第一步：**$i_{W'}(W'^*) \subseteq \textbf{Ann}(W)$ 任取 $f \in i_{W'}(W'^*)$，则存在 $g \in W'^*$ 使得 $f = i_{W'}(g)$。对任意 $w \in W$，将 w 分解为 $w = w + 0$（其中 $0 \in W'$），则：$f(w) = i_{W'}(g)(w) = g(0) = 0$（线性泛函在零向量上的作用必为 0）。

因此 f 在 W 上恒为 0，即 $f \in \text{Ann}(W)$。故 $i_{W'}(W'^*) \subseteq \text{Ann}(W)$。

第二步：$\textbf{Ann}(W) \subseteq i_{W'}(W'^*)$ 任取 $f \in \text{Ann}(W)$（即 $f \in V^*$ 且 $f(w) = 0, \forall w \in W$）。需要构造 $g \in W'^*$，使得 $f = i_{W'}(g)$：

对 $w' \in W'$，定义 $g(w') = f(w')$。由于 f 是线性泛函，g 也满足线性性（W' 中线性组合的作用保持线性），因此 $g \in W'^*$。

验证 $f = i_{W'}(g)$：对任意 $v = w + w' \in V$（$w \in W, w' \in W'$），有：$i_{W'}(g)(v) = g(w') = f(w')$，而 $f(v) = f(w + w') = f(w) + f(w') = 0 + f(w') = f(w')$（因 $f \in \text{Ann}(W)$，故 $f(w) = 0$）。

因此 $f(v) = i_{W'}(g)(v)$ 对所有 $v \in V$ 成立，即 $f = i_{W'}(g)$。故 $f \in i_{W'}(W'^*)$，即 $\text{Ann}(W) \subseteq i_{W'}(W'^*)$。

□

包含关系的“反向性”若子空间满足 $W_1 \subset W_2 \subset V$，则它们的零化子满足反向包含：$\text{Ann}(W_1) \supset \text{Ann}(W_2)$（子空间越大，能“零化”它的对偶泛函越少）

证明. □

零化子的“自反性”(有限维情形) 若 W 是有限维子空间 ($\dim_{\mathbb{F}}(W) < \infty$), 则“零化子的零化子”等于 W 本身：$\mathrm{Ann}(\mathrm{Ann}(W)) = W$ 更精确地，$\mathrm{Ann}_V(\mathrm{Ann}(W)) = W$ (强调在 V 中取两次零化子)。

证明. □

直和分解的推广 $W \oplus (\mathrm{Ann}_V(i_{W'}(W^*))) = V$ (W 与由零化子定义的子空间也构成 V 的直和)

证明. □

这些性质体现了“零化子”在对偶空间与原空间的子空间之间的对应关系，以及对包含、直和等结构的刻画作用。

此外，我们还有两个推论：

Corollary & Secondary Conclusion

$\mathrm{Ann}(W_1 + W_2) = \mathrm{Ann}(W_1) \cap \mathrm{Ann}(W_2)$

证明. 第一步：$\mathrm{Ann}(W_1+W_2) \subseteq \mathrm{Ann}(W_1)\cap\mathrm{Ann}(W_2)$ 任取 $f \in \mathrm{Ann}(W_1+W_2)$，根据零化子定义，对所有 $v \in W_1+W_2$，有 $f(v) = 0$。由于 $W_1 \subseteq W_1+W_2$ (任取 $w_1 \in W_1$，可表为 $w_1 + 0 \in W_1 + W_2$)，因此对所有 $w_1 \in W_1$，$f(w_1) = 0$，故 $f \in \mathrm{Ann}(W_1)$。

同理，$W_2 \subseteq W_1 + W_2$，对所有 $w_2 \in W_2$，$f(w_2) = 0$，故 $f \in \mathrm{Ann}(W_2)$。

因此 $f \in \mathrm{Ann}(W_1) \cap \mathrm{Ann}(W_2)$，即 $\mathrm{Ann}(W_1 + W_2) \subseteq \mathrm{Ann}(W_1) \cap \mathrm{Ann}(W_2)$。

第二步：$\mathrm{Ann}(W_1)\cap\mathrm{Ann}(W_2) \subseteq \mathrm{Ann}(W_1+W_2)$ 任取 $f \in \mathrm{Ann}(W_1)\cap\mathrm{Ann}(W_2)$，则 $f \in \mathrm{Ann}(W_1)$ 且 $f \in \mathrm{Ann}(W_2)$。

