Chapter 2

线性代数基础

Matrix Theory

November 27, 2014

黄正华, 数学与统计学院, 武汉大学

2.1

Contents

2.2

1 线性空间

1.1 线性空间的定义及例子

概念的源起

- 向量在加法运算、数乘运算所满足的运算规律, 例如
  - 加法交换律: $\boldsymbol{x}+\boldsymbol{y}=\boldsymbol{y}+\boldsymbol{x}$,
  - 加法结合律: $(\boldsymbol{x}+\boldsymbol{y})+\boldsymbol{z}=\boldsymbol{x}+(\boldsymbol{y}+\boldsymbol{z})$,
  - 数乘分配律: $(\lambda+\mu)\boldsymbol{x}=\lambda\boldsymbol{x}+\mu\boldsymbol{x}$;

  矩阵、多项式、连续函数等, 也都满足完全相同的运算规律.
- 研究不同对象在线性运算方面的所表现的共性, 导致线性空间的公理化定义.
- 方法: 代数与几何的结合.

2.3

4 个组成部分

线性空间 (亦称向量空间), 有 4 个组成部分: 两个集合 V 和 F, 两个运算——一个称为向量加法, 一个称为数量乘法.

- V: 一些被称为向量的对象的集合. 比如 n 维向量, 矩阵.
- F: 一个数域 ——实数域 $\mathbb{R}$ 或者复数域 $\mathbb{C}$.
- 向量加法 (记为 $\boldsymbol{x}+\boldsymbol{y}$): 集合 V 中两个元素之间的一种运算.
  运算要满足封闭性: $\boldsymbol{x}+\boldsymbol{y}\in V$, $\forall\boldsymbol{x},\ \boldsymbol{y}\in V$.
- 数量乘法 (记为 $\alpha\boldsymbol{x}$): 集合 F 和 V 中元素之间的一种运算.
  运算要满足封闭性: $\alpha\boldsymbol{x}\in V$, $\forall\alpha\in F$, $\forall\boldsymbol{x}\in V$.

2.4

线性空间的定义

Definition 1 (线性空间)**.** 设 V 是一个非空集合, F 是一个数域.

- 在集合 V 的元素之间定义了一种代数运算, 叫做加法; 这就是说给出了一个法则, 对于 V 中任意两个元素 $\boldsymbol{x}$ 与 $\boldsymbol{y}$, 在 V 中都有唯一的一个元素 $\boldsymbol{z}$ 与它们对应, 称为 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 的和, 记为 $\boldsymbol{z}=\boldsymbol{x}+\boldsymbol{y}$.
- 在数域 F 与集合 V 的元素之间还定义了一种运算, 叫做数量乘法; 这就是说, 对于数域 F 中任一个数 λ 与 V 中任一个元素 $\boldsymbol{x}$, 在 V 中都有唯一的一个元素 $\boldsymbol{y}$ 与它们对应, 称为 λ 与 $\boldsymbol{x}$ 的数量乘积, 记为 $\boldsymbol{y}=\lambda\boldsymbol{x}$.

如果加法与数量乘法满足下述 8 条规则, 那么 V 称为数域 F 上的线性空间.

加法满足下面四条规则:

(1) $\boldsymbol{x}+\boldsymbol{y}=\boldsymbol{y}+\boldsymbol{x}$;

(2) $(\boldsymbol{x}+\boldsymbol{y})+\boldsymbol{z}=\boldsymbol{x}+(\boldsymbol{y}+\boldsymbol{z})$;

(3) 在 V 中有一个元素 $\boldsymbol{\theta}$, $\forall\boldsymbol{x}\in V$, 都有

$$\boldsymbol{x}+\boldsymbol{\theta}=\boldsymbol{x}$$

(具有这个性质的元素 $\boldsymbol{\theta}$ 称为 V 的零元素, 记为 $\mathbf{0}$);

(4) $\forall \boldsymbol{x} \in V$, $\exists \boldsymbol{y} \in V$, 使得

$$\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{0}$$

($\boldsymbol{y}$ 称为 $\boldsymbol{x}$ 的负元素, 记为 $-\boldsymbol{x}$).

数量乘法满足下面两条规则:

(5) $1\boldsymbol{x} = \boldsymbol{x}$;

(6) $\lambda(\mu\boldsymbol{x}) = (\lambda\mu)\boldsymbol{x}$;

数量乘法与加法满足下面两条规则:

(7) $(\lambda + \mu)\boldsymbol{x} = \lambda\boldsymbol{x} + \mu\boldsymbol{x}$;

(8) $\lambda(\boldsymbol{x} + \boldsymbol{y}) = \lambda\boldsymbol{x} + \lambda\boldsymbol{y}$;

在以上规则中, λ, μ 等表示数域 F 中任意数; $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{z}$ 等表示集合 V 中任意元素.
☞ 借助几何语言, 把线性空间中的元素称为向量, 线性空间又可称为向量空间. F 中的数称为数或标量. V 称为线性空间的基集. 注意: 第 (7) 条 $(\lambda + \mu)\boldsymbol{x} = \lambda\boldsymbol{x} + \mu\boldsymbol{x}$, 左侧 $+$ 号是普通数的加法, 而右侧 $+$ 号是向量的加法. 这两个加法的含义是不同的. 2.5

线性空间的实例

Example 2 (n 维向量空间).

- $\mathbb{R}^n$: n 元实坐标向量空间. 运算: (1) 向量的加法, (2) 数与向量的数量乘法.
  $V = \left\{\boldsymbol{x} \mid \boldsymbol{x} = (\xi_1, \xi_2, \cdots, \xi_n)^{\mathrm{T}},\ \xi_i \in \mathbb{R}, i = 1, 2, \cdots, n\right\}$, $F = \mathbb{R}$.
- $\mathbb{C}^n$: n 元复坐标向量空间. 运算: (1) 向量的加法, (2) 数与向量的数量乘法.
  $V = \left\{\boldsymbol{x} \mid \boldsymbol{x} = (\xi_1, \xi_2, \cdots, \xi_n)^{\mathrm{T}},\ \xi_i \in \mathbb{C}, i = 1, 2, \cdots, n\right\}$, $F = \mathbb{C}$.

注意

- 若 $V = \left\{\boldsymbol{x} \mid \boldsymbol{x} = (\xi_1, \xi_2, \cdots, \xi_n)^{\mathrm{T}},\ \xi_i \in \mathbb{R}, i = 1, 2, \cdots, n\right\}$, 而 $F = \mathbb{C}$, 则 V 不是线性空间. 因为此时数量乘法不满足封闭性.
- $V = \left\{\boldsymbol{x} \mid \boldsymbol{x} = (\xi_1, \xi_2, \cdots, \xi_n)^{\mathrm{T}},\ \xi_i \in \mathbb{C}, i = 1, 2, \cdots, n\right\}$, $F = \mathbb{R}$, 此时 V 任是一个线性空间. 记为 $\mathbb{C}^n_{\mathbb{R}}$. 2.6

Example 3. 元素属于数域 F 的 $m \times n$ 矩阵, 按 (1) 矩阵的加法, (2) 数与矩阵的数量乘法, 构成数域 F 上的一个线性空间, 用 $F^{m \times n}$ 表示.

- $\mathbb{R}^{m \times n}$: $m \times n$ 阶实矩阵空间.
- $\mathbb{C}^{m \times n}$: $m \times n$ 阶复矩阵空间. 2.7

所有系数在数域 F 中的一元多项式的全体, 称为数域 F 上的一元多项式环, 记为 $F[x]$. F 称为 $F[x]$ 的系数域.

Example 4.

- 数域 F 上一元多项式环 $F[x]$, 按 (1) 通常的多项式加法, (2) 数与多项式的乘法, 构成一个数域 F 上的线性空间.
- 如果只考虑其中次数不超过 n 的多项式, 再添上零多项式, 也构成数域 F 上的一个线性空间, 用 $F[x]_n$ 表示.

1. $\mathbb{R}[x]_n$: 实数域 $\mathbb{R}$ 上的次数不超过 n 的多项式空间,

$$\mathbb{R}[x]_n = \{p(x) \mid p(x) = \alpha_n x^n + \alpha_{n-1}x^{n-1} + \cdots + \alpha_1 x + \alpha_0,\ \alpha_i \in \mathbb{R}\}.$$

2. $\mathbb{C}[x]_n$: 复数域 $\mathbb{C}$ 上的次数不超过 n 的多项式空间,

$$\mathbb{C}[x]_n = \{p(x) \mid p(x) = \alpha_n x^n + \alpha_{n-1}x^{n-1} + \cdots + \alpha_1 x + \alpha_0,\ \alpha_i \in \mathbb{C}\}.$$

2.8

Example 5. 全体实函数, 按 (1) 函数加法, (2) 数与函数的数量乘法, 构成一个实数域上的线性空间.

Example 6. 记 $C[a,b]$ 为区间 $[a,b]$ 上所有连续函数的集合. 则 $C[a,b]$ 按 (1) 函数加法, (2) 数与函数的数量乘法, 构成一个实数域上的线性空间.

Example 7. 数域 F 按照本身的加法与乘法, 即构成一个自身上的线性空间.

2.9

Example 8 (重要概念). 设 $\boldsymbol{A} \in F^{m\times n}$, 齐次线性方程组 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ 的解集合

$$V = \{\boldsymbol{x} \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}\},$$

又取 $F = \mathbb{C}$, 并规定 V 中的向量加法和数乘运算与 $\mathbb{C}^m$ 中相应的运算相同, 则 V 是 $\mathbb{C}$ 上的线性空间, 并称 V 为齐次线性方程组 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ 的解空间; 也称为矩阵 $\boldsymbol{A}$ 的零空间(nullspace) 或核空间(kernel), 记作 $N(\boldsymbol{A})$ 或 $\ker(\boldsymbol{A})$.

2.10

Example 9 (重要概念). 给定矩阵 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$, 设

$$V = \{\boldsymbol{y} \in \mathbb{C}^m \mid \boldsymbol{y} = \boldsymbol{A}\boldsymbol{x},\ \boldsymbol{x} \in \mathbb{C}^n\},$$

又取 $F = \mathbb{C}$, 并规定 V 中的向量加法和数乘运算与 $\mathbb{C}^m$ 中相应的运算相同, 则 V 是 $\mathbb{C}$ 上的线性空间, 并称 V 为矩阵 $\boldsymbol{A}$ 的列空间, 或值域, 或值空间, 并记为 $R(\boldsymbol{A})$.

2.11

注意

- 对同样的基集 V 和数域 F, 对它们规定不同的加法和数乘, 若构成线性空间, 则得到不同的线性空间.
- 线性空间的零向量, 也与该空间中向量加法和数量乘法的规定方法有关. 比如下面例题中的零元素是常数 1.

Exercise 10 (P.35 习题二 3, 经典例题)**.** 设 V 是正实数集, $\mathbb{R}$ 为实数域. 定义加法 $\oplus$ 和数乘 $\odot$:

$$\alpha \oplus \beta = \alpha\beta, \qquad \text{(即 } \alpha \text{ 与 } \beta \text{ 的积)}$$
$$k \odot \alpha = \alpha^k, \qquad \text{(即 } \alpha \text{ 的 } k \text{ 次幂)}$$

其中 $\alpha, \beta \in V$, $k \in \mathbb{R}$. 问: V 对于加法 $\oplus$ 和数乘 $\odot$ 是否构成 $\mathbb{R}$ 上的线性空间?

2.12

解: 加法 $\oplus$ 封闭: $\forall \alpha, \beta \in V = \mathbb{R}^+$, $\alpha \oplus \beta = \alpha\beta \in V$. 且满足:

(1) $\alpha \oplus \beta = \alpha\beta = \beta\alpha = \beta \oplus \alpha$;

(2) $(\alpha \oplus \beta) \oplus \gamma = (\alpha\beta)\gamma = \alpha(\beta\gamma) = \alpha \oplus (\beta \oplus \gamma)$;

(3) 零元为常数 1: $1 \oplus \alpha = \alpha$;

(4) $\forall \alpha \in V$ 的负元为 $\frac{1}{\alpha}$: $\frac{1}{\alpha} \oplus \alpha = 1$;

数乘 $\odot$ 封闭: $\forall \alpha \in V = \mathbb{R}^+$, $k \in \mathbb{R}$, $k \odot \alpha = \alpha^k \in V = \mathbb{R}^+$. 且满足:

(5) $1 \odot \alpha = \alpha^1 = \alpha$;

(6) $(k_1k_2) \odot \alpha = \alpha^{k_1k_2} = (\alpha^{k_2})^{k_1} = k_1 \odot (\alpha^{k_2}) = k_1 \odot (k_2 \odot \alpha)$;

(7) $(k_1 + k_2) \odot \alpha = \alpha^{k_1+k_2} = \alpha^{k_1}\alpha^{k_2} = \alpha^{k_1} \oplus \alpha^{k_2} = k_1 \odot \alpha \oplus k_2 \odot \alpha$;

(8) $k \odot (\alpha \oplus \beta) = k \odot (\alpha\beta) = (\alpha\beta)^k = \alpha^k\beta^k = \alpha^k \oplus \beta^k = (k \odot \alpha) \oplus (k \odot \beta)$.

所以 V 对于加法 $\oplus$ 和数乘 $\odot$ 构成 $\mathbb{R}$ 上的线性空间. □

☞ 这是一个经典的例子. 从中我们可以看到, 线性空间中加法、数量乘法的含义已经拓展了. 此例中的“加法”其实是普通的乘法, 数量乘法是幂运算. 另外, 还要注意到第 7 条中等式两边的加法是不同的. 2.13

我们再次强调, 零向量不一定是形如 0, $(0, 0, \cdots, 0)^{\mathrm{T}}$, $\left[\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right]$ 之类的对象, 而必须是符合零向量的定义的那个元素. 2.14

Example 11. 集合 $C = \{(x_1, x_2) \mid x_1, x_2 \in \mathbb{C}\}$, 向量加法:

$$(x_1, x_2) + (y_1, y_2) = (x_1 + y_1 + 1, x_2 + y_2 + 1),$$

数量乘法:

$$\alpha(x_1, x_2) = (\alpha x_1 + \alpha - 1, \alpha x_2 + \alpha - 1).$$

容易证明该集合是数域 $\mathbb{C}$ 上的线性空间.

问题: 零向量是什么? (x_1, x_2) 的负向量呢?

☞ 零向量: $\mathbf{0} = (-1, -1)$. 因为

$$\boldsymbol{u} + \mathbf{0} = (x_1, x_2) + (-1, -1) = \big(x_1 + (-1) + 1, x_2 + (-1) + 1\big) = (x_1, x_2) = \boldsymbol{u}.$$

负向量: $-(x_1, x_2) = (-x_1 - 2, -x_2 - 2)$. 事实上

$$(x_1, x_2) + (-x_1 - 2, -x_2 - 2) = \big(x_1 + (-x_1 - 2) + 1, -x_2 + (x_2 - 2) + 1\big) = (-1, -1) = \mathbf{0}.$$

2.15

Example 12. 设集合 $Z = \{\boldsymbol{z}\}$. 向量加法: $\boldsymbol{z} + \boldsymbol{z} = \boldsymbol{z}$. 数量乘法: $\lambda\boldsymbol{z} = \boldsymbol{z}$, $\lambda \in F$. 可以证明: 集合 Z 是数域 F 上的线性空间.

这个线性空间的零元素是什么? 显然, 只能是 $\boldsymbol{z}$.

但有趣的是, 我们并没有指明 $\boldsymbol{z}$ 的具体内容. 事实上, $\boldsymbol{z}$ 可以是任何一个向量、矩阵、函数等等. 2.16

小结

- 线性空间是 n 维向量空间的抽象和推广.
- 为了几何直观, 线性空间也叫做向量空间.
- 但这里的向量不一定是有序数组, 而是广义的向量, 例如函数、矩阵等.
- 加法、数量乘法、零元的含义都得到了拓展.

2.17

线性空间的简单性质

1. 零元素是唯一的.
2. 负元素是唯一的.
3. $0\boldsymbol{x} = \mathbf{0}$; $k\,\mathbf{0} = \mathbf{0}$; $(-1)\boldsymbol{x} = -\boldsymbol{x}$.
4. 如果 $k\boldsymbol{x} = \mathbf{0}$, 那么 $k = 0$ 或者 $\boldsymbol{x} = \mathbf{0}$. 2.18

发展历程

- The idea of defining a vector space by using a set of abstract axioms was contained in a general theory published in 1844 by Hermann Grassmann (1808–1887), a theologian and philosopher from Stettin, Poland, who was a self-taught mathematician.
- The Italian mathematician Giuseppe Peano (1858–1932) was one of the few people who noticed Grassmann's work, and in 1888 Peano published a condensed interpretation of it.
- The current definition is derived from the 1918 work of the German mathematician Hermann Weyl (1885–1955). Weyl's success with the idea was due in part to the fact that he thought of vector spaces in terms of geometry, whereas Grassmann and Peano treated them as abstract algebraic structures. 2.19

课后作业

P. 34　习题 3(完成例 12 的验证工作), 7 2.20

1.2　子空间的概念

平面 $\mathbb{R}^2$ 是 $\mathbb{R}^3$ 的子集, 而 $\mathbb{R}^2$ 也构成线性空间, 称 $\mathbb{R}^2$ 是 $\mathbb{R}^3$ 的线性子空间.

过原点的一条直线或一个平面都是 $\mathbb{R}^3$ 的子集, 而且它们关于向量加法和数乘分别构成一个一维和二维的线性空间.

考虑一般的情形: 线性空间的子集, 关于原线性空间的加法和数乘, 是否构成一个线性空间? 2.21

子空间

Definition 13. 数域 F 上的线性空间 V 的一个非空子集合 S 称为 V 的一个线性子空间(或简称子空间, subspace) , 如果 S 对于 V 的两种运算也构成数域 F 上的线性空间.

Theorem 14. *如果线性空间 V 的一个非空集合 S 对于 V 的两种运算是封闭的, 即*

1. *$\forall \boldsymbol{x}, \boldsymbol{y} \in S$, 都有 $\boldsymbol{x} + \boldsymbol{y} \in S$;*
2. *$\forall \boldsymbol{x} \in S$, $\forall \lambda \in F$, 都有 $\lambda\boldsymbol{x} \in S$.*

那么 S 就是一个子空间.

☞ 若 $\boldsymbol{x} \in S$, 则 $-\boldsymbol{x} = (-1)\boldsymbol{x} \in S$; 且 $\boldsymbol{x} + (-\boldsymbol{x}) = \boldsymbol{0} \in S$. 又 S 中的元素也是 V 的元素, 故满足 V 中的结合律、交换律、分配律等. 2.22

Theorem 15. 数域 F 上的线性空间 V 的一个非空子集合 S 为 V 的一个线性子空间 $\iff \forall \boldsymbol{x}, \boldsymbol{y} \in S, \forall \lambda \in F$, 都有

$$\lambda \boldsymbol{x} + \boldsymbol{y} \in S.$$

证: (1) 充分性. $\forall \boldsymbol{x}, \boldsymbol{y} \in S, \forall \lambda \in F$, 因

$$\begin{aligned} \boldsymbol{x} + \boldsymbol{y} &= 1\boldsymbol{x} + \boldsymbol{y} \in S, \\ \lambda \boldsymbol{x} &= \lambda \boldsymbol{x} + \boldsymbol{0} \in S, \end{aligned}$$

故 S 对于 V 的两种运算是封闭的. 即 S 为 V 的一个线性子空间.

(2) 必要性. $\forall \boldsymbol{x}, \boldsymbol{y} \in S, \forall \lambda \in F$, 由数乘封闭性, 有

$$\lambda \boldsymbol{x} \in S,$$

又由加法封闭性, 有

$$\lambda \boldsymbol{x} + \boldsymbol{y} \in S. \quad \square$$ 2.23

Example 16. 在线性空间中, 由单个的零向量所组成的子集合是一个线性子空间, 它叫做零子空间.

Example 17. 线性空间 V 本身也是 V 的一个子空间.

每个线性空间至少有两个子空间: 零子空间和线性空间本身. 这两个子空间通常叫做 V 的平凡子空间(trivial subspace), 而其它的子空间叫做非平凡子空间.

Example 18. 在全体实函数组成的空间中, 所有的实系数多项式组成一个子空间.

Example 19. $F[x]_n$ 是线性空间 $F[x]$ 的子空间. 2.24

Example 20. 设 $\boldsymbol{A} \in F^{m \times n}$, 则齐次线性方程组 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ 的解集合

$$S = \{\boldsymbol{x} \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}\}$$

是 F^n 的一个子空间.

☞ 但是, 非齐次线性方程组 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ 的解集合 W 不是 F^n 的子空间.

事实上, 若 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ 无解, 则 W 是空集, 当然 W 不是线性空间.

当 W 非空, 由于 $\boldsymbol{b} \neq \boldsymbol{0}$, 则 $\boldsymbol{0} \notin W$, 因此 W 不是线性空间. 2.25

Example 21. 设 $\mathbb{R}^3$ 的子集合

$$\begin{aligned} V_1 &= \{(x, 0, 0) \mid x \in \mathbb{R}\}, \\ V_2 &= \{(1, 0, z) \mid z \in \mathbb{R}\}, \end{aligned}$$

则 V_1 是 $\mathbb{R}^3$ 的子空间; 而 V_2 不是 $\mathbb{R}^3$ 的子空间.

☞ V_1 是 x 轴上的全体向量; V_2 是过点 $(1, 0, 0)$ 与 z 轴平行的直线上的全体向量. 显然 V_2 关于加法和数乘不封闭. 2.26

线性组合, 线性表示

Definition 22. 设 V 为数域 F 上的线性空间, $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k$ 为 V 中的向量, $\alpha_1, \alpha_2, \cdots, \alpha_k$ 为 F 中的数, 令

$$\boldsymbol{y} = \alpha_1\boldsymbol{x}_1 + \alpha_2\boldsymbol{x}_2 + \cdots + \alpha_k\boldsymbol{x}_k,$$

则称 $\boldsymbol{y}$ 为 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k$ 的线性组合(linear combination), $\alpha_1, \alpha_2, \cdots, \alpha_k$ 为该线性组合的组合系数, 而称和式

$$\alpha_1\boldsymbol{x}_1 + \alpha_2\boldsymbol{x}_2 + \cdots + \alpha_k\boldsymbol{x}_k$$

为 $\boldsymbol{y}$ 的线性表示或线性表出. 2.27

Example 23. 设 V 为数域 F 上的线性空间, $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k$ 为 V 中 k 个固定向量, 作 V 的子集 S 如下:

$$S = \{\boldsymbol{y} \mid \boldsymbol{y} = \alpha_1\boldsymbol{x}_1 + \alpha_2\boldsymbol{x}_2 + \cdots + \alpha_k\boldsymbol{x}_k,\ \alpha_i \in F\}.$$

则 S 是 V 的子空间.

证: 设 $\boldsymbol{y} = \alpha_1\boldsymbol{x}_1 + \alpha_2\boldsymbol{x}_2 + \cdots + \alpha_k\boldsymbol{x}_k \in S$, $\boldsymbol{z} = \beta_1\boldsymbol{x}_1 + \beta_2\boldsymbol{x}_2 + \cdots + \beta_k\boldsymbol{x}_k \in S$.

- 因 $\boldsymbol{y} + \boldsymbol{z} = \sum_{i=1}^{k}(\alpha_i + \beta_i)\boldsymbol{x}_i \in S$, 故对加法封闭.
- $\forall \lambda \in F$, $\lambda\boldsymbol{y} = \sum_{i=1}^{k}(\lambda\alpha_i)\boldsymbol{x}_i \in S$, 故对数乘封闭.

得证 S 是 V 的子空间. □ 2.28

Definition 24. 设 V 为数域 F 上的线性空间, $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k$ 为 V 中 k 个固定向量, 作 V 的子集 S 如下:

$$S = \{\boldsymbol{y} \mid \boldsymbol{y} = \alpha_1\boldsymbol{x}_1 + \alpha_2\boldsymbol{x}_2 + \cdots + \alpha_k\boldsymbol{x}_k,\ \alpha_i \in F\}.$$

则 S 称为由 V 中向量 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k$ 张成的线性子空间, 简称为 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k$ 的张空间, 并记为

$$S = \operatorname{span}[\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k].$$

For example, if $\boldsymbol{u} \neq \boldsymbol{0}$ is a vector in $\mathbb{R}^3$, then $\operatorname{span}[\boldsymbol{u}]$ is the straight line passing through the origin and $\boldsymbol{u}$.

If $S = \operatorname{span}[\boldsymbol{u}, \boldsymbol{v}]$, where $\boldsymbol{u}$ and $\boldsymbol{v}$ are two nonzero vectors in $\mathbb{R}^3$ not lying on the same line, then $S = \operatorname{span}[\boldsymbol{u}, \boldsymbol{v}]$ is the plane passing through the origin and the points $\boldsymbol{u}$ and $\boldsymbol{v}$. 2.29

Example 25.

- $\left\{\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}2\\2\end{bmatrix}\right\}$ spans the line $y = x$ in $\mathbb{R}^2$.
- The unit vectors $\left\{\boldsymbol{e}_1 = \begin{bmatrix}1\\0\\0\end{bmatrix}, \boldsymbol{e}_2 = \begin{bmatrix}0\\1\\0\end{bmatrix}, \boldsymbol{e}_3 = \begin{bmatrix}0\\0\\1\end{bmatrix}\right\}$ spans $\mathbb{R}^3$.

- The unit vectors $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n\}$ in $\mathbb{R}^n$ form a spanning set for $\mathbb{R}^n$.
- The finite set $\{1, x, x^2, \cdots, x^{n-1}\}$ spans the space of all polynomials such that $\deg p(x) \leqslant n-1$, and the infinite set $\{1, x, x^2, \cdots\}$ spans the space of all polynomials. 2.30

Example 26. 设 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$, 证明

$$R(\boldsymbol{A}) = \{\boldsymbol{y} \in \mathbb{C}^m \mid \boldsymbol{y} = \boldsymbol{A}\boldsymbol{x},\ \boldsymbol{x} \in \mathbb{C}^n\}$$

是 $\boldsymbol{A}$ 的所有列向量张成的 $\mathbb{C}^m$ 的子空间.

证: 首先, $R(\boldsymbol{A})$ 是 $\mathbb{C}^m$ 的子集, 且按 $\mathbb{C}^m$ 的向量加法和数量乘法构成线性空间, 故 $R(\boldsymbol{A})$ 是 $\mathbb{C}^m$ 的子空间.

其次, 记 $\boldsymbol{A} = [\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n]$, $\boldsymbol{x} = (\xi_1, \xi_2, \cdots, \xi_n)^{\mathrm{T}}$, 则

$$\begin{aligned} R(\boldsymbol{A}) &= \{\boldsymbol{y} \in \mathbb{C}^m \mid \boldsymbol{y} = \boldsymbol{A}\boldsymbol{x},\ \boldsymbol{x} \in \mathbb{C}^n\} \\ &= \{\boldsymbol{y} \in \mathbb{C}^m \mid \boldsymbol{y} = \xi_1\boldsymbol{a}_1 + \xi_2\boldsymbol{a}_2 + \cdots + \xi_n\boldsymbol{a}_n,\ \boldsymbol{x} = (\xi_1, \xi_2, \cdots, \xi_n)^{\mathrm{T}} \in \mathbb{C}^n\}, \end{aligned}$$

即 $R(\boldsymbol{A})$ 中任一向量都是 $\boldsymbol{A}$ 的列向量 $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\cdots$, $\boldsymbol{a}_n$ 的线性组合, 因此是它们的张空间. □ 2.31

结论

对于一个给定的矩阵 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$, 若 $\boldsymbol{A}$ 的列向量记为 $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\cdots$, $\boldsymbol{a}_n$, 即

$$\boldsymbol{A} = [\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n],$$

则 $\boldsymbol{A}$ 的列空间可以表示为

$$R(\boldsymbol{A}) = \mathrm{span}[\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n].$$

2.32

Theorem 27. *如果 V_1, V_2 是线性空间 V 的两个子空间, 那么它们的交 $V_1 \cap V_2$ 也是 V 的子空间.*

由集合的交的定义有, 子空间的交适合下列运算规律:

$$V_1 \cap V_2 = V_2 \cap V_1, \qquad \text{(交换律)}$$

$$(V_1 \cap V_2) \cap V_3 = V_1 \cap (V_2 \cap V_3). \qquad \text{(结合律)}$$

由结合律, 可以定义多个子空间的交:

$$V_1 \cap V_2 \cap \cdots \cap V_s = \bigcap_{i=1}^{s} V_i,$$

它也是子空间. 2.33

Definition 28 (子空间的和)**.** 设 V_1, V_2 是线性空间 V 的子空间, 所谓 V_1 与 V_2 的和, 是指由所有能表示成 $\alpha_1 + \alpha_2$, 而 $\alpha_1 \in V_1$, $\alpha_2 \in V_2$ 的向量组成的子集合, 记作 $V_1 + V_2$.

Theorem 29. 如果 V_1, V_2 是线性空间 V 的子空间, 那么它们的和 V_1+V_2 也是 V 的子空间. 2.34

由定义有, 子空间的和适合下列运算规律:

$$V_1+V_2=V_2+V_1, \qquad \text{(交换律)}$$

$$(V_1+V_2)+V_3=V_1+(V_2+V_3). \qquad \text{(结合律)}$$

由结合律, 可以定义多个子空间的和

$$V_1+V_2+\cdots+V_s=\sum_{i=1}^{s} V_i.$$

它是由所有表示成

$$\alpha_1+\alpha_2+\cdots+\alpha_s, \qquad \alpha_i \in V_i \quad (i=1,2,\cdots,s)$$

的向量组成的子空间. 2.35

直和, direct sum

Definition 30 (直和)**.** 设 V 为数域 F 上的线性空间, S_1, S_2 为 V 的两个子空间, 如果有

$$S=S_1+S_2, \qquad \text{且} \qquad S_1\cap S_2=\{\mathbf{0}\},$$

则称 S 为 S_1, S_2 的直和空间, 并记为

$$S=S_1\oplus S_2, \qquad \text{或} \qquad S=S_1\dotplus S_2.$$

此时又称 S 有直和分解 $S_1\oplus S_2$.

更进一步, 如果 $V=S_1\oplus S_2$, 则称 S_1 为 S_2 的(代数) 补空间, 也称 S_2 为 S_1 的 (代数) 补空间. 或简单地称 S_1 与 S_2 为 V 的互补子空间, 并以 S^{C} 表示 V 的任一子空间 S 的代数补空间. 2.36

Example 31. 在三维几何空间 $\mathbb{R}^3$ 中, 用 V_1 表示过坐标原点的直线, V_2 表示一个通过坐标原点而且与 V_1 垂直的平面, 那么

$$V_1\cap V_2=\{\mathbf{0}\},$$

$$V_1+V_2=\mathbb{R}^3.$$

故 $\mathbb{R}^3=V_1\oplus V_2$, 且 V_1 与 V_2 为互补子空间. 即 $V_1^{\mathrm{C}}=V_2$, $V_2^{\mathrm{C}}=V_1$.

显然, $\mathbb{R}^3$ 的直和分解是不唯一的. 等价地, 一个真子空间的补空间是不唯一的. 比如上例中的 V_1 与 V_2 相互垂直的条件, 减弱为不共面, 它们仍然是互补的.

注意

"子空间的交" 与 "集合的交" 的概念是一致的, 但是 "子空间的和" 与 "集合的并" 的概念是不一致的.

上例中, $V_1+V_2=\mathbb{R}^3$, 而 $V_1\cup V_2=\{$过原点的直线与垂直平面上的点$\}$.

☞ $V_1\cup V_2$ 一般不构成子空间. 可以证明: $V_1\cup V_2$ 是 V 的子空间 $\Leftrightarrow$ $V_1\subseteq V_2$, 或 $V_2\subseteq V_1$. 2.37

Exercise 32 (P.39, 习题 (一) 3)**.** 设 S_1, S_2 是线性空间 V 的子空间, 则 S_1+S_2 是直和的充要条件是 S_1+S_2 中每个向量 $\boldsymbol{x}$ 的分解式

$$\boldsymbol{x}=\boldsymbol{x}_1+\boldsymbol{x}_2 \qquad (\boldsymbol{x}_1\in S_1,\ \boldsymbol{x}_2\in S_2)$$

是唯一的.

2.38

证: 充分性. 已知 S_1+S_2 是直和, 则 $S_1\cap S_2=\{\mathbf{0}\}$. $\forall \boldsymbol{x}\in S_1+S_2$, 设它有两个分解式

$$\boldsymbol{x}=\boldsymbol{x}_1+\boldsymbol{x}_2=\boldsymbol{y}_1+\boldsymbol{y}_2 \qquad (\boldsymbol{x}_1,\boldsymbol{y}_1\in S_1,\ \boldsymbol{x}_2,\boldsymbol{y}_2\in S_2).$$

则

$$\boldsymbol{x}_1-\boldsymbol{y}_1=\boldsymbol{y}_2-\boldsymbol{x}_2,$$

而 $\boldsymbol{x}_1-\boldsymbol{y}_1\in S_1$, $\boldsymbol{y}_2-\boldsymbol{x}_2\in S_2$, 故

$$\boldsymbol{x}_1-\boldsymbol{y}_1=\boldsymbol{y}_2-\boldsymbol{x}_2\in S_1\cap S_2=\{\mathbf{0}\}.$$

得 $\boldsymbol{x}_1=\boldsymbol{y}_1$, $\boldsymbol{y}_2=\boldsymbol{x}_2$. 故 $\boldsymbol{x}$ 的表示法是唯一的. 2.39

必要性. 任取向量 $\boldsymbol{x}\in S_1\cap S_2$, 于是零向量可以表示为

$$\mathbf{0}=\boldsymbol{x}+(-\boldsymbol{x}) \qquad (\boldsymbol{x}\in S_1,\ -\boldsymbol{x}\in S_2).$$

而 $\mathbf{0}=\mathbf{0}+\mathbf{0}$, 由表示法唯一, 得

$$\boldsymbol{x}=-\boldsymbol{x}=\mathbf{0},$$

得证

$$S_1\cap S_2=\{\mathbf{0}\}.$$

即 S_1+S_2 是直和. □ 2.40

课后作业

P.41, 习题 8

设 S_1 与 S_2 分别是齐次线性方程组

$$x_1+x_2+\cdots+x_n=0$$

和

$$x_1=x_2=\cdots=x_n$$

的解空间, 证明:

$$\mathbb{R}^n=S_1\oplus S_2.$$

2.41

证: 第一步, 证明 $\mathbb{R}^n=S_1+S_2$. 任取 $\boldsymbol{x}=(a_1,a_2,\cdots,a_n)^{\mathrm{T}}\in\mathbb{R}^n$, 设

$$\boldsymbol{x}=\boldsymbol{x}_1+\boldsymbol{x}_2,\quad \boldsymbol{x}_1\in S_1,\ \boldsymbol{x}_2\in S_2.$$

由 S_2 的定义, 可设 $\boldsymbol{x}_2=(b,b,\cdots,b)^{\mathrm{T}}$, 则

$$\boldsymbol{x}_1=\boldsymbol{x}-\boldsymbol{x}_2=(a_1-b,a_2-b,\cdots,a_n-b)^{\mathrm{T}}.$$

由 S_1 的定义, 有

$$(a_1-b)+(a_2-b)+\cdots+(a_n-b)=0,$$

故

$$b=\frac{1}{n}(a_1+a_2+\cdots+a_n).$$

上述讨论表明, 对任意 $\boldsymbol{x}=(a_1,a_2,\cdots,a_n)^{\mathrm{T}}\in\mathbb{R}^n$, 总有 $\boldsymbol{x}_1=(a_1-b,a_2-b,\cdots,a_n-b)^{\mathrm{T}}\in S_1$, $\boldsymbol{x}_2=(b,b,\cdots,b)^{\mathrm{T}}\in S_2$, 其中 $b=\frac{1}{n}(a_1+a_2+\cdots+a_n)$, 使得 $\boldsymbol{x}=\boldsymbol{x}_1+\boldsymbol{x}_2$. 故 $\mathbb{R}^n=S_1+S_2$. 2.42

第二步, 证明 S_1+S_2 是直和. 下证 $S_1\cap S_2=\{\mathbf{0}\}$.