对任意 $v \in W_1 + W_2$，由和空间定义，$v = w_1 + w_2$ (其中 $w_1 \in W_1$，$w_2 \in W_2$)。由线性泛函的线性性，$f(v) = f(w_1 + w_2) = f(w_1) + f(w_2)$。

因 $f \in \mathrm{Ann}(W_1)$，故 $f(w_1) = 0$；因 $f \in \mathrm{Ann}(W_2)$，故 $f(w_2) = 0$。

因此 $f(v) = 0 + 0 = 0$。对所有 $v \in W_1 + W_2$，$f(v) = 0$，故 $f \in \mathrm{Ann}(W_1 + W_2)$，即 $\mathrm{Ann}(W_1) \cap \mathrm{Ann}(W_2) \subseteq \mathrm{Ann}(W_1 + W_2)$。

结合两步，$\mathrm{Ann}(W_1 + W_2) = \mathrm{Ann}(W_1) \cap \mathrm{Ann}(W_2)$，命题成立。

□

【Remark】:

命题 $\mathrm{Ann}(W_1 \cap W_2) = \mathrm{Ann}(W_1) + \mathrm{Ann}(W_2)$ 不成立！

反例构造 (无限维情形) 设 $V = \mathbb{R}^\infty$ (基为 $\{e_n\}_{n=1}^\infty$), 取: $W_1 = \mathrm{span}\{e_{2k-1} \mid k \geq 1\}$ (奇数下标基向量生成的子空间)，$W_2 = \mathrm{span}\{e_{2k} \mid k \geq 1\}$ (偶数下标基向量生成的子空间)。

则 $W_1 \cap W_2 = \{0\}$, 故 $\mathrm{Ann}(W_1 \cap W_2) = V^*$ (零子空间的零化子是整个对偶空间)。

分析 $\mathrm{Ann}(W_1)+\mathrm{Ann}(W_2)$

$\mathrm{Ann}(W_1)+\mathrm{Ann}(W_2)$ 中的元素形如 $f+g$（$f\in\mathrm{Ann}(W_1)$，$g\in\mathrm{Ann}(W_2)$），即：$f+g=\sum_{k=1}^{\infty}a_k e_{2k}^*+\sum_{k=1}^{\infty}b_k e_{2k-1}^*$

注意到 V^* 中的泛函可由任意实序列（$c_1,c_2,\dots$）表示（因为 V 中向量是有限支撑的，线性泛函对 e_n 的作用 $c_n=f(e_n)$ 可任意选取）。但 $\mathrm{Ann}(W_1)+\mathrm{Ann}(W_2)$ 中的泛函具有"奇数位置系数来自 g、偶数位置系数来自 f"的限制，无法覆盖所有实序列。

2.5.8 伴随

有限维内积空间中，我们大致会讨论两种伴随，这两种伴随是同一概念的"代数版"与"内积版"，通过里斯同构自洽。

现在我们先来介绍代数版：从"向量映射"到"泛函映射"

设 $T:V\to W$ 是域 $\mathbb{F}$ 上的线性映射（向量空间间的线性映射）：W^* 是 W 的对偶空间：由所有 $W\to\mathbb{F}$ 的线性泛函（即"向量 → 数"的线性映射）构成的向量空间；V^* 是 V 的对偶空间，同理。

接着，我们给出伴随映射 $T^*:W^*\to V^*$ 的定义：

Definition

对任意 $f\in W^*$（W 上的泛函）和 $v\in V$，规定 $T^*(f)(v):=f(T(v))$..

逻辑：T 将 V 中向量 v 送到 W 中向量 $T(v)$；然后 f 将 W 中向量 $T(v)$ 送到数 $f(T(v))$。因此，"先作用 T，再作用 f"的复合，定义了一个 $V\to\mathbb{F}$ 的映射 $T^*(f)$。由于 T 和 f 都是线性的，复合后 $T^*(f)$ 仍为线性泛函，故 $T^*(f)\in V^*$。

即 T^* 把 W 上的泛函 f，转化为 V 上的泛函 $f\circ T$（先作用 T，再作用 f）。

Theorem

线性映射 $T:V\to W$ 与其伴随 $T^*:W^*\to V^*$ 的单射（injective）、满射（surjective）性质**反向等价**，即：

如果 T 是满射 $\iff$ T^* 是单射；

如果 T 是单射 $\iff$ T^* 是满射。

证明. **1. T 满射 $\implies$ T^* 单射**

目标：证明 T^* 的核为 $\{0\}$（单射的定义：核仅含零元）。

步骤：取 $f\in\ker(T^*)$（即 $T^*(f)=0$）。由伴随定义，对所有 $v\in V$，$T^*(f)(v)=0$，即 $f(T(v))=0$。

因 T 满射，对任意 $w\in W$，存在 $v\in V$ 使得 $T(v)=w$。因此，对任意 $w\in W$，$f(w)=f(T(v))=0$，即 f 是 W^* 中的零泛函（$f=0$）。

故 $\ker(T^*)=\{0\}$，即 T^* 是单射。

2.T^* 单射 $\implies$ T 满射

目标：证明 $\operatorname{Im}(T) = W$（满射的定义：像等于目标空间）。

步骤（反证法）：假设 T 不满射，则 $\operatorname{Im}(T) \subsetneq W$。由零化子性质（子空间的零化子非空）：

> *Claim*
>
> 根据零化子的基本性质：若 S 是 W 的真子空间，则 $\operatorname{Ann}(S)$ 中存在 * 非零泛函（子空间的零化子不会只有零泛函）

存在非零泛函 $f \in W^*$，使得 f 在 $\operatorname{Im}(T)$ 上恒为 0（即 $f(T(v)) = 0$ 对所有 $v \in V$ 成立）。

由伴随定义，$T^*(f)(v) = f(T(v)) = 0$ 对所有 $v \in V$ 成立，故 $T^*(f) = 0$，即 $f \in \ker(T^*)$。

但 $f \neq 0$，与 T^* 单射（$\ker(T^*) = \{0\}$）矛盾。因此假设不成立，T 必满射。

3. T 单射 $\implies$ T^* 满射

目标：证明对任意 $g \in V^*$，存在 $f \in W^*$ 使得 $T^*(f) = g$（满射的定义：像覆盖目标空间）。

步骤：因 T 单射，$T : V \to \operatorname{Im}(T)$ 是同构（单射且像为 $\operatorname{Im}(T)$），故存在逆映射 $T^{-1} : \operatorname{Im}(T) \to V$。对任意 $g \in V^*$，定义 $f : W \to \mathbb{F}$：若 $w \in \operatorname{Im}(T)$，则 $f(w) = g(T^{-1}(w))$；若 $w \in W \setminus \operatorname{Im}(T)$，利用“子空间补空间”（$W = \operatorname{Im}(T) \oplus U$），对 U 中元素任意赋值（不影响线性性）。易证 $f \in W^*$（线性泛函），且对任意 $v \in V$，$T^*(f)(v) = f(T(v)) = g(T^{-1}(T(v))) = g(v)$，故 $T^*(f) = g$。因此 T^* 是满射。

4. T^* 满射 $\implies$ T 单射

目标：证明 $\ker(T) = \{0\}$（单射的定义：核仅含零元）。

步骤（反证法）：假设 T 不单射，则 $\ker(T) \neq \{0\}$，故 $\dim \operatorname{Im}(T) < \dim V$。

由对偶空间维度性质（$\dim X^* = \dim X$），$\dim W^* = \dim W$，$\dim V^* = \dim V$。因 T^* 满射，$\dim \operatorname{Im}(T^*) = \dim V^* = \dim V$。但 $\operatorname{Im}(T^*) \subseteq V^*$，且 $\dim \operatorname{Im}(T^*) \leq \dim W^* = \dim W$，结合 $\dim \operatorname{Im}(T) < \dim V$，得 $\dim W \geq \dim V > \dim \operatorname{Im}(T)$。

由零化子性质，存在非零泛函 $f \in W^*$ 使 f 在 $\operatorname{Im}(T)$ 上恒为 0（即 $T^*(f) = 0$），与 T^* 满射（$\operatorname{Im}(T^*) = V^*$，无“非零核”）矛盾。因此假设不成立，T 必单射。 □

> *Claim*
>
> $\ker(T^*) \simeq (W/\operatorname{Im}(T))^*$

解释：商空间 $W/\operatorname{Im}(T)$ 的元素是“W 中模 $\operatorname{Im}(T)$ 的等价类”。

其对偶空间 $(W/\operatorname{Im}(T))^*$ 中的泛函，必须“在 $\operatorname{Im}(T)$ 上恒为 0”（商空间泛函的“良定义性”要求）。

而 $\ker(T^*)$ 中的泛函 $f \in W^*$ 满足 $T^*(f) = 0$，即 $f(T(v)) = 0$ 对所有 $v \in V$ 成立，等价于 f 在 $\mathrm{Im}(T)$ 上恒为 0。