任取 $\boldsymbol{x}=(b_1,b_2,\cdots,b_n)^{\mathrm{T}}\in S_1\cap S_2$, 则 $\boldsymbol{x}\in S_1$ 且 $\boldsymbol{x}\in S_2$. 由 S_1 和 S_2 的定义有

$$b_1+b_2+\cdots+b_n=0,$$

且

$$b_1=b_2=\cdots=b_n.$$

故

$$b_1=b_2=\cdots=b_n=0,$$

即 $\boldsymbol{x}=\mathbf{0}$. 从而 $S_1\cap S_2=\{\mathbf{0}\}$.

综上得证 $\mathbb{R}^n=S_1\oplus S_2$. □ 2.43

另证: 方程 $x_1+x_2+\cdots+x_n=0$ 的基础解系为

$$\boldsymbol{\alpha}_1=(-1,1,0,\cdots,0)^{\mathrm{T}},\boldsymbol{\alpha}_2=(-1,0,1,0,\cdots,0)^{\mathrm{T}},\cdots,\boldsymbol{\alpha}_{n-1}=(-1,0,\cdots,0,1)^{\mathrm{T}},$$

方程 $x_1=x_2=\cdots=x_n$ 的基础解系为

$$\boldsymbol{\beta}=(1,1,\cdots,1)^{\mathrm{T}}.$$

因为

$$|\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{n-1},\boldsymbol{\beta}|=\begin{vmatrix}-1&-1&\cdots&-1&1\\1&0&\cdots&0&1\\0&1&\cdots&0&1\\\vdots&\vdots&&\vdots&\vdots\\0&0&\cdots&1&1\end{vmatrix}=(-1)^{n+1}n\neq0,$$

故 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{n-1},\boldsymbol{\beta}$ 线性无关, 从而为 $\mathbb{R}^n$ 的一组基. 由教材 P.48 推论 2.1.25, 知 $\mathbb{R}^n=S_1\oplus S_2$. □ 2.44

1.3 基底和维数

线性无关 (Linear independence)

Definition 33 (线性无关)**.** 设 V 为数域 F 上的线性空间, $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k$ 为 V 中的一组向量. 若存在 F 中的一组不全为零的数 $\alpha_1, \alpha_2, \cdots, \alpha_k$ 使得

$$\alpha_1\boldsymbol{x}_1 + \alpha_2\boldsymbol{x}_2 + \cdots + \alpha_k\boldsymbol{x}_k = \boldsymbol{0}, \tag{1}$$

则称向量组 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k$ 线性相关(linearly dependent), 否则称为线性无关(linearly independent).

☞ 换句话说, 向量组 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k$ 称为线性无关, 如果等式 (1) 只有在

$$\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$$

时才成立. 2.45

Example 34. 讨论 $\mathbb{R}^{2\times 2}$ 的矩阵组

$$\boldsymbol{A}_1 = \begin{bmatrix} a & 1 \\ 1 & 1 \end{bmatrix}, \quad \boldsymbol{A}_2 = \begin{bmatrix} 1 & a \\ 1 & 1 \end{bmatrix}, \quad \boldsymbol{A}_3 = \begin{bmatrix} 1 & 1 \\ a & 1 \end{bmatrix}, \quad \boldsymbol{A}_4 = \begin{bmatrix} 1 & 1 \\ 1 & a \end{bmatrix},$$

的线性相关性.

解: 设 $k_1\boldsymbol{A}_1 + k_2\boldsymbol{A}_2 + k_3\boldsymbol{A}_3 + k_4\boldsymbol{A}_4 = \boldsymbol{O}$, 即

$$k_1\begin{bmatrix} a & 1 \\ 1 & 1 \end{bmatrix} + k_2\begin{bmatrix} 1 & a \\ 1 & 1 \end{bmatrix} + k_3\begin{bmatrix} 1 & 1 \\ a & 1 \end{bmatrix} + k_4\begin{bmatrix} 1 & 1 \\ 1 & a \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

故

$$\begin{cases} ak_1 + k_2 + k_3 + k_4 = 0, \\ k_1 + ak_2 + k_3 + k_4 = 0, \\ k_1 + k_2 + ak_3 + k_4 = 0, \\ k_1 + k_2 + k_3 + ak_4 = 0, \end{cases}$$

系数行列式为 2.46

$$\begin{vmatrix} a & 1 & 1 & 1 \\ 1 & a & 1 & 1 \\ 1 & 1 & a & 1 \\ 1 & 1 & 1 & a \end{vmatrix} = (a-1)^3(a+3),$$

- 当 $a \neq 1$ 且 $a \neq -3$ 时, 方程组只有零解, 从而 $\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{A}_3, \boldsymbol{A}_4$ 线性无关;
- 当 $a = 1$ 或 $a = -3$ 时, 方程组有非零解, 从而 $\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{A}_3, \boldsymbol{A}_4$ 线性相关. □ 2.47

常用结论

1. 单个向量 $\boldsymbol{x}$ 线性相关的充要条件是 $\boldsymbol{x}=\mathbf{0}$. 两个以上的向量 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_r$ 线性相关的充要条件是: 其中有一个向量是其余向量的线性组合.
2. 如果向量组 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_r$ 线性无关, 而且可以被 $\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_s$ 线性表出, 那么 $r \leqslant s$.
   由此推出, 两个等价的线性无关的向量组, 必含有相同个数的向量.
3. 如果向量组 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_r$ 线性无关, 但 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_r, \boldsymbol{y}$ 线性相关, 那么 $\boldsymbol{y}$ 可以被 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_r$ 线性表出, 而且表示法是唯一的.

2.48

Definition 35 (维数, dimension)**.** 如果在线性空间 V 中有 n 个线性无关的向量, 但是没有更多数目的线性无关的向量, 那么就称 V 为 n 维的; 记为 $\dim V=n$. 如果在 V 中可以找到任意多个线性无关的向量, 那么就称 V 为无限维的.

☞ 单个的零向量总是线性相关的. 事实上 $\forall \lambda \neq 0$, 都有

$$\lambda \mathbf{0}=\mathbf{0}.$$

故零子空间 $\{\mathbf{0}\}$ 中不存在线性无关的向量, 即

$$\dim\{\mathbf{0}\}=0.$$

2.49

Definition 36. 在 n 维线性空间 V 中, n 个线性无关的向量 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 称为 V 的一组基(basis). 记为

$$\mathscr{B}_V=\{\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n\}.$$

设 $\boldsymbol{x}$ 是 V 中任一向量, 于是 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n, \boldsymbol{x}$ 线性相关, 因此 $\boldsymbol{x}$ 可以被基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 线性表出:

$$\boldsymbol{x}=a_1\varepsilon_1+a_2\varepsilon_2+\cdots+a_n\varepsilon_n$$

其中系数 $a_1, a_2, \cdots, a_n$ 是被向量 $\boldsymbol{x}$ 和基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 唯一确定的, 这组数就称为 $\boldsymbol{x}$ 在基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 下的坐标, 记为 $(a_1, a_2, \cdots, a_n)^{\mathrm{T}}$.

2.50

Example 37. 在线性空间 $\mathbb{R}[x]_n$ 中,

$$1, x, x^2, \cdots, x^{n-1}, x^n$$

是 $n+1$ 个线性无关的向量, 而且每一个次数不超过 n 的数域 $\mathbb{R}$ 上的多项式都可以被它们线性表出, 所以 $\mathbb{R}[x]_n$ 是 $n+1$ 维的, 而 $1, x, x^2, \cdots, x^{n-1}, x^n$ 就是它的一组基.

在这组基下, 多项式 $f(x)=a_0+a_1x+\cdots+a_{n-1}x^{n-1}+a_nx^n$ 的坐标就是其系数

$$(a_0, a_1, \cdots, a_n)^{\mathrm{T}}.$$

☞ 如果取 $\mathbb{R}[x]_n$ 中的另外一组基

$$1, (x-a), \cdots, (x-a)^{n-1}, (x-a)^n,$$

按泰勒展开公式 $f(x) = f(a) + f'(a)(x-a) + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n$, 则 $f(x)$ 在此组基下的坐标是

$$\left(f(a), f'(a), \cdots, \frac{f^{(n)}(a)}{n!}\right)^{\mathrm{T}}.$$

2.51

Example 38. 考虑 $m \times n$ 矩阵

$$\boldsymbol{E}_{ij} = \begin{array}{cc} \begin{bmatrix} 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \end{bmatrix} & \begin{array}{c} \\ \\ \\ i \\ \\ \\ \\ \end{array} \\ j & \end{array}$$

即矩阵 $\boldsymbol{E}_{ij}$ 的元素在第 i 行第 j 列位置上为 1, 其他位置皆为 0.

称 $\boldsymbol{E}_{11}, \boldsymbol{E}_{12}, \cdots, \boldsymbol{E}_{1n}, \boldsymbol{E}_{21}, \boldsymbol{E}_{22}, \cdots, \boldsymbol{E}_{2n}, \cdots, \boldsymbol{E}_{m1}, \boldsymbol{E}_{m2}, \cdots, \boldsymbol{E}_{mn}$ 为空间 $\mathbb{C}^{m\times n}$ 的标准基.

☞ 显然 $\dim \mathbb{C}^{m\times n} = mn$. 2.52

Example 39. 已知 $\boldsymbol{E}_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $\boldsymbol{E}_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $\boldsymbol{E}_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$, $\boldsymbol{E}_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ 是 $\mathbb{R}^{2\times 2}$ 的标准基, 求矩阵

$$\boldsymbol{A} = \begin{bmatrix} 2 & 3 \\ -4 & 5 \end{bmatrix}$$

在该组基底下的坐标.

解: 由

$$\begin{aligned} \boldsymbol{A} = \begin{bmatrix} 2 & 3 \\ -4 & 5 \end{bmatrix} &= 2\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + 3\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} - 4\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + 5\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \\ &= 2\boldsymbol{E}_{11} + 3\boldsymbol{E}_{12} - 4\boldsymbol{E}_{21} + 5\boldsymbol{E}_{22}, \end{aligned}$$

故矩阵 $\boldsymbol{A}$ 在该组基底下的坐标为 $(2, 3, -4, 5)^{\mathrm{T}}$. □

☞ 一般地, 矩阵 $\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ 在标准基 $\boldsymbol{E}_{11}, \boldsymbol{E}_{12}, \boldsymbol{E}_{21}, \boldsymbol{E}_{22}$ 下的坐标为

$$(a_{11}, a_{12}, a_{21}, a_{22})^{\mathrm{T}}.$$

2.53

Example 40.
- 如果把复数域 $\mathbb{C}$ 看作是自身上的线性空间, 那么它是一维的, 数 1 就是一组基.
- 如果把复数域 $\mathbb{C}$ 看作是实数域 $\mathbb{R}$ 上的线性空间, 那么就是二维的, 数 1 与 i 就是一组基.

☞ 维数与所考虑的数域有关. 2.54

Theorem 41. 设 V 是数域 F 上的线性空间, $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k$ 为 V 中的向量组, 则它为线性相关向量组的充要条件是, 其中必有 (至少有一个) 向量能由其余的向量线性表出.

Theorem 42. 向量组 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k$ 线性无关的充要条件是, 其中任一向量均不能由其余向量线性表出. 2.55

Theorem 43. 设 V 是数域 F 上的线性空间, $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n\} = \mathscr{B}_V$ 是 V 的基, $\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_k$ 为 V 的线性无关的向量组, 则在 $\mathscr{B}_V$ 中至少可以找到一个向量, 不妨设为 $\boldsymbol{x}_1$, 使得向量组 $\boldsymbol{x}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_k$ 仍然线性无关.

证: 反证法. 假设对任意的 $\boldsymbol{x}_i$ $(i = 1, 2, \cdots, n)$, 都有 $\boldsymbol{x}_i, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_k$ 线性相关.

已知 $\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_k$ 线性无关, 故 $\boldsymbol{y}_2, \boldsymbol{y}_3, \cdots, \boldsymbol{y}_k$ 也线性无关.

从而 $\boldsymbol{x}_i$ 可以由 $\boldsymbol{y}_2, \boldsymbol{y}_3, \cdots, \boldsymbol{y}_k$ 线性表示.

则向量组 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n$ 可以由向量组 $\boldsymbol{y}_2, \boldsymbol{y}_3, \cdots, \boldsymbol{y}_k$ 线性表示.

而 $\boldsymbol{y}_1$ 可以由基 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n$ 线性表示, 导致 $\boldsymbol{y}_1$ 可以由向量组 $\boldsymbol{y}_2, \boldsymbol{y}_3, \cdots, \boldsymbol{y}_k$ 线性表示.

与题设 "$\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_k$ 线性无关" 矛盾. 假设不成立. 得证. □ 2.56

Theorem 44. 线性空间 V 的任意两组基

$$\mathscr{B}_V^{(1)} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k\}, \qquad \mathscr{B}_V^{(2)} = \{\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_n\},$$

所含向量个数相等.

Corollary 45. n 维线性空间 V 中任意 n 个线性无关的向量

$$\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_n$$

均可以作为 V 的基. 2.57

Corollary 46. 设 V 是 n 维向量空间, 则 V 中任意 k $(k < n)$ 个线性无关的向量

$$\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k$$

必可以扩充成 V 的一组基.

证: 由 $k < n$, 知向量组 $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k$ 不是 V 的基, 故 V 中必有一个向量 $\boldsymbol{\beta}_1$ 不能由 $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k$ 线性表出, 从而向量组

$$\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k, \boldsymbol{\beta}_1$$

线性无关. 若 $k+1=n$, 则向量组 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_k,\boldsymbol{\beta}_1$ 为 V 的一组基.

若 $k+1<n$, 则 V 中必有一个向量 $\boldsymbol{\beta}_2$ 不能由 $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_k,\boldsymbol{\beta}_1$ 线性表出, 从而 $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_k,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2$ 线性无关.

依次下去, 可得到线性无关的向量组

$$\boldsymbol{x}_1,\cdots,\boldsymbol{x}_k,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s,$$

其中 $r+s=n$, 从而把 $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_k$ 扩充成了 V 的一组基. □ 2.58

基变换和坐标变换

设 V 是数域 F 上的线性空间, 给定 V 的两组基底:

$$\mathscr{B}_V^{(1)}=\{\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n\},\qquad \mathscr{B}_V^{(2)}=\{\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_n\},$$

设向量 $\boldsymbol{\alpha}\in V$ 在两组基底下的坐标分别为

$$\boldsymbol{x}=(a_1,a_2,\cdots,a_n)^{\mathrm{T}}\quad 和 \quad \boldsymbol{y}=(b_1,b_2,\cdots,b_n)^{\mathrm{T}}.$$

试问: $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 有什么关系? 2.59

(1) 两组基底之间的关系.

由于 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$ 是 V 的一组基底, 故有

$$\begin{cases}\boldsymbol{y}_1=p_{11}\boldsymbol{x}_1+p_{21}\boldsymbol{x}_2+\cdots+p_{n1}\boldsymbol{x}_n,\\ \boldsymbol{y}_2=p_{12}\boldsymbol{x}_1+p_{22}\boldsymbol{x}_2+\cdots+p_{n2}\boldsymbol{x}_n,\\ \qquad\vdots\\ \boldsymbol{y}_n=p_{1n}\boldsymbol{x}_1+p_{2n}\boldsymbol{x}_2+\cdots+p_{nn}\boldsymbol{x}_n.\end{cases}\tag{2}$$

引进如下形式的写法:

$$a_1\boldsymbol{x}_1+a_2\boldsymbol{x}_2+\cdots+a_n\boldsymbol{x}_n\triangleq[\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n]\begin{bmatrix}a_1\\a_2\\\vdots\\a_n\end{bmatrix},$$

2.60

则 (2) 式可以写成

$$[\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_n]=[\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n]\begin{bmatrix}p_{11}&p_{12}&\cdots&p_{1n}\\p_{21}&p_{22}&\cdots&p_{2n}\\\vdots&\vdots&&\vdots\\p_{n1}&p_{n2}&\cdots&p_{nn}\end{bmatrix}.$$

记

$$\boldsymbol{P}=\begin{bmatrix}p_{11}&p_{12}&\cdots&p_{1n}\\p_{21}&p_{22}&\cdots&p_{2n}\\\vdots&\vdots&&\vdots\\p_{n1}&p_{n2}&\cdots&p_{nn}\end{bmatrix},$$

即

$$[\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_n]=[\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n]\boldsymbol{P},$$

则称矩阵 $\boldsymbol{P}$ 为由基底 $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n$ 到基底 $\boldsymbol{y}_1,\cdots,\boldsymbol{y}_n$ 的过渡矩阵. 2.61

Theorem 47. 设 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$ 是 V 的一组基底, 且向量组 $\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_n$ 满足

$$[\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_n]=[\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n]\boldsymbol{P},\tag{3}$$

则向量组 $\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_n$ 是 V 的一组基底当且仅当 $\boldsymbol{P}$ 是可逆矩阵.

证:

$$\begin{aligned}
& \boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_n \text{ 是 } V \text{ 的一组基底}\\
\Longleftrightarrow\quad & \boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_n \text{ 线性无关}\\
\Longleftrightarrow\quad & \text{从 } k_1\boldsymbol{y}_1+k_2\boldsymbol{y}_2+\cdots+k_n\boldsymbol{y}_n=\boldsymbol{0} \text{ 可推出 } k_1=k_2=\cdots=k_n=0\\
\Longleftrightarrow\quad & \text{从 } [\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_n]\begin{bmatrix}k_1\\k_2\\\vdots\\k_n\end{bmatrix}=\boldsymbol{0} \text{ 可推出 } k_1=k_2=\cdots=k_n=0\\
\Longleftrightarrow\quad & \text{从 } [\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n]\boldsymbol{P}\begin{bmatrix}k_1\\k_2\\\vdots\\k_n\end{bmatrix}=\boldsymbol{0} \text{ 可推出 } k_1=k_2=\cdots=k_n=0\\
\Longleftrightarrow\quad & \text{从 } \boldsymbol{P}\begin{bmatrix}k_1\\k_2\\\vdots\\k_n\end{bmatrix}=\boldsymbol{0} \text{ 可推出 } k_1=k_2=\cdots=k_n=0\\
\Longleftrightarrow\quad & \text{齐次线性方程组 } \boldsymbol{P}\boldsymbol{x}=\boldsymbol{0} \text{ 只有零解}\\
\Longleftrightarrow\quad & \boldsymbol{P} \text{ 是可逆矩阵.}
\end{aligned}$$

2.62

在上述证明过程中, 倒数第 3 个 "$\Longleftrightarrow$" 用到了 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$ 线性无关的条件. □ 2.63

(2) 坐标转换

Theorem 48. 设向量 $\boldsymbol{\alpha}\in V$ 在两组基底 $\mathscr{B}_V^{(1)}=\{\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n\}$ 和 $\mathscr{B}_V^{(2)}=\{\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_n\}$ 下的坐标分别为

$$\boldsymbol{x}=(a_1,a_2,\cdots,a_n)^{\mathrm{T}}\quad \text{和}\quad \boldsymbol{y}=(b_1,b_2,\cdots,b_n)^{\mathrm{T}}.$$

基 $\mathscr{B}_V^{(1)}$ 到 $\mathscr{B}_V^{(2)}$ 的过渡矩阵为 $\boldsymbol{P}$, 则

$$\boldsymbol{P}\boldsymbol{y}=\boldsymbol{x}\quad \text{或}\quad \boldsymbol{y}=\boldsymbol{P}^{-1}\boldsymbol{x}.$$

2.64

证:

$$
\boldsymbol{\alpha} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n] \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = [\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_n] \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}
$$

$$
= [\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n]\boldsymbol{P} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n] \left(\boldsymbol{P} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \right)
$$

由于 $\boldsymbol{\alpha}$ 在基底 $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n\}$ 下的坐标是唯一的, 故

$$
\boldsymbol{P}\boldsymbol{y} = \boldsymbol{x} \quad \text{或} \quad \boldsymbol{y} = \boldsymbol{P}^{-1}\boldsymbol{x}. \quad \square
$$

2.65

Example 49. 证明 $\boldsymbol{G}_1 = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$, $\boldsymbol{G}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\boldsymbol{G}_3 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, $\boldsymbol{G}_4 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ 是 $\mathbb{R}^{2\times 2}$ 的一组基, 并求 $\boldsymbol{A} = \begin{bmatrix} 0 & 1 \\ 3 & -3 \end{bmatrix}$ 在此基下的坐标.

2.66

证: 矩阵 $\boldsymbol{G}_1$, $\boldsymbol{G}_2$, $\boldsymbol{G}_3$, $\boldsymbol{G}_4$ 在标准基 $\boldsymbol{E}_{11}$, $\boldsymbol{E}_{12}$, $\boldsymbol{E}_{21}$, $\boldsymbol{E}_{22}$ 下的坐标分别为

$$
\boldsymbol{x}_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \boldsymbol{x}_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \quad \boldsymbol{x}_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \quad \boldsymbol{x}_4 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix},
$$

记 $\boldsymbol{P} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4]$, 则

$$
[\boldsymbol{G}_1, \boldsymbol{G}_2, \boldsymbol{G}_3, \boldsymbol{G}_4] = [\boldsymbol{E}_{11}, \boldsymbol{E}_{12}, \boldsymbol{E}_{21}, \boldsymbol{E}_{22}]\boldsymbol{P}.
$$

而

$$
\det \boldsymbol{P} = -2 \neq 0,
$$

即矩阵 $\boldsymbol{P}$ 可逆, 因此向量组 $\boldsymbol{G}_1$, $\boldsymbol{G}_2$, $\boldsymbol{G}_3$, $\boldsymbol{G}_4$ 也是 $\mathbb{R}^{2\times 2}$ 的一组基. 2.67

矩阵 $\boldsymbol{A} = \begin{bmatrix} 0 & 1 \\ 3 & -3 \end{bmatrix}$ 在标准基 $\boldsymbol{E}_{11}$, $\boldsymbol{E}_{12}$, $\boldsymbol{E}_{21}$, $\boldsymbol{E}_{22}$ 下的坐标为

$$
\boldsymbol{x} = (0, 1, 3, -3)^{\mathrm{T}}.
$$

设矩阵 $\boldsymbol{A}$ 在基 $\boldsymbol{G}_1$, $\boldsymbol{G}_2$, $\boldsymbol{G}_3$, $\boldsymbol{G}_4$ 下的坐标为 $\boldsymbol{y}$, 则

$$
\boldsymbol{y} = \boldsymbol{P}^{-1}\boldsymbol{x} = (0, -1, -2, 3)^{\mathrm{T}}. \quad \square
$$

2.68

维数定理

Theorem 50. 设 S_1 和 S_2 均为数域 F 上的线性空间 V 的子空间, 则有

$$\dim(S_1+S_2)=\dim S_1+\dim S_2-\dim(S_1\cap S_2).$$

或记为

$$\dim S_1+\dim S_2=\dim(S_1+S_2)+\dim(S_1\cap S_2).$$

☞ 和的维数往往比维数的和来得小. 例如在 $\mathbb{R}^3$ 空间中, 两张通过原点的不同的平面之和是整个 $\mathbb{R}^3$ 空间, 而其维数之和为 4. 2.69

证: 设 $\dim S_1=n_1$, $\dim S_2=n_2$, $\dim(S_1\cap S_2)=m$. 取 $S_1\cap S_2$ 的一组基

$$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m.$$

它可以扩充成 S_1 的一组基

$$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m,\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_{n_1-m}.$$

也可以扩充成 S_2 的一组基

$$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m,\boldsymbol{\gamma}_1,\cdots,\boldsymbol{\gamma}_{n_2-m}.$$

下面来证明, 向量组

$$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m,\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_{n_1-m},\boldsymbol{\gamma}_1,\cdots,\boldsymbol{\gamma}_{n_2-m}$$

是 S_1+S_2 的一组基，这样 $\dim(S_1+S_2)=n_1+n_2-m$, 因而维数公式成立. 2.70

现在来证明向量组是线性无关的. 假设有等式

$$\begin{aligned}k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_m\boldsymbol{\alpha}_m+p_1\boldsymbol{\beta}_1+\cdots+p_{n_1-m}\boldsymbol{\beta}_{n_1-m}&\\+q_1\boldsymbol{\gamma}_1+\cdots+q_{n_2-m}\boldsymbol{\gamma}_{n_2-m}&=\mathbf{0}.\end{aligned}$$

令

$$\begin{aligned}\boldsymbol{\alpha}&=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_m\boldsymbol{\alpha}_m+p_1\boldsymbol{\beta}_1+\cdots+p_{n_1-m}\boldsymbol{\beta}_{n_1-m} &(4)\\&=-q_1\boldsymbol{\gamma}_1-\cdots-q_{n_2-m}\boldsymbol{\gamma}_{n_2-m}. &(5)\end{aligned}$$

由等式 (4), 有 $\boldsymbol{\alpha}\in S_1$; 而由等式 (5), 有 $\boldsymbol{\alpha}\in S_2$. 于是 $\boldsymbol{\alpha}\in S_1\cap S_2$, 则 $\boldsymbol{\alpha}$ 可以由 $\boldsymbol{\alpha}_1$, $\boldsymbol{\alpha}_2$, $\cdots$, $\boldsymbol{\alpha}_m$ 线性表示. 令 $\boldsymbol{\alpha}=l_1\boldsymbol{\alpha}_1+l_2\boldsymbol{\alpha}_2+\cdots+l_m\boldsymbol{\alpha}_m$ ，则

$$l_1\boldsymbol{\alpha}_1+l_2\boldsymbol{\alpha}_2+\cdots+l_m\boldsymbol{\alpha}_m+q_1\boldsymbol{\gamma}_1+\cdots+q_{n_2-m}\boldsymbol{\gamma}_{n_2-m}=\mathbf{0}.$$

由于 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m,\boldsymbol{\gamma}_1,\cdots,\boldsymbol{\gamma}_{n_2-m}$ 线性无关，得

$$l_1=l_2=\cdots=l_m=q_1=\cdots=q_{n_2-m}=0,$$

因而 $\boldsymbol{\alpha}=\mathbf{0}$. 2.71

从而有

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_m\boldsymbol{\alpha}_m+p_1\boldsymbol{\beta}_1+\cdots+p_{n_1-m}\boldsymbol{\beta}_{n_1-m}=\mathbf{0}.$$

由于 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m,\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_{n_1-m}$ 线性无关，故

$$k_1=k_2=\cdots=k_m=p_1=\cdots=p_{n_1-m}=0.$$

从而 $\boldsymbol{\alpha}_1$, $\boldsymbol{\alpha}_2$, $\cdots$, $\boldsymbol{\alpha}_m$, $\boldsymbol{\beta}_1$, $\cdots$, $\boldsymbol{\beta}_{n_1-m}$, $\boldsymbol{\gamma}_1$, $\cdots$, $\boldsymbol{\gamma}_{n_2-m}$ 线性无关.

另一方面, S_1+S_2 中任一向量都可以由上述向量组线性表示.

事实上, 设 $\boldsymbol{x}\in S_1+S_2$, 令 $\boldsymbol{x}=\boldsymbol{x}_1+\boldsymbol{x}_2$, $\boldsymbol{x}_1\in S_1$, $\boldsymbol{x}_2\in S_2$. 则 $\boldsymbol{x}_1$ 可以用 $\boldsymbol{\alpha}_1$, $\boldsymbol{\alpha}_2$, $\cdots$, $\boldsymbol{\alpha}_m$, $\boldsymbol{\beta}_1$, $\cdots$, $\boldsymbol{\beta}_{n_1-m}$, 线性表示; 且 $\boldsymbol{x}_2$ 可以用 $\boldsymbol{\alpha}_1$, $\boldsymbol{\alpha}_2$, $\cdots$, $\boldsymbol{\alpha}_m$, $\boldsymbol{\gamma}_1$, $\cdots$, $\boldsymbol{\gamma}_{n_2-m}$ 线性表示.

故 $\boldsymbol{x}$ 可以用 $\boldsymbol{\alpha}_1$, $\boldsymbol{\alpha}_2$, $\cdots$, $\boldsymbol{\alpha}_m$, $\boldsymbol{\beta}_1$, $\cdots$, $\boldsymbol{\beta}_{n_1-m}$, $\boldsymbol{\gamma}_1$, $\cdots$, $\boldsymbol{\gamma}_{n_2-m}$ 线性表示.

因而向量组 $\boldsymbol{\alpha}_1$, $\boldsymbol{\alpha}_2$, $\cdots$, $\boldsymbol{\alpha}_m$, $\boldsymbol{\beta}_1$, $\cdots$, $\boldsymbol{\beta}_{n_1-m}$, $\boldsymbol{\gamma}_1$, $\cdots$, $\boldsymbol{\gamma}_{n_2-m}$ 是 S_1+S_2 的一组基. 故维数公式成立. □ 2.72

Corollary 51. *设 S_1, S_2 是线性空间 V 的子空间, 则下列叙述是等价的:*

1. *$V=S_1\oplus S_2$.*
2. *$\dim(S_1+S_2)=\dim S_1+\dim S_2$.*
3. *S_1 的基底 $\mathscr{B}_{S_1}=\{\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_k\}$ 和 S_2 的基底 $\mathscr{B}_{S_2}=\{\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_l\}$ 组成的集合 $\{\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_k,\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_l\}$ 是和空间 S_1+S_2 的基底.*
4. *任意 $\boldsymbol{x}\in V$ 有唯一的分解式:*

$$\boldsymbol{x}=\boldsymbol{x}_1+\boldsymbol{x}_2,\qquad \boldsymbol{x}_1\in S_1,\ \boldsymbol{x}_2\in S_2.$$

5. *零向量 $\mathbf{0}$ 的分解式是唯一的.*

2.73

Corollary 52. *设 S 为 n 维线性空间 V 的一个真子空间, 则 V 中必有真子空间 M 成为 S 的代数补空间.*

2.74

Theorem 53. *两个向量组 $\{\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_s\}$ 及 $\{\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_t\}$ 张成相同子空间的充要条件是这两个向量组等价, 即两个向量组可以相互线性表示.*

2.75

Theorem 54. *设向量组 $\{\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_s\}$ 的秩为 r, 则*

$$\dim\operatorname{span}[\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_s]=r.$$

2.76

Exercise 55 (P.50 习题 (一) 7)**.** 设 V_1, V_2 都是线性空间 V 的子空间, 若 $\dim V_1=\dim V_2$, 且 $V_1\subseteq V_2$, 试证 $V_1=V_2$.

证: 设 $\dim V_1 = r$. 若 $r=0$, 则 V_1, V_2 都是零子空间 $\{\mathbf{0}\}$, 故 $V_1 = V_2$.

当 $r \neq 0$ 时, 任取 V_1 的一组基底 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_r$.

由于 $V_1 \subseteq V_2$, 则 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_r$ 也是 V_2 中的一组线性无关向量.

又 $\dim V_2 = r$, 故 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_r$ 是 V_2 的一组基底. 故

$$V_2 = \operatorname{span}[\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_r] = V_1. \quad \square$$

2.77

Exercise 56 (P.50 习题 (一) 8)**.** 给定矩阵

$$\boldsymbol{A} = \begin{bmatrix} 1 & 3 & 0 & 0 \\ 2 & 1 & 5 & 1 \\ 3 & 7 & 2 & 0 \end{bmatrix},$$

试求 $\dim R(\boldsymbol{A})$, $\dim N(\boldsymbol{A})$.

解: $\dim R(\boldsymbol{A}) = \operatorname{rank} \boldsymbol{A}$, $\dim N(\boldsymbol{A}) = n - \operatorname{rank} \boldsymbol{A}$, 其中 n 为矩阵 $\boldsymbol{A}$ 的列数.

对矩阵 $\boldsymbol{A}$ 进行列变换, 易知 $\operatorname{rank} \boldsymbol{A} = 3$, 故

$$\dim R(\boldsymbol{A}) = 3, \qquad \dim N(\boldsymbol{A}) = 1.$$

☞ $\dim R(\boldsymbol{A}) + \dim N(\boldsymbol{A}) = n$. $\square$

2.78

Exercise 57 (P.50 习题 (一) 9)**.** 试证多项式组 $1, (x-1), (x-1)^2, \cdots, (x-1)^n$ 也是线性空间 $\mathbb{R}[x]_n$ 的一个基底, 并求由基底 $\{1, x, x^2, \cdots, x^n\}$ 过渡到基底 $\{1, (x-1), (x-1)^2, \cdots, (x-1)^n\}$ 的过渡矩阵.

解: $\forall k \in \{0, 1, 2, \cdots, n\}$, 有

$$\begin{aligned}(x-1)^k &= (-1)^k + \mathrm{C}_k^1(-1)^{k-1}x + \mathrm{C}_k^2(-1)^{k-2}x^2 + \cdots + \mathrm{C}_k^{k-1}(-1)x^{k-1} + x^k \\ &= (1, x, x^2, \cdots, x^k, x^{k+1}, \cdots, x^n)\begin{bmatrix} (-1)^k \\ \mathrm{C}_k^1(-1)^{k-1} \\ \mathrm{C}_k^2(-1)^{k-2} \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.\end{aligned}$$

2.79

则

$$\big(1, (x-1), (x-1)^2, \cdots, (x-1)^n\big) = (1, x, x^2, \cdots, x^n)\boldsymbol{P},$$

其中

$$\boldsymbol{P} = \begin{bmatrix} 1 & -1 & 1 & \cdots & (-1)^n \\ 0 & 1 & -2 & \cdots & \mathrm{C}_n^1(-1)^{n-1} \\ 0 & 0 & 1 & \cdots & \mathrm{C}_n^2(-1)^{n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}_{(n+1)\times(n+1)}.$$

由于 $\det \boldsymbol{P} = 1 \neq 0$, 所以 1, $(x-1)$, $(x-1)^2$, $\cdots$, $(x-1)^n$ 也是线性空间 $\mathbb{R}[x]_n$ 的一个基底, 且所求过渡矩阵为 $\boldsymbol{P}$. □ 2.80

Exercise 58 (P.51 习题 (二) 7)**.** 设 V 为数域 F 上的 n 维线性空间, 向量 $\boldsymbol{u}_1$, $\boldsymbol{u}_2$, $\cdots$, $\boldsymbol{u}_l \in V$ 线性无关, $\boldsymbol{A} \in F_r^{l\times m}$, 而

$$[\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_m] = [\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_l]\boldsymbol{A},$$

试证 $\dim \operatorname{span}[\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_m] = \operatorname{rank} \boldsymbol{A} = r$.

证: 记 $\boldsymbol{A} = [\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_m]$, 已知 $\operatorname{rank} \boldsymbol{A} = r$, 不妨设 $\boldsymbol{A}$ 的列向量的极大无关组为 $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\cdots$, $\boldsymbol{a}_r$, 下证 $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, $\cdots$, $\boldsymbol{v}_r$ 线性无关.

$$\begin{aligned} 0 &= k_1\boldsymbol{v}_1 + k_2\boldsymbol{v}_2 + \cdots + k_r\boldsymbol{v}_r \\ &= k_1[\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_l]\boldsymbol{a}_1 + \cdots + k_r[\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_l]\boldsymbol{a}_r \\ &= [\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_l](k_1\boldsymbol{a}_1 + \cdots + k_r\boldsymbol{a}_r), \end{aligned}$$

因 $\boldsymbol{u}_1$, $\boldsymbol{u}_2$, $\cdots$, $\boldsymbol{u}_l$ 线性无关, 故

$$k_1\boldsymbol{a}_1 + \cdots + k_r\boldsymbol{a}_r = \boldsymbol{0}.$$

又 $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\cdots$, $\boldsymbol{a}_r$ 线性无关, 所以 $k_1 = k_2 = \cdots = k_r = 0$, 得 $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, $\cdots$, $\boldsymbol{v}_r$ 线性无关. 2.81

再证任取 $\boldsymbol{v}_i$ 均可以由 $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, $\cdots$, $\boldsymbol{v}_r$ 线性表示.