因此，$\ker(T^*)$ 与 $(W/\mathrm{Im}(T))^*$ 作为"在 $\mathrm{Im}(T)$ 上零化的 W 泛函"集合，是同构的。这也是 $W/\mathrm{Im}(T)$ 被称为 T 的余核 (cokernel) 的原因 (对偶于 T 的核)。

接下来我们讨论伴随映射的矩阵表示和转置的来源:

要找 T^* 在对偶基下的矩阵，与 T 在原基下的矩阵的关系，步骤如下：

1. 原映射 T 的矩阵: 设 V 的基 $\mathcal{A} = \{v_1, \ldots, v_n\}$，$W$ 的基 $\mathcal{B} = \{w_1, \ldots, w_m\}$。

T 在这组基下的矩阵为 $T_{\mathcal{B},\mathcal{A}} = (a_{ij})$，即 $T(v_j) = \sum_{i=1}^m a_{ij} w_i$ (矩阵第 j 列是 $T(v_j)$ 在 $\mathcal{B}$ 下的坐标)。

2. 设伴随映射 T^* 的矩阵形式对偶基为 $\mathcal{A}^* = \{v_1^*, \ldots, v_n^*\}$（$V^*$ 的基），$\mathcal{B}^* = \{w_1^*, \ldots, w_m^*\}$（$W^*$ 的基）。

要确定 $T^* : W^* \to V^*$ 在基 $\mathcal{B}^*$（定义域基）和 $\mathcal{A}^*$（值域基）下的矩阵 $T^*_{\mathcal{A}^*,\mathcal{B}^*}$，设：$T^*(w_k^*) = \sum_{l=1}^n b_{lk} v_l^*$ 需确定系数 b_{lk}。

用对偶基的性质求系数 b_{lk} 对 $v_j \in \mathcal{A}$，计算 $T^*(w_k^*)(v_j)$：由伴随定义：$T^*(w_k^*)(v_j) = (w_k^* \circ T)(v_j) = w_k^*(T(v_j))$。

分析 $T^*(w_k^*)(v_j)$ 的结构 $w_k^* \in W^*$：w_k^* 是 W 的对偶基 $\mathcal{B}^* = \{w_1^*, \ldots, w_m^*\}$ 中的元素（因为 $\mathcal{B} = \{w_1, \ldots, w_m\}$ 是 W 的基)。$T^*(w_k^*) \in V^*$：由于 $T^* : W^* \to V^*$，将 W^* 中的 w_k^* 输入 T^*，输出是 V^* 中的一个线性泛函。$T^*(w_k^*)(v_j) \in \mathbb{F}$：将 V^* 中的泛函 $T^*(w_k^*)$ 作用在 V 的基向量 v_j 上，结果是域 $\mathbb{F}$ 中的一个标量。

回到主题，

代入 $T(v_j) = \sum_{i=1}^m a_{ij} w_i$，得到：$w_k^* \left(\sum_{i=1}^m a_{ij} w_i\right) = \sum_{i=1}^m a_{ij} \cdot w_k^*(w_i)$

由对偶基性质 $w_k^*(w_i) = \delta_{ki}$（仅 $i = k$ 时为 1)，故上式化简为 a_{kj}。

另一方面，若 $T^*(w_k^*) = \sum_{l=1}^n b_{lk} v_l^*$，则作用在 v_j 上：$T^*(w_k^*)(v_j) = \sum_{l=1}^n b_{lk} v_l^*(v_j) = \sum_{l=1}^n b_{lk} \delta_{lj} = b_{jk}$

比较系数，得到转置关系由 $T^*(w_k^*)(v_j) = a_{kj}$ 且 $T^*(w_k^*)(v_j) = b_{jk}$，得 $b_{jk} = a_{kj}$。

这说明：T^* 在对偶基下的矩阵 (b_{jk})，恰好是 T 在原基下的矩阵 (a_{ij}) 的转置。即：$T^*_{\mathcal{A}^*,\mathcal{B}^*} = {}^tT_{\mathcal{B},\mathcal{A}}$

我们可以得到如下推论：

Corollary & Secondary Conclusion

${}^t(AB) = {}^tB\,{}^tA$（矩阵转置的"乘积法则"）

要证明这个"反转律"，需结合线性映射的复合与对偶映射的复合，再利用"对偶映射的矩阵是原矩阵的转置"。

证明. 步骤 1：线性映射的复合与矩阵乘法的关系设：U, V, W 是有限维向量空间；线性映射 $S : U \to V$，在基 $\mathcal{C}$（U 的基）和 $\mathcal{A}$（V 的基）下的矩阵为 $S_{\mathcal{A},\mathcal{C}}$；线性映射 $T : V \to W$，在基 $\mathcal{A}$（V 的基）和 $\mathcal{B}$（W 的基）下的矩阵为 $T_{\mathcal{B},\mathcal{A}}$。线性映射的复合

$T \circ S : U \to W$（先作用 S，再作用 T），在基 $\mathcal{C}$（U 的基）和 $\mathcal{B}$（W 的基）下的矩阵是矩阵乘法：$(T \circ S)_{\mathcal{B},\mathcal{C}} = T_{\mathcal{B},\mathcal{A}} \cdot S_{\mathcal{A},\mathcal{C}}$

步骤 2：对偶映射的复合与转置矩阵的关系对偶映射是反变的（即 "复合的对偶 = 对偶的反向复合"）：对线性映射 S, T，有 $(T \circ S)^* = S^* \circ T^*$ 其中：$(T \circ S)^* : W^* \to U^*$ 是 $T \circ S$ 的对偶映射；$S^* : V^* \to U^*$ 是 S 的对偶映射，$T^* : W^* \to V^*$ 是 T 的对偶映射，故 $S^* \circ T^* : W^* \to U^*$。再结合对偶映射的矩阵是原矩阵的转置：$(T \circ S)^*$ 在对偶基 $\mathcal{C}^*$（U^* 的基）和 $\mathcal{B}^*$（W^* 的基）下的矩阵为 ${}^t(T_{\mathcal{B},\mathcal{A}} \cdot S_{\mathcal{A},\mathcal{C}})$（因为 $T \circ S$ 的矩阵是 $T_{\mathcal{B},\mathcal{A}} \cdot S_{\mathcal{A},\mathcal{C}}$，对偶映射的矩阵是其转置）；$S^*$ 在对偶基 $\mathcal{C}^*$（U^* 的基）和 $\mathcal{A}^*$（V^* 的基）下的矩阵为 ${}^tS_{\mathcal{A},\mathcal{C}}$；$T^*$ 在对偶基 $\mathcal{A}^*$（V^* 的基）和 $\mathcal{B}^*$（W^* 的基）下的矩阵为 ${}^tT_{\mathcal{B},\mathcal{A}}$。

步骤 3：复合对偶映射的矩阵 = 对偶映射矩阵的乘积因为 $(T \circ S)^* = S^* \circ T^*$，所以它们的矩阵也相等：${}^t(T_{\mathcal{B},\mathcal{A}} \cdot S_{\mathcal{A},\mathcal{C}}) = {}^tS_{\mathcal{A},\mathcal{C}} \cdot {}^tT_{\mathcal{B},\mathcal{A}}$ 若将线性映射对应的矩阵一般化为 "任意矩阵"（令 $A = S_{\mathcal{A},\mathcal{C}}$，$B = T_{\mathcal{B},\mathcal{A}}$），则上式变为：${}^t(BA) = {}^tA\,{}^tB$ 对 "任意性" 稍作调整（交换 A, B 的角色，或重新命名映射），即可得到更常用的形式：对任意两个矩阵 A（尺寸 $n \times p$）和 B（尺寸 $p \times m$），有 ${}^t(AB) = {}^tB\,{}^tA$ 这就是矩阵转置的乘积法则（反转律）。

$\square$

参考文献

[1] Friedberg. Fruedberg linear algebra.

[2] 谢启鸿，姚慕生，吴泉水. 高等代数学. In 高等代数学（复旦大学出版社第四版），pages 373–429, 2024.

[3] Michael Artin. 代数（algebra）. In 代数（机械工业出版社翻译，第 *2* 版）.

[4] 莊重（國立交通大學）. La freidberg. In *Friedberg* 线性代數课程.

[5] 上海交通大学数学组. 线性代数解题方法与技巧. In 线性代数（第三版）.

[6] 倪維明（國立臺灣大學）. Mat2042 linear algebra. In 线性代數 *2042*, 2025.

[7] 彭勇寧（臺灣成功大學）. Linear algebra. In 线性代數.

[8] 罗才华（香港中文大学（深圳））. Mat3040 advanced linear algebra. In 高等线性代數讲义, 2025.

[9] 同济大学. 工程数学 · 线性代数.