设 $\boldsymbol{a}_i = \lambda_1\boldsymbol{a}_1 + \cdots + \lambda_r\boldsymbol{a}_r$, 则

$$\begin{aligned} \boldsymbol{v}_i &= [\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_l]\boldsymbol{a}_i \\ &= [\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_l](\lambda_1\boldsymbol{a}_1 + \cdots + \lambda_r\boldsymbol{a}_r) \\ &= \lambda_1[\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_l]\boldsymbol{a}_1 + \cdots + \lambda_r[\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_l]\boldsymbol{a}_r \\ &= \lambda_1\boldsymbol{v}_1 + \cdots + \lambda_r\boldsymbol{v}_r. \end{aligned}$$

从而 $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, $\cdots$, $\boldsymbol{v}_r$ 为 $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, $\cdots$, $\boldsymbol{v}_m$ 的一个极大线性无关组. 故

$$\dim \operatorname{span}[\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_m] = \dim[\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_m] = r = \operatorname{rank} \boldsymbol{A}.$$

□ 2.82

Exercise 59 (P.51 习题 (二) 9)**.** 设 S_1 和 S_2 均为 n 维线性空间 V 的真子空间, 则 V 中至少有一个向量 $\boldsymbol{u}$ 既不属于 S_1, 也不属于 S_2.

证: 因 S_1 是真子空间, 故存在 $\boldsymbol{u} \notin S_1$.

- 若 $\boldsymbol{u} \notin S_2$, 则结论已成立.
- 若 $\boldsymbol{u} \in S_2$, 因 S_2 是真子空间, 故存在 $\boldsymbol{v} \notin S_2$.
  - 若 $\boldsymbol{v} \notin S_1$, 则结论已成立.
  - 若 $\boldsymbol{v} \in S_1$, 则便有

$$\boldsymbol{u} \notin S_1, \boldsymbol{v} \in S_1, \qquad \boldsymbol{u} \in S_2, \boldsymbol{v} \notin S_2.$$

于是存在

$$\boldsymbol{w} = \boldsymbol{u} + \boldsymbol{v} \notin S_1, \quad \boldsymbol{w} \notin S_2.$$

□ 2.83

1.4 和空间与直和空间概念的推广

Definition 60. 设 $S_1, S_2, \cdots, S_n$ 是线性空间 V 的一组子空间, 且满足:

1. $S = S_1 + S_2 + \cdots + S_n = \sum\limits_{i=1}^{n} S_i$;
2. $S_j \cap \Big(\sum\limits_{\substack{i=1 \\ i\neq j}}^{n} S_i\Big) = \{\mathbf{0}\}$, $1 \leqslant j \leqslant n$, 则称 S 是 $S_1, S_2, \cdots, S_n$ 的直和, 记为

$$S = S_1 \oplus S_2 \oplus \cdots \oplus S_n.$$

特别地, 当

$$V = S_1 \oplus S_2 \oplus \cdots \oplus S_n$$

时, 称 $S_1 \oplus S_2 \oplus \cdots \oplus S_n$ 为 V 的一个直和分解.

☞ 定义中的条件 2 不能简单地换为条件 $S_i \cap S_j = \{\mathbf{0}\}$, $i \neq j$. 2.84

Theorem 61. *设 $S_1, S_2, \cdots, S_n$ 是线性空间 V 的一组子空间, $S = \sum\limits_{i=1}^{n} S_i$. 则下列命题互相等价:*

1. *$S = S_1 \oplus S_2 \oplus \cdots \oplus S_n$;*
2. *对任意 $\boldsymbol{x} \in S$, 它有唯一的分解式:*

$$\boldsymbol{x} = \sum_{i=1}^{n} \boldsymbol{x}_i, \qquad \boldsymbol{x}_i \in S_i, \qquad i = 1, 2, \cdots, n.$$

3. *$S = \sum\limits_{i=1}^{n} S_i$ 中零向量有唯一的分解式.*
4. *$\dim S = \sum\limits_{i=1}^{n} \dim S_i$.*
5. *取每个 S_i 的基底, 其全体构成 S 的一个基底.*
6. *$\Big(\sum\limits_{i=1}^{j-1} S_i\Big) \cap S_j = \{\mathbf{0}\}$, $2 \leqslant j \leqslant n$.*

2.85

Example 62. 考虑空间 $\mathbb{C}^3$ 的自然基底 $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$, 令

$$\begin{aligned} S_1 &= \operatorname{span}[\boldsymbol{e}_1], \\ S_2 &= \operatorname{span}[\boldsymbol{e}_1 + \boldsymbol{e}_2], \\ S_3 &= \operatorname{span}[\boldsymbol{e}_2], \end{aligned}$$

则 S_1, S_2, S_3 都是 $\mathbb{C}^3$ 的一维子空间. 作和 $S = S_1 + S_2 + S_3$, 则有

$$S = \operatorname{span}[\boldsymbol{e}_1, \boldsymbol{e}_1 + \boldsymbol{e}_2, \boldsymbol{e}_2].$$

S 不是 S_1, S_2, S_3 的直和, 因为

$$2 = \dim S \neq \dim S_1 + \dim S_2 + \dim S_3 = 3.$$

但满足 $S_i \cap S_j = \{\mathbf{0}\}$, $i \neq j$. 这说明定义 60 中的条件 2 不能简单地换为条件 $S_i \cap S_j = \{\mathbf{0}\}$, $i \neq j$.

2.86

2 内积空间

2.1 内积空间的定义及例子

问题: 向量的长度、夹角在线性空间中如何定义?

在解析几何中, 设 $\boldsymbol{\alpha}=(x_1,y_1,z_1)$, $\boldsymbol{\beta}=(x_2,y_2,z_2)$, 它们的数量积 (又称内积) 为

$$(\boldsymbol{\alpha},\boldsymbol{\beta})=\boldsymbol{\alpha}\cdot\boldsymbol{\beta}=|\boldsymbol{\alpha}||\boldsymbol{\beta}|\cos\theta=x_1x_2+y_1y_2+z_1z_2,$$

其中 θ 是 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 的夹角.

有了数量积 (内积) 的概念, 向量的长度和夹角就可以表示为

$$|\boldsymbol{\alpha}|=\sqrt{(\boldsymbol{\alpha},\boldsymbol{\alpha})},$$
$$\theta=\arccos\frac{(\boldsymbol{\alpha},\boldsymbol{\beta})}{|\boldsymbol{\alpha}||\boldsymbol{\beta}|}.$$

☞ 数量积 (内积) 的概念蕴含着长度和夹角的概念. 为了给抽象的线性空间引进长度、夹角等度量, 我们先以数量积所具备的 4 条代数性质为依据, 在抽象的线性空间引入与数量积相类似的功能, 这就是内积的概念, 并把定义了内积的线性空间叫做内积空间. 2.87

内积概念的公理化

Definition 63. 设 V 是实数域 $\mathbb{R}$ 上的线性空间. $\forall \boldsymbol{x}$, $\boldsymbol{y}\in V$, 如能给定某种规则使 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 对应于一个实数, 记为 $(\boldsymbol{x},\boldsymbol{y})$, 且满足下列条件:

1. $(\boldsymbol{x},\boldsymbol{y})=(\boldsymbol{y},\boldsymbol{x})$;
2. $(\lambda\boldsymbol{x},\boldsymbol{y})=\lambda(\boldsymbol{x},\boldsymbol{y})$;
3. $(\boldsymbol{x}+\boldsymbol{y},\boldsymbol{z})=(\boldsymbol{x},\boldsymbol{z})+(\boldsymbol{y},\boldsymbol{z})$;
4. $(\boldsymbol{x},\boldsymbol{x})\geqslant 0$, 当且仅当 $\boldsymbol{x}=\boldsymbol{0}$ 时, $(\boldsymbol{x},\boldsymbol{x})=0$.

则该实数 $(\boldsymbol{x},\boldsymbol{y})$ 称为向量 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 的内积. 其中 $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{z}\in V$, $\lambda\in\mathbb{R}$.

定义了内积的实线性空间 V 叫做欧几里德空间, 简称欧氏空间或实内积空间
.

欧氏空间与实线性空间的差别, 在于欧氏空间比实线性空间多定义了内积. 或者说, 欧氏空间是一个特殊的实线性空间. 2.88

Example 64. 在 n 维向量空间 $\mathbb{R}^n$ 中, 对于 $\boldsymbol{x}=(a_1,a_2,\cdots,a_n)^{\mathrm{T}}$, $\boldsymbol{y}=(b_1,b_2,\cdots,b_n)^{\mathrm{T}}$, 我们规定

$$(\boldsymbol{x},\boldsymbol{y})=a_1b_1+a_2b_2+\cdots+a_nb_n=\sum_{k=1}^{n}a_kb_k. \tag{6}$$

不难验证, 这样确定的实数满足内积的 4 个条件, 所以 (6) 是 $\mathbb{R}^n$ 中的内积, 从而 $\mathbb{R}^n$ 关于上述内积构成 n 维欧氏空间.

上述内积 (6) 称为 $\mathbb{R}^n$ 上的标准内积. 2.89

在同一个实线性空间中, 定义不同的内积, 将构成不同的欧氏空间.

Example 65. 在空间 $\mathbb{R}^n$ 中, 规定

$$(\boldsymbol{x},\boldsymbol{y}) = a_1b_1 + 2a_2b_2 + \cdots + na_nb_n = \sum_{k=1}^{n} k a_k b_k.$$

这样确定的实数也满足内积的 4 个条件, 从而 $\mathbb{R}^n$ 关于上述内积也构成欧氏空间. 2.90

Example 66. 在线性空间 $\mathbb{R}^2$ 中, 对于向量对于 $\boldsymbol{x} = (a_1, a_2)^{\mathrm{T}}$, $\boldsymbol{y} = (b_1, b_2)^{\mathrm{T}}$, 我们规定

$$(\boldsymbol{x},\boldsymbol{y}) = a_1b_1 - a_2b_1 - a_1b_2 + 4a_2b_2, \tag{7}$$

这样确定的实数满足内积的 4 个条件, 所以 (7) 是 $\mathbb{R}^2$ 中的内积. 故 $\mathbb{R}^2$ 在上述内积定义下构成一个新的欧氏空间.

问: $(\boldsymbol{x},\boldsymbol{y}) = a_1 + a_2 + b_1 + b_2$ 是否构成内积?

不是内积, 非负性不满足. 2.91

Exercise 67 (P.56 习题 (一) 7)**.** 考虑线性空间 $C[a,b]$, 任取 $f(x), g(x) \in C[a,b]$, 定义函数

$$(f,g) = \int_a^b f(x)g(x)\,\mathrm{d}x,$$

试证明函数为 $C[a,b]$ 上的内积.

解: 对于连续函数 $f(x)$, $g(x)$, 积分 $\int_a^b f(x)g(x)\,\mathrm{d}x$ 是唯一确定的实数. 并满足:

1. $(f,g) = \int_a^b f(x)g(x)\,\mathrm{d}x = \int_a^b g(x)f(x)\,\mathrm{d}x = (g,f)$;
2. $(\lambda f,g) = \int_a^b \lambda f(x)g(x)\,\mathrm{d}x = \lambda\int_a^b f(x)g(x)\,\mathrm{d}x = \lambda(f,g)$;
3. $(f+g,h) = \int_a^b \big(f(x)+g(x)\big)h(x)\,\mathrm{d}x = \int_a^b f(x)h(x)\,\mathrm{d}x + \int_a^b g(x)h(x)\,\mathrm{d}x = (f,h)+(g,h)$;
4. $(f,f) = \int_a^b \big[f(x)\big]^2\,\mathrm{d}x \geqslant 0$, 且 $(f,f)=0$ 当且仅当 $f(x)\equiv 0$.

得证函数 $(f,g) = \int_a^b f(x)g(x)\,\mathrm{d}x$ 为 $C[a,b]$ 上的内积 □ 2.92

复向量的内积

对于 n 维实向量 $\boldsymbol{x} = (a_1, a_2, \cdots, a_n)^{\mathrm{T}}$, $\boldsymbol{y} = (b_1, b_2, \cdots, b_n)^{\mathrm{T}}$, $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 的内积为

$$(\boldsymbol{x},\boldsymbol{y}) = \sum_{k=1}^{n} a_k b_k.$$

直接以上式作为复向量的内积将导致不合理的结果. 例如复向量 $\boldsymbol{x} = (3, 4, 5\mathrm{i})^{\mathrm{T}}$,

$$(\boldsymbol{x},\boldsymbol{x}) = 3^2 + 4^2 + (5\mathrm{i})^2 = 0.$$

☞ 而对实向量, 当且仅当 $\boldsymbol{x} = \boldsymbol{0}$ 时, $(\boldsymbol{x},\boldsymbol{x}) = 0$. 2.93

Definition 68. 设 $\boldsymbol{x} = (a_1, a_2, \cdots, a_n)^{\mathrm{T}} \in \mathbb{C}^n$, $\boldsymbol{y} = (b_1, b_2, \cdots, b_n)^{\mathrm{T}} \in \mathbb{C}^n$, 令

$$\begin{aligned}(\boldsymbol{x},\boldsymbol{y}) &= a_1\overline{b_1} + a_2\overline{b_2} + \cdots + a_n\overline{b_n} \\ &= \sum_{k=1}^{n} a_k\overline{b_k} = \boldsymbol{y}^{\mathrm{H}}\boldsymbol{x},\end{aligned}$$

称 $(\boldsymbol{x},\boldsymbol{y})$ 为向量 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 的内积.

该内积称为复线性空间 $\mathbb{C}^n$ 的标准内积.

比如 $\boldsymbol{x} = (3, 4, 5\mathrm{i})^{\mathrm{T}}$, $\boldsymbol{y} = (4\mathrm{i}, -3, 1 + 2\mathrm{i})^{\mathrm{T}}$,

$$(\boldsymbol{x}, \boldsymbol{x}) = 3^2 + 4^2 + (5\mathrm{i}) \times (-5\mathrm{i}) = 50.$$

$$(\boldsymbol{x}, \boldsymbol{y}) = 3 \times (-4\mathrm{i}) + 4 \times (-3) + (5\mathrm{i}) \times (1 - 2\mathrm{i}) = -2 - 7\mathrm{i}.$$

$$(\boldsymbol{y}, \boldsymbol{x}) = 4\mathrm{i} \times 3 + (-3) \times 4 + (1 + 2\mathrm{i}) \times (-5\mathrm{i}) = -2 + 7\mathrm{i}.$$

☞ $(\boldsymbol{x}, \boldsymbol{y}) \neq (\boldsymbol{y}, \boldsymbol{x})$. 2.94

复向量的内积具有以下性质:

Theorem 69. *设 $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{z} \in \mathbb{C}^n$, $\lambda \in \mathbb{C}$, 则*

1. $(\boldsymbol{x}, \boldsymbol{y}) = \overline{(\boldsymbol{y}, \boldsymbol{x})}$;
2. $(\lambda \boldsymbol{x}, \boldsymbol{y}) = \lambda(\boldsymbol{x}, \boldsymbol{y})$, $(\boldsymbol{x}, \lambda \boldsymbol{y}) = \overline{\lambda}(\boldsymbol{x}, \boldsymbol{y})$;
3. $(\boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z}) = (\boldsymbol{x}, \boldsymbol{z}) + (\boldsymbol{y}, \boldsymbol{z})$;
4. $(\boldsymbol{x}, \boldsymbol{x}) \geqslant 0$, 当且仅当 $\boldsymbol{x} = \boldsymbol{0}$ 时, $(\boldsymbol{x}, \boldsymbol{x}) = 0$.

例如:

$$\begin{aligned}(\boldsymbol{x}, \lambda \boldsymbol{y}) &= (\lambda \boldsymbol{y})^{\mathrm{H}} \boldsymbol{x} \\ &= \overline{\lambda} \boldsymbol{y}^{\mathrm{H}} \boldsymbol{x} \\ &= \overline{\lambda}(\boldsymbol{x}, \boldsymbol{y}).\end{aligned}$$

2.95

Definition 70. 设 V 是复数域 $\mathbb{C}$ 上的线性空间. $\forall \boldsymbol{x}$, $\boldsymbol{y} \in V$, 若能给定某种规则使 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 对应于一个复数 $(\boldsymbol{x}, \boldsymbol{y})$, 它满足下列条件:

1. $(\boldsymbol{x}, \boldsymbol{y}) = \overline{(\boldsymbol{y}, \boldsymbol{x})}$;
2. $(\lambda \boldsymbol{x}, \boldsymbol{y}) = \lambda(\boldsymbol{x}, \boldsymbol{y})$;
3. $(\boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z}) = (\boldsymbol{x}, \boldsymbol{z}) + (\boldsymbol{y}, \boldsymbol{z})$;
4. $(\boldsymbol{x}, \boldsymbol{x}) \geqslant 0$, 当且仅当 $\boldsymbol{x} = \boldsymbol{0}$ 时, $(\boldsymbol{x}, \boldsymbol{x}) = 0$.

则该复数 $(\boldsymbol{x}, \boldsymbol{y})$ 称为向量 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 的内积. 其中 $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{z} \in V$, $\lambda \in \mathbb{C}$.

定义了内积的复线性空间 V 叫做酉空间(或 U 空间, 或复内积空间). 2.96

内积的一般化定义

Definition 71. 设 V 是数域 F 上的线性空间. 所谓 V 上的一个内积, 是定义在 V 上并取值在 F 上的二元函数 $(\boldsymbol{x} \mid \boldsymbol{y})$, 其中 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 为 V 中的任意向量, 且对一切 $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{z} \in V$ 及一切 $\lambda \in F$ 满足

1. 正定性: $(\boldsymbol{x} \mid \boldsymbol{x}) \geqslant 0$, 当且仅当 $\boldsymbol{x} = \boldsymbol{0}$ 时, $(\boldsymbol{x} \mid \boldsymbol{x}) = 0$.
2. 共轭对称性: $(\boldsymbol{x} \mid \boldsymbol{y}) = \overline{(\boldsymbol{y} \mid \boldsymbol{x})}$, 其中 $\overline{(\boldsymbol{y} \mid \boldsymbol{x})}$ 表示 $(\boldsymbol{x} \mid \boldsymbol{y})$ 的共轭复数.
3. 关于第一变元的线性性: $(\lambda \boldsymbol{x} + \boldsymbol{y} \mid \boldsymbol{z}) = \lambda(\boldsymbol{x} \mid \boldsymbol{z}) + (\boldsymbol{y} \mid \boldsymbol{z})$. 2.97

注 1. 对 $(\lambda \boldsymbol{x} + \boldsymbol{y} \mid \boldsymbol{z}) = \lambda(\boldsymbol{x} \mid \boldsymbol{z}) + (\boldsymbol{y} \mid \boldsymbol{z})$, *1.* $\lambda = 1$ 时, 得

$$(\boldsymbol{x}+\boldsymbol{y} \mid \boldsymbol{z}) = (\boldsymbol{x} \mid \boldsymbol{z}) + (\boldsymbol{y} \mid \boldsymbol{z}).$$

2. 上式中 $\boldsymbol{y} = \mathbf{0}$ 时, 得

$$(\boldsymbol{x} \mid \boldsymbol{z}) = (\boldsymbol{x}+\mathbf{0} \mid \boldsymbol{z}) = (\boldsymbol{x} \mid \boldsymbol{z}) + (\mathbf{0} \mid \boldsymbol{z}).$$

故

$$(\mathbf{0} \mid \boldsymbol{z}) = 0.$$

进而

$$(\boldsymbol{z} \mid \mathbf{0}) = \overline{(\mathbf{0} \mid \boldsymbol{z})} = \overline{0} = 0.$$

3. $(\lambda\boldsymbol{x} \mid \boldsymbol{z}) = (\lambda\boldsymbol{x}+\mathbf{0} \mid \boldsymbol{z}) = \lambda(\boldsymbol{x} \mid \boldsymbol{z}) + (\mathbf{0} \mid \boldsymbol{z}) = \lambda(\boldsymbol{x} \mid \boldsymbol{z})$.

4. $(\boldsymbol{x} \mid \mu\boldsymbol{y}) = \overline{(\mu\boldsymbol{y} \mid \boldsymbol{x})} = \overline{\mu(\boldsymbol{y} \mid \boldsymbol{x})} = \overline{\mu}\,\overline{(\boldsymbol{y} \mid \boldsymbol{x})} = \overline{\mu}\,(\boldsymbol{x} \mid \boldsymbol{y})$.

2.98

内积的性质

1. $\left(\sum\limits_{i=1}^{n} \lambda_i \boldsymbol{x}_i \mid \boldsymbol{y}\right) = \sum\limits_{i=1}^{n} \lambda_i \left(\boldsymbol{x}_i \mid \boldsymbol{y}\right)$.
2. $\left(\boldsymbol{x} \mid \sum\limits_{j=1}^{m} \mu_j \boldsymbol{y}_j\right) = \sum\limits_{j=1}^{m} \overline{\mu_j} \left(\boldsymbol{x} \mid \boldsymbol{y}_j\right)$.
3. $\left(\sum\limits_{i=1}^{n} \lambda_i \boldsymbol{x}_i \mid \sum\limits_{j=1}^{m} \mu_j \boldsymbol{y}_j\right) = \sum\limits_{i=1}^{n} \sum\limits_{j=m}^{n} \lambda_i \overline{\mu_j} \left(\boldsymbol{x}_i \mid \boldsymbol{y}_j\right)$.
4. $(\boldsymbol{x} \mid \mathbf{0}) = (\mathbf{0} \mid \boldsymbol{x}) = 0$.

2.99

Definition 72 (内积空间)**.** 数域 F 上线性空间 V 如果定义了内积 $(\cdot \mid \cdot)$, 则称之为内积空间, 记为 $\big(V, (\cdot \mid \cdot)\big)$, 也简记为 V.

2.100

Exercise 73 (P.55, 习题 (一) 4)**.** 给定某 n 阶正定矩阵$\boldsymbol{A} = [a_{ij}]$, $a_{ij} \in \mathbb{R}$, 而 $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \cdots, \alpha_n)^{\mathrm{T}} \in \mathbb{R}^n$, $\boldsymbol{\beta} = (\beta_1, \beta_2, \cdots, \beta_n)^{\mathrm{T}} \in \mathbb{R}^n$, 定义函数 $(\boldsymbol{\alpha} \mid \boldsymbol{\beta}) = \boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha}$, 证明此函数是 $\mathbb{R}^n$ 上的一个内积.

☞ 对称矩阵 $\boldsymbol{A}$ 为正定的 $\iff$ 二次型 $f(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 为正定的 $\iff$ $\boldsymbol{A}$ 的特征值全为正 $\iff$ $\boldsymbol{A}$ 的各阶主子式都为正.

Proof. (1) 正定性. 当 $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \cdots, \alpha_n)^{\mathrm{T}} \neq \mathbf{0}$ 时, 有

$$(\boldsymbol{\alpha} \mid \boldsymbol{\alpha}) = \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha} = f(\alpha_1, \alpha_2, \cdots, \alpha_n),$$

因 $\boldsymbol{A}$ 为正定矩阵, 故二次型 $f(\alpha_1, \alpha_2, \cdots, \alpha_n)$ 为正定的. 由于 $\boldsymbol{\alpha} \neq \mathbf{0}$, 即 α_1, α_2, $\cdots$, α_n 是不全为零的实数, 故

$$(\boldsymbol{\alpha} \mid \boldsymbol{\alpha}) = f(\alpha_1, \alpha_2, \cdots, \alpha_n) > 0.$$

当且仅当 $\boldsymbol{\alpha} = \mathbf{0}$ 时, 有 $(\boldsymbol{\alpha} \mid \boldsymbol{\alpha}) = f(\alpha_1, \alpha_2, \cdots, \alpha_n) = 0$.

(2) 对称性. 正定矩阵 $\boldsymbol{A}$ 是对称的, 又注意到 $\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha}$ 为实数. 故

$$(\boldsymbol{\alpha} \mid \boldsymbol{\beta}) = \boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha} = \left(\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha}\right)^{\mathrm{T}} = \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\beta} = \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\beta} = (\boldsymbol{\beta} \mid \boldsymbol{\alpha}).$$

(3) 关于第一变元的线性性.

$$(\lambda\boldsymbol{\alpha}+\boldsymbol{\beta}\mid\gamma)=\gamma^{\mathrm{T}}\boldsymbol{A}(\lambda\boldsymbol{\alpha}+\boldsymbol{\beta})=\lambda\gamma^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha}+\gamma^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\beta}=\lambda(\boldsymbol{\alpha}\mid\gamma)+(\boldsymbol{\beta}\mid\gamma).$$

得证函数 $(\boldsymbol{\alpha}\mid\boldsymbol{\beta})=\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha}$ 是 $\mathbb{R}^n$ 上的一个内积. □

2.101

Exercise 74 (P.56 习题 (一) 10)**.** 给定 $C[-\pi,\pi]$ 中的函数 $\sin x$, 并取第 7 题中所定义的内积, 试求一 $C[-\pi,\pi]$ 中的函数 $f(x)$, 使其满足 $\big(\sin x\mid f(x)\big)=0$.

解: 因

$$\big(\sin x\mid f(x)\big)=\int_{-\pi}^{\pi}\sin x f(x)\,\mathrm{d}x=0,$$

故可以取 $f(x)$ 为定义在 $[-\pi,\pi]$ 上的任意偶函数. □

☞ 以下函数

$$1,\cos x,\sin x,\cos 2x,\sin 2x,\cdots,\cos nx,\sin nx,\cdots$$

在 $C[-\pi,\pi]$ 内是两两正交的.

2.102

Exercise 75 (P.56, 习题 (二) 2)**.** 设 V 是一个内积空间, 若对一切 $\boldsymbol{x}\in V$ 都有

$$(\boldsymbol{x}\mid\boldsymbol{u})=(\boldsymbol{x}\mid\boldsymbol{v}),$$

试证 $\boldsymbol{u}=\boldsymbol{v}$.

Proof. 即对一切 $\boldsymbol{x}\in V$ 都有

$$(\boldsymbol{u}\mid\boldsymbol{x})=(\boldsymbol{v}\mid\boldsymbol{x}),$$

故

$$(\boldsymbol{u}-\boldsymbol{v}\mid\boldsymbol{x})=(\boldsymbol{u}\mid\boldsymbol{x})+(-\boldsymbol{v}\mid\boldsymbol{x})=(\boldsymbol{u}\mid\boldsymbol{x})-(\boldsymbol{v}\mid\boldsymbol{x})=0.$$

取 $\boldsymbol{x}=\boldsymbol{u}-\boldsymbol{v}$, 得

$$(\boldsymbol{u}-\boldsymbol{v}\mid\boldsymbol{u}-\boldsymbol{v})=0.$$

故 $\boldsymbol{u}-\boldsymbol{v}=\mathbf{0}$, 从而 $\boldsymbol{u}=\boldsymbol{v}$. □

2.103

Exercise 76 (P.56, 习题 (二) 4)**.** 已知 E 是欧氏空间, $\dim E=n$. 设 $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, $\cdots$, $\boldsymbol{x}_k$, $\boldsymbol{y}\in E$, 且 $\boldsymbol{y}$ 可用 $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, $\cdots$, $\boldsymbol{x}_k$ 线性表示. 试证: 当且仅当 $\boldsymbol{y}=\mathbf{0}$ 时有 $(\boldsymbol{x}_i\mid\boldsymbol{y})=0$ 成立, $i=1,2,\cdots,k$.

Proof. (1) 若 $(\boldsymbol{x}_i\mid\boldsymbol{y})=0$. 设 $\boldsymbol{y}=\lambda_1\boldsymbol{x}_1+\lambda_2\boldsymbol{x}_2+\cdots+\lambda_k\boldsymbol{x}_k$, 则

$$(\boldsymbol{y}\mid\boldsymbol{y})=\Big(\sum_{i=1}^{k}\lambda_i\boldsymbol{x}_i\mid\boldsymbol{y}\Big)=\sum_{i=1}^{k}\lambda_i\big(\boldsymbol{x}_i\mid\boldsymbol{y}\big)=0.$$

(2) 当 $\boldsymbol{y}=\mathbf{0}$ 时, 显然有 $(\boldsymbol{x}_i\mid\boldsymbol{y})=0$ 成立. □

2.104

本节内容强调

- 内积的公理化定义是抽象于经典内积的性质;
- 一个线性空间可以定义多种内积;
- 同一个线性空间定义不同的内积, 将得到不同的内积空间.

2.105

2.2 由内积诱导出的几何概念

Definition 77. 设 $\big(V,(\cdot\mid\cdot)\big)$ 为内积空间, $\boldsymbol{x}\in V$, 则称

$$\|\boldsymbol{x}\|=\sqrt{(\boldsymbol{x}\mid\boldsymbol{x})}$$

为 $\boldsymbol{x}$ 的由内积 $(\cdot\mid\cdot)$ 诱导的范数(norm). 又若 $\boldsymbol{x}$, $\boldsymbol{y}\in V$, 则称

$$\|\boldsymbol{x}-\boldsymbol{y}\|$$

为向量 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 的距离.

2.106

Theorem 78. *设 $\big(V,(\cdot\mid\cdot)\big)$ 为内积空间, F 为 V 的标量域, 则对任意 $\boldsymbol{x}$, $\boldsymbol{y}\in V$ 和任意 $\lambda\in F$ 有:*

1. $\|\lambda\boldsymbol{x}\|=|\lambda|\|\boldsymbol{x}\|$;
2. $|(\boldsymbol{x}\mid\boldsymbol{y})|\leqslant\|\boldsymbol{x}\|\|\boldsymbol{y}\|$ *(Cauchy–Bunyakovskii–Schwarz 不等式);*
3. $\|\boldsymbol{x}+\boldsymbol{y}\|\leqslant\|\boldsymbol{x}\|+\|\boldsymbol{y}\|$ *(三角不等式).*

2.107

Definition 79. 设 V 为 n 维欧氏空间, $\boldsymbol{x}$, $\boldsymbol{y}$ 为 V 中的非零向量, 则称实数

$$\theta=\arccos\frac{(\boldsymbol{x}\mid\boldsymbol{y})}{\|\boldsymbol{x}\|\|\boldsymbol{y}\|}$$

为向量 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 的交角.

☞ 上式只能定义欧氏空间的向量交角, 而对一般的酉空间, 上式无意义. 2.108

Definition 80. 设 $\boldsymbol{x}$, $\boldsymbol{y}$ 为内积空间 V 中的向量, 若

$$(\boldsymbol{x}\mid\boldsymbol{y})=0,$$

则称向量 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 正交(orthogonal). 记为 $\boldsymbol{x}\perp\boldsymbol{y}$.

- For $\mathbb{R}^n$ with the standard inner product, $\boldsymbol{x}\perp\boldsymbol{y}\Leftrightarrow\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y}=0$.
- For $\mathbb{C}^n$ with the standard inner product, $\boldsymbol{x}\perp\boldsymbol{y}\Leftrightarrow\boldsymbol{y}^{\mathrm{H}}\boldsymbol{x}=0$.

2.109

Example 81. 设 $V=C[-1,1]$, 计算交角 $\langle 1,x^2\rangle$. (内积按通常的定义: $\forall f(x)$, $g(x)\in C[a,b]$, $\big(f(x)\mid g(x)\big)=\displaystyle\int_a^b f(x)g(x)\,\mathrm{d}x$.)

解: 由 $\cos\langle 1, x^2\rangle = \dfrac{(1 \mid x^2)}{\|1\|\,\|x^2\|}$. 其中

$$(1 \mid x^2) = \int_{-1}^{1} x^2\,\mathrm{d}x = \frac{2}{3},$$
$$\|1\|^2 = (1 \mid 1) = \int_{-1}^{1} \mathrm{d}x = 2,$$
$$\|x^2\|^2 = (x^2 \mid x^2) = \int_{-1}^{1} x^4\,\mathrm{d}x = \frac{2}{5}.$$

故

$$\langle 1, x^2\rangle = \arccos\frac{(1 \mid x^2)}{\|1\|\,\|x^2\|} = \arccos\frac{\sqrt{5}}{3}. \quad \square$$

☞ 试指出一个与 x^2 正交的向量? 2.110

Definition 82. 若内积空间 V 中的不含零向量的向量组 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k$ 两两互相正交, 则称之为一个正交向量组.

Theorem 83. *设 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k$ 为内积空间 V 的一组正交向量, 且 $\boldsymbol{x} = \sum_{i=1}^{k} \boldsymbol{x}_i$, 则*

$$\|\boldsymbol{x}\|^2 = \sum_{i=1}^{k} \|\boldsymbol{x}_i\|^2.$$

特别地, 若有 $\boldsymbol{y} \perp \boldsymbol{z}$, 则

$$\|\boldsymbol{y} + \boldsymbol{z}\|^2 = \|\boldsymbol{y}\|^2 + \|\boldsymbol{z}\|^2.$$

2.111

Exercise 84 (P.64 习题 (一) 6)**.** 在欧氏空间 E^n 中, 试证: $\boldsymbol{x} \perp \boldsymbol{y}$ 的充分必要条件为 $\|\boldsymbol{x} + \boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2$; 而在酉空间中, 由 $\boldsymbol{x} \perp \boldsymbol{y}$ 可以推得 $\|\boldsymbol{x} + \boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2$, 但反过来却未必成立, 请举一反例.

证: 因为

$$\begin{aligned}\|\boldsymbol{x} + \boldsymbol{y}\|^2 &= (\boldsymbol{x} + \boldsymbol{y} \mid \boldsymbol{x} + \boldsymbol{y}) = (\boldsymbol{x} \mid \boldsymbol{x}) + (\boldsymbol{x} \mid \boldsymbol{y}) + (\boldsymbol{y} \mid \boldsymbol{x}) + (\boldsymbol{y} \mid \boldsymbol{y}) \\ &= \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2 + (\boldsymbol{x} \mid \boldsymbol{y}) + (\boldsymbol{y} \mid \boldsymbol{x}),\end{aligned}$$

(1) 在欧氏空间中, $(\boldsymbol{x} \mid \boldsymbol{y}) = (\boldsymbol{y} \mid \boldsymbol{x})$, 故 $\|\boldsymbol{x} + \boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2 + 2(\boldsymbol{x} \mid \boldsymbol{y})$. 从而

$$\boldsymbol{x} \perp \boldsymbol{y} \Leftrightarrow (\boldsymbol{x} \mid \boldsymbol{y}) = 0 \Leftrightarrow \|\boldsymbol{x} + \boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2.$$

(2) 在酉空间中, $\overline{(\boldsymbol{x} \mid \boldsymbol{y})} = (\boldsymbol{y} \mid \boldsymbol{x})$, 故 $\|\boldsymbol{x} + \boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2 + 2\mathrm{Re}(\boldsymbol{x} \mid \boldsymbol{y})$. 从而

$$\boldsymbol{x} \perp \boldsymbol{y} \Leftrightarrow (\boldsymbol{x} \mid \boldsymbol{y}) = 0 \Rightarrow \mathrm{Re}(\boldsymbol{x} \mid \boldsymbol{y}) = 0 \Leftrightarrow \|\boldsymbol{x} + \boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2.$$

其中 $\mathrm{Re}(\boldsymbol{x} \mid \boldsymbol{y}) = 0$ 不能推得 $(\boldsymbol{x} \mid \boldsymbol{y}) = 0$, 因为 $(\boldsymbol{x} \mid \boldsymbol{y})$ 还可能是纯虚数. 例如 $\boldsymbol{x} = (0, \mathrm{i})^{\mathrm{T}}$, $\boldsymbol{y} = (0, 1)^{\mathrm{T}}$. 因 $(\boldsymbol{x} \mid \boldsymbol{y}) = \mathrm{i} \neq 0$,

故两者不正交, 且满足 $\|\boldsymbol{x} + \boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2$. $\square$ 2.112

2.3 标准正交基底与 Gram-Schmidt 过程

Theorem 85. *内积空间 V 中任意正交向量组 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k$ $(k \leqslant \dim V)$ 都是 V 中线性无关的向量组.*

Proof. 设 $\lambda_1\boldsymbol{x}_1 + \lambda_2\boldsymbol{x}_2 + \cdots + \lambda_k\boldsymbol{x}_k = \boldsymbol{0}$, 则

$$\begin{aligned}0 &= (\boldsymbol{0} \mid \boldsymbol{x}_i) = (\lambda_1\boldsymbol{x}_1 + \lambda_2\boldsymbol{x}_2 + \cdots + \lambda_k\boldsymbol{x}_k \mid \boldsymbol{x}_i) \\ &= \lambda_1(\boldsymbol{x}_1 \mid \boldsymbol{x}_i) + \cdots + \lambda_i(\boldsymbol{x}_i \mid \boldsymbol{x}_i) + \cdots + \lambda_k(\boldsymbol{x}_k \mid \boldsymbol{x}_i) = \lambda_i\|\boldsymbol{x}_i\|^2,\end{aligned}$$

得 $\lambda_i = 0,\ i = 1, 2, \cdots, k$. 故 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k$ 线性无关. □

2.113

Definition 86. 设有 n 维内积空间 V 的 k 个向量 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k$. 若 $(\boldsymbol{x}_i \mid \boldsymbol{x}_j) = \delta_{ij},\ i, j = 1, 2, \cdots, k$, 其中

$$\delta_{ij} = \begin{cases} 1, & 若\ i = j, \\ 0, & 若\ i \neq j. \end{cases}$$

则称 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k$ 为标准正交向量组.

显然, 当 $k = n$ 时, 向量组 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n$ 成为 V 的一个基底, 称为标准正交基底.

δ_{ij} 称为克罗内克符号(Kronecker Delta). 克罗内克 (Leopold Kronecker, 1823–1891), 德国数学家.

这是个常用的符号, 例如单位矩阵可以表示为

$$\boldsymbol{I} = [\delta_{ij}].$$

2.114

Gram-Schmidt 正交化过程

Theorem 87. *每个 n 维内积空间 V 一定存在标准正交基底.*

证: 设向量组 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n$ 是 V 的一个基底, 下面逐个求出标准正交向量 $\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_n$, 并满足:

$$\operatorname{span}[\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_i] = \operatorname{span}[\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_i], \quad i = 1, 2, \cdots, n.$$

首先, 可取 $\boldsymbol{y}_1 = \frac{1}{\|\boldsymbol{x}_1\|}\boldsymbol{x}_1$. 一般地, 假定已经求出 $\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_m$, 它们是单位正交的, 具有性质

$$\operatorname{span}[\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_i] = \operatorname{span}[\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_i], \quad i = 1, 2, \cdots, m.$$

下求 $\boldsymbol{y}_{m+1}$. 2.115

因为 $\operatorname{span}[\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_i] = \operatorname{span}[\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_i]$, 所以 $\boldsymbol{x}_{m+1}$ 不能被 $\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_m$ 线性表示, 作向量

$$\boldsymbol{u}_{m+1} = \boldsymbol{x}_{m+1} - (\lambda_1\boldsymbol{y}_1 + \lambda_2\boldsymbol{y}_2 + \cdots + \lambda_m\boldsymbol{y}_m) \neq \boldsymbol{0}, \tag{8}$$

这里 $\lambda_1, \lambda_2, \cdots, \lambda_m$ 是待定的系数, 用 $\boldsymbol{y}_i$ 与 $\boldsymbol{u}_{m+1}$ 作内积得

$$(\boldsymbol{u}_{m+1} \mid \boldsymbol{y}_i) = (\boldsymbol{x}_{m+1} \mid \boldsymbol{y}_i) - \lambda_i(\boldsymbol{y}_i \mid \boldsymbol{y}_i),$$

取

$$\lambda_i = (\boldsymbol{x}_{m+1} \mid \boldsymbol{y}_i), \tag{9}$$

有 $(\boldsymbol{u}_{m+1} \mid \boldsymbol{y}_i) = 0$. 令

$$\boldsymbol{y}_{m+1} = \frac{\boldsymbol{u}_{m+1}}{\|\boldsymbol{u}_{m+1}\|}, \tag{10}$$

$\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_m, \boldsymbol{y}_{m+1}$ 就是一个标准正交向量组, 同时

$$\mathrm{span}[\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_{m+1}] = \mathrm{span}[\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_{m+1}].$$

由归纳法原理, 定理得证. □ 2.116

Gram-Schmidt 正交化过程的步骤

1. $\boldsymbol{u}_1 = \boldsymbol{x}_1$,
   $\boldsymbol{y}_1 = \frac{1}{\|\boldsymbol{u}_1\|}\boldsymbol{u}_1$;
2. $\boldsymbol{u}_{m+1} = \boldsymbol{x}_{m+1} - (\boldsymbol{x}_{m+1} \mid \boldsymbol{y}_1)\boldsymbol{y}_1 - (\boldsymbol{x}_{m+1} \mid \boldsymbol{y}_2)\boldsymbol{y}_2 - \cdots - (\boldsymbol{x}_{m+1} \mid \boldsymbol{y}_m)\boldsymbol{y}_m$,
   $\boldsymbol{y}_{m+1} = \frac{1}{\|\boldsymbol{u}_{m+1}\|}\boldsymbol{u}_{m+1}$.

上述从线性无关向量组 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n$ 导出标准正交向量组 $\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_n$ 的过程, 称为 Gram-Schmidt 正交化过程(Gram–Schmidt Orthogonalization Procedure). 2.117

Jørgen P. Gram

Jørgen P. Gram (1850–1916) was a Danish actuary who implicitly presented the essence of orthogonalization procedure in 1883. Gram was apparently unaware that Pierre-Simon Laplace (1749–1827) had earlier used the method. Today, Gram is remembered primarily for his development of this process, but in earlier times his name was also associated with the matrix product $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ that historically was referred to as the *Gram matrix* of $\boldsymbol{A}$.

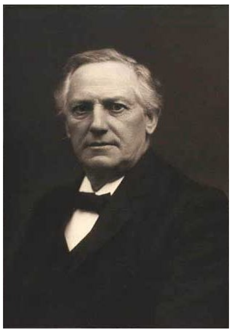


Figure 1: Jørgen P. Gram

2.118

Erhard Schmidt

Erhard Schmidt (1876–1959) was a student of Hermann Schwarz and the great German mathematician David Hilbert. Schmidt explicitly employed the orthogonalization process in 1907 in his study of integral equations, which in turn led to the development of what are now called *Hilbert spaces*. Schmidt made significant use of the orthogonalization process to develop the geometry of Hilbert Spaces, and thus it came to bear Schmidt's name.

Figure 2: Erhard Schmidt

2.119

Example 88. 在具有标准内积的空间 $\mathbb{C}^3$ 中, 设 $\boldsymbol{x}_1 = (1, \mathrm{i}, 0)^{\mathrm{T}}$, $\boldsymbol{x}_2 = (1, 0, \mathrm{i})^{\mathrm{T}}$, $\boldsymbol{x}_3 = (0, 0, 1)^{\mathrm{T}}$, 试用 Gram-Schmidt 正交化方法把这组向量单位正交化.

解: $\boldsymbol{u}_1 = \boldsymbol{x}_1 = (1, \mathrm{i}, 0)^{\mathrm{T}}$, $\boldsymbol{y}_1 = \frac{1}{\|\boldsymbol{u}_1\|}\boldsymbol{u}_1 = \frac{1}{\sqrt{2}}(1, \mathrm{i}, 0)^{\mathrm{T}}$.

$$\boldsymbol{u}_2 = \boldsymbol{x}_2 - (\boldsymbol{x}_2 \mid \boldsymbol{y}_1)\boldsymbol{y}_1 = (1, 0, \mathrm{i})^{\mathrm{T}} - \frac{1}{\sqrt{2}}(1, -\mathrm{i}, 0)\begin{pmatrix} 1 \\ 0 \\ \mathrm{i} \end{pmatrix}\boldsymbol{y}_1 = (\frac{1}{2}, -\frac{\mathrm{i}}{2}, \mathrm{i})^{\mathrm{T}}.$$

$$\boldsymbol{y}_2 = \frac{1}{\|\boldsymbol{u}_2\|}\boldsymbol{u}_2 = \frac{1}{\sqrt{6}}(1, -\mathrm{i}, 2\mathrm{i})^{\mathrm{T}}.$$

$$\boldsymbol{u}_3 = \boldsymbol{x}_3 - (\boldsymbol{x}_3 \mid \boldsymbol{y}_1)\boldsymbol{y}_1 - (\boldsymbol{x}_3 \mid \boldsymbol{y}_2)\boldsymbol{y}_2 = \boldsymbol{x}_3 - 0\boldsymbol{y}_1 + \frac{2\mathrm{i}}{\sqrt{6}}\boldsymbol{y}_2 = \frac{1}{3}(\mathrm{i}, 1, 1)^{\mathrm{T}}.$$

$$\boldsymbol{y}_3 = \frac{1}{\|\boldsymbol{u}_3\|}\boldsymbol{u}_3 = \frac{1}{\sqrt{3}}(\mathrm{i}, 1, 1)^{\mathrm{T}}.$$

向量 $\boldsymbol{y}_1$, $\boldsymbol{y}_2$, $\boldsymbol{y}_3$ 即为所求. □ 2.120

Gram 矩阵

Definition 89. 设 $(V, (\cdot \mid \cdot))$ 为内积空间, $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, $\cdots$, $\boldsymbol{x}_k$ $(0 < k \leqslant \dim V)$ 为 V 中一组向量, 则称矩阵

$$\boldsymbol{G}(\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k) = \begin{bmatrix} (\boldsymbol{x}_1 \mid \boldsymbol{x}_1) & (\boldsymbol{x}_1 \mid \boldsymbol{x}_2) & \cdots & (\boldsymbol{x}_1 \mid \boldsymbol{x}_k) \\ (\boldsymbol{x}_2 \mid \boldsymbol{x}_1) & (\boldsymbol{x}_2 \mid \boldsymbol{x}_2) & \cdots & (\boldsymbol{x}_2 \mid \boldsymbol{x}_k) \\ \vdots & \vdots & & \vdots \\ (\boldsymbol{x}_k \mid \boldsymbol{x}_1) & (\boldsymbol{x}_k \mid \boldsymbol{x}_2) & \cdots & (\boldsymbol{x}_k \mid \boldsymbol{x}_k) \end{bmatrix}$$

为向量组 $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, $\cdots$, $\boldsymbol{x}_k$ 的格拉姆矩阵(Gram Matrix), 而

$$g(\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k) = \det \boldsymbol{G}(\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k)$$

称为格拉姆行列式.

2.121

格拉姆矩阵为 Hermite 矩阵, 即

$$\boldsymbol{G}^{\mathrm{H}}(\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k) = \boldsymbol{G}(\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k).$$

2.122

格拉姆矩阵可以方便地表达两个向量的内积.

设 $(V, (\cdot \mid \cdot))$ 为内积空间, $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, $\cdots$, $\boldsymbol{x}_n$ 为 V 的一组基底. 设向量 $\boldsymbol{x}$, $\boldsymbol{y}$ 在该组基底下的坐标分别为

$$\boldsymbol{u} = (\xi_1, \xi_2, \cdots, \xi_n)^{\mathrm{T}}, \qquad \boldsymbol{v} = (\eta_1, \eta_2, \cdots, \eta_n)^{\mathrm{T}}.$$

即 $\boldsymbol{x} = \xi_1\boldsymbol{x}_1 + \xi_2\boldsymbol{x}_2 + \cdots + \xi_n\boldsymbol{x}_n$, $\boldsymbol{y} = \eta_1\boldsymbol{x}_1 + \eta_2\boldsymbol{x}_2 + \cdots + \eta_n\boldsymbol{x}_n$. 则

$$\begin{aligned}
(\boldsymbol{x} \mid \boldsymbol{y}) &= \Big(\sum_{i=1}^{n} \xi_i\boldsymbol{x}_i \mid \sum_{j=1}^{n} \eta_j\boldsymbol{x}_j\Big) = \sum_{i=1}^{n}\sum_{j=1}^{n} \xi_i\,\overline{\eta_j}\,(\boldsymbol{x}_i \mid \boldsymbol{x}_j) \\
&= (\xi_1, \xi_2, \cdots, \xi_n)\begin{bmatrix} (\boldsymbol{x}_1 \mid \boldsymbol{x}_1) & (\boldsymbol{x}_1 \mid \boldsymbol{x}_2) & \cdots & (\boldsymbol{x}_1 \mid \boldsymbol{x}_k) \\ (\boldsymbol{x}_2 \mid \boldsymbol{x}_1) & (\boldsymbol{x}_2 \mid \boldsymbol{x}_2) & \cdots & (\boldsymbol{x}_2 \mid \boldsymbol{x}_k) \\ \vdots & \vdots & & \vdots \\ (\boldsymbol{x}_k \mid \boldsymbol{x}_1) & (\boldsymbol{x}_k \mid \boldsymbol{x}_2) & \cdots & (\boldsymbol{x}_k \mid \boldsymbol{x}_k) \end{bmatrix}\begin{pmatrix} \overline{\eta_1} \\ \overline{\eta_2} \\ \vdots \\ \overline{\eta_n} \end{pmatrix} \\
&= \boldsymbol{u}^{\mathrm{T}}\boldsymbol{G}(\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n)\overline{\boldsymbol{v}} \\
&= \boldsymbol{v}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{T}}(\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n)\boldsymbol{u}.
\end{aligned}$$

2.123

即

$$(\boldsymbol{x} \mid \boldsymbol{y}) = \boldsymbol{v}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{T}}(\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n)\boldsymbol{u}.$$

上式表明, 向量 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 的内积可以由 $\boldsymbol{x}$, $\boldsymbol{y}$ 在某组基底下的坐标和 Gram 矩阵来表示. 因而 Gram 矩阵完全确定了内积.

若 V 为实内积空间, 则

$$(\boldsymbol{x} \mid \boldsymbol{y}) = \boldsymbol{v}^{\mathrm{T}}\boldsymbol{G}(\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n)\boldsymbol{u}.$$

特别地, 当 $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, $\cdots$, $\boldsymbol{x}_n$ 为标准正交基时, $\boldsymbol{G}(\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n) = \boldsymbol{I}_n$, 从而

$$(\boldsymbol{x} \mid \boldsymbol{y}) = \boldsymbol{v}^{\mathrm{H}}\boldsymbol{u} = \sum_{i=1}^{n} \xi_i\,\overline{\eta_i}.$$

即当取内积空间的标准正交基底时, 向量 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 的内积等于它们的坐标向量 $\boldsymbol{u}$, $\boldsymbol{v}$ 的标准内积

$$(\boldsymbol{x} \mid \boldsymbol{y}) = \boldsymbol{v}^{\mathrm{H}}\boldsymbol{u}.$$

2.124

Exercise 90 (P.65 习题 (二) 1)**.** 设 $(V, (\cdot \mid \cdot))$ 为实内积空间, 若 $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, $\cdots$, $\boldsymbol{x}_n$ 为 V 的基底, 则 $\boldsymbol{G}(\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k)$ 为正定矩阵.

Proof. 设向量 $\boldsymbol{\alpha}$ 在该组基底下的坐标为

$$\boldsymbol{x} = (a_1, a_2, \cdots, a_n)^{\mathrm{T}},$$

则

$$(\boldsymbol{\alpha} \mid \boldsymbol{\alpha}) = \boldsymbol{x}^{\mathrm{T}} \boldsymbol{G} \boldsymbol{x},$$

对任意向量 $\boldsymbol{\alpha} \neq \mathbf{0}$, 有 $(\boldsymbol{\alpha} \mid \boldsymbol{\alpha}) > 0$. 从而有 $\boldsymbol{x} \neq \mathbf{0}$, 即 $\boldsymbol{x} \neq (0, 0, \cdots, 0)^{\mathrm{T}}$, 且二次型

$$\boldsymbol{x}^{\mathrm{T}} \boldsymbol{G} \boldsymbol{x} > 0.$$

故矩阵 $\boldsymbol{G}(\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k)$ 为正定矩阵. □

2.125

子空间的正交关系

Definition 91. 设 V 是一个内积空间, W 是 V 中向量集合. 给定 $\boldsymbol{x} \in V$, 若 $\forall \boldsymbol{y} \in W$, 都有

$$(\boldsymbol{x} \mid \boldsymbol{y}) = 0,$$

则称 $\boldsymbol{x}$ 与W 是正交的, 记为 $\boldsymbol{x} \perp W$, 或 $W \perp \boldsymbol{x}$.

设 W_1 也是 V 中向量集合, 若 $\forall \boldsymbol{y} \in W$, $\forall \boldsymbol{y}_1 \in W_1$, 都有

$$(\boldsymbol{y} \mid \boldsymbol{y}_1) = 0,$$

则称 W_1 与W 是正交的, 记为 $W_1 \perp W$, 或 $W \perp W_1$.

特别地, 若 W, W_1 是 V 中的子空间, 则上述转化为向量与子空间正交和子空间与子空间正交的定义.

2.126

Theorem 92. *设 $(V, (\cdot \mid \cdot))$ 为内积空间, W 为 V 的任一子集, 则集合*

$$S = \{\boldsymbol{y} \mid \boldsymbol{y} \perp W,\ \boldsymbol{y} \in V\}$$

为 V 的子空间.

Proof. 验证集合 S 中加法和数乘满足封闭性即可. 记 V 的标量域为 F, $\forall \lambda \in F$, $\forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in S$, 有

$$\boldsymbol{x}_1 \perp W, \qquad \boldsymbol{x}_2 \perp W,$$

则 $\forall \boldsymbol{z} \in W$, 有 $(\boldsymbol{x}_1 \mid \boldsymbol{z}) = 0$, $(\boldsymbol{x}_2 \mid \boldsymbol{z}) = 0$. 故

$$(\lambda \boldsymbol{x}_1 \mid \boldsymbol{z}) = \lambda (\boldsymbol{x}_1 \mid \boldsymbol{z}) = 0,$$
$$(\boldsymbol{x}_1 + \boldsymbol{x}_2 \mid \boldsymbol{z}) = (\boldsymbol{x}_1 \mid \boldsymbol{z}) + (\boldsymbol{x}_2 \mid \boldsymbol{z}) = 0.$$

即 $\lambda \boldsymbol{x}_1 \in S$, $\boldsymbol{x}_1 + \boldsymbol{x}_2 \in S$. 得证 S 为 V 的子空间. □

2.127

Corollary 93. *设 $(V, (\cdot \mid \cdot))$ 为内积空间, W 为 V 的任一子空间, 则集合*

$$S = \{\boldsymbol{y} \mid \boldsymbol{y} \perp W,\ \boldsymbol{y} \in V\}$$

为 V 的子空间, 并且 $\dim S = \dim V - \dim W$.

2.128

Definition 94. 设 $(V,(\cdot\mid\cdot))$ 为内积空间, W 为 V 的任一子空间, 则集合

$$S=\{\boldsymbol{y}\mid \boldsymbol{y}\perp W,\ \boldsymbol{y}\in V\}$$

为子空间 W 的正交补空间, 记为 $W^{\perp}$.

Corollary 95. 设 $(V,(\cdot\mid\cdot))$ 为内积空间, W 为 V 的任一子空间, 则 W 在 V 中一定有正交补空间 $W^{\perp}$ 使

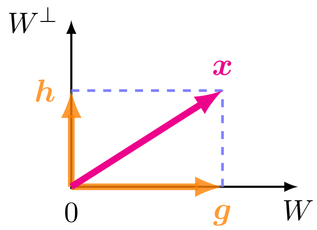
$$V=W\oplus W^{\perp}. \tag{11}$$

式 (11) 右端的表达式 $W\oplus W^{\perp}$ 称为空间 V 的正交直和分解式. 2.129

Example 96. 设 W 为内积空间 $(V,(\cdot\mid\cdot))$ 的子空间, $\forall \boldsymbol{x}\in V$, 试证: $\exists!\boldsymbol{g}\in W$, $\exists!\boldsymbol{h}\in W^{\perp}$, 使得 $\boldsymbol{x}=\boldsymbol{g}+\boldsymbol{h}$.

Proof. 由前述推论知 V 有正交直和分解 $V=W\oplus W^{\perp}$, 因此, $\forall \boldsymbol{x}\in V$, $\exists!\boldsymbol{g}\in W$, $\exists!\boldsymbol{h}\in W^{\perp}$, 使得 $\boldsymbol{x}=\boldsymbol{g}+\boldsymbol{h}$. □

☞ 称 $\boldsymbol{g}$ 为 $\boldsymbol{x}$ 的沿 $W^{\perp}$ 向 W 的正交投影, $\|\boldsymbol{h}\|$ 为点 $\boldsymbol{x}$ 到子空间 W 的距离.
称 $\boldsymbol{h}$ 为 $\boldsymbol{x}$ 的沿 W 向 $W^{\perp}$ 的正交投影, $\|\boldsymbol{g}\|$ 为点 $\boldsymbol{x}$ 到子空间 $W^{\perp}$ 的距离.



2.130

Exercise 97 (P.66 习题 (二) 9)**.** 设 $(V,(\cdot\mid\cdot))$ 为内积空间, $\{\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_n\}$ 为 V 的任一基底. 试证 $\{\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_n\}$ 为 V 的标准正交基底的充分必要条件为: 对 V 中任意两个向量 $\boldsymbol{x}=\sum\limits_{i=1}^{n}\xi_i\boldsymbol{y}_i$ 和 $\boldsymbol{y}=\sum\limits_{i=1}^{n}\eta_i\boldsymbol{y}_i$ 都有 $(\boldsymbol{x}\mid\boldsymbol{y})=\sum\limits_{i=1}^{n}\xi_i\overline{\eta_i}$.

证: 充分性. 设对任意 $\boldsymbol{x}$, $\boldsymbol{y}\in V$, 有

$$(\boldsymbol{x}\mid\boldsymbol{y})=\xi_1\overline{\eta_1}+\xi_2\overline{\eta_2}+\cdots+\xi_n\overline{\eta_n},$$

因为

$$\boldsymbol{y}_i=0\boldsymbol{y}_1+\cdots+1\boldsymbol{y}_i+\cdots+0\boldsymbol{y}_n,$$

故

$$(\boldsymbol{y}_i\mid\boldsymbol{y}_j)=\begin{cases}1, & 若\ i=j,\\ 0, & 若\ i\neq j.\end{cases}$$

故 $\{\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_n\}$ 为 V 的标准正交基底. 必要性是显然的 (或见 P.61 例 5). □ 2.131

3 线性变换

3.1 映射和线性变换

Definition 98. 设有映射 $T\colon X \to Y$, $\boldsymbol{x} \mapsto \boldsymbol{y}$.

- 若任意 $\boldsymbol{y} \in Y$ 都有原像 $\boldsymbol{x} \in X$, 则称 T 为从 X 到 Y 的满映射(onto mapping, surjection), 简称满射.
- 若对 X 中任意两个不同的元素 $\boldsymbol{x}_1, \boldsymbol{x}_2$, 都有 $T\boldsymbol{x}_1 \neq T\boldsymbol{x}_2$, 则称 T 为一对一的映射(one-to-one mapping), 或单射(injection).
- 若 T 既是满射, 又是一对一的映射, 则称 T 是一一到上的映射, 或双射(bijection).

2.132

线性变换的定义

Definition 99. 设 X 和 Y 均为数域 F 上的线性空间, 映射 $T : X \to Y$, $\boldsymbol{x} \mapsto \boldsymbol{y}$ 满足:

1. $T(\boldsymbol{x}_1 + \boldsymbol{x}_2) = T(\boldsymbol{x}_1) + T(\boldsymbol{x}_2)$, $\forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in X$;
2. $T(\lambda \boldsymbol{x}) = \lambda T(\boldsymbol{x})$, $\forall \boldsymbol{x} \in X$, $\lambda \in F$,

则称 T 为从 X 到 Y 的线性映射. 线性映射常称为线性变换(linear transformation) 或线性算子(linear operator).

☞ 上述两个条件或者等价地表示为

$$T(\lambda \boldsymbol{x}_1 + \boldsymbol{x}_2) = \lambda T(\boldsymbol{x}_1) + T(\boldsymbol{x}_2), \quad \forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in X,\ \lambda \in F.$$

两个向量的和变换得到的向量是这两个向量变换得到的向量的和, 数 λ 与向量的数乘变换得到的向量是 λ 与该向量变换得到的向量的数乘.

像这样向量之间加法与数乘关系都不受影响的变换, 它与线性空间的运算相适应, 能够反映线性空间中向量的内在联系, 是线性空间的重要变换. 2.133

Example 100. 设 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$, $\boldsymbol{x} \in \mathbb{C}^n$, 求证 $T(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}$ 是从空间 $\mathbb{C}^n$ 到空间 $\mathbb{C}^m$ 的线性变换.

Proof. $\forall \boldsymbol{x} \in \mathbb{C}^n$, $T(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x} \in \mathbb{C}^m$, 故 $T : \boldsymbol{x} \mapsto \boldsymbol{A}\boldsymbol{x}$ 是从 $\mathbb{C}^n$ 到 $\mathbb{C}^m$ 的一个映射.

$\forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{C}^n$, $\lambda \in \mathbb{C}$,

$$T(\lambda \boldsymbol{x}_1 + \boldsymbol{x}_2) = A(\lambda \boldsymbol{x}_1 + \boldsymbol{x}_2) = \lambda \boldsymbol{A}\boldsymbol{x}_1 + \boldsymbol{A}\boldsymbol{x}_2 = \lambda T(\boldsymbol{x}_1) + T(\boldsymbol{x}_2),$$

得证 $T : \boldsymbol{x} \mapsto \boldsymbol{A}\boldsymbol{x}$ 是从空间 $\mathbb{C}^n$ 到空间 $\mathbb{C}^m$ 的线性变换. □

2.134

Example 101. 积分运算是线性变换. 设 $J\colon C[a,b] \mapsto C[a,b]$, 且 J 定义为

$$J\big(f(x)\big) = \int_a^x f(t)\,\mathrm{d}t,$$

因 $\forall f(x), g(x) \in C[a,b]$, $\lambda \in \mathbb{R}$,

$$J\big(\lambda f(x) + g(x)\big) = \int_a^x \big(\lambda f(t) + g(t)\big)\,\mathrm{d}t = \lambda \int_a^x f(t)\,\mathrm{d}t + \int_a^x g(t)\,\mathrm{d}t$$

$$= \lambda J\big(f(x)\big) + J\big(g(x)\big).$$

故 J 是 $C[a,b]$ 上的线性变换. 2.135

Example 102. 微分运算是线性变换. 记 $C^1[a,b]$ 表示在区间 $[a,b]$ 上可微函数所构成的线性空间. 设 D: $C^1[a,b] \to C[a,b]$, 且 D 定义为

$$D\big(f(x)\big) = f'(x).$$

因 $\forall f(x), g(x) \in C^1[a,b]$, $\lambda \in \mathbb{R}$,

$$D\big(\lambda f(x) + g(x)\big) = \lambda f'(x) + g'(x) = \lambda D\big(f(x)\big) + D\big(g(x)\big),$$

故 D 是 $C^1[a,b]$ 到 $C[a,b]$ 的线性变换. 2.136

非线性变换的例子

Example 103. 设 T 定义为: $T(\boldsymbol{A}) = \det \boldsymbol{A}$, $\boldsymbol{A} \in \mathbb{R}^{n\times n}$. 则 T 是由 $\mathbb{R}^{n\times n}$ 到 $\mathbb{R}$ 的一个映射, 但不是线性变换.

因为一般情况下, $\det(\boldsymbol{A}+\boldsymbol{B}) \neq \det \boldsymbol{A} + \det \boldsymbol{B}$.

Example 104. 共轭转置运算不是线性变换. 设 T: $\mathbb{C}^{m\times n} \to \mathbb{C}^{n\times m}$, 且 T 定义为

$$T(\boldsymbol{A}) = \boldsymbol{A}^{\mathrm{H}}, \qquad \forall \boldsymbol{A} \in \mathbb{C}^{m\times n}.$$

取 $\lambda = \mathrm{i}$, 因

$$\begin{aligned} T(\mathrm{i}\boldsymbol{A}) &= (\mathrm{i}\boldsymbol{A})^{\mathrm{H}} = -\mathrm{i}(\boldsymbol{A})^{\mathrm{H}} \\ &\neq \mathrm{i}T(\boldsymbol{A}). \end{aligned}$$

2.137

Definition 105. 设 T 为从空间 X 到空间 X 的线性变换, 则称 T 为 X 上的线性变换.

Definition 106. 设 X 和 Y 均为数域 F 上的线性空间, 称映射 $T: X \to Y$, $\boldsymbol{x} \mapsto \boldsymbol{0}$ 为零变换(zero transformation), 记为 0^*. 即 $\forall \boldsymbol{x} \in X$, 有

$$0^*(\boldsymbol{x}) = \boldsymbol{0}.$$

2.138

Definition 107. 设 X 为数域 F 上的线性空间, 称映射 $T: X \to X$, $\boldsymbol{x} \mapsto \boldsymbol{x}$ 为 X 上的恒等变换或单位变换(identity operator), 记为 E. 即 $\forall \boldsymbol{x} \in X$, 恒有

$$E(\boldsymbol{x}) = \boldsymbol{x}.$$

Definition 108. 设 X 为数域 F 上的线性空间, $\alpha \in F$ 为固定的数, 且 $\alpha \neq 0$. 称映射 $T: X \to X$, $\boldsymbol{x} \mapsto \alpha\boldsymbol{x}$ 为 X 上的相似映射或数乘变换, 记为 α^*. 即 $\forall \boldsymbol{x} \in X$, 恒有

$$\alpha^*(\boldsymbol{x}) = \alpha\boldsymbol{x}.$$

☞ 当 $\alpha = 1$ 时, 便得恒等变换. 当 $\alpha = 0$ 时, 便得零变换.

恒等变换 E, 零变换 0^* 和相似变换 α^* 都是线性变换. 2.139

Theorem 109. *设 T 为从 X 到 Y 的线性变换, 则*

1. *$T(\mathbf{0}) = \mathbf{0}$, 其中等式左边的 $\mathbf{0} \in X$, 右边的 $\mathbf{0} \in Y$.*
2. *若 $\boldsymbol{x} = \sum_{i=1}^{k} \alpha_i \boldsymbol{x}_i \in X$, 则 $T(\boldsymbol{x}) = \sum_{i=1}^{k} \alpha_i T(\boldsymbol{x}_i)$.*
3. *若 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k$ 为 X 中线性相关的向量组, 则 $T(\boldsymbol{x}_1), T(\boldsymbol{x}_2), \cdots, T(\boldsymbol{x}_k)$ 为 Y 中线性相关的向量组.*
4. *若 $T(\boldsymbol{x}_1), T(\boldsymbol{x}_2), \cdots, T(\boldsymbol{x}_k)$ 为 Y 中线性无关的向量组, 则 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k$ 为 X 中线性无关的向量组.*

Proof. (1) $T(\mathbf{0}) = T(0\boldsymbol{x}) = 0T(\boldsymbol{x}) = \mathbf{0}$. 即, 若 $\mathbf{0}$ 是 X 的零元素, 则 $T(\mathbf{0})$ 是 Y 的零元素.

(2) 归纳法可得.

(3) 设存在不全为零的 $\lambda_1, \lambda_2, \cdots, \lambda_k$ 使得 $\lambda_1\boldsymbol{x}_1 + \lambda_2\boldsymbol{x}_2 + \cdots + \lambda_k\boldsymbol{x}_k = \mathbf{0}$, 由结论 (1), (2), 则

$$\lambda_1 T(\boldsymbol{x}_1) + \lambda_2 T(\boldsymbol{x}_2) + \cdots + \lambda_k T(\boldsymbol{x}_k) = \mathbf{0}.$$

则 $T(\boldsymbol{x}_1), T(\boldsymbol{x}_2), \cdots, T(\boldsymbol{x}_k)$ 线性相关.

(4) 此命题为 (3) 的逆否命题.

但要注意, (3) 的逆命题是不对的, 线性变换可能把线性无关的向量组也变成线性相关的向量组. 比如零变换 0^*. □ 2.140

线性变换保持线性组合与线性关系式不变

(1) 如果 $\boldsymbol{x}$ 是 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k$ 的线性组合:

$$\boldsymbol{x} = \lambda_1\boldsymbol{x}_1 + \lambda_2\boldsymbol{x}_2 + \cdots + \lambda_k\boldsymbol{x}_k,$$

那么经过线性变换 T 之后, $T(\boldsymbol{x})$ 是 $T(\boldsymbol{x}_1), T(\boldsymbol{x}_2), \cdots, T(\boldsymbol{x}_k)$ 同样的线性组合:

$$T(\boldsymbol{x}) = \lambda_1 T(\boldsymbol{x}_1) + \lambda_2 T(\boldsymbol{x}_2) + \cdots + \lambda_k T(\boldsymbol{x}_k).$$

☞ 这表明: 只要知道了 X 的一组基底 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n$ 在 T 下的像, 那么 X 中任一向量在 T 下的像就确定了. 即, n 维线性空间 X 到线性空间 Y 的线性变换, 完全被它在 X 的一组基底上的作用所决定. 2.141

(2) 如果 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k$ 之间有一线性关系式

$$\lambda_1\boldsymbol{x}_1 + \lambda_2\boldsymbol{x}_2 + \cdots + \lambda_k\boldsymbol{x}_k = \mathbf{0},$$

那么它们的像之间也有同样的关系

$$\lambda_1 T(\boldsymbol{x}_1) + \lambda_2 T(\boldsymbol{x}_2) + \cdots + \lambda_k T(\boldsymbol{x}_k) = \mathbf{0}'.$$

这里 $\mathbf{0}, \mathbf{0}'$ 分别是线性空间 X 和 Y 的零向量. 2.142

Exercise 110 (P.68 习题 (一) 1)**.** 判定下列映射哪些是线性变换?

(1) 设 $\boldsymbol{x}_0$ 为空间 X 中的一个固定向量, 映射 T: $X \to X$, $\boldsymbol{x} \mapsto \boldsymbol{x} + \boldsymbol{x}_0$.

解: 当 $\boldsymbol{x}_0 = \mathbf{0}$ 时, $T\boldsymbol{x} = \boldsymbol{x} + \boldsymbol{x}_0 = \boldsymbol{x}$ 显然是 X 上的线性变换.

当 $\boldsymbol{x}_0 \neq \mathbf{0}$ 时,

$$T(\boldsymbol{x}_1 + \boldsymbol{x}_2) = \boldsymbol{x}_1 + \boldsymbol{x}_2 + \boldsymbol{x}_0, \qquad T(\boldsymbol{x}_1) + T(\boldsymbol{x}_2) = \boldsymbol{x}_1 + \boldsymbol{x}_2 + 2\boldsymbol{x}_0.$$

则 $T(\boldsymbol{x}_1 + \boldsymbol{x}_2) \neq T(\boldsymbol{x}_1) + T(\boldsymbol{x}_2)$, 即此时 T 不是 X 上的线性变换. □

(2) T: $\mathbb{C} \to \mathbb{C}$, $\boldsymbol{\xi} \mapsto \overline{\boldsymbol{\xi}}$.

解: T 不是线性变换. 因为取 $\lambda = \mathrm{i}$, $\boldsymbol{\xi} = 1$ 时, 有

$$T(\lambda\boldsymbol{\xi}) = \overline{\lambda\boldsymbol{\xi}} = -\mathrm{i}, \qquad \lambda T(\boldsymbol{\xi}) = \lambda\overline{\boldsymbol{\xi}} = \mathrm{i}.$$

故 $T(\lambda\boldsymbol{\xi}) \neq \lambda T(\boldsymbol{\xi})$. □ 2.143

(3) 把复数域 $\mathbb{C}$ 看作实数域 $\mathbb{R}$ 上的线性空间, 记为 $\mathbb{C}_{\mathbb{R}}$. 映射 T: $\mathbb{C}_{\mathbb{R}} \to \mathbb{C}_{\mathbb{R}}$, $\boldsymbol{\xi} \mapsto \mathrm{i}\,\mathrm{Re}\boldsymbol{\xi}$.

解: $\forall \boldsymbol{\xi}, \boldsymbol{\eta} \in \mathbb{C}_{\mathbb{R}}$, $\lambda \in \mathbb{R}$, 有

$$\begin{aligned}
T(\lambda\boldsymbol{\xi} + \boldsymbol{\eta}) &= \mathrm{i}\,\mathrm{Re}(\lambda\boldsymbol{\xi} + \boldsymbol{\eta}) \\
&= \mathrm{i}\,\mathrm{Re}(\lambda\boldsymbol{\xi}) + \mathrm{i}\,\mathrm{Re}\boldsymbol{\eta} \\
&= \lambda(\mathrm{i}\,\mathrm{Re}\boldsymbol{\xi}) + \mathrm{i}\,\mathrm{Re}\boldsymbol{\eta} \\
&= \lambda T(\boldsymbol{\xi}) + T(\boldsymbol{\eta}).
\end{aligned}$$

故 T 是线性变换. □ 2.144

(4) 映射 T: $\mathbb{C}_{\mathbb{R}} \to \mathbb{C}$, $\boldsymbol{\xi} \mapsto \overline{\boldsymbol{\xi}}$.

解: T 是线性变换. $\forall \boldsymbol{\xi}, \boldsymbol{\eta} \in \mathbb{C}_{\mathbb{R}}$, $\lambda \in \mathbb{R}$, 有

$$\begin{aligned}
T(\lambda\boldsymbol{\xi} + \boldsymbol{\eta}) &= \overline{\lambda\boldsymbol{\xi} + \boldsymbol{\eta}} \\
&= \lambda\overline{\boldsymbol{\xi}} + \overline{\boldsymbol{\eta}} \\
&= \lambda T(\boldsymbol{\xi}) + T(\boldsymbol{\eta}).
\end{aligned}$$

□ 2.145

(5) T: $\mathbb{R}^3 \to \mathbb{R}^3$, $(x_1, x_2, x_3)^{\mathrm{T}} \mapsto (2x_1 - x_2, x_2 + x_3, x_1)^{\mathrm{T}}$.

解: T 是线性变换. $\forall \boldsymbol{x} = (x_1, x_2, x_3)^{\mathrm{T}}$, $\boldsymbol{y} = (y_1, y_2, y_3)^{\mathrm{T}} \in \mathbb{R}^3$, $\lambda \in \mathbb{R}$, 有

$$\begin{aligned}
T(\boldsymbol{x}) + T(\boldsymbol{y}) &= T\big((x_1, x_2, x_3)^{\mathrm{T}}\big) + T\big((y_1, y_2, y_3)^{\mathrm{T}}\big) \\
&= (2x_1 - x_2, x_2 + x_3, x_1)^{\mathrm{T}} + (2y_1 - y_2, y_2 + y_3, y_1)^{\mathrm{T}} \\
&= \big(2(x_1 + y_1) - (x_2 + y_2), (x_2 + y_2) + (x_3 + y_3), (x_1 + y_1)\big)^{\mathrm{T}} \\
&= T\big(x_1 + y_1, x_2 + y_2, x_3 + y_3\big) = T(\boldsymbol{x} + \boldsymbol{y}). \\
T(\lambda\boldsymbol{x}) &= T\big((\lambda x_1, \lambda x_2, \lambda x_3)^{\mathrm{T}}\big)
\end{aligned}$$

$$
\begin{aligned}
&= (2\lambda x_1 - \lambda x_2, \lambda x_2 + \lambda x_3, \lambda x_1)^{\mathrm{T}} \\
&= \lambda(2x_1 - x_2, x_2 + x_3, x_1)^{\mathrm{T}} = \lambda T(\boldsymbol{x}).
\end{aligned}
$$

□ 2.146

(6) T: $\mathbb{R}^{n\times n} \to \mathbb{C}^{n\times n}$, $\boldsymbol{Z} \mapsto \boldsymbol{BZC}$, 其中 $\boldsymbol{B}$, $\boldsymbol{C}$ 是 $\mathbb{C}^{n\times n}$ 中固定的矩阵.

解: $\forall \boldsymbol{Z}_1$, $\boldsymbol{Z}_2 \in \mathbb{R}^{n\times n}$, $\lambda \in \mathbb{R}$, 有

$$
\begin{aligned}
T(\lambda \boldsymbol{Z}_1 + \boldsymbol{Z}_2) &= \boldsymbol{B}(\lambda \boldsymbol{Z}_1 + \boldsymbol{Z}_2)\boldsymbol{C} \\
&= \lambda \boldsymbol{B}\boldsymbol{Z}_1\boldsymbol{C} + \boldsymbol{B}\boldsymbol{Z}_2\boldsymbol{C} \\
&= \lambda T(\boldsymbol{Z}_1) + T(\boldsymbol{Z}_2).
\end{aligned}
$$

故 T 是线性变换. □ 2.147

(7) T: $\mathbb{R}[x]_n \to \mathbb{R}[x]_n$, $p(x) \mapsto p(x+1)$.

解: T 是线性变换. 设 $p(x)$, $q(x) \in \mathbb{R}[x]_n$, $\lambda \in \mathbb{R}$, 并令

$$
r(x) \triangleq p(x) + q(x), \qquad s(x) \triangleq \lambda p(x).
$$

则 $r(x+1) = p(x+1) + q(x+1)$, $s(x+1) = \lambda p(x+1)$. 故

$$
\begin{aligned}
T\big(p(x) + q(x)\big) &= T\big(r(x)\big) = r(x+1) \\
&= p(x+1) + q(x+1) = T\big(p(x)\big) + T\big(q(x)\big), \\
T\big(\lambda p(x)\big) &= T\big(s(x)\big) = s(x+1) \\
&= \lambda p(x+1) = \lambda T\big(p(x)\big).
\end{aligned}
$$

□ 2.148

3.2 线性变换的运算

Definition 111. 设空间 X 和空间 Y 为数域 F 上的线性空间. T_1, T_2 都是从 X 到 Y 的映射, $\lambda \in F$, 若

1. 对任意 $\boldsymbol{x} \in X$ 均有 $T_1(\boldsymbol{x}) = T_2(\boldsymbol{x})$, 则称 T_1 与 T_2 相等, 记为 $T_1 = T_2$.
2. 对任意 $\boldsymbol{x} \in X$ 均有 $T_1(\boldsymbol{x}) + T_2(\boldsymbol{x}) = T(\boldsymbol{x})$, 则称 T 为 T_1 与 T_2 的和, 记为 $T = T_1 + T_2$.
3. 对任意 $\boldsymbol{x} \in X$ 均有 $T(\boldsymbol{x}) = \lambda\big(T_1(\boldsymbol{x})\big)$, 则称 T 为 T_1 与 λ 的标量乘积, 记为 $T = \lambda T_1$.

☞ 若 T_1 与 T_2 为线性变换, 则 $T_1 + T_2$ 及 λT_1 仍是线性变换. 2.149

Theorem 112. *设 $L(X,Y)$ 为从 X 到 Y 的所有线性变换所构成的集合, 则 $L(X,Y)$ 按照定义 111 中的加法与标量乘法构成数域 F 上的一个线性空间, 称为 X, Y 所诱导的变换空间.*

2.150

Definition 113. 设 $T \in L(X,Y)$, $S \in L(Y,Z)$, 若线性变换 $G \in L(X,Z)$ 对任意 $\boldsymbol{x} \in X$ 都满足

$$G(\boldsymbol{x}) = S\big(T(\boldsymbol{x})\big),$$

则称 G 为 T 与 S 的积, 并记为 ST, 即 $G = ST$.

☞ 线性变换的乘积一般是不可交换的. 即 $TS = ST$ 一般不成立. 例如下面的习题. 2.151

Exercise 114 (P.71 习题 (一) 4)**.** 设 T, S 为线性空间 $\mathbb{R}[x]$ 的如下的两个线性变换:

$$Tp(x) = p'(x), \qquad Sp(x) = xp(x).$$

试问等式 $TS = ST$ 是否成立? 并证明 $TS - ST = E$.

解: (1) 因为

$$\begin{aligned}
(TS)p(x) &= T\big(Sp(x)\big) = T\big(xp(x)\big) = p(x) + xp'(x), \\
(ST)p(x) &= S\big(Tp(x)\big) = S\big(p'(x)\big) = xp'(x),
\end{aligned}$$

可见 $p(x) \neq 0$ 时, $(TS)p(x) \neq (ST)p(x)$, 故

$$TS \neq ST.$$

2.152

(2) 由上述讨论知, 对任意 $p(x) \in \mathbb{R}[x]$, 有

$$(TS)p(x) = p(x) + (ST)p(x),$$

即

$$(TS - ST)p(x) = p(x),$$

故 $TS - ST = E$. □ 2.153

Definition 115. 设 $T \in L(X,X)$, 则称 TT 为 T 的平方, 并记为 $T^2 = TT$. 一般地, 以下述递推式来表示 T 的 k 次方($k \geqslant 0$):

$$\begin{cases} T^0 = E, \\ T^k = T(T^{k-1}), \qquad k = 1, 2, \cdots \end{cases} \tag{12}$$

若 T 是可逆变换, 则上式中的 k 可以取任何整数.

☞ 可逆变换: 设 T, S 为空间 X 上的变换, 若

$$TS = ST = E,$$

则称 S 为 T 的逆变换, 记为 T^{-1}.

比如取 $k = -2$, 则 $T^{-2} = T(T^{-3})$. 2.154

Definition 116. 设 X 为数域 F 上的线性空间, $T \in L(X, X)$, 又设 $g(\lambda)$ 为关于 λ 的多项式, 其系数属于 F, 即

$$g(\lambda) = \alpha_0 + \alpha_1 \lambda + \cdots + \alpha_m \lambda^m,$$

则表达式

$$g(T) = \alpha_0 E + \alpha_1 T + \cdots + \alpha_m T^m,$$

称为线性变换 T 的多项式.

2.155

Exercise 117 (P.71 习题 (一) 6)**.** 设 T, $S \in L(X, X)$, 并且 $TS - ST = E$, 试证 $T^m S - ST^m = mT^{m-1}$, $m = 1, 2, \cdots$.

Proof. 对 m 用数学归纳法.

当 $m = 1$ 时, 即 $TS - ST = T^0 = E$, 由题设成立.

假定等式对 m 成立, 即有 $T^m S - ST^m = mT^{m-1}$. 下面证明等式对 $m + 1$ 也成立.

$$\begin{aligned}
&T^{m+1} S - ST^{m+1} = T^m (TS) - ST^{m+1} \\
&= T^m (E + ST) - ST^{m+1} = T^m + T^m ST - ST^{m+1} \\
&= T^m + (T^m S - ST^m) T = T^m + (mT^{m-1}) T = (m + 1) T^m,
\end{aligned}$$

即等式对 $m + 1$ 也成立. 从而对任意正整数都成立. □

2.156

3.3 与线性变换有关的子空间

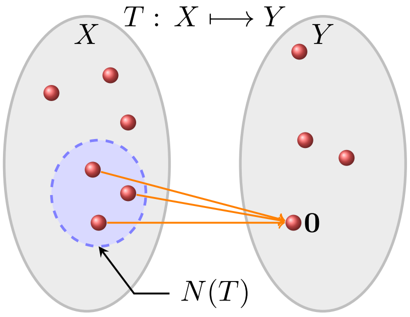
线性变换的值域与核

Definition 118. 给定线性变换 T: $X \to Y$, 记

$$\begin{aligned}
R(T) &= \{\boldsymbol{y} \in Y \mid \boldsymbol{y} = T(\boldsymbol{x}),\ \boldsymbol{x} \in X\}, \\
N(T) &= \{\boldsymbol{x} \in X \mid T(\boldsymbol{x}) = \mathbf{0}\},
\end{aligned}$$

则 $R(T)$ 称为 T 的值空间(或值域) ; $N(T)$ 称为 T 的零空间(或核).

☞ 另见教材 P.33: 矩阵 $\boldsymbol{A}$ 的值空间 $R(\boldsymbol{A})$ 和零空间 $N(\boldsymbol{A})$. 2.157



2.158

Example 119. 证明: $R(T)$ 是 Y 的一个子空间, $N(T)$ 是 X 的一个子空间.

证: (1) $R(T)$ 是 Y 的非空子集. 下证 $R(T)$ 对加法和数乘是封闭的.

$\forall \boldsymbol{y}_1 = T(\boldsymbol{x}_1)$, $\boldsymbol{y}_2 = T(\boldsymbol{x}_2) \in R(T)$, $\forall k \in F$, 有

$$\begin{aligned} \boldsymbol{y}_1 + \boldsymbol{y}_2 &= T(\boldsymbol{x}_1) + T(\boldsymbol{x}_2) = T(\boldsymbol{x}_1 + \boldsymbol{x}_2) \in R(T), \\ k\boldsymbol{y}_1 &= kT(\boldsymbol{x}_1) = T(k\boldsymbol{x}_1) \in R(T). \end{aligned}$$

故 $R(T)$ 是 Y 的一个子空间.

(2) $N(T)$ 是 X 的非空子集. 下证 $N(T)$ 对加法和数乘是封闭的.

$\forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in N(T)$, 有 $T(\boldsymbol{x}_1) = \mathbf{0}$, $T(\boldsymbol{x}_2) = \mathbf{0}$, 则

$$T(\boldsymbol{x}_1 + \boldsymbol{x}_2) = T(\boldsymbol{x}_1) + T(\boldsymbol{x}_2) = \mathbf{0} + \mathbf{0} = \mathbf{0}.$$

得 $\boldsymbol{x}_1 + \boldsymbol{x}_2 \in N(T)$. 又 $\forall \lambda \in F$, 有

$$T(\lambda \boldsymbol{x}_1) = \lambda T(\boldsymbol{x}_1) = \mathbf{0},$$

得 $\lambda \boldsymbol{x}_1 \in N(T)$. 故 $N(T)$ 是 X 的一个子空间. □ 2.159

Example 120. 在线性空间 $\mathbb{R}[x]_n$ 中, 令

$$T\big(p(x)\big) = p'(x).$$

则 T 的值域就是 $\mathbb{R}[x]_{n-1}$, T 的核就是子空间 $\mathbb{R}$. 2.160

Theorem 121. 设 X 和 Y 均为数域 F 上线性空间, $\dim X = n$, $\dim Y = m$, 又设 $T \in L(X, Y)$, 则有

$$\dim R(T) + \dim N(T) = \dim X.$$

证: 记 $\dim N(T) = k$, 设 $N(T)$ 的基底为

$$\{\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k\}.$$

则 $T(\boldsymbol{x}_1) = \cdots = T(\boldsymbol{x}_k) = \mathbf{0}$. 将其扩充为 X 的基底 $\{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k, \boldsymbol{x}_{k+1}, \cdots, \boldsymbol{x}_n\}$.

$\forall \boldsymbol{y} \in R(T)$, $\exists \boldsymbol{x} \in X$, 满足 $\boldsymbol{y} = T(\boldsymbol{x})$. 设 $\boldsymbol{x} = \sum\limits_{i=1}^{n} \alpha_i \boldsymbol{x}_i$, 则

$$\begin{aligned} \boldsymbol{y} = T(\boldsymbol{x}) &= T\big(\alpha_1 \boldsymbol{x}_1 + \cdots + \alpha_k \boldsymbol{x}_k + \alpha_{k+1} \boldsymbol{x}_{k+1} + \cdots + \alpha_n \boldsymbol{x}_n\big) \\ &= \alpha_{k+1} T(\boldsymbol{x}_{k+1}) + \cdots + \alpha_n T(\boldsymbol{x}_n), \end{aligned}$$

由 $\boldsymbol{y}$ 的任意性, 下证 $T(\boldsymbol{x}_{k+1}), \cdots, T(\boldsymbol{x}_n)$ 线性无关, 从而是 $R(T)$ 的一组基底, 则 $\dim R(T) = n - k$, 得到 $\dim R(T) + \dim N(T) = \dim X$. 2.161

设有 $\xi_{k+1} T(\boldsymbol{x}_{k+1}) + \cdots + \xi_n T(\boldsymbol{x}_n) = \mathbf{0}$, 即

$$T(\xi_{k+1} \boldsymbol{x}_{k+1} + \cdots + \xi_n \boldsymbol{x}_n) = \mathbf{0},$$

从而 $\xi_{k+1} \boldsymbol{x}_{k+1} + \cdots + \xi_n \boldsymbol{x}_n \in N(T)$. 故可设

$$\xi_{k+1} \boldsymbol{x}_{k+1} + \cdots + \xi_n \boldsymbol{x}_n = \lambda_1 \boldsymbol{x}_1 + \cdots + \lambda_k \boldsymbol{x}_k,$$

即

$$-\lambda_1 \boldsymbol{x}_1 - \cdots - \lambda_k \boldsymbol{x}_k + \xi_{k+1}\boldsymbol{x}_{k+1} + \cdots + \xi_n \boldsymbol{x}_n = \mathbf{0}.$$

而 $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k, \boldsymbol{x}_{k+1}, \cdots, \boldsymbol{x}_n$ 为 X 的基底, 故

$$\lambda_1 = \cdots = \lambda_k = \xi_{k+1} = \cdots = \xi_n = 0.$$

故 $T(\boldsymbol{x}_{k+1}), \cdots, T(\boldsymbol{x}_n)$ 线性无关. □ 2.162

☞ **n 元齐次线性方程组的解空间的维数公式, 本质上是线性变换的核与值域的维数公式.**

工科线性代数的核心理论是 "$n-r$" 问题, 即 n 元齐次线性方程组 $\boldsymbol{A}\boldsymbol{x} = \mathbf{0}$ 的解空间的维数为

$$n - \operatorname{rank}(\boldsymbol{A}).$$

换言之, $\boldsymbol{A}$ 的核空间的维数为

$$\dim N(\boldsymbol{A}) = n - \operatorname{rank}(\boldsymbol{A}).$$

又 $\dim R(\boldsymbol{A}) = \operatorname{rank}(\boldsymbol{A})$, 上式即

$$\dim N(\boldsymbol{A}) + \dim R(\boldsymbol{A}) = n.$$

2.163

对矩阵 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$, $\boldsymbol{x} \in \mathbb{C}^n$,

$$\boldsymbol{A} : \boldsymbol{x} \longmapsto \boldsymbol{A}\boldsymbol{x}$$

是从空间 $\mathbb{C}^n$ 到空间 $\mathbb{C}^m$ 的线性变换. 有

$$\dim R(\boldsymbol{A}) + \dim N(\boldsymbol{A}) = \dim \mathbb{C}^n.$$

2.164

Example 122. 给定矩阵

$$\boldsymbol{A} = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 1 & 7 \\ 0 & 1 & 2 \\ 8 & 4 & 3 \end{pmatrix},$$

将 $\boldsymbol{A}$ 看作线性变换 $\boldsymbol{A} : \mathbb{C}^3 \longrightarrow \mathbb{C}^4$. 求 $R(\boldsymbol{A})$ 和 $N(\boldsymbol{A})$ 的维数.

解: 由

$$\begin{vmatrix} 1 & 3 & 5 \\ 2 & 1 & 7 \\ 0 & 1 & 2 \end{vmatrix} = \begin{vmatrix} 1 & 3 & 5 \\ 0 & -5 & -3 \\ 0 & 1 & 2 \end{vmatrix} = -7 \neq 0,$$

知 $\operatorname{rank}(\boldsymbol{A}) = 3$. 故

$$\begin{aligned} \dim R(\boldsymbol{A}) &= \operatorname{rank}(\boldsymbol{A}) = 3, \\ \dim N(\boldsymbol{A}) &= \dim \mathbb{C}^3 - \operatorname{rank}(\boldsymbol{A}) = 0. \end{aligned}$$

2.165

Definition 123. 给定线性变换 $T\colon X \to Y$, 则称 $R(T)$ 的维数为 T 的秩, 记为 $\mathrm{rank}(T)$. 又称 $N(T)$ 的维数为 T 的零度, 记为 $\mathrm{null}(T)$.

☞ 前述定理的结论也可以表达为

$$\mathrm{rank}(T) + \mathrm{null}(T) = \dim X.$$

2.166

不变子空间

Definition 124. 设 W 为空间 X 的子空间, T 为 X 上的线性变换. 如果对任意 $\boldsymbol{w} \in W$ 均有 $T(\boldsymbol{w}) \in W$, 则称 W 为 T 的不变子空间(invariant subspace), 记为 $TW \subset W \subset X$.

Example 125. 整个空间 X 和零子空间 $\{\mathbf{0}\}$, 对于每个线性变换 T 来说, 都是 T 的不变子空间. 2.167

Example 126. 线性变换 $T\colon X \to X$ 的值域 $R(T)$ 和核 $N(T)$, 都是 T 的不变子空间.

事实上, 对线性变换 $T\colon X \to X$, 因

$$\begin{aligned} R(T) &= \{\boldsymbol{y} \in X \mid \boldsymbol{y} = T(\boldsymbol{x}),\ \boldsymbol{x} \in X\}, \\ N(T) &= \{\boldsymbol{x} \in X \mid T(\boldsymbol{x}) = \mathbf{0}\}, \end{aligned}$$

值域 $R(T)$ 是 X 中向量在 T 下的像的集合, 它当然也包含 $R(T)$ 中向量的像. 所以 $R(T)$ 是 T 的不变子空间.

核 $N(T)$ 是被 T 变成零的向量的集合, 核中向量的像是零, 自然在核中, 因此核 $N(T)$ 是 T 的不变子空间. 2.168

Exercise 127 (P.74 习题 (一) 4)**.** 设 $T, S \in L(X, X)$, 且 $TS = ST$, 试证 $R(T)$ 与 $N(T)$ 都是 S 的不变子空间.

Proof. $\forall \boldsymbol{\xi} \in N(T)$,

$$T\big(S(\boldsymbol{\xi})\big) = TS(\boldsymbol{\xi}) = S\big(T(\boldsymbol{\xi})\big) = S(\mathbf{0}) = \mathbf{0}.$$

即 $S(\boldsymbol{\xi}) \in N(T)$. 故 $N(T)$ 是 S 的不变子空间.

在 T 的值域 $R(T)$ 中任取一向量 $T(\boldsymbol{\eta})$, 则

$$S\big(T(\boldsymbol{\eta})\big) = T\big(S(\boldsymbol{\eta})\big) \in R(T),$$

因此 $R(T)$ 也是 S 的不变子空间. □

☞ 若线性变换 T 与 S 是可交换的, 则 T 的值域与核都是 S 的不变子空间. 2.169

Definition 128. 设 T 为空间 X 上的线性变换, 且 $TW \subset W \subset X$, 则 T 可以看作 W 上的线性变换, 此变换称为 T 在 W 上的限制, 记为 T_W. 又称 T 为 T_W 在 X 上的扩张.

In such a situation, T can be considered as a linear operator on W by forgetting about everything else in X and restricting T to act only on vectors from W. Hereafter, this **restricted operator** will be denoted by T_W. 2.170

Exercise 129 (P.74 习题 (一) 8)**.** 设 $W = \text{span}[\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_m]$, 则 W 为 T 的不变子空间的充分必要条件是 $T(\boldsymbol{u}_1)$, $T(\boldsymbol{u}_2)$, $\cdots$, $T(\boldsymbol{u}_m)$ 全属于 W.

Proof. 必要性是显然的. 现在来证充分性.

如果 $T(\boldsymbol{u}_1)$, $T(\boldsymbol{u}_2)$, $\cdots$, $T(\boldsymbol{u}_m)$ 全属于 W, 由于 W 中每个向量 $\boldsymbol{\xi}$ 都可以被 $\boldsymbol{u}_1$, $\boldsymbol{u}_2$, $\cdots$, $\boldsymbol{u}_m$ 线性表示, 即有

$$\boldsymbol{\xi} = k_1\boldsymbol{u}_1 + k_2\boldsymbol{u}_2 + \cdots + k_m\boldsymbol{u}_m,$$

所以

$$T(\boldsymbol{\xi}) = k_1T(\boldsymbol{u}_1) + k_2T(\boldsymbol{u}_2) + \cdots + k_mT(\boldsymbol{u}_m) \in W.$$

故 W 为 T 的不变子空间. □

2.171

4 线性变换的矩阵表示和空间的同构

4.1 线性变换的矩阵表示

设 X 和 Y 均为数域 F 上的线性空间, $\dim X = n$, $\dim Y = m$. 又设 T: $X \to Y$ 为给定的从 X 到 Y 的线性变换. 任取 X 和 Y 的基底分别为

$$\mathscr{B}_X = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n\}, \qquad \mathscr{B}_Y = \{\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_m\}.$$

则基底 $\mathscr{B}_X$ 的像 $T(\boldsymbol{x}_1)$, $T(\boldsymbol{x}_2)$, $\cdots$, $T(\boldsymbol{x}_n)$ 可以由基底 $\mathscr{B}_Y$ 线性表示:

$$\begin{cases} T(\boldsymbol{x}_1) = a_{11}\boldsymbol{y}_1 + a_{21}\boldsymbol{y}_2 + \cdots + a_{m1}\boldsymbol{y}_m, \\ T(\boldsymbol{x}_2) = a_{12}\boldsymbol{y}_1 + a_{22}\boldsymbol{y}_2 + \cdots + a_{m2}\boldsymbol{y}_m, \\ \qquad\qquad \vdots \\ T(\boldsymbol{x}_n) = a_{1n}\boldsymbol{y}_1 + a_{2n}\boldsymbol{y}_2 + \cdots + a_{mn}\boldsymbol{y}_m, \end{cases}$$

即

$$\big[T(\boldsymbol{x}_1), T(\boldsymbol{x}_2), \cdots, T(\boldsymbol{x}_n)\big] = [\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_m]\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

2.172

记

$$\boldsymbol{P} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \in F^{m\times n},$$

则

$$\big[T(\boldsymbol{x}_1), T(\boldsymbol{x}_2), \cdots, T(\boldsymbol{x}_n)\big] = [\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_m]\boldsymbol{P}. \tag{13}$$

若记

$$T\big[\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n\big] \triangleq \big[T(\boldsymbol{x}_1), T(\boldsymbol{x}_2), \cdots, T(\boldsymbol{x}_n)\big],$$

则式 (13) 还可以表达为

$$T\big[\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n\big] = [\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_m]\boldsymbol{P},$$

或者

$$T\mathscr{B}_X = \mathscr{B}_Y\boldsymbol{P}. \tag{14}$$

2.173

Definition 130. 表达式 $T\mathscr{B}_X = \mathscr{B}_Y\boldsymbol{P}$ 中的矩阵 $\boldsymbol{P}$ 称为线性变换 T 的关于基底 $\mathscr{B}_X$ 和基底 $\mathscr{B}_Y$ 的矩阵表示, 记为

$$m_{\mathscr{B}_X,\mathscr{B}_Y}(T).$$

2.174

利用线性变换的矩阵可以直接计算一个向量的像.

Theorem 131. 设线性变换 $T\colon X \to Y$ 的关于基底 $\mathscr{B}_X$ 和基底 $\mathscr{B}_Y$ 的矩阵表示为 $\boldsymbol{P}$, 向量 $\boldsymbol{x}$ 在基底 $\mathscr{B}_X$ 下的坐标为 $\boldsymbol{\xi} = (a_1, a_2, \cdots, a_n)^{\mathrm{T}}$, 且 $\boldsymbol{y} = T(\boldsymbol{x})$ 在基底 $\mathscr{B}_Y$ 下的坐标 $\boldsymbol{\eta} = (b_1, b_2, \cdots, b_m)^{\mathrm{T}}$. 则

$$\boldsymbol{\eta} = \boldsymbol{P}\boldsymbol{\xi}.$$

证: 记 $\mathscr{B}_X = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n\}$, $\mathscr{B}_Y = \{\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_m\}$.

由假设 $\boldsymbol{x} = a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2 + \cdots + a_n\boldsymbol{x}_n$, 于是

$$\begin{aligned} T(\boldsymbol{x}) &= a_1T(\boldsymbol{x}_1) + a_2T(\boldsymbol{x}_2) + \cdots + a_nT(\boldsymbol{x}_n) \\ &= \big[T(\boldsymbol{x}_1), T(\boldsymbol{x}_2), \cdots, T(\boldsymbol{x}_n)\big]\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = [\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_m]\boldsymbol{P}\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}. \end{aligned}$$

2.175

另一方面, 由假设

$$T(\boldsymbol{x}) = [\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_m]\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix},$$

因为向量 $\boldsymbol{y}_1$, $\boldsymbol{y}_2$, $\cdots$, $\boldsymbol{y}_m$ 线性无关, 所以

$$\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} = \boldsymbol{P}\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}, \qquad 即 \quad \boldsymbol{\eta} = \boldsymbol{P}\boldsymbol{\xi}.$$

$\square$

☞ 线性空间中的向量可以用坐标来表示, 抽象的线性变换也能同具体的数发生联系: 用矩阵来表示. 2.176

线性变换的矩阵是与空间中的一组基底联系在一起的. 一般说来, 随着基底的改变, 同一个线性变换就有不同的矩阵. 为了利用矩阵来研究线性变换, 我们有必要弄清楚线性变换的矩阵是如何随着基底的改变而改变的.

线性变换在不同基底下的矩阵之间的关系

Theorem 132. 设 X 和 Y 为数域 F 上的线性空间, 对给定的线性变换 T: $X \to Y$, 设其关于 X 和 Y 的两组不同基底 $\mathscr{B}_X$, $\mathscr{B}_Y$ 和 $\mathscr{B}_X^{(1)}$, $\mathscr{B}_Y^{(1)}$ 的矩阵表示分别为 $\boldsymbol{P}$ 和 $\boldsymbol{P}_1$, 即

$$T\mathscr{B}_X = \mathscr{B}_Y \boldsymbol{P}, \qquad T\mathscr{B}_X^{(1)} = \mathscr{B}_Y^{(1)} \boldsymbol{P}_1.$$

又设有过渡矩阵 $\boldsymbol{A}$, $\boldsymbol{B}$ 使得 $\mathscr{B}_X = \mathscr{B}_X^{(1)}\boldsymbol{A}$, $\mathscr{B}_Y = \mathscr{B}_Y^{(1)}\boldsymbol{B}$. 则

$$\boldsymbol{P}_1 = \boldsymbol{B}\boldsymbol{P}\boldsymbol{A}^{-1}.$$

证:

$$\begin{aligned} T\mathscr{B}_X = \mathscr{B}_Y \boldsymbol{P} \quad &\Rightarrow \quad T\mathscr{B}_X^{(1)}\boldsymbol{A} = \mathscr{B}_Y^{(1)}\boldsymbol{B}\boldsymbol{P}, \\ &\Rightarrow \quad \mathscr{B}_Y^{(1)}\boldsymbol{P}_1\boldsymbol{A} = \mathscr{B}_Y^{(1)}\boldsymbol{B}\boldsymbol{P}, \\ &\Rightarrow \quad \boldsymbol{P}_1\boldsymbol{A} = \boldsymbol{B}\boldsymbol{P}, \end{aligned}$$

即 $\boldsymbol{P}_1 = \boldsymbol{B}\boldsymbol{P}\boldsymbol{A}^{-1}$. $\square$ 2.177

即线性变换在不同基底下的矩阵是相抵的.

Corollary 133. 设 T 是线性空间 X 上的线性变换, 即 T: $X \to X$. 若 T 在 X 的两组基底 $\mathscr{B}$ 和 $\mathscr{B}^{(1)}$ 下的矩阵表示分别为 $\boldsymbol{P}$ 和 $\boldsymbol{P}_1$, 且从基 $\mathscr{B}$ 到基 $\mathscr{B}^{(1)}$ 的过渡矩阵为 $\boldsymbol{A}$, 则

$$\boldsymbol{P}_1 = \boldsymbol{A}^{-1}\boldsymbol{P}\boldsymbol{A}.$$

☞ 即: 同一线性变换在线性空间 X 的不同基底下的矩阵表示是相似的. 这就是线性代数教材用相当多的篇幅讨论 n 阶矩阵的相似关系的主要原因. 2.178

Exercise 134 (P.82 习题 (一) 2)**.** 已知 $\mathbb{R}^3$ 中线性变换 T 在基底 $\mathscr{B} = \{\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\eta}_3\}$ 下的矩阵表示为

$$\boldsymbol{P} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ -1 & 2 & 1 \end{bmatrix}$$

其中 $\boldsymbol{\eta}_1 = (-1, 1, 1)^{\mathrm{T}}$, $\boldsymbol{\eta}_2 = (1, 0, -1)^{\mathrm{T}}$, $\boldsymbol{\eta}_3 = (0, 1, 1)^{\mathrm{T}}$. 求 T 在自然基底 $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ 下的矩阵表示.

解: 已知

$$T[\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\eta}_3] = [\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\eta}_3]\boldsymbol{P}$$

设所求矩阵为 $\boldsymbol{P}_1$, 即满足

$$T[\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3] = [\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3]\boldsymbol{P}_1.$$

2.179

由

$$[\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\eta}_3] = [\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3]\begin{bmatrix} -1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 1 \end{bmatrix} \triangleq [\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3]\boldsymbol{A},$$

则

$$\begin{aligned} \boldsymbol{P}_1 &= \boldsymbol{A}\boldsymbol{P}\boldsymbol{A}^{-1} \\ &= \begin{bmatrix} -1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ -1 & 2 & 1 \end{bmatrix}\begin{bmatrix} -1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 1 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} -1 & 1 & -2 \\ 2 & 2 & 0 \\ 3 & 0 & 2 \end{bmatrix}. \end{aligned}$$

□

2.180

或者不使用公式, 用下述表达方式.

已知

$$T[\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\eta}_3] = [\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\eta}_3]\boldsymbol{P}$$

代入

$$[\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\eta}_3] = [\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3]\begin{bmatrix} -1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 1 \end{bmatrix}$$

得

$$T[\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3]\begin{bmatrix} -1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 1 \end{bmatrix} = [\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3]\begin{bmatrix} -1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 1 \end{bmatrix}\boldsymbol{P}.$$

故

$$T[\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3] = [\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3]\begin{bmatrix} -1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 1 \end{bmatrix}\boldsymbol{P}\begin{bmatrix} -1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 1 \end{bmatrix}^{-1}.$$

2.181

Exercise 135 (P.83 习题 (一) 3)**.** 在 $\mathbb{R}^3$ 中线性变换 T 定义如下:

$$\begin{aligned} T(-1, 0, 2)^{\mathrm{T}} &= (-5, 0, 3)^{\mathrm{T}}, \\ T(0, 1, 1)^{\mathrm{T}} &= (0, -1, 6)^{\mathrm{T}}, \end{aligned}$$

$$T(3,-1,0)^{\mathrm{T}}=(-5,-1,9)^{\mathrm{T}}.$$

求 T 在自然基底 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3\}$ 下的矩阵表示, 并求 $R(T)$ 及 $\dim R(T)$.

解: 记 $\boldsymbol{\eta}_1=(-1,0,2)^{\mathrm{T}}$, $\boldsymbol{\eta}_2=(0,1,1)^{\mathrm{T}}$, $\boldsymbol{\eta}_3=(3,-1,0)^{\mathrm{T}}$. 由

$$\begin{vmatrix} -1 & 0 & 3 \\ 0 & 1 & -1 \\ 2 & 1 & 0 \end{vmatrix}=5\neq 0,$$

知 $\boldsymbol{\eta}_1$, $\boldsymbol{\eta}_2$, $\boldsymbol{\eta}_3$ 是 $\mathbb{R}^3$ 的一组基底. 2.182

由题设

$$T[\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\boldsymbol{\eta}_3]=[\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3]\begin{bmatrix} -5 & 0 & -5 \\ 0 & -1 & -1 \\ 3 & 6 & 9 \end{bmatrix},$$

又

$$[\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\boldsymbol{\eta}_3]=[\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3]\begin{bmatrix} -1 & 0 & 3 \\ 0 & 1 & -1 \\ 2 & 1 & 0 \end{bmatrix},$$

故

$$\begin{aligned} T[\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3] &= [\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3]\begin{bmatrix} -5 & 0 & -5 \\ 0 & -1 & -1 \\ 3 & 6 & 9 \end{bmatrix}\begin{bmatrix} -1 & 0 & 3 \\ 0 & 1 & -1 \\ 2 & 1 & 0 \end{bmatrix}^{-1} \\ &= [\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3]\cdot\frac{1}{7}\begin{bmatrix} -5 & 20 & -20 \\ -4 & -5 & -2 \\ 27 & 18 & 24 \end{bmatrix}. \end{aligned}$$

2.183

得 T 在基底 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3\}$ 下的矩阵表示为

$$\boldsymbol{A}=\frac{1}{7}\begin{bmatrix} -5 & 20 & -20 \\ -4 & -5 & -2 \\ 27 & 18 & 24 \end{bmatrix}.$$

由

$$\begin{vmatrix} -5 & 20 & -20 \\ -4 & -5 & -2 \\ 27 & 18 & 24 \end{vmatrix}=0,$$

知矩阵不满秩, 而列向量 $\boldsymbol{\alpha}_1=(-5,-4,27)^{\mathrm{T}}$, $\boldsymbol{\alpha}_2=(20,-5,18)^{\mathrm{T}}$ 不成比例, $\boldsymbol{A}$ 的列向量的一个极大无关组为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$, 故

$$R(\boldsymbol{A})=\operatorname{span}[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2],$$

从而

$$R(T)=\operatorname{span}[-5\boldsymbol{e}_1-4\boldsymbol{e}_2+27\boldsymbol{e}_3,\ 20\boldsymbol{e}_1-5\boldsymbol{e}_2+18\boldsymbol{e}_3].$$

且 $\dim R(T)=2$. □ 2.184

Theorem 136. 设给定数域 F 上的 n 维线性空间 X 和 m 维线性空间 Y, 其基底分别为 $\mathscr{B}_X$ 和 $\mathscr{B}_Y$, 则所有从 X 到 Y 的线性变换的集合

$$L(X,Y)=\{T \mid T \text{ 为从 } X \text{ 到 } Y \text{ 的线性变换}\}$$

与矩阵空间 $F^{m\times n}$ 之间存在一一到上的映射. 2.185

4.2 线性空间的同构

Definition 137. 设 X 和 Y 为数域 F 上的线性空间. 如果存在从 X 到 Y 的映射 Γ 满足:

1. Γ 是 X 到 Y 的双射;
2. Γ 是 X 到 Y 上的线性映射,

则称线性空间 X 同构于线性空间 Y, 记为 $X \simeq Y$, 并称映射 Γ 是 X 到 Y 上的同构映射(isomorphism).

特别地, 当 $Y=X$ 时, 则称 X 是自同构的. 并称 Γ 为 X 上的自同构映射, 简称为 X 上的同构映射. 2.186

研究同构的目的

如果忽略同构对象的属性或运算的具体定义, 单从结构上讲, 同构的对象是等价的.

研究同构的主要目的是为了把理论应用于不同的领域. 如果两个结构是同构的, 那么其上的对象会有相似的属性和运算, 对某个结构成立的命题在另一个结构上也就成立. 因此, 如果在某个数学领域发现了一个对象结构同构于某个结构, 且对于该结构已经证明了很多定理, 那么这些定理马上就可以应用到该领域. 如果某些数学方法可以用于该结构, 那么这些方法也可以用于新领域的结构. 这就使得理解和处理该对象结构变得容易, 并往往可以让人们对该领域有更深刻的理解. 2.187

Theorem 138. 给定数域 F 上的 n 维线性空间 X 和 m 维线性空间 Y, 则线性空间 $L(X,Y)$ 与矩阵空间 $F^{m\times n}$ 是同构的.

Proof. 在给定 X 与 Y 的基底 $\mathscr{B}_X$ 和 $\mathscr{B}_Y$ 之下, $L(X,Y)$ 与 $F^{m\times n}$ 之间存在一个一一到上的映射 Γ. 下证 Γ 是线性的.

$\forall T_1, T_2 \in L(X,Y)$, 设它们关于基底 $\mathscr{B}_X$, $\mathscr{B}_Y$ 的矩阵表示分别为 $\boldsymbol{A}_1$, $\boldsymbol{A}_2$, 即 $T_1\mathscr{B}_X=\mathscr{B}_Y\boldsymbol{A}_1$, $T_2\mathscr{B}_X=\mathscr{B}_Y\boldsymbol{A}_2$. $\forall\lambda\in F$, 有

$$(\lambda T_1+T_2)\mathscr{B}_X=\mathscr{B}_Y(\lambda\boldsymbol{A}_1+\boldsymbol{A}_2).$$

由 Γ 的定义可知 $\Gamma T_1=\boldsymbol{A}_1$, $\Gamma T_2=\boldsymbol{A}_2$, 故

$$\Gamma(\lambda T_1+T_2)=(\lambda\boldsymbol{A}_1+\boldsymbol{A}_2)=\lambda\Gamma T_1+\Gamma T_2.$$

知 Γ 是线性映射, 从而是同构映射. 得证空间 $L(X,Y)$ 与空间 $F^{m\times n}$ 同构. $\square$

2.188

Exercise 139 (P.83 习题 (二) 2)**.** 设 X 和 Y 均为数域 F 上的有限维线性空间, 且 Γ 为从 X 到 Y 的同构映射. 则 X 中的向量组 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_t$ 线性无关当且仅当 $\Gamma(\boldsymbol{x}_1), \Gamma(\boldsymbol{x}_2), \cdots, \Gamma(\boldsymbol{x}_t)$ 线性无关.

证: 记 $\mathbf{0}_X, \mathbf{0}_Y$ 分别为 X 和 Y 的零向量, 则 $\Gamma(\mathbf{0}_X) = \mathbf{0}_Y$. 因为 Γ 是 X 到 Y 的一个单射, 所以若 $\Gamma(\boldsymbol{\alpha}) = \Gamma(\boldsymbol{\beta})$, 则 $\boldsymbol{\alpha} = \boldsymbol{\beta}$. 于是有

$$\begin{aligned} & k_1\boldsymbol{x}_1 + k_2\boldsymbol{x}_2 + \cdots + k_t\boldsymbol{x}_t = \mathbf{0}_X \\ \Longleftrightarrow\ & \Gamma\big(k_1\boldsymbol{x}_1 + k_2\boldsymbol{x}_2 + \cdots + k_t\boldsymbol{x}_t\big) = \Gamma\big(\mathbf{0}_X\big) \\ \Longleftrightarrow\ & k_1\Gamma(\boldsymbol{x}_1) + k_2\Gamma(\boldsymbol{x}_2) + \cdots + k_t\Gamma(\boldsymbol{x}_t) = \mathbf{0}_Y. \end{aligned}$$

从而 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_t$ 线性无关当且仅当 $\Gamma(\boldsymbol{x}_1), \Gamma(\boldsymbol{x}_2), \cdots, \Gamma(\boldsymbol{x}_t)$ 线性无关. □

☞ 对于一般的线性变换, 只能保证 "若 $\Gamma(\boldsymbol{x}_1), \Gamma(\boldsymbol{x}_2), \cdots, \Gamma(\boldsymbol{x}_t)$ 线性无关, 则 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_t$ 线性无关", 反之不成立 (见教材 P.67 定理 2.3.4, 或本文定理 109).

2.189

Example 140. 如果 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n$ 是 X 的一组基底, 那么 $\Gamma(\boldsymbol{x}_1), \Gamma(\boldsymbol{x}_2), \cdots, \Gamma(\boldsymbol{x}_n)$ 是 Y 的一组基底. 其中 Γ 为从 X 到 Y 的同构映射.

证: 因 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n$ 线性无关, 由上题得, $\Gamma(\boldsymbol{x}_1), \Gamma(\boldsymbol{x}_2), \cdots, \Gamma(\boldsymbol{x}_n)$ 线性无关. 下证 Y 中的任意向量 $\boldsymbol{y}$ 都可以由 $\Gamma(\boldsymbol{x}_1), \Gamma(\boldsymbol{x}_2), \cdots, \Gamma(\boldsymbol{x}_n)$ 线性表示.

由于 Γ 是 X 到 Y 的一个满射, 因此存在 $\boldsymbol{x} \in X$, 使得 $\Gamma(\boldsymbol{x}) = \boldsymbol{y}$. 而 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n$ 是 X 的一组基底, 故

$$\boldsymbol{x} = k_1\boldsymbol{x}_1 + k_2\boldsymbol{x}_2 + \cdots + k_n\boldsymbol{x}_n.$$

从而

$$\boldsymbol{y} = k_1\Gamma(\boldsymbol{x}_1) + k_2\Gamma(\boldsymbol{x}_2) + \cdots + k_n\Gamma(\boldsymbol{x}_n).$$

得证 $\Gamma(\boldsymbol{x}_1), \Gamma(\boldsymbol{x}_2), \cdots, \Gamma(\boldsymbol{x}_n)$ 是 Y 的一组基底. □

☞ 由此立即得到, 若 $\dim X = n$, 且 $X \simeq Y$, 则 $\dim Y = \dim X = n$. 即言, 同构的有限维线性空间其维数相等. 反之是否成立? 回答是肯定的. 2.190

Theorem 141. 数域 F 上的两个有限维的线性空间 X 和 Y 同构的充分必要条件是它们的维数相等.

证: 上面已证必要性, 现在来证充分性.

设 X 和 Y 均为数域 F 上的 n 维线性空间. 在 X 和 Y 中各取一组基底: $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n$; $\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_n$. 令

$$\begin{aligned} \Gamma : X &\longrightarrow Y \\ \boldsymbol{x} = \sum_{i=1}^{n} k_i\boldsymbol{x}_i &\longmapsto \sum_{i=1}^{n} k_i\boldsymbol{y}_i. \end{aligned}$$

对 Y 中任一向量 $\boldsymbol{y}^* = \sum\limits_{i=1}^{n} \eta_i\boldsymbol{y}_i$, 可以在 X 中找到唯一的 $\boldsymbol{x}^* = \sum\limits_{i=1}^{n} \eta_i\boldsymbol{x}_i$ 作为其原像, 故 Γ 是双射. 2.191

任取 $\boldsymbol{\alpha}=\sum\limits_{i=1}^{n}a_i\boldsymbol{x}_i$, $\boldsymbol{\beta}=\sum\limits_{i=1}^{n}b_i\boldsymbol{x}_i\in X$, 又任取 $\lambda\in F$, 则由 Γ 的定义有

$$\begin{aligned}\Gamma(\lambda\boldsymbol{\alpha}+\boldsymbol{\beta})&=\Gamma\Big[\sum_{i=1}^{n}(\lambda a_i+b_i)\boldsymbol{x}_i\Big]\\&=\sum_{i=1}^{n}(\lambda a_i+b_i)\boldsymbol{y}_i\\&=\lambda\sum_{i=1}^{n}a_i\boldsymbol{y}_i+\sum_{i=1}^{n}b_i\boldsymbol{y}_i\\&=\lambda\Gamma(\boldsymbol{\alpha})+\Gamma(\boldsymbol{\beta}).\end{aligned}$$

因此, Γ 是线性变换. 综上所述, Γ 是从 X 到 Y 的同构映射, 得证 X 和 Y 同构. □ 2.192

Corollary 142. *设数域 F 上的线性空间 X 和 Y 的维数分别为 n 和 m, 则空间 $L(X,Y)$ 的维数为 mn.*

Corollary 143. *数域 F 上任意 n 维线性空间都与空间 F^n 同构. 特别地, 复数域 $\mathbb{C}$ 上的 n 维线性空间都与空间 $\mathbb{C}^n$ 同构. 实数域 $\mathbb{R}$ 上的 n 维线性空间都与空间 $\mathbb{R}^n$ 同构.*

2.193

☞ 数域 F 上任意 n 维线性空间 V 都与空间 F^n 同构, 并且可以如下建立 V 到 F^n 的一个同构映射. 取 V 的一个基底 $\boldsymbol{\alpha}_1$, $\boldsymbol{\alpha}_2$, $\cdots$, $\boldsymbol{\alpha}_n$, 取 F^n 的标准基 $\boldsymbol{\varepsilon}_1$, $\boldsymbol{\varepsilon}_2$, $\cdots$, $\boldsymbol{\varepsilon}_n$. 令

$$\Gamma:\ \boldsymbol{\alpha}=\sum_{i=1}^{n}a_i\boldsymbol{\alpha}_i\ \longmapsto\ \sum_{i=1}^{n}a_i\boldsymbol{\varepsilon}_i=(a_1,a_2,\cdots,a_n)^{\mathrm{T}},$$

即把 V 中每一个向量 $\boldsymbol{\alpha}$ 对应到它在 V 的一个基底下的坐标, 这就是 V 到 F^n 的一个同构映射.

由于数域 F 上任意 n 维线性空间 V 都与空间 F^n 同构, 因此可以利用 F^n 的性质来研究 V 的性质, 这是研究数域 F 上有限维线性空间的重要途径. 2.194

Corollary 144. *设 $T\in L(X,Y)$, X 和 Y 均为数域 F 上的线性空间, $\dim X=n$, $\dim Y=m$. 又设 $\mathscr{B}_X$ 和 $\mathscr{B}_Y$ 分别为 X 和 Y 的基底, T 关于 $\mathscr{B}_X$ 和 $\mathscr{B}_Y$ 的矩阵表示为 $m_{\mathscr{B}_X,\mathscr{B}_Y}(T)=\boldsymbol{A}$, 则有*

$$R(T)\simeq R(\boldsymbol{A}),\tag{15}$$

$$N(T)\simeq N(\boldsymbol{A}).\tag{16}$$

2.195

Exercise 145 (P.82 习题 (一) 1)**.** 求线性变换在给定基底下的矩阵表示. 给定空间

$$V^6=\mathrm{span}[\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\boldsymbol{\xi}_3,\boldsymbol{\xi}_4,\boldsymbol{\xi}_5,\boldsymbol{\xi}_6],$$

其中

$$\boldsymbol{\xi}_1=\mathrm{e}^{\alpha x}\cos\beta x,\qquad\qquad \boldsymbol{\xi}_2=\mathrm{e}^{\alpha x}\sin\beta x$$

$$\boldsymbol{\xi}_3 = x\mathrm{e}^{\alpha x}\cos\beta x, \qquad \boldsymbol{\xi}_4 = x\mathrm{e}^{\alpha x}\sin\beta x$$
$$\boldsymbol{\xi}_5 = \frac{1}{2}x^2\mathrm{e}^{\alpha x}\cos\beta x, \qquad \boldsymbol{\xi}_6 = \frac{1}{2}x^2\mathrm{e}^{\alpha x}\sin\beta x.$$

$T = \frac{\mathrm{d}}{\mathrm{d}x}$. 求 T 在基底 $\{\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \boldsymbol{\xi}_3, \boldsymbol{\xi}_4, \boldsymbol{\xi}_5, \boldsymbol{\xi}_6\}$ 下的矩阵表示, 并求 $\dim R(T)$.

解: 计算基底的导数, 并用基底线性表示, 得: 2.196

$$\begin{aligned}
T(\boldsymbol{\xi}_1) &= (\mathrm{e}^{\alpha x}\cos\beta x)' = \alpha\boldsymbol{\xi}_1 - \beta\boldsymbol{\xi}_2,\\
T(\boldsymbol{\xi}_2) &= (\mathrm{e}^{\alpha x}\sin\beta x)' = \beta\boldsymbol{\xi}_1 + \alpha\boldsymbol{\xi}_2,\\
T(\boldsymbol{\xi}_3) &= (x\mathrm{e}^{\alpha x}\cos\beta x)' = \boldsymbol{\xi}_1 + \alpha\boldsymbol{\xi}_3 - \beta\boldsymbol{\xi}_4,\\
T(\boldsymbol{\xi}_4) &= (x\mathrm{e}^{\alpha x}\sin\beta x)' = \boldsymbol{\xi}_2 + \beta\boldsymbol{\xi}_3 + \alpha\boldsymbol{\xi}_4\\
T(\boldsymbol{\xi}_5) &= (\frac{1}{2}x^2\mathrm{e}^{\alpha x}\cos\beta x)' = \boldsymbol{\xi}_3 + \alpha\boldsymbol{\xi}_5 - \beta\boldsymbol{\xi}_6,\\
T(\boldsymbol{\xi}_6) &= (\frac{1}{2}x^2\mathrm{e}^{\alpha x}\sin\beta x)' = \boldsymbol{\xi}_4 + \beta\boldsymbol{\xi}_5 + \alpha\boldsymbol{\xi}_6.
\end{aligned}$$

故 T 在基底 $\{\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \boldsymbol{\xi}_3, \boldsymbol{\xi}_4, \boldsymbol{\xi}_5, \boldsymbol{\xi}_6\}$ 下的矩阵表示为

$$\boldsymbol{A} = \begin{bmatrix} \alpha & \beta & 1 & 0 & 0 & 0\\ -\beta & \alpha & 0 & 1 & 0 & 0\\ 0 & 0 & \alpha & \beta & 1 & 0\\ 0 & 0 & -\beta & \alpha & 0 & 1\\ 0 & 0 & 0 & 0 & \alpha & \beta\\ 0 & 0 & 0 & 0 & -\beta & \alpha \end{bmatrix}.$$

2.197

由 $\alpha \neq 0$, $\beta \neq 0$ (否则 $\boldsymbol{\xi}_1$, $\boldsymbol{\xi}_2$ 线性相关), 得

$$\det \boldsymbol{A} = (\alpha^2 + \beta^2)^3 \neq 0,$$

知 $\dim R(T) = \operatorname{rank}(\boldsymbol{A}) = 6$. □ 2.198

Exercise 146 (P.83 习题 (一) 4)**.** 设 $\boldsymbol{\xi}_1$, $\boldsymbol{\xi}_2$, $\boldsymbol{\xi}_3$, $\boldsymbol{\xi}_4$ 是 4 维线性空间 V 的一组基, 已知线性变换 T 在这组基底之下的矩阵表示为

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 2 & 1\\ -1 & 2 & 1 & 3\\ 1 & 2 & 5 & 5\\ 2 & -2 & 1 & -2 \end{bmatrix}.$$

1. 求 T 在基 $\boldsymbol{\eta}_1 = \boldsymbol{\xi}_1 - 2\boldsymbol{\xi}_2 + \boldsymbol{\xi}_4$, $\boldsymbol{\eta}_2 = 3\boldsymbol{\xi}_2 - \boldsymbol{\xi}_3 - \boldsymbol{\xi}_4$, $\boldsymbol{\eta}_3 = \boldsymbol{\xi}_3 + \boldsymbol{\xi}_4$, $\boldsymbol{\eta}_4 = 2\boldsymbol{\xi}_4$ 下的矩阵表示.
2. 求 $N(T)$ 和 $R(T)$.
3. 在 $N(T)$ 中选一组基底, 把它扩充成 V 的一组基底, 并求 T 在这组基底下的矩阵表示.
4. 在 $R(T)$ 中选一组基底, 把它扩充成 V 的一组基底, 并求 T 在这组基底下的矩阵表示.

2.199

解: (1)

$$[\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\boldsymbol{\eta}_3,\boldsymbol{\eta}_4]=[\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\boldsymbol{\xi}_3,\boldsymbol{\xi}_4]\begin{bmatrix}1&0&0&0\\-2&3&0&0\\0&-1&1&0\\1&-1&1&2\end{bmatrix},$$

记上式右端的矩阵为 $\boldsymbol{B}$. 则 T 在基底 $\boldsymbol{\eta}_1$, $\boldsymbol{\eta}_2$, $\boldsymbol{\eta}_3$, $\boldsymbol{\eta}_4$ 下的矩阵表示为

$$\boldsymbol{Q}=\boldsymbol{B}^{-1}\boldsymbol{A}\boldsymbol{B}=\begin{bmatrix}2&-3&3&2\\\frac{2}{3}&-\frac{4}{3}&\frac{10}{3}&\frac{10}{3}\\\frac{8}{3}&-\frac{16}{3}&\frac{40}{3}&\frac{40}{3}\\0&1&-7&-8\end{bmatrix}.$$

2.200

(2) 把矩阵 $\boldsymbol{A}$ 经过初等行变换化成行阶梯形矩阵:

$$\boldsymbol{A}=\begin{bmatrix}1&0&2&1\\-1&2&1&3\\1&2&5&5\\2&-2&1&-2\end{bmatrix}\longrightarrow\begin{bmatrix}1&0&2&1\\0&1&\frac{3}{2}&2\\0&0&0&0\\0&0&0&0\end{bmatrix}.$$

于是 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 的基础解系为

$$\boldsymbol{x}_1=(4,3,-2,0)^{\mathrm{T}},\qquad \boldsymbol{x}_2=(1,2,0,-1)^{\mathrm{T}}.$$

从而 $N(\boldsymbol{A})=\mathrm{span}[\boldsymbol{x}_1,\boldsymbol{x}_2]$. 于是

$$N(T)=\mathrm{span}[4\boldsymbol{\xi}_1+3\boldsymbol{\xi}_2-2\boldsymbol{\xi}_3,\ \boldsymbol{\xi}_1+2\boldsymbol{\xi}_2-\boldsymbol{\xi}_4].$$

$\boldsymbol{A}$ 的列向量的一个极大无关组是

$$\boldsymbol{\alpha}_1=(1,-1,1,2)^{\mathrm{T}},\qquad \boldsymbol{\alpha}_2=(0,2,2,-2)^{\mathrm{T}},$$

从而 $R(\boldsymbol{A})=\mathrm{span}[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2]$. 故

$$R(T)=\mathrm{span}[\boldsymbol{\xi}_1-\boldsymbol{\xi}_2+\boldsymbol{\xi}_3+2\boldsymbol{\xi}_4,\ 2\boldsymbol{\xi}_2+2\boldsymbol{\xi}_3-2\boldsymbol{\xi}_4].$$

2.201

(3) 先把 $N(\boldsymbol{A})$ 的基底 $\boldsymbol{x}_1,\boldsymbol{x}_2$ 扩充成 $\mathbb{R}^4$ 的一组基底:

$$\boldsymbol{x}_1=(4,3,-2,0)^{\mathrm{T}},\quad \boldsymbol{x}_2=(1,2,0,1)^{\mathrm{T}},\quad \boldsymbol{x}_3=(0,1,0,0)^{\mathrm{T}},\qquad \boldsymbol{x}_4=(1,0,0,0)^{\mathrm{T}}.$$

于是 $N(T)$ 的基底扩充为 V 的一个基底为

$$4\boldsymbol{\xi}_1+3\boldsymbol{\xi}_2-2\boldsymbol{\xi}_3,\ \boldsymbol{\xi}_1+2\boldsymbol{\xi}_2-\boldsymbol{\xi}_4,\ \boldsymbol{\xi}_2,\ \boldsymbol{\xi}_1.$$

V 的基底 $\boldsymbol{\xi}_1$, $\boldsymbol{\xi}_2$, $\boldsymbol{\xi}_3$, $\boldsymbol{\xi}_4$ 到上述基底的过渡矩阵为

$$\boldsymbol{P}=\begin{bmatrix}4&1&0&1\\3&2&1&0\\-2&0&0&0\\0&-1&0&0\end{bmatrix}.$$

从而 T 在基底 $4\boldsymbol{\xi}_1+3\boldsymbol{\xi}_2-2\boldsymbol{\xi}_3$, $\boldsymbol{\xi}_1+2\boldsymbol{\xi}_2-\boldsymbol{\xi}_4$, $\boldsymbol{\xi}_2$, $\boldsymbol{\xi}_1$ 下的矩阵 $\boldsymbol{C}$ 为

$$\boldsymbol{C}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{bmatrix} 0 & 0 & -1 & -\frac{1}{2} \\ 0 & 0 & 2 & -2 \\ 0 & 0 & 1 & \frac{9}{2} \\ 0 & 0 & 2 & 5 \end{bmatrix}.$$

2.202

(4) 先把 $R(A)$ 的基底扩充为 $\mathbb{R}^4$ 上的一个基底:

$$\boldsymbol{\alpha}_1=(1,-1,1,2)^{\mathrm{T}},\quad \boldsymbol{\alpha}_2=(0,2,2,-2)^{\mathrm{T}},\quad \boldsymbol{\alpha}_3=(0,1,0,0)^{\mathrm{T}},\quad \boldsymbol{\alpha}_4=(1,0,0,0)^{\mathrm{T}},$$

则 $R(T)$ 的基底扩充为 V 的一个基底:

$$\boldsymbol{\xi}_1-\boldsymbol{\xi}_2+\boldsymbol{\xi}_3+2\boldsymbol{\xi}_4,\ 2\boldsymbol{\xi}_2+2\boldsymbol{\xi}_3-2\boldsymbol{\xi}_4,\ \boldsymbol{\xi}_2,\ \boldsymbol{\xi}_1.$$

V 的基底 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\boldsymbol{\xi}_3,\boldsymbol{\xi}_4$ 到上述基底的过渡矩阵为

$$\boldsymbol{Q}=\begin{bmatrix} 1 & 0 & 0 & 1 \\ -1 & 2 & 1 & 0 \\ 1 & 2 & 0 & 0 \\ 2 & -2 & 0 & 0 \end{bmatrix}.$$

于是 T 在基底 $\boldsymbol{\xi}_1-\boldsymbol{\xi}_2+\boldsymbol{\xi}_3+2\boldsymbol{\xi}_4, 2\boldsymbol{\xi}_2+2\boldsymbol{\xi}_3-2\boldsymbol{\xi}_4, \boldsymbol{\xi}_2, \boldsymbol{\xi}_1$ 下的矩阵表示为

$$\boldsymbol{D}=\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}=\begin{bmatrix} 5 & 2 & 0 & 1 \\ \frac{9}{2} & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \quad \square$$

2.203

5 线性变换的最简矩阵表示

从这一节开始, 我们将针对数域 F 上 n 维线性空间 X 上的线性变换 T, 研究如何找出 X 的一组适当的基底, 使得 T 在这个基底下的矩阵具有最简单的形式.

首先研究对于线性变换 T, 能不能找到 X 的一组基底, 使得 T 在这个基底下的矩阵为对角矩阵.

T 在 X 的基底 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$ 下的矩阵为对角矩阵

$$\boldsymbol{\Lambda}=\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \lambda_3 & \\ & & & \lambda_4 \end{pmatrix}$$

$\iff T(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n)=(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n)\boldsymbol{\Lambda}=(\lambda_1\boldsymbol{x}_1,\lambda_2\boldsymbol{x}_2,\cdots,\lambda_n\boldsymbol{x}_n)$

$\iff T\boldsymbol{x}_i=\lambda_i\boldsymbol{x}_i,\ i=1,2,\cdots,n.$ 2.204

5.1 线性变换的特征值与特征向量

Definition 147. 设 X 为数域 F 上的 n 维线性空间, $T \in L(X,X)$. 如果存在 $\lambda \in F$, $\boldsymbol{x} \in X$, 且 $\boldsymbol{x} \neq \mathbf{0}$, 满足

$$T\boldsymbol{x} = \lambda \boldsymbol{x},$$

则称 λ 为 T 的一个特征值, 而 $\boldsymbol{x}$ 称为 T 的属于 λ 的特征向量. 所有满足方程 $T\boldsymbol{x} = \lambda\boldsymbol{x}$ 的向量 (包括零向量) 的集合, 构成 X 的一个子空间, 称为 T 关于 λ 的特征子空间, 记为 $E_T(\lambda)$.

☞ 特征向量不是零向量; 或者说, 零向量一定不是特征向量. 2.205

设 X 为数域 F 上的 n 维线性空间, $T \in L(X,X)$, T 在 X 的某一基底 $\mathscr{B}_X$ 下的矩阵表示为 $\boldsymbol{A}$. 向量 $\boldsymbol{x} \in X$ 在基底 $\mathscr{B}_X$ 下的坐标为 $\boldsymbol{\xi}$, $\boldsymbol{y} = T\boldsymbol{x}$ 在基底 $\mathscr{B}_X$ 下的坐标为 $\boldsymbol{\eta}$. 则

$$T\boldsymbol{x} = \boldsymbol{y} \quad \Longleftrightarrow \quad \boldsymbol{A}\boldsymbol{\xi} = \boldsymbol{\eta}.$$

若 $T\boldsymbol{x} = \lambda\boldsymbol{x}$, 因 $\lambda\boldsymbol{x}$ 在基底 $\mathscr{B}_X$ 下的坐标为 $\lambda\boldsymbol{\xi}$, 故

$$T\boldsymbol{x} = \lambda\boldsymbol{x} \quad \Longleftrightarrow \quad \boldsymbol{A}\boldsymbol{\xi} = \lambda\boldsymbol{\xi}.$$

由此得出下述结论:

(1) λ 是 T 的特征值 $\Longleftrightarrow$ λ 是 $\boldsymbol{A}$ 的特征值.

(2) $\boldsymbol{x}$ 是 T 的属于特征值 λ 的特征向量 $\Longleftrightarrow$ $\boldsymbol{\xi}$ 是 $\boldsymbol{A}$ 的属于特征值 λ 的特征向量.

2.206

其完整叙述为下面两个定理.

Theorem 148. *设 X 为数域 F 上的 n 维线性空间, $T \in L(X,X)$. 则 λ 为 T 的特征值的充分必要条件是: λ 为 T 的任一矩阵表示 $\boldsymbol{A}$ 的特征值.*

"任一矩阵"? —— 这是因为 T 的不同矩阵表示是相似的, 它们有完全相同的特征值.

Theorem 149. *设 X 为数域 F 上的 n 维线性空间, $T \in L(X,X)$, T 关于基底 $\mathscr{B}_X$ 的矩阵表示为 $\boldsymbol{A}$, 又非零向量 $\boldsymbol{x} \in X$ 关于基底 $\mathscr{B}_X$ 的坐标为 $\boldsymbol{\xi} \in F^n$. 则 $\boldsymbol{\xi}$ 为 $\boldsymbol{A}$ 的属于特征值 λ 的特征向量的充分必要条件是: $\boldsymbol{x}$ 为 T 的属于特征值 λ 的特征向量.*

2.207

确定一个线性变换 T 的特征值与特征向量的方法可以分成以下几步:

1. 在线性空间 V 中取一组基 $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n$, 写出 T 在这组基下的矩阵表示 $\boldsymbol{A}$;
2. 求出 $\boldsymbol{A}$ 的全部特征值和特征向量;
3. 矩阵 $\boldsymbol{A}$ 的全部特征值就是线性变换 T 的全部特征值; 矩阵 $\boldsymbol{A}$ 的特征向量是线性变换 T 的特征向量在基底 $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n$ 下的坐标.

   即, 如果 $\boldsymbol{A}$ 关于特征值 λ_i 的特征向量为 $\xi_{i_1}, \cdots, \xi_{i_k}$, 则 T 关于特征值 λ_i 的特征向量为

$$(\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n)\xi_{i_1}, \ \cdots, \ (\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n)\xi_{i_k}.$$

2.208

Example 150. 设线性变换 T 在基 $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3$ 下的矩阵表示为

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix},$$

求 T 的特征值与特征向量.

解: 因为特征多项式为

$$|\lambda \boldsymbol{I} - \boldsymbol{A}| = \begin{vmatrix} \lambda - 1 & -2 & -2 \\ -2 & \lambda - 1 & -2 \\ -2 & -2 & \lambda - 1 \end{vmatrix} = (\lambda + 1)^2(\lambda - 5).$$

所以特征值是 $\lambda_1 = \lambda_2 = -1$, $\lambda_3 = 5$. 2.209

将 -1 代入方程 $(\lambda \boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$, 得基础解系为

$$\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}.$$

因此, T 的属于 -1 的两个线性无关的特征向量就是

$$\boldsymbol{\eta}_1 = \boldsymbol{x}_1 - \boldsymbol{x}_3, \qquad \boldsymbol{\eta}_2 = \boldsymbol{x}_2 - \boldsymbol{x}_3.$$

而 T 的属于 -1 的全部特征向量就是 $k_1\boldsymbol{\eta}_1 + k_2\boldsymbol{\eta}_2$, $k_1, k_2 \in F$ 且不同时为零.

再用特征值 5 代入 $(\lambda \boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$, 它的基础解系是

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

因此, T 的属于 5 的一个线性无关的特征向量就是

$$\boldsymbol{\eta}_3 = \boldsymbol{x}_1 + \boldsymbol{x}_2 + \boldsymbol{x}_3,$$

而 T 的属于 5 的全部特征向量就是 $k\boldsymbol{\eta}_3$, $k \in F$ 且 $k \neq 0$. $\square$ 2.210

Corollary 151. *设 T 的关于某一特征值 λ 的特征子空间为 $E_T(\lambda)$, 则 $E_T(\lambda)$ 中所有向量关于基底 $\mathscr{B}_X$ 的坐标向量的全体, 构成 $\boldsymbol{A}$ 的关于 λ 的特征子空间 $E_{\boldsymbol{A}}(\lambda)$. 反之亦然.*

2.211

☞ 求线性变换 T 的特征值、特征向量及特征子空间的问题, 完全转换为其矩阵表示 $\boldsymbol{A}$ 的相应问题.

可以证明: 对于一个固定的特征值 λ, 空间 $E_{\boldsymbol{A}}(\lambda)$ 与空间 $E_T(\lambda)$ 是同构的. 2.212

线性变换的特征多项式

以下总假设 X 是复数域 $\mathbb{C}$ 上的线性空间, $T \in L(X,X)$, $\boldsymbol{A} = m_{\mathscr{B}_X}(T)$, $\dim X = n$. 于是 $\boldsymbol{A}$ 为 n 阶复矩阵, 并且其特征多项式为

$$f(\lambda) = \det[\lambda \boldsymbol{I} - \boldsymbol{A}] = \lambda^n + \sum_{i=1}^{n-1} \alpha_i \lambda^{n-i}, \qquad \alpha_i \in \mathbb{C}.$$

$f(\lambda)$ 的全部根就是 $\boldsymbol{A}$ 的特征值, 因而也是 T 的特征值. 因此也把 $\boldsymbol{A}$ 的特征多项式称为 T 的特征多项式. 2.213

如果特征多项式 $f(\lambda)$ 恰有 k 个互异的根, 则线性变换 T 恰有 k 个互异的特征子空间. 相应地, 其矩阵表示 $\boldsymbol{A}$ 也恰有 k 个互异的特征子空间. 2.214

Definition 152. 设 $f(\lambda)$ 为线性变换 T 的特征多项式, λ_i $(i = 1, 2, \cdots, k)$ 为 $f(\lambda)$ 的 k 个互异的根, 其重数为 d_i $(i = 1, 2, \cdots, k)$,

$$f(\lambda) = (\lambda - \lambda_1)^{d_1} (\lambda - \lambda_2)^{d_2} \cdots (\lambda - \lambda_k)^{d_k},$$

则称 d_i 为 λ_i 的代数重数, 而称 $t_i = \dim E_T(\lambda_i)$ 为 λ_i 的几何重数.

Theorem 153. *设 $T \in L(X,X)$, λ_1, λ_2, $\cdots$, λ_k 为其 k 个互异的特征值, 则其任一特征值 λ_i $(i = 1, 2, \cdots, k)$ 的几何重数 t_i 不大于其代数重数 d_i, 即 $t_i \leqslant d_i$.* 2.215

Theorem 154. *设 $T \in L(X,X)$, 则 T 的属于不同特征值的特征向量必线性无关.* 2.216

Corollary 155. *设 $T \in L(X,X)$, 则 T 的矩阵表示 $\boldsymbol{A}$ 的属于不同特征值的特征向量必线性无关.* 2.217

Corollary 156. *设 $T \in L(X,X)$, λ_1, λ_2, $\cdots$, λ_k 为 T 的所有互异的特征值, 则 T 的矩阵表示 $\boldsymbol{A}$ 相似于对角矩阵的充分必要条件为: T 的各个特征值的几何重数均等于其代数重数.* 2.218

Corollary 157. *设 $T \in L(X,X)$, $\dim X = n$, 则 T 的矩阵表示 $\boldsymbol{A}$ 相似于对角矩阵的充分必要条件为: T 有 n 个线性无关的特征向量.* 2.219

Definition 158. 设 $T \in L(X,X)$, 且 T 的矩阵表示与对角矩阵相似, 则称 T 为 X 上可对角化的线性变换. 相应地, 与对角矩阵相似的方阵, 称为可对角化的矩阵. 2.220

Corollary 159. *$T \in L(X,X)$ 可对角化的充分必要条件为: 空间 X 可分解为 T 的一维不变子空间的直和.* 2.221

Corollary 160. *$T \in L(X,X)$ 可对角化的充分必要条件为: $\mathbb{C}^n$ 可分解为 T 的矩阵表示 $\boldsymbol{A}$ 的一维不变子空间 (将 $\boldsymbol{A}$ 看作 $\mathbb{C}^n$ 上的线性变换) 的直和.*

☞ 一个命题如果关于线性变换 $T \in L(X, X)$ 成立, 则它关于 T 的任一矩阵表示也必然成立, 只是在叙述时把空间 X 换成它的同构空间 $\mathbb{C}^n$ (如果 $\dim X = n$), 并把 X 中的向量 $\boldsymbol{x}$ 换成 $\mathbb{C}^n$ 中的向量 $\boldsymbol{\xi}$, 其中 $\boldsymbol{\xi}$ 是 $\boldsymbol{x}$ 的关于 X 的基底 $\mathscr{B}_X$ 的坐标向量. 2.222

反之亦然.

在以后的叙述中, 有时就只限于对线性变换 T 的矩阵表示 $\boldsymbol{A}$ 来进行, 而不每次都提及 T. 2.223

Corollary 161. *$T \in L(X, X)$ 的矩阵表示 $\boldsymbol{A}$ 可对角化的充分必要条件为: $\mathbb{C}^n$ 可分解为 $\boldsymbol{A}$ 全部特征子空间 $E_{\boldsymbol{A}}(\lambda_1)$, $E_{\boldsymbol{A}}(\lambda_2)$, $\cdots$, $E_{\boldsymbol{A}}(\lambda_k)$ 的直和. 这里 λ_1, λ_2, $\cdots$, λ_k 为 $\boldsymbol{A}$ 的全部互异特征值.* 2.224

Example 162. 设 T 是数域 F 上的 4 维线性空间 X 上的一个线性变换, 它在 X 的一个基底 $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, $\boldsymbol{x}_3$, $\boldsymbol{x}_4$ 下的矩阵表示为

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

求 X 的一个基底, 使得 T 在这个基底下的矩阵为对角矩阵, 并且写出这个矩阵. 2.225

解: (1) 先求 T 的特征值和特征向量.

$$|\lambda \boldsymbol{I} - \boldsymbol{A}| = \begin{vmatrix} \lambda - 1 & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ -1 & 0 & \lambda & 0 \\ 0 & 0 & 0 & \lambda - 1 \end{vmatrix} = \lambda^2(\lambda - 1)2.$$

故 $\boldsymbol{A}$ 的全部特征值为 $\lambda_1 = \lambda_2 = 0$, $\lambda_3 = \lambda_4 = 1$. 这也就是 T 的全部特征值.

对于特征值 $\lambda_1 = \lambda_2 = 0$, 解方程组 $(-\boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$, 求出一个基础解系:

$$(0, 1, 0, 0)^{\mathrm{T}}, \qquad (0, 0, 1, 0)^{\mathrm{T}}.$$

得到 T 的属于特征值 $\lambda_1 = \lambda_2 = 0$ 的两个线性无关的特征向量为

$$(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4)(0, 1, 0, 0)^{\mathrm{T}} = \boldsymbol{x}_2,$$
$$(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4)(0, 0, 1, 0)^{\mathrm{T}} = \boldsymbol{x}_3.$$

2.226

对于特征值 $\lambda_3 = \lambda_4 = 1$, 解方程组 $(\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$, 求出一个基础解系:

$$(1, 0, 1, 0)^{\mathrm{T}}, \qquad (0, 0, 0, 1)^{\mathrm{T}}.$$

得到 T 的属于特征值 $\lambda_3 = \lambda_4 = 1$ 的两个线性无关的特征向量为

$$(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4)(1, 0, 1, 0)^{\mathrm{T}} = \boldsymbol{x}_1 + \boldsymbol{x}_3,$$

$$(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4)(0,0,0,1)^{\mathrm{T}} = \boldsymbol{x}_4.$$

(2) 故线性变换 T 在 X 的一组基底

$$\boldsymbol{x}_2,\ \boldsymbol{x}_3,\ \boldsymbol{x}_1 + \boldsymbol{x}_3,\ \boldsymbol{x}_4$$

下的矩阵表示为对角阵

$$\boldsymbol{\Lambda} = \mathrm{diag}(0,0,1,1).$$

2.227

Exercise 163 (P.92 习题 (一) 1)**.** 设 X 为复数域 $\mathbb{C}$ 上的 4 维线性空间, 给定一组基底 $\mathscr{B}_X = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4\}$, X 上的线性变换 T 在基底 $\mathscr{B}_X$ 之下的矩阵表示为

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ -1 & -1 & 0 & 1 \end{bmatrix}.$$

求 T 的特征值及特征向量.

2.228

Exercise 164 (P.92 习题 (一) 2)**.** 设 T 为 $\mathbb{C}^3$ 上的线性变换, 给定 $\mathbb{C}^3$ 上的一组基底

$$\mathscr{B}_{\mathbb{C}^3} = \left\{(\mathrm{i}, 1, -1)^{\mathrm{T}},\ (0, -\mathrm{i}, 1,)^{\mathrm{T}},\ (1, 1, 2+\mathrm{i})^{\mathrm{T}}\right\},$$

T 关于基底 $\mathscr{B}_{\mathbb{C}^3}$ 的矩阵表示为

$$\boldsymbol{A} = \begin{bmatrix} 3 & -2 & -4 \\ -2 & 6 & -2 \\ -4 & -2 & 3 \end{bmatrix}.$$

求 T 的特征值及特征向量.

2.229

5.2 线性变换的零化多项式及最小多项式

零化多项式

Definition 165. 设 X 是数域 F 上的线性空间, T 是 X 上的一个线性变换. 如果 F 上的一元多项式 $g(x)$ 使得

$$g(T) = 0^*,$$

那么称 $g(x)$ 是 T 的一个零化多项式.

Definition 166. 设 $\boldsymbol{A}$ 是数域 F 上的 n 阶矩阵, 如果 $F[x]$ 中的一个多项式 $f(x)$ 使得 $f(\boldsymbol{A}) = \boldsymbol{O}$, 则称 $f(x)$ 是 $\boldsymbol{A}$ 的一个零化多项式.

2.230

Theorem 167. 设 $T \in L(X,X)$, $\boldsymbol{A}$ 为 T 的一个矩阵表示, 则多项式 $g(\lambda)$ 为 T 的零化多项式的 $\Leftrightarrow$ $g(\lambda)$ 为 $\boldsymbol{A}$ 的零化多项式.

证: 设 $g(\lambda) = \sum\limits_{i=0}^{m} \alpha_i \lambda^{m-i}$, 且 $T\mathscr{B} = \mathscr{B}\boldsymbol{A}$, 其中 $\mathscr{B}$ 为 T 的基底. 则

$$T^2\mathscr{B} = T\mathscr{B}\boldsymbol{A} = \mathscr{B}\boldsymbol{A}^2,$$
$$\vdots$$
$$T^m\mathscr{B} = \mathscr{B}\boldsymbol{A}^m,$$

又

$$E\mathscr{B} = \mathscr{B}\boldsymbol{I}$$

得

$$\begin{aligned}&(\alpha_0 T^m + \alpha_1 T^{m-1} + \cdots + \alpha_{m-1} T + \alpha_m E)\mathscr{B}\\ =&\mathscr{B}(\alpha_0 \boldsymbol{A}^m + \alpha_1 \boldsymbol{A}^{m-1} + \cdots + \alpha_{m-1}\boldsymbol{A} + \alpha_m \boldsymbol{I}).\end{aligned}$$

2.231

即

$$g(T)\mathscr{B} = \mathscr{B}g(\boldsymbol{A}).$$

若 $g(\lambda)$ 为 $\boldsymbol{A}$ 的零化多项式 (即 $g(\boldsymbol{A}) = \boldsymbol{O}$), 则

$$g(T)\mathscr{B} = \boldsymbol{0}.$$

任意 $\boldsymbol{x} \in X$, 设 $\boldsymbol{x} = \mathscr{B}\boldsymbol{\alpha}$, 有

$$g(T)\boldsymbol{x} = g(T)\mathscr{B}\boldsymbol{\alpha} = \boldsymbol{0}.$$

由 $\boldsymbol{x}$ 的任意性, 得

$$g(T) = 0^*.$$

反过来, 若已知 $g(T) = 0^*$, 则

$$\mathscr{B}g(\boldsymbol{A}) = g(T)\mathscr{B} = \boldsymbol{0},$$

故 $g(\boldsymbol{A}) = \boldsymbol{O}$. □ 2.232

Theorem 168. 设 T 为 n 维空间 X 上的线性变换, 则 T 必有零化多项式.

证: 因为 $T \in L(X,X)$, 又 $\dim L(X,X) = n^2$, 故 $L(X,X)$ 的 n^2+1 个向量

$$E, T, T^2, \cdots, T^{n^2}$$

必线性相关. 因而存在不全为零的 $\alpha_0, \alpha_1, \cdots, \alpha_{n^2}$, 使得

$$\sum_{i=0}^{n^2} \alpha_i T^{n^2-i} = 0^*.$$

则

$$g(\lambda) = \sum_{i=0}^{n^2} \alpha_i \lambda^{n^2-i}$$

为 T 的一个零化多项式. □ 2.233

Theorem 169 (哈密尔顿 — 凯莱 (Hamilton-Cayley) 定理)**.** 设 $\boldsymbol{A}$ 是数域 F 上一个 $n\times n$ 矩阵, $f(\lambda)=|\lambda\boldsymbol{I}-\boldsymbol{A}|$ 是 $\boldsymbol{A}$ 的特征多项式, 则

$$\begin{aligned} f(\boldsymbol{A}) &= \boldsymbol{A}^n-(a_{11}+a_{22}+\cdots+a_{nn})\boldsymbol{A}^{n-1}+\cdots+(-1)^n|\boldsymbol{A}|\boldsymbol{I} \\ &= \boldsymbol{O}. \end{aligned}$$

证: 设 $\boldsymbol{B}(\lambda)$ 是 λ-矩阵 $\lambda\boldsymbol{I}-\boldsymbol{A}$ 的伴随矩阵, 则

$$\boldsymbol{B}(\lambda)(\lambda\boldsymbol{I}-\boldsymbol{A})=|\lambda\boldsymbol{I}-\boldsymbol{A}|\boldsymbol{I}=f(\lambda)\boldsymbol{I}. \tag{17}$$

因为 $\boldsymbol{B}(\lambda)$ 的元素是矩阵 $\lambda\boldsymbol{I}-\boldsymbol{A}$ 的元素 (一次或零次多项式) 的代数余子式, $\boldsymbol{B}(\lambda)$ 的元素都是次数不超过 $n-1$ 的一元多项式. 从而 $\boldsymbol{B}(\lambda)$ 可以写成

$$\boldsymbol{B}(\lambda)=\lambda^{n-1}\boldsymbol{B}_{n-1}+\lambda^{n-2}\boldsymbol{B}_{n-2}+\cdots+\lambda\boldsymbol{B}_1+\boldsymbol{B}_0,$$

其中 $\boldsymbol{B}_{n-1}$, $\boldsymbol{B}_{n-2}$, $\cdots$, $\boldsymbol{B}_1$, $\boldsymbol{B}_0$ 都是数域 F 上的 n 阶矩阵. 2.234

从而

$$\begin{aligned} \boldsymbol{B}(\lambda)(\lambda\boldsymbol{I}-\boldsymbol{A}) &= \big(\lambda^{n-1}\boldsymbol{B}_{n-1}+\lambda^{n-2}\boldsymbol{B}_{n-2}+\cdots+\lambda\boldsymbol{B}_1+\boldsymbol{B}_0\big)(\lambda\boldsymbol{I}-\boldsymbol{A}) \\ &= \lambda^n\boldsymbol{B}_{n-1}+\lambda^{n-1}(\boldsymbol{B}_{n-2}-\boldsymbol{B}_{n-1}\boldsymbol{A})+\cdots+\lambda(\boldsymbol{B}_0-\boldsymbol{B}_1\boldsymbol{A})-\boldsymbol{B}_0\boldsymbol{A}. \end{aligned} \tag{18}$$

记 $\boldsymbol{A}$ 的特征多项式为 $f(\lambda)=\lambda^n+a_{n-1}\lambda^{n-1}+\cdots+a_1\lambda+a_0$, 则

$$f(\lambda)\boldsymbol{I}=\lambda^n\boldsymbol{I}+a_{n-1}\lambda^{n-1}\boldsymbol{I}+\cdots+a_1\lambda\boldsymbol{I}+a_0\boldsymbol{I}. \tag{19}$$

根据 (17), (18), (19) 式得

$$\begin{cases} \boldsymbol{B}_{n-1}=\boldsymbol{I}, \\ \boldsymbol{B}_{n-2}-\boldsymbol{B}_{n-1}\boldsymbol{A}=a_{n-1}\boldsymbol{I}, \\ \quad\vdots \\ \boldsymbol{B}_0-\boldsymbol{B}_1\boldsymbol{A}=a_1\boldsymbol{I}, \\ -\boldsymbol{B}_0\boldsymbol{A}=a_0\boldsymbol{I}. \end{cases} \tag{20}$$

2.235

用 $\boldsymbol{A}^n$, $\boldsymbol{A}^{n-1}$, $\cdots$, $\boldsymbol{A}$, $\boldsymbol{I}$ 依次从右边乘以 (20) 式中的第 1 式, 第 2 式, $\cdots$, 第 n 式, 第 $n+1$ 式, 得

$$\begin{cases} \boldsymbol{B}_{n-1}\boldsymbol{A}^n=\boldsymbol{A}^n, \\ \boldsymbol{B}_{n-2}\boldsymbol{A}^{n-1}-\boldsymbol{B}_{n-1}\boldsymbol{A}^n=a_{n-1}\boldsymbol{A}^{n-1}, \\ \quad\vdots \\ \boldsymbol{B}_0\boldsymbol{A}-\boldsymbol{B}_1\boldsymbol{A}^2=a_1\boldsymbol{A}, \\ -\boldsymbol{B}_0\boldsymbol{A}=a_0\boldsymbol{I}. \end{cases} \tag{21}$$

把 (21) 式中的 $n+1$ 个式子叠加起来, 左边为零矩阵, 右边为 $f(\boldsymbol{A})$, 因此 $f(\boldsymbol{A})=\boldsymbol{O}$. □ 2.236

Theorem 170. 设 T 为 n 维空间 X 上的线性变换, 则 T 的特征多项式是 T 的一个零化多项式.

证: 设 $\boldsymbol{A}$ 为 T 的一个矩阵表示, $f(\lambda)$ 为 $\boldsymbol{A}$ 的特征多项式, 由哈密尔顿 –凯莱定理知

$$f(\boldsymbol{A})=\boldsymbol{O},$$

即 $f(\lambda)$ 为 $\boldsymbol{A}$ 的一个零化多项式, 从而 $f(\lambda)$ 为 T 的一个零化多项式. □ 2.237

最小多项式

Definition 171. 设 T 是 n 维线性空间 X 上的一个线性变换, $\varphi(\lambda)$ 是 T 的所有零化多项式中次数最低且首项系数为 1 的多项式, 则称 $\varphi(\lambda)$ 为 T 的最小多项式. 记为 $m_T(\lambda)$, 即

$$m_T(\lambda)=\varphi(\lambda).$$

引入记号: 多项式 $f(x)$ 的次数记为

$$\partial\big(f(x)\big).$$

☞ 最小多项式 $m_T(\lambda)$ 的次数 $\leqslant$ 特征多项式 $f(\lambda)$ 的次数. 即

$$\partial\big(m_T(\lambda)\big)\leqslant\partial\big(f(\lambda)\big)=n.$$

2.238

Theorem 172. 设 T 是 n 维线性空间 X 上的一个线性变换, 则 T 的最小多项式 $m_T(\lambda)$ 存在且唯一.

证: 存在性是显然的. 下证唯一性.

设 $\varphi_1(\lambda)$ 与 $\varphi_2(\lambda)$ 都是 T 的最小多项式, 则 $\varphi_1(\lambda)$ 与 $\varphi_2(\lambda)$ 的首项系数都为 1, 且次数相同.

假设 $\varphi_1(\lambda)\neq\varphi_2(\lambda)$, 记

$$\varphi(\lambda)=\varphi_1(\lambda)-\varphi_2(\lambda),$$

则

$$\partial\big(\varphi(\lambda)\big)<\partial\big(\varphi_1(\lambda)\big),$$

且 $\varphi(\lambda)$ 为 T 的零化多项式, 这与 “$\varphi_1(\lambda)$ 是 T 的最小多项式” 相矛盾. 假设不成立, 得 $\varphi_1(\lambda)=\varphi_2(\lambda)$. 唯一性得证. □ 2.239

Theorem 173. 设 T 为 n 维线性空间 X 上的线性变换, 则 T 的最小多项式必整除 T 的任一零化多项式.

证: 设 $\varphi(\lambda)$ 为 T 的任一零化多项式, $m_T(\lambda)$ 为 T 的最小多项式, 则

$$\partial\big(m_T(\lambda)\big)\leqslant\partial\big(\varphi(\lambda)\big).$$

根据带余除法, $\varphi(\lambda)$ 可以表示成

$$\varphi(\lambda) = g(\lambda)m_T(\lambda) + r(\lambda),$$

其中 $r(\lambda) = 0$ 或 $\partial\big(r(\lambda)\big) < \partial\big(m_T(\lambda)\big)$. 则

$$0^* = \varphi(T) = g(T)m_T(T) + r(T) = g(T)0^* + r(T),$$

故 $r(T) = 0^*$. 由最小多项式的定义, $r(\lambda) = 0$, 故 $m_T(\lambda) \mid \varphi(\lambda)$. □

☞ 整除记号: $m_T(\lambda)$ 整除 $\varphi(\lambda)$, 记为

$$m_T(\lambda) \mid \varphi(\lambda).$$

2.240

Corollary 174. *n 维线性空间 X 上的线性变换 T 的最小多项式 $m_T(\lambda)$, 与 T 的特征多项式 $f(\lambda)$ 有相同的根.*

证: 因 $f(\lambda)$ 是 T 的零化多项式, 故

$$m_T(\lambda) \mid f(\lambda),$$

从而 $m_T(\lambda)$ 的根都是 $f(\lambda)$ 的根.

反之, 设 λ_0 为 $f(\lambda)$ 的任一根, 则 λ_0 为 T 的一个特征值, 从而存在 $\boldsymbol{x} \in X$, 且 $\boldsymbol{x} \neq \boldsymbol{0}$, 使得 $T\boldsymbol{x} = \lambda_0\boldsymbol{x}$.

设 $m_T(\lambda) = \lambda^k + \alpha_1\lambda^{k-1} + \cdots + \alpha_{k-1}\lambda + \alpha_k$, 则

$$\begin{aligned}\boldsymbol{0} &= m_T(T)\boldsymbol{x} = (T^k + \alpha_1 T^{k-1} + \cdots + \alpha_{k-1}T + \alpha_k E)\boldsymbol{x}\\ &= \lambda_0^k\boldsymbol{x} + \alpha_1\lambda_0^{k-1}\boldsymbol{x} + \cdots + \alpha_{k-1}\lambda_0\boldsymbol{x} + \alpha_k\boldsymbol{x} = m_T(\lambda_0)\boldsymbol{x},\end{aligned}$$

得 $m_T(\lambda_0) = 0$, 故 λ_0 也是 $m_T(\lambda)$ 的根. □ 2.241

☞ 如果 T 的特征多项式为

$$f(\lambda) = (\lambda - \lambda_1)^{d_1}(\lambda - \lambda_2)^{d_2}\cdots(\lambda - \lambda_k)^{d_k},$$

则必有

$$m_T(\lambda) = (\lambda - \lambda_1)^{e_1}(\lambda - \lambda_2)^{e_2}\cdots(\lambda - \lambda_k)^{e_k},$$

且 $1 \leqslant e_i \leqslant d_i$, $i = 1, 2, \cdots, k$.

e_i 称为 λ_i 的指数, $i = 1, 2, \cdots, k$. 2.242

Example 175. 设 T 为 $\mathbb{R}^3$ 上的线性变换, T 关于自然基底 $\mathscr{B} = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ 的矩阵表示为

$$\boldsymbol{A} = \begin{bmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{bmatrix},$$

试求 T 的最小多项式.

解: T 的特征多项式为

$$f(\lambda) = \det(\lambda \boldsymbol{I} - \boldsymbol{A}) = (\lambda - 1)(\lambda - 2)^2,$$

设 $m_T(\lambda)$ 为 T 的最小多项式, 则 $m_T(\lambda)$ 只能是下列形式之一:

$$(\lambda - 1)(\lambda - 2), \quad 或 \quad (\lambda - 1)(\lambda - 2)^2.$$

由于

$$(\boldsymbol{A} - \boldsymbol{I})(\boldsymbol{A} - 2\boldsymbol{I}) = \boldsymbol{O},$$

故 $(\lambda - 1)(\lambda - 2)$ 为 T 的最小多项式. □ 2.243

T 可对角化的又一充要条件

Theorem 176. *设 T 为 n 维线性空间 X 上的线性变换, 则 T 可对角化的充分必要条件是, T 的最小多项式 $m_T(\lambda)$ 可分解为不同一次因式的乘积. 此时若 λ_1, λ_2, $\cdots$, λ_k 为 T 的所有互异特征值, 则*

$$m_T(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_k).$$

☞ $T \in L(X, X)$ 可对角化 $\Longleftrightarrow$ T 的各特征值指数均为 1. 2.244

Corollary 177. *若 n 维线性空间 X 上的线性变换 T 的某一零化多项式可以分解为不同一次因式的积, 则 T 可对角化.*

事实上, 因最小多项式 $m_T(\lambda)$ 整除该零化多项式, 故 $m_T(\lambda)$ 可分解为不同一次因式的乘积. 从而 T 可对角化. 2.245

Example 178. 设 $T \in L(X, X)$ 满足 $T^n = E$, 试证 T 为可对角化的线性变换.

证: 由 $T^n = E$ 得 $T^n - E = 0^*$, 故 $g(\lambda) = \lambda^n - 1$ 是 T 的一个零化多项式. 下证 $g(\lambda)$ 可以分解为不同一次因式的乘积.

$$\begin{aligned} & c = r(\cos\theta + \mathrm{i}\sin\theta) \text{ 是 } x^n - 1 \text{ 的根} \\ \Longleftrightarrow\ & c^n = 1 \\ \Longleftrightarrow\ & r^n(\cos n\theta + \mathrm{i}\sin n\theta) = \cos 0 + \mathrm{i}\sin 0 \\ \Longleftrightarrow\ & r^n = 1 \text{ 且 } n\theta = 0 + 2k\pi,\ k \in \mathbb{Z} \\ \Longleftrightarrow\ & r = 1 \text{ 且 } \theta = \frac{2k\pi}{n},\ k \in \mathbb{Z} \\ \Longleftrightarrow\ & c = \cos\frac{2k\pi}{n} + \mathrm{i}\sin\frac{2k\pi}{n} = \mathrm{e}^{\mathrm{i}\frac{2k\pi}{n}},\ k \in \mathbb{Z}. \end{aligned}$$

记 $\xi = \mathrm{e}^{\mathrm{i}\frac{2\pi}{n}}$, 则 1, ξ, ξ^2, $\cdots$, ξ^{n-1} 是 $x^n - 1$ 的所有的根, 且两两不等. 故 $x^n - 1 = (x-1)(x-\xi)(x-\xi^2)\cdots(x-\xi^{n-1})$, 即 $g(\lambda)$ 可以分解为不同因式的乘积, 得证 T 为可对角化的线性变换. □ 2.246

Exercise 179 (P.100 习题 (二) 7)**.** 设 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$, 且满足 $\boldsymbol{A}^3 - \boldsymbol{A} = 2\boldsymbol{A}^2 - 2\boldsymbol{I}$. 试证 $\boldsymbol{A}$ 可对角化.

解: 由于 $\boldsymbol{A}^3-\boldsymbol{A}-2\boldsymbol{A}^2+2\boldsymbol{I}=\boldsymbol{O}$, 故 $g(\lambda)=\lambda^3-2\lambda^2-\lambda+2$ 是 $\boldsymbol{A}$ 的零化多项式. 由于

$$g(\lambda)=\lambda^2(\lambda-2)-\lambda+2=(\lambda-2)(\lambda+1)(\lambda-1),$$

且 $\boldsymbol{A}$ 的最小多项式 $m(\lambda)\big|g(\lambda)$, 因此, $m(\lambda)$ 可以分解为不同一次因式的乘积. 从而 $\boldsymbol{A}$ 可对角化.

或者, 由零化多项式 $g(\lambda)$ 分解为一次因式的乘积, 直接使用推论 177 知 $\boldsymbol{A}$ 可对角化. □ 2.247

5.3 不可对角化线性变换的最简矩阵表示

这一小节的主要结论是: 不可对角化矩阵 $\boldsymbol{A}$ 总可以相似于一个若当矩阵, 即相似于一个形如

$$\left[\begin{array}{cc|c|ccc}
\lambda_1 & 1 & 0 & 0 & 0 & 0\\
0 & \lambda_1 & 0 & 0 & 0 & 0\\
\hline
0 & 0 & \lambda_2 & 0 & 0 & 0\\
\hline
0 & 0 & 0 & \lambda_3 & 1 & 0\\
0 & 0 & 0 & 0 & \lambda_3 & 1\\
0 & 0 & 0 & 0 & 0 & \lambda_3
\end{array}\right]$$

的分块对角矩阵, 其主对角线元素是 $\boldsymbol{A}$ 的全部特征值.

为这个结论做准备, 先讨论以下几个概念: λ-矩阵, 行列式因子, 不变因子, 初等因子. 2.248

一. λ-矩阵

Definition 180. 若一个矩阵的元素是 λ 的复系数多项式, 则称这种矩阵为 λ-矩阵或多项式矩阵. 今后总以 $\boldsymbol{A}(\lambda)$, $\boldsymbol{B}(\lambda)$, $\cdots$ 来表示 λ-矩阵.

- 特征矩阵 $\lambda\boldsymbol{I}-\boldsymbol{A}$ 就是一个常见的 λ-矩阵.
- λ-矩阵也包括数字矩阵.
- λ-矩阵也可以定义加法、乘法, 它们与数字矩阵的运算有相同的运算规律.
- λ-矩阵也可以定义行列式 (比如 $|\lambda\boldsymbol{I}-\boldsymbol{A}|$), 它与数字矩阵的行列式有相同的性质. 例如

$$\det\boldsymbol{A}(\lambda)\boldsymbol{B}(\lambda)=\det\boldsymbol{A}(\lambda)\det\boldsymbol{B}(\lambda).$$ 2.249

Definition 181.

1. λ-矩阵 $\boldsymbol{A}(\lambda)$ 中不恒等于零的子式的最高阶数 r, 称为 $\boldsymbol{A}(\lambda)$ 的秩, 记为 $\operatorname{rank}\boldsymbol{A}(\lambda)$.
2. 若 $\boldsymbol{A}(\lambda)$ 为正方 λ-矩阵, 且 $\det\boldsymbol{A}(\lambda)$ 不恒等于零, 则称 $\boldsymbol{A}(\lambda)$ 为满秩的, 否则称之为降秩的. 2.250

λ-矩阵 $\boldsymbol{A}$ 的秩的概念, 与常数矩阵的秩的概念, 是一致的. 但要注意, $\boldsymbol{A}$ 的子式都是 λ 的多项式, 一个 λ 的多项式等于零是指任意数都是它的根, 因此它的系数都是零, 这时我们又常说它恒等于零. 任意 n 次多项式只能有 n 根, 所以它不为零, 即它不恒等于零.

假如 $\boldsymbol{A}$ 是 n 阶矩阵, 因为 $|\lambda \boldsymbol{I}-\boldsymbol{A}|$ 是 λ 的 n 次多项式, 不恒为零, 所以 $\boldsymbol{A}$ 的特征矩阵 $|\lambda \boldsymbol{I}-\boldsymbol{A}|$ 的秩是 n. 也就是说, **不论 $\boldsymbol{A}$ 是满秩还是降秩的, $\boldsymbol{A}$ 的特征矩阵总是满秩的.** 2.251

Definition 182. 一个 $n\times n$ 的 λ-矩阵 $\boldsymbol{A}(\lambda)$ 称为可逆的, 如果有一个 $n\times n$ 的 λ-矩阵 $\boldsymbol{B}(\lambda)$ 使

$$\boldsymbol{A}(\lambda)\boldsymbol{B}(\lambda)=\boldsymbol{B}(\lambda)\boldsymbol{A}(\lambda)=\boldsymbol{I},$$

这里 $\boldsymbol{I}$ 是 n 阶单位矩阵. $\boldsymbol{B}(\lambda)$ 称为 $\boldsymbol{A}(\lambda)$ 的逆矩阵, 记为 $\boldsymbol{A}^{-1}(\lambda)$. 2.252

Theorem 183. *一个 $n\times n$ 的 λ-矩阵 $\boldsymbol{A}(\lambda)$ 可逆的充分必要条件为 $|\boldsymbol{A}(\lambda)|$ 是一个非零的数.*

证: 先证充分性. 设 $|\boldsymbol{A}(\lambda)|$ 是一个非零的数. $\boldsymbol{A}^*(\lambda)$ 是 $\boldsymbol{A}(\lambda)$ 的伴随矩阵, 它也是一个 λ-矩阵. 当 $|\boldsymbol{A}(\lambda)|$ 是一个非零的数, $\frac{1}{|\boldsymbol{A}(\lambda)|}\boldsymbol{A}^*(\lambda)$ 的元素才是 λ 的多项式, 从而 $\frac{1}{|\boldsymbol{A}(\lambda)|}\boldsymbol{A}^*(\lambda)$ 也是一个 λ-矩阵. 而

$$\boldsymbol{A}(\lambda)\frac{1}{|\boldsymbol{A}(\lambda)|}\boldsymbol{A}^*(\lambda)=\frac{1}{|\boldsymbol{A}(\lambda)|}\boldsymbol{A}^*(\lambda)\boldsymbol{A}(\lambda)=\boldsymbol{I},$$

因此, $\boldsymbol{A}(\lambda)$ 可逆.

反过来, 如果 $\boldsymbol{A}(\lambda)$ 可逆, 则

$$|\boldsymbol{A}(\lambda)\boldsymbol{A}^{-1}(\lambda)|=|\boldsymbol{A}(\lambda)||\boldsymbol{A}^{-1}(\lambda)|=1,$$

而 $|\boldsymbol{A}(\lambda)|$, $|\boldsymbol{A}^{-1}(\lambda)|$ 都是 λ 的多项式, 故两者都只能是零次多项式, 即常数. 得证 $|\boldsymbol{A}(\lambda)|$ 是非零常数. □ 2.253

☞ 满秩 λ-矩阵不一定是可逆矩阵.

例如

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

是满秩的, 但不是可逆的.

Definition 184. 行列式为非零常数的正方 λ-矩阵, 称为初等 λ-矩阵.

☞ $\boldsymbol{A}(\lambda)$ 为初等 λ-矩阵 $\Leftrightarrow$ $\boldsymbol{A}(\lambda)$ 可逆 $\Leftrightarrow$ $\det \boldsymbol{A}(\lambda)$ 为非零的数. 2.254

λ-矩阵的初等变换

Definition 185. λ-矩阵的初等变换是指下列三种变换:

1. 用数 $\alpha\neq 0$ 去乘某行 (列);
2. 用 λ 的多项式 $\varphi(\lambda)$ 去乘某行 (列) 后, 加到另一行 (列) 上去;
3. 互换 i, j 两行 (列).

☞ (1) 与数字矩阵的初等变换比较, 只有第二种变换作了合理的改变. *注意只有乘法, 没有除法: 把每一行除以多项式 $\varphi(\lambda)$.*

(2) 第一种变换没有变化. 可否改为 “用多项式 $\varphi(\lambda)$ 去乘某行 (列)”? 2.255

Definition 186. 初等行 (列) 变换 λ-矩阵是指下列三种 λ-矩阵:

1. 用数 $\alpha \neq 0$ 去乘单位矩阵 I 的某行 (列) 而得的矩阵;
2. 用 λ 的多项式 $\varphi(\lambda)$ 去乘单位矩阵的某行 (列) 后, 加到另一行 (列) 上去;
3. 互换单位矩阵的某两行 (列).

初等行变换 λ-矩阵分别记为:

$$\boldsymbol{J}_1(\alpha \boldsymbol{e}_i^{\mathrm{T}}) = \left[\begin{array}{ccccccc} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \alpha & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{array}\right] \begin{array}{l} \\ \\ \\ \longleftarrow \text{第 } i \text{ 行} \\ \\ \\ \\ \end{array} \tag{22}$$

☞ $\det \boldsymbol{J}_1(\alpha \boldsymbol{e}_i^{\mathrm{T}}) = \alpha \neq 0$, 矩阵 $\boldsymbol{J}_1(\alpha \boldsymbol{e}_i^{\mathrm{T}})$ 可逆. 2.256

$$\boldsymbol{J}_2(\varphi(\lambda)\boldsymbol{e}_i^{\mathrm{T}} \xrightarrow{+} \boldsymbol{e}_j^{\mathrm{T}}) = \left[\begin{array}{cccccc} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & \vdots & \ddots & & \\ & & \varphi(\lambda) & \cdots & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{array}\right] \begin{array}{l} \\ \\ i \\ \\ j \\ \\ \\ \end{array} \tag{23}$$

☞ $\det \boldsymbol{J}_2(\varphi(\lambda)\boldsymbol{e}_i^{\mathrm{T}} \xrightarrow{+} \boldsymbol{e}_j^{\mathrm{T}}) = 1$, 矩阵 $\boldsymbol{J}_2(\varphi(\lambda)\boldsymbol{e}_i^{\mathrm{T}} \xrightarrow{+} \boldsymbol{e}_j^{\mathrm{T}})$ 可逆. 2.257

$$\boldsymbol{J}_3(\boldsymbol{e}_i^{\mathrm{T}} \leftrightarrow \boldsymbol{e}_j^{\mathrm{T}}) = \left[\begin{array}{ccccccccccc} 1 & & & & & & & & & & \\ & \ddots & & & & & & & & & \\ & & 1 & & & & & & & & \\ & & & 0 & \cdots & \cdots & \cdots & 1 & & & \\ & & & \vdots & 1 & & & \vdots & & & \\ & & & \vdots & & \ddots & & \vdots & & & \\ & & & \vdots & & & 1 & \vdots & & & \\ & & & 1 & \cdots & \cdots & \cdots & 0 & & & \\ & & & & & & & & 1 & & \\ & & & & & & & & & \ddots & \\ & & & & & & & & & & 1 \end{array}\right] \begin{array}{l} \\ \\ \\ i \\ \\ \\ \\ j \\ \\ \\ \\ \end{array} \tag{24}$$

☞ $\det \boldsymbol{J}_3(\boldsymbol{e}_i^{\mathrm{T}} \leftrightarrow \boldsymbol{e}_j^{\mathrm{T}}) = -1$, 矩阵 $\boldsymbol{J}_3(\boldsymbol{e}_i^{\mathrm{T}} \leftrightarrow \boldsymbol{e}_j^{\mathrm{T}})$ 可逆. 2.258

$\boldsymbol{J}_1(\alpha\boldsymbol{e}_i^{\mathrm{T}})$, $\boldsymbol{J}_2(\varphi(\lambda)\boldsymbol{e}_i^{\mathrm{T}} \xrightarrow{+} \boldsymbol{e}_j^{\mathrm{T}})$, $\boldsymbol{J}_3(\boldsymbol{e}_i^{\mathrm{T}} \leftrightarrow \boldsymbol{e}_j^{\mathrm{T}})$ 分别称为第 (1)、(2)、(3) 种初等行变换 λ-矩阵.

类似地, 第 (1)、(2)、(3) 种初等列变换 λ-矩阵分别记为

1. $\boldsymbol{J}_1^*(\alpha\boldsymbol{e}_i)$,
2. $\boldsymbol{J}_2^*(\varphi(\lambda)\boldsymbol{e}_i \xrightarrow{+} \boldsymbol{e}_j)$,
3. $\boldsymbol{J}_3^*(\boldsymbol{e}_i \leftrightarrow \boldsymbol{e}_j)$.

同样地, 对一个 $s \times n$ 的 λ-矩阵 $\boldsymbol{A}(\lambda)$ 作一次初等行变换, 相当于在 $\boldsymbol{A}(\lambda)$ 的左边乘以相应的 $s \times s$ 初等变换矩阵; 对 $\boldsymbol{A}(\lambda)$ 作一次初等列变换, 相当于在 $\boldsymbol{A}(\lambda)$ 的右边乘以相应的 $n \times n$ 初等变换矩阵. 2.259

初等变换 λ-矩阵都是可逆的, 其逆矩阵仍是初等变换 λ-矩阵.

$$\begin{aligned}
\boldsymbol{J}_1(\alpha\boldsymbol{e}_i^{\mathrm{T}})^{-1} &= \boldsymbol{J}_1(\alpha^{-1}\boldsymbol{e}_i^{\mathrm{T}}), \\
\boldsymbol{J}_2\big(\varphi(\lambda)\boldsymbol{e}_i^{\mathrm{T}} \xrightarrow{+} \boldsymbol{e}_j^{\mathrm{T}}\big)^{-1} &= \boldsymbol{J}_2\big(-\varphi(\lambda)\boldsymbol{e}_i^{\mathrm{T}} \xrightarrow{+} \boldsymbol{e}_j^{\mathrm{T}}\big), \\
\boldsymbol{J}_3(\boldsymbol{e}_i^{\mathrm{T}} \leftrightarrow \boldsymbol{e}_j^{\mathrm{T}})^{-1} &= \boldsymbol{J}_3(\boldsymbol{e}_i^{\mathrm{T}} \leftrightarrow \boldsymbol{e}_j^{\mathrm{T}}).
\end{aligned}$$

由此得出初等变换具有可逆性.

☞ 在第一种初等变换中, 规定只能乘以一个非零常数, 而不能 “用多项式 $\varphi(\lambda)$ 去乘某行 (列)”, 这也是为了使 $\boldsymbol{J}_1(\alpha\boldsymbol{e}_i^{\mathrm{T}})$ 可逆的缘故.

具体而言, 初等变换是可逆变换, 假如允许 “用多项式 $\varphi(\lambda)$ 去乘某行”, 则其逆变换就需要将该行除以多项式 $\varphi(\lambda)$. 初等变换的逆变换还是初等变换, 这意味着 “将某行除以多项式 $\varphi(\lambda)$” 也是初等变换, 但这种做法一般会产生分式, 其结果不再是 λ-矩阵. 2.260

λ-矩阵的等价

Definition 187. 若 λ-矩阵 $\boldsymbol{A}(\lambda)$ 经过有限次初等变换化为 λ-矩阵 $\boldsymbol{B}(\lambda)$, 则称 $\boldsymbol{B}(\lambda)$ 与 $\boldsymbol{A}(\lambda)$ 等价, 记为

$$\boldsymbol{A}(\lambda) \cong \boldsymbol{B}(\lambda).$$

λ-矩阵的等价关系满足下面的等价律:

1. 反身性: $\boldsymbol{A}(\lambda) \cong \boldsymbol{A}(\lambda)$;
2. 对称性: 若 $\boldsymbol{A}(\lambda) \cong \boldsymbol{B}(\lambda)$, 则 $\boldsymbol{B}(\lambda) \cong \boldsymbol{A}(\lambda)$;
3. 传递性: 若 $\boldsymbol{A}(\lambda) \cong \boldsymbol{B}(\lambda)$, $\boldsymbol{B}(\lambda) \cong \boldsymbol{C}(\lambda)$, 则 $\boldsymbol{A}(\lambda) \cong \boldsymbol{C}(\lambda)$. 2.261

Theorem 188. *如果 $\boldsymbol{A}_1(\lambda) \cong \boldsymbol{B}_1(\lambda)$, $\boldsymbol{A}_2(\lambda) \cong \boldsymbol{B}_2(\lambda)$, 则*

$$\boldsymbol{A}_1(\lambda) \oplus \boldsymbol{A}_2(\lambda) \cong \boldsymbol{B}_1(\lambda) \oplus \boldsymbol{B}_2(\lambda),$$

即

$$\begin{bmatrix} \boldsymbol{A}_1(\lambda) & \\ & \boldsymbol{A}_2(\lambda) \end{bmatrix} \cong \begin{bmatrix} \boldsymbol{B}_1(\lambda) & \\ & \boldsymbol{B}_2(\lambda) \end{bmatrix}.$$

☞ 符号 $\oplus$ 表示矩阵的直和: 设 $\boldsymbol{A}_i$ 为 n_i 阶方阵 $(i=1,2,\cdots,k)$, 则

$$\boldsymbol{A}_1\oplus\boldsymbol{A}_2\oplus\cdots\oplus\boldsymbol{A}_k=\begin{bmatrix}\boldsymbol{A}_1 & & & \\ & \boldsymbol{A}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{A}_k\end{bmatrix}.$$

2.262

Exercise 189 (P.109 习题 (一) 2)**.** 证明: 若 $\boldsymbol{P}_1(\lambda)$, $\boldsymbol{P}_2(\lambda)$, $\cdots$, $\boldsymbol{P}_k(\lambda)$ 均为初等 λ-矩阵, 且皆为 n 阶方阵, 则它们的积也为初等 λ-矩阵.

证: $\det\big(\boldsymbol{P}_1(\lambda)\cdots\boldsymbol{P}_l(\lambda)\big)=\det\boldsymbol{P}_1(\lambda)\cdots\det\boldsymbol{P}_l(\lambda)\neq 0$. 即证. □ 2.263

Exercise 190 (P.109 习题 (一) 3)**.** 设 λ-矩阵 $\boldsymbol{A}(\lambda)$ 与 $\boldsymbol{B}(\lambda)$ 等价, 试证存在初等 λ-矩阵 $\boldsymbol{P}(\lambda)$, $\boldsymbol{Q}(\lambda)$, 使得

$$\boldsymbol{A}(\lambda)=\boldsymbol{P}(\lambda)\boldsymbol{B}(\lambda)\boldsymbol{Q}(\lambda).$$

证: 由初等变换与初等矩阵的关系即得, 矩阵 $\boldsymbol{A}(\lambda)$ 与 $\boldsymbol{B}(\lambda)$ 等价的充分必要条件为: 有一系列的初等变换 λ-矩阵 $\boldsymbol{P}_1(\lambda)$, $\boldsymbol{P}_2(\lambda)$, $\cdots$, $\boldsymbol{P}_s(\lambda)$, $\boldsymbol{Q}_1(\lambda)$, $\boldsymbol{Q}_2(\lambda)$, $\cdots$, $\boldsymbol{Q}_t(\lambda)$ 使

$$\boldsymbol{A}(\lambda)=\boldsymbol{P}_s(\lambda)\cdots\boldsymbol{P}_2(\lambda)\boldsymbol{P}_1(\lambda)\boldsymbol{B}(\lambda)\boldsymbol{Q}_1(\lambda)\boldsymbol{Q}_2(\lambda)\cdots\boldsymbol{Q}_t(\lambda),$$

记 $\boldsymbol{P}(\lambda)=\boldsymbol{P}_s(\lambda)\cdots\boldsymbol{P}_2(\lambda)\boldsymbol{P}_1(\lambda)$, $\boldsymbol{Q}(\lambda)=\boldsymbol{Q}_1(\lambda)\boldsymbol{Q}_2(\lambda)\cdots\boldsymbol{Q}_t(\lambda)$, 则 $\boldsymbol{P}(\lambda)$, $\boldsymbol{Q}(\lambda)$ 为初等 λ-矩阵. 即证. □

☞ 一个小结论:

两个同型 λ-矩阵 $\boldsymbol{A}(\lambda)$ 与 $\boldsymbol{B}(\lambda)$ 等价的充要条件是存在两个可逆的 λ-矩阵 $\boldsymbol{P}(\lambda)$, $\boldsymbol{Q}(\lambda)$, 使得

$$\boldsymbol{A}(\lambda)=\boldsymbol{P}(\lambda)\boldsymbol{B}(\lambda)\boldsymbol{Q}(\lambda).$$

2.264

Theorem 191. *等价的 λ-矩阵, 秩相等.*

证: 只需证明一次初等变换不改变矩阵的秩. (略)

☞ 初等变换不改变矩阵的秩. 2.265

Theorem 192. *秩为 r 的正方的 λ-矩阵恒可由若干次初等变换变成一个如下的对角 λ-矩阵:*

$$\boldsymbol{B}(\lambda)=\begin{bmatrix}\varphi_1(\lambda) & & & & & & \\ & \varphi_2(\lambda) & & & & & \\ & & \ddots & & & & \\ & & & \varphi_r(\lambda) & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0\end{bmatrix}, \tag{25}$$

其中 $\varphi_i(\lambda)$ $(i=1,2,\cdots,r)$ 是首项系数为 1 的多项式, 且

$$\varphi_1(\lambda)\mid\varphi_2(\lambda),\ \varphi_2(\lambda)\mid\varphi_3(\lambda),\ \cdots,\ \varphi_{r-1}(\lambda)\mid\varphi_r(\lambda).$$

☞ $\boldsymbol{B}(\lambda)$ 称为 $\boldsymbol{A}(\lambda)$ 的史密斯 (Smith) 标准形, 简称标准形.

上述结论对一般的 $m\times n$ 阶 λ-矩阵也成立. 2.266

Example 193. 用初等变换化 λ-矩阵 $\boldsymbol{A}(\lambda)=\begin{bmatrix}1-\lambda & 2\lambda-1 & \lambda\\ \lambda & \lambda^2 & -\lambda\\ 1+\lambda^2 & \lambda^3+\lambda-1 & -\lambda^2\end{bmatrix}$ 为标准形.

解: $$\boldsymbol{A}(\lambda)\xrightarrow{c_1+c_3}\begin{bmatrix}1 & 2\lambda-1 & \lambda\\ 0 & \lambda^2 & -\lambda\\ 1 & \lambda^3+\lambda-1 & -\lambda^2\end{bmatrix}\xrightarrow{r_3-r_1}\begin{bmatrix}1 & 2\lambda-1 & \lambda\\ 0 & \lambda^2 & -\lambda\\ 0 & \lambda^3-\lambda & -\lambda^2-\lambda\end{bmatrix}$$
$$\xrightarrow[c_3-\lambda c_1]{c_2-(2\lambda-1)c_1}\begin{bmatrix}1 & 0 & 0\\ 0 & \lambda^2 & -\lambda\\ 0 & \lambda^3-\lambda & -\lambda^2-\lambda\end{bmatrix}\xrightarrow[c_2+\lambda c_3]{c_3\times(-1)}\begin{bmatrix}1 & 0 & 0\\ 0 & 0 & \lambda\\ 0 & -\lambda^2-\lambda & \lambda^2+\lambda\end{bmatrix}$$
$$\xrightarrow{c_2\leftrightarrow c_3}\begin{bmatrix}1 & 0 & 0\\ 0 & \lambda & 0\\ 0 & \lambda^2+\lambda & -\lambda^2-\lambda\end{bmatrix}\xrightarrow[c_2+c_3]{c_3\times(-1)}\begin{bmatrix}1 & 0 & 0\\ 0 & \lambda & 0\\ 0 & 0 & \lambda^2+\lambda\end{bmatrix}.$$ □ 2.267

二. 不变因子

Definition 194. 设 λ-矩阵 $\boldsymbol{A}(\lambda)$ 的秩为 r, 对正整数 k, $1\leqslant k\leqslant r$, $\boldsymbol{A}(\lambda)$ 中必有非零的 k 阶子式. $\boldsymbol{A}(\lambda)$ 中全部 k 阶子式的首项系数为 1 的最大公因式 $D_k(\lambda)$, 称为 $\boldsymbol{A}(\lambda)$ 的 k 阶行列式因子.

由定义可知, 对于秩为 r 的 λ-矩阵, 行列式因子一共有 r 个. 行列式因子的意义在于, 它在初等变换下是不变的. 2.268

Theorem 195. 等价的 λ-矩阵有相同的秩与相同的各阶行列式因子.

证: 只需要证明, λ-矩阵经过一次初等变换, 秩与行列式因子是不变的.

设 λ-矩阵 $\boldsymbol{A}(\lambda)$ 经过一次初等行变换变成 $\boldsymbol{B}(\lambda)$, $f(\lambda)$ 与 $g(\lambda)$ 分别是 $\boldsymbol{A}(\lambda)$ 与 $\boldsymbol{B}(\lambda)$ 的 k 阶行列式因子. 我们证明 $f=g$. 下面分三种情况讨论:

(1) $\boldsymbol{A}(\lambda)\xrightarrow{r_i\leftrightarrow r_j}\boldsymbol{B}(\lambda)$. 这时, $\boldsymbol{B}(\lambda)$ 的每个 k 阶子式或者等于 $\boldsymbol{A}(\lambda)$ 的某个 k 阶子式, 或者与 $\boldsymbol{A}(\lambda)$ 的某个 k 阶子式反号, 因此 $f(\lambda)$ 是 $\boldsymbol{B}(\lambda)$ 的 k 阶子式的公因式, 从而 $f(\lambda)\mid g(\lambda)$.

(2) $\boldsymbol{A}(\lambda)\xrightarrow{r_i\times\alpha}\boldsymbol{B}(\lambda)$. 这时, $\boldsymbol{B}(\lambda)$ 的每个 k 阶子式或者等于 $\boldsymbol{A}(\lambda)$ 的某个 k 阶子式, 或者等于 $\boldsymbol{A}(\lambda)$ 的某个 k 阶子式的 α 倍, 因此 $f(\lambda)$ 是 $\boldsymbol{B}(\lambda)$ 的 k 阶子式的公因式, 从而 $f(\lambda)\mid g(\lambda)$. 2.269

(3) $\boldsymbol{A}(\lambda)\xrightarrow{r_i+\varphi(\lambda)r_j}\boldsymbol{B}(\lambda)$. 这时 $\boldsymbol{B}(\lambda)$ 中那些包含 i 行与 j 行的 k 阶子式和那些不包含 i 行的 k 阶子式, 都等于 $\boldsymbol{A}(\lambda)$ 中对应的 k 阶子式; $\boldsymbol{B}(\lambda)$ 中那些包含 i 行但不包含 j 行的 k 阶子式, 按 i 行分成两部分, 而等于 $\boldsymbol{A}(\lambda)$ 的一个 k 阶子

式与另一个 k 阶子式的 $\pm\varphi(\lambda)$ 倍的和, 也就是 $\boldsymbol{A}(\lambda)$ 的两个 k 阶子式和组合. 因此 $f(\lambda)$ 是 $\boldsymbol{B}(\lambda)$ 的 k 阶子式的公因式, 从而 $f(\lambda) \mid g(\lambda)$.

对于列变换, 可以完全一样地讨论. 总之, 如果 $\boldsymbol{A}(\lambda)$ 经过一次初等行变换变成 $\boldsymbol{B}(\lambda)$, 那么 $f(\lambda) \mid g(\lambda)$. 但又初等变换的可逆性, $\boldsymbol{B}(\lambda)$ 经过一次初等行变换变成 $\boldsymbol{A}(\lambda)$. 由上面的讨论, 同样应有 $g(\lambda) \mid f(\lambda)$, 于是 $f(\lambda) = g(\lambda)$.

当 $\boldsymbol{A}(\lambda)$ 的全部 k 阶子式为零时, $\boldsymbol{B}(\lambda)$ 全部 k 阶子式也就等于零; 反之亦然. 因此, $\boldsymbol{A}(\lambda)$ 与 $\boldsymbol{B}(\lambda)$ 即有相同的各级行列式因子, 又有相同的秩. □

☞ 求 λ-矩阵的行列式因子可以利用任何一个与之等价的 λ-矩阵来求, 特别地,
可以利用其 Smith 标准形来求. 2.270

Theorem 196. *λ-矩阵 $\boldsymbol{A}(\lambda)$ 的 Smith 标准形*

$$\boldsymbol{B}(\lambda) = \operatorname{diag}\left[\varphi_1(\lambda), \varphi_2(\lambda), \cdots, \varphi_r(\lambda), 0, \cdots, 0\right]$$

是唯一的, 且

$$\varphi_k(\lambda) = \frac{D_k(\lambda)}{D_{k-1}(\lambda)}, \qquad k = 1, 2, \cdots, r. \tag{26}$$

而当 $k > r$ 时, $\varphi_k(\lambda) = 0$. 其中 $r = \operatorname{rank} \boldsymbol{A}(\lambda)$.

证: 标准形 $\boldsymbol{B}(\lambda)$ 中, 非零的一阶子式共有 r 个, 它们的最高公因式显然是 $\varphi_1(\lambda)$; 非零的二阶子式共有 C_r^2 个, 它们的最高公因式显然是 $\varphi_1(\lambda)\varphi_2(\lambda)$; 非零的 k $(k \leqslant r)$ 阶子式有 C_r^k 个, 它们的最高公因式显然是 $\varphi_1(\lambda)\varphi_2(\lambda)\cdots\varphi_k(\lambda)$.

等价的矩阵有相同的各阶行列式因子, 故 $\boldsymbol{A}(\lambda)$ 的各阶行列式因子为

$$D_k(\lambda) = \varphi_1(\lambda)\varphi_2(\lambda)\cdots\varphi_k(\lambda), \quad k = 1, 2, \cdots, r.$$

从而 (26) 式成立. □ 2.271

☞ 求 λ-矩阵的 Smith 标准形时, 可先求其各阶行列式因子, 然后利用公式

$$\varphi_k(\lambda) = \frac{D_k(\lambda)}{D_{k-1}(\lambda)}, \qquad k = 1, 2, \cdots, r \tag{27}$$

来求其 Smith 标准形中主对角线上的非零多项式 $\varphi_k(\lambda)$, $k = 1, 2, \cdots, r$. 2.272

Definition 197. λ-矩阵 $\boldsymbol{A}(\lambda)$ 的 Smith 标准形中的非零多项式 $\varphi_k(\lambda)(k = 1, 2, \cdots, r)$ 称为 $\boldsymbol{A}(\lambda)$ 的第 k 个不变因子.

Theorem 198. *两个 λ-矩阵等价 $\Leftrightarrow$ 它们有相同的行列式因子或不变因子.*

证: 等式 (27) 说明行列式因子与不变因子是相互确定的. 因此, 说两个矩阵有相同的各阶行列式因子, 就等于说它们有相同的各阶不变因子.

必要性已经由定理 195 证明.

充分性. 若 λ-矩阵 $\boldsymbol{A}(\lambda)$ 与 $\boldsymbol{B}(\lambda)$ 有相同的不变因子, 则 $\boldsymbol{A}(\lambda)$ 与 $\boldsymbol{B}(\lambda)$ 和同
一个标准形等价, 因而 $\boldsymbol{A}(\lambda)$ 与 $\boldsymbol{B}(\lambda)$ 等价. □ 2.273

等式 (27) 表明, 行列式因子之间满足关系

$$D_k(\lambda) \mid D_{k+1}(\lambda), \qquad k = 1, 2, \cdots, r-1. \tag{28}$$

在计算行列式因子时, 常常是先计算最高阶的行列式因子.

Example 199. 求 n 阶可逆矩阵 $\boldsymbol{A}(\lambda)$ 的标准形.

解: $\boldsymbol{A}(\lambda)$ 可逆, 则行列式 $\det \boldsymbol{A}(\lambda)$ 是一个非零的数. 从而

$$D_n(\lambda) = 1.$$

于是由 (28) 知,

$$D_1(\lambda) = D_2(\lambda) = \cdots = D_{n-1}(\lambda) = 1.$$

由 $\varphi_k(\lambda) = \frac{D_k(\lambda)}{D_{k-1}(\lambda)}$, 所以

$$\varphi_1(\lambda) = \varphi_2(\lambda) = \cdots = \varphi_n(\lambda) = 1.$$

因此, 可逆矩阵的标准形是单位矩阵 $\boldsymbol{I}$. □

☞ λ-矩阵可逆的充分必要条件是它与单位矩阵等价. 2.274

Exercise 200 (P.110 习题 (一) 5)**.** 用行列式因子来求下列矩阵的 Smith 标准形:

(2) $\begin{bmatrix} 0 & 0 & 1 & \lambda+2 \\ 0 & 1 & \lambda+2 & 0 \\ 1 & \lambda+2 & 0 & 0 \\ \lambda+2 & 0 & 0 & 0 \end{bmatrix}$.

解: $D_4(\lambda) = (\lambda+2)^4$. 存在 3 阶子式

$$\begin{vmatrix} 0 & 0 & 1 \\ 0 & 1 & \lambda+2 \\ 1 & \lambda+2 & 0 \end{vmatrix} = -1,$$

故 $D_3(\lambda) = 1$, 由 $D_k(\lambda) \mid D_{k+1}(\lambda)$, 从而 $D_1(\lambda) = D_2(\lambda) = 1$. 于是, 由 $\varphi_k(\lambda) = \frac{D_k(\lambda)}{D_{k-1}(\lambda)}$ 得

$$\varphi_1(\lambda) = \varphi_2(\lambda) = \varphi_3(\lambda) = 1,\ \varphi_4(\lambda) = (\lambda+2)^4.$$

因此其 Smith 标准形为 $\operatorname{diag}\left(1, 1, 1, (\lambda+2)^4\right)$. □ 2.275

三. 初等因子

前面我们看到, 两个同型 λ-矩阵等价的充要条件是它们有相同的行列式因子 $D_k(\lambda)$ 或不变因子 $\varphi_k(\lambda)$. 对此可以这样来理解, 引用行列式因子 $D_k(\lambda)$ 我们得到 λ-矩阵等价的充要条件; 把 $D_k(\lambda)$ 用某种方法分解因子, 得出不变因子 $\varphi_k(\lambda)$, 引用 $\varphi_k(\lambda)$, 我们得到 λ-矩阵等价的另一个充要条件. 假如我们再把 $\varphi_k(\lambda)$ 分解为最简的因式, 是否也能得到 λ-矩阵等价的充要条件? 下面就是讨论这个问题. 2.276

Definition 201. 设 λ-矩阵 $\boldsymbol{A}(\lambda)$ 的各个不变因子的分解式为

$$\begin{cases} \varphi_1(\lambda) = (\lambda-\lambda_1)^{e_{11}}(\lambda-\lambda_2)^{e_{12}} \cdots (\lambda-\lambda_k)^{e_{1k}}, \\ \varphi_2(\lambda) = (\lambda-\lambda_1)^{e_{21}}(\lambda-\lambda_2)^{e_{22}} \cdots (\lambda-\lambda_k)^{e_{2k}}, \\ \qquad \vdots \\ \varphi_r(\lambda) = (\lambda-\lambda_1)^{e_{r1}}(\lambda-\lambda_2)^{e_{r2}} \cdots (\lambda-\lambda_k)^{e_{rk}}, \end{cases} \tag{29}$$

这里 $\lambda_1, \lambda_2, \cdots, \lambda_k$ 互不相等, $e_{1j} \leqslant e_{2j} \leqslant \cdots \leqslant e_{rj}$, 而 $j = 1, 2, \cdots, k$, 则式 (29) 中所有 $e_{ij} \neq 0$ 的因式

$$(\lambda - \lambda_j)^{e_{ij}}, \qquad 1 \leqslant i \leqslant r,\ 1 \leqslant j \leqslant k$$

称为 $\boldsymbol{A}(\lambda)$ 的初等因子(相同的必须按出现的次数计算), 所有这些初等因子的集合称为 $\boldsymbol{A}(\lambda)$ 的初等因子组.

2.277

Example 202. 设 12 阶矩阵的不变因子是

$$\underbrace{1, 1, \cdots, 1}_{9\text{ 个}}, (\lambda-1)^2, (\lambda-1)^2(\lambda+1), (\lambda-1)^2(\lambda+1)(\lambda^2+1)^2.$$

按定义, 它的初等因子有 7 个, 即

$$(\lambda-1)^2, (\lambda-1)^2, (\lambda-1)^2, (\lambda+1), (\lambda+1), (\lambda-\mathrm{i})^2, (\lambda+\mathrm{i})^2.$$

其中 $(\lambda-1)^2$ 出现三次, $\lambda+1$ 出现两次. □

2.278

我们知道, $\boldsymbol{A}(\lambda)$ 结果初等变换后, 它的秩不变, 不变因子 $\varphi_k(\lambda)$ 不变, 所以它的初等因子也不变. 因此两个 λ-矩阵等价, 则它们的秩相等并且有相同的初等因子.

反过来也成立, 假如两个 λ-矩阵的秩相等并且初等因子相同, 那么它们等价. 例如有两个 6 阶 λ-矩阵 $\boldsymbol{A}(\lambda)$, $\boldsymbol{B}(\lambda)$, 它们的秩都是 4, 初等因子是

$$\lambda, \lambda^3, \lambda-1, \lambda-1, \lambda+1, (\lambda+1)^2.$$

因为秩是 4, 故它们都有 4 个不变因子. 又根据 $\varphi_{k-1}(\lambda) | \varphi_k(\lambda)$, 它们的 4 个不变因子是

$$\begin{aligned}
\varphi_4(\lambda) &= \lambda^3(\lambda-1)(\lambda+1)^2, \\
\varphi_3(\lambda) &= \lambda(\lambda-1)(\lambda+1), \\
\varphi_2(\lambda) &= \varphi_1(\lambda) = 1.
\end{aligned}$$

2.279

要注意的是, 两个同型 λ-矩阵只初等因子相同, 不一定等价. 比如前述的 $\lambda, \lambda^3, \lambda-1, \lambda-1, \lambda+1, (\lambda+1)^2$ 可能是某个秩为 3 的矩阵 $\boldsymbol{C}(\lambda)$ 的初等因子, 则矩阵 $\boldsymbol{C}(\lambda)$ 的不变因子是

$$\begin{aligned}
\varphi_3(\lambda) &= \lambda^3(\lambda-1)(\lambda+1)^2, \\
\varphi_2(\lambda) &= \lambda(\lambda-1)(\lambda+1), \\
\varphi_1(\lambda) &= 1.
\end{aligned}$$

显然矩阵 $\boldsymbol{C}(\lambda)$ 与 $\boldsymbol{A}(\lambda)$ 或 $\boldsymbol{B}(\lambda)$ 不是等价的.

于是我们得到 λ-矩阵等价的又一个充要条件.

Theorem 203. *两个同型 λ-矩阵等价的充要条件是它们的秩相等并且有相同的初等因子.*

后面化约当标准形时, 重要步骤在求特征矩阵的初等因子, 所以求初等因子是一个重要问题.

初等因子也可由对角阵或分块对角阵求得, 不必求标准形. 这就是下面的定理. 2.280

Theorem 204. 设 λ-矩阵 $\boldsymbol{A}(\lambda)$ 等价于下面的分块对角 λ-矩阵

$$\boldsymbol{B}(\lambda)=\begin{bmatrix} \boldsymbol{B}_1(\lambda) & & & & \\ & \boldsymbol{B}_2(\lambda) & & & \\ & & \ddots & & \\ & & & \boldsymbol{B}_t(\lambda) & \\ & & & & \boldsymbol{O} \end{bmatrix},$$

则各子块 $\boldsymbol{B}_i(\lambda)$, $i=1,2,\cdots,t$ 的各个初等因子均为 $\boldsymbol{A}(\lambda)$ 的初等因子, 且 $\boldsymbol{A}(\lambda)$ 不再有其他初等因子.

证: 初等变换把 $\boldsymbol{B}(\lambda)$ 中某个 $\boldsymbol{B}_i(\lambda)$ 变化时, 不影响其他 $\boldsymbol{B}_j(\lambda)$, $i\neq j$. 所以可以用初等变换把 $\boldsymbol{B}_1(\lambda)$, $\cdots$, $\boldsymbol{B}_t(\lambda)$ 都化成对角线矩阵, 从而把 $\boldsymbol{B}(\lambda)$ 化为对角线. 所以 $\boldsymbol{B}_1(\lambda)$, $\cdots$, $\boldsymbol{B}_t(\lambda)$ 的初等因子就是 $\boldsymbol{B}(\lambda)$ 的全部初等因子, 进而也是 $\boldsymbol{A}(\lambda)$ 的全部初等因子. □ 2.281

Example 205. 求 λ-矩阵

$$\boldsymbol{A}(\lambda)=\begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & \lambda-1 \end{bmatrix}$$

的初等因子.

解: 因为 $\boldsymbol{A}(\lambda)$ 是分块对角阵

$$\boldsymbol{A}(\lambda)=\begin{bmatrix} \lambda & 1 & & \\ 0 & \lambda & & \\ & & \lambda & 0 \\ & & 0 & \lambda-1 \end{bmatrix},$$

其中 $\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$ 的初等因子是 λ^2, $\begin{bmatrix} \lambda & 0 \\ 0 & \lambda-1 \end{bmatrix}$ 的初等因子是 λ, $\lambda-1$. 故 $\boldsymbol{A}(\lambda)$ 的全部初等因子是 λ^2, λ, $\lambda-1$. □ 2.282

Example 206. 求

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

的特征矩阵的初等因子.

2.283

解:

$$\lambda \boldsymbol{I}-\boldsymbol{A}=\begin{bmatrix}\lambda-4 & -6 & 0\\ 3 & \lambda+5 & 0\\ 3 & 6 & \lambda-1\end{bmatrix}\xrightarrow{r_3+r_1}\begin{bmatrix}\lambda-4 & -6 & 0\\ 3 & \lambda+5 & 0\\ \lambda-1 & 0 & \lambda-1\end{bmatrix}$$

$$\xrightarrow{c_1-c_3}\begin{bmatrix}\lambda-4 & -6 & 0\\ 3 & \lambda+5 & 0\\ 0 & 0 & \lambda-1\end{bmatrix}.$$

记 $\boldsymbol{B}_1(\lambda)=\begin{bmatrix}\lambda-4 & -6\\ 3 & \lambda+5\end{bmatrix}$, 则其行列式因子为

$$D_1(\lambda)=1, D_2(\lambda)=|\boldsymbol{B}_1(\lambda)|=(\lambda-1)(\lambda+2).$$

从而其不变因子为 $\varphi_1(\lambda)=1$, $\varphi_2(\lambda)=(\lambda-1)(\lambda+2)$, 故 $\boldsymbol{B}_1(\lambda)$ 的初等因子为 $\lambda-1$, $\lambda+2$. 得 $\lambda\boldsymbol{I}-\boldsymbol{A}$ 的全部初等因子为 $\lambda-1$, $\lambda-1$, $\lambda+2$. □ 2.284

四. 若当标准形

Definition 207. 形式为

$$\boldsymbol{J}=\begin{bmatrix}\lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1\\ & & & & \lambda\end{bmatrix}$$

的矩阵称为若当 (Jordan) 矩阵块, 或简称若当块, 其中 λ 是复数. 由若干个若当块组成的准对角矩阵称为若当矩阵.

若当块可以是一阶的. 2.285

例如

$$\begin{bmatrix}2 & 1 & 0\\ 0 & 2 & 1\\ 0 & 0 & 2\end{bmatrix},\quad \begin{bmatrix}0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 0\end{bmatrix},\quad \begin{bmatrix}\mathrm{i} & 1\\ 0 & \mathrm{i}\end{bmatrix},\quad \begin{bmatrix}3\end{bmatrix}$$

都是若当块 (当然也是若当矩阵). 而

$$\left[\begin{array}{cc|c|ccc}1 & 1 & 0 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0 & 0\\ \hline 0 & 0 & 4 & 0 & 0 & 0\\ \hline 0 & 0 & 0 & -\mathrm{i} & 1 & 0\\ 0 & 0 & 0 & 0 & -\mathrm{i} & 1\\ 0 & 0 & 0 & 0 & 0 & -\mathrm{i}\end{array}\right]$$

是一个若当矩阵.

☞ 对角矩阵是若当矩阵 (因若当块可以是一阶的). 2.286

若当块也可以定义为

$$
\boldsymbol{J}=\begin{bmatrix} \lambda & & & & \\ 1 & \lambda & & & \\ & \ddots & \ddots & & \\ & & 1 & \lambda & \\ & & & 1 & \lambda \end{bmatrix}.
$$

相应地, 若当矩阵形如:

$$
\left[\begin{array}{cc:c:ccc} 1 & 0 & 0 & 0 & 0 & 0 \\ \color{red}{1} & 1 & 0 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 4 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & -\mathrm{i} & 0 & 0 \\ 0 & 0 & 0 & \color{red}{1} & -\mathrm{i} & 0 \\ 0 & 0 & 0 & 0 & \color{red}{1} & -\mathrm{i} \end{array}\right].
$$

2.287

Theorem 208. 设 n 阶正方矩阵 $\boldsymbol{A}\in\mathbb{C}^{n\times n}$ 的特征矩阵 $\lambda\boldsymbol{I}-\boldsymbol{A}$ 的初等因子组为

$$
(\lambda-\lambda_1)^{e_1},\ (\lambda-\lambda_2)^{e_2},\ \cdots,\ (\lambda-\lambda_k)^{e_k},
$$

其中 $\lambda_1, \lambda_2, \cdots, \lambda_k$ 及 $e_1, e_2, \cdots, e_k$ 中可能有相同的, 且 $\sum\limits_{i=1}^{k} e_i=n$, $e_i\geqslant 1$, 则 $\boldsymbol{A}$ 相似于若当矩阵

$$
\boldsymbol{J}=\begin{bmatrix} \boldsymbol{J}_1 & & & \\ & \boldsymbol{J}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{J}_k \end{bmatrix}\in\mathbb{C}^{n\times n},
$$

其中 $\boldsymbol{J}_i$ 是一个 e_i 阶若当块, 其对角线元素为特征值 λ_i.

☞ 任意一个 n 阶复矩阵都相似于一个 n 阶若当矩阵, 且该矩阵的主对角线上的元素均为 $\boldsymbol{A}$ 的特征值. 2.288

Example 209. 求矩阵 $\boldsymbol{A}=\begin{bmatrix} -1 & -2 & 6 \\ -1 & 0 & 3 \\ -1 & -1 & 4 \end{bmatrix}$ 的若当标准形.

解: 首先求 $\lambda\boldsymbol{I}-\boldsymbol{A}$ 的初等因子.

$$
\lambda\boldsymbol{I}-\boldsymbol{A}=\begin{bmatrix} \lambda+1 & 2 & -6 \\ 1 & \lambda & -3 \\ 1 & 1 & \lambda-4 \end{bmatrix}\xrightarrow[r_2-r_3]{r_1-(\lambda+1)r_3}\begin{bmatrix} 0 & -\lambda+1 & -\lambda^2+3\lambda-2 \\ 0 & \lambda-1 & -\lambda+1 \\ 1 & 1 & \lambda-4 \end{bmatrix}
$$

$$\xrightarrow[r_1 \leftrightarrow r_3]{c_2 - c_1, c_3 - (\lambda-4)c_1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \lambda - 1 & -\lambda + 1 \\ 0 & -\lambda + 1 & -\lambda^2 + 3\lambda - 2 \end{bmatrix}$$

$$\xrightarrow{r_3 + r_2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \lambda - 1 & -\lambda + 1 \\ 0 & 0 & -\lambda^2 + 2\lambda - 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & \lambda - 1 & 0 \\ 0 & 0 & (\lambda - 1)^2 \end{bmatrix}.$$

2.289

因此 $\boldsymbol{A}$ 的初等因子是 $\lambda - 1$, $(\lambda - 1)^2$. $\boldsymbol{A}$ 的若当标准形是

$$\boldsymbol{J} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

即存在可逆矩阵 $\boldsymbol{P}$ 使得, $\boldsymbol{A} = \boldsymbol{P}\boldsymbol{J}\boldsymbol{P}^{-1}$. □ 2.290

下面介绍利用特征矩阵 $\lambda\boldsymbol{I} - \boldsymbol{A}$ 的标准形或不变因子, 如何去求出 $\boldsymbol{A}$ 的最小多项式.

Theorem 210. 设方阵 $\boldsymbol{A}$ 的特征矩阵 $\lambda\boldsymbol{I} - \boldsymbol{A}$ 的不变因子依次为 $\varphi_1(\lambda)$, $\varphi_2(\lambda)$, $\cdots$, $\varphi_n(\lambda)$, 其中

$$\varphi_l(\lambda) \mid \varphi_{l+1}(\lambda), \qquad l = 1, 2, 3, \cdots, n-1,$$

则 $\varphi_n(\lambda)$ 即为 $\boldsymbol{A}$ 的最小多项式.

2.291

五. 小结

本小节主要的结论: 矩阵 $\boldsymbol{A}$ 即便不能对角化, 它也可以相似于一个若当矩阵.

考虑到对角矩阵也是若当矩阵, 故 $\boldsymbol{A}$ 可对角化的结论可以和上述结论统一起来, 即:

任何矩阵 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ 都与一个若当矩阵 $\boldsymbol{J}$ 相似. 即存在可逆阵 $\boldsymbol{P} \in \mathbb{C}^{n\times n}$, 使得

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{J},$$

其中 $\boldsymbol{J}$ 主对角线上的元素是 $\boldsymbol{A}$ 的全部特征值. 不考虑 $\boldsymbol{J}$ 中若当块的排列次序, $\boldsymbol{J}$ 是被 $\boldsymbol{A}$ 唯一确定的. 2.292