高等代数 (II) 习题解答

2023 至 2024 学年春季学期

作者：韩勇

组织：深圳大学

时间：2024 年 7 月 1 日

好好学习，天天向上。——毛泽东

目录

第 1 章 线性空间

内容提要

- 向量空间的定义和例子
- 子空间
- 线性相关性
- 极大线性无关组
- 矩阵的秩
- 基与维数
- 坐标与坐标变换
- 线性空间的同构
- 线性方程组的解和商空间

1.1 2023 年 3 月 5 日第一周星期二作业-向量空间的定义和例子

设 V 是一个非空集合, 它的元素称为向量. 设 $\mathbb{F}$ 是一个数域[1], 它的元素称为纯量. 假定我们在 V 上定义了如下两个运算:

- (加法运算) 对任意的 $\boldsymbol{\alpha}, \boldsymbol{\beta} \in V$, V 中有唯一的向量 $\boldsymbol{\alpha}+\boldsymbol{\beta}$ 与之对应, 向量 $\boldsymbol{\alpha}+\boldsymbol{\beta}$ 称为向量 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 的和;
- (数量乘法的运算) 对任意纯量 $\lambda \in \mathbb{F}$, 向量 $\boldsymbol{\alpha} \in V$, V 中有唯一的向量 $\lambda\boldsymbol{\alpha}$ 与之对应, 向量 $\lambda\boldsymbol{\alpha}$ 称为纯量 λ 与向量 $\boldsymbol{\alpha}$ 的积.

定义 1.1 (Vector Space)

如果还满足如下的 8 条运算律, 称 V 为域 $\mathbb{F}$ 上的一个向量空间或者线性空间.

(1) $\boldsymbol{\alpha}+\boldsymbol{\beta}=\boldsymbol{\beta}+\boldsymbol{\alpha}$ (加法交换律)

(2) $\boldsymbol{\alpha}+(\boldsymbol{\beta}+\boldsymbol{\gamma})=(\boldsymbol{\alpha}+\boldsymbol{\beta})+\boldsymbol{\gamma}$ (加法结合律)

(3) $\exists o$, 使得 $\forall \boldsymbol{\alpha} \in V$, 有 $\boldsymbol{\alpha}+o=\boldsymbol{\alpha}$(零元素的存在性)

(4) 对 $\forall \boldsymbol{\alpha} \in V, \exists -\boldsymbol{\alpha} \in V$, 使得 $\boldsymbol{\alpha}+(-\boldsymbol{\alpha})=o$

(5) $1\boldsymbol{\alpha}=\boldsymbol{\alpha}$

(6) $(kl)\boldsymbol{\alpha}=k(l\boldsymbol{\alpha})$

(7) $(k+l)\boldsymbol{\alpha}=k\boldsymbol{\alpha}+l\boldsymbol{\alpha}$

(8) $k(\boldsymbol{\alpha}+\boldsymbol{\beta})=k\boldsymbol{\alpha}+k\boldsymbol{\beta}$

其中, $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}$ 是 V 中任意元素; k, l 是 $\mathbb{F}$ 中任意元素.

♣

- 有时候我们仍用 0 或者黑体的 $\mathbf{0}$ 来表示零向量 o.
- 我们通常用 $\mathbb{R}$ 或者 $\mathbf{R}$ 表示实数域，用 $\mathbb{C}$ 或者 $\mathbf{C}$ 表示复数域。

例题 1.1 特别, 如果数域 $\mathbb{F}$ 为实数域 $\mathbb{R}$, 则 V 称为实线性空间; 如果数域 $\mathbb{F}$ 为复数域 $\mathbb{C}$, 则 V 称为复线性空间. 对所有复数的集合 $\mathbb{C}$, 取数域 $\mathbb{F}$ 为复数域 $\mathbb{C}$, 向量加法取为通常复数的加法, 纯量与向量的乘法取为通常复数的乘法. 这就形成一个复线性空间.

[1] 回忆什么是数域。

例题 1.2 向量空间的例子

(1) 数域 $\mathbb{F}$ 上所有 n 维向量的全体 $\mathbb{F}^n$ 关于向量加法和数乘向量;

(2) 数域 $\mathbb{F}$ 上多项式环 $\mathbb{F}[x]$ 关于多项式加法和数乘多项式;

(3) 数域 $\mathbb{F}$ 上所有次数小于 n 的多项式集合 $P[x]_n$ 关于多项式加法和数乘多项式;

(4) 数域 $\mathbb{F}$ 本身关于加法和乘法;

(5) 数域 $\mathbb{F}$ 上所有 $m \times n$ 级矩阵的全体 $M(\mathbb{F})_{m\times n}$ 关于矩阵的加法和数乘矩阵.

例题 1.3 所有正实数的集合记为 $\mathbb{R}_+$. 取数域 $\mathbb{F}$ 为实数域 $\mathbb{R}$. 在集合 $\mathbb{R}_+$ 中规定向量加法为通常实数的加法, 纯量与向量的乘法规定为通常实数的乘法.

在这两种运算下, 集合 $\mathbb{R}_+$ 不构成实线性空间, 因为非负数与正实数的乘积为非负数, 即纯量与 $\mathbb{R}_+$ 中的向量的乘积并不封闭.

例题 1.4 对集合 $\mathbb{R}_+$, 仍取数域 $\mathbb{F}$ 为实数域 $\mathbb{R}$, 在集合 $\mathbb{R}_+$ 中规定向量加法为通常实数的乘法, 即对任意 $\alpha, \beta \in \mathbb{R}_+$, 规定向量 α 与 β 的和 $\alpha \oplus \beta = \alpha\beta$. 并规定纯量 $\lambda \in \mathbb{R}$ 与向量 $\alpha \in \mathbb{R}_+$ 的乘积 $\lambda \otimes \alpha = \alpha^\lambda$. 可以验证, 集合 $\mathbb{R}_+$ 的这两种运算满足线性空间定义中的八条公理, 其中零向量是实数 1 , 向量 α 的负向量是 α^{-1}. 所以 $\mathbb{R}_+$ 是实线性空间.

例子 1.3 和例子 1.4 说明, 一个集合 V 是否成为线性空间与集合 V 的向量加法以及纯量与向量的乘法如何规定密切相关.

命题 1.1

对任意有限多个向量作加法时, 其和与向量的结和方式以及向量的先后次序无关。

在线性空间的定义中, 只规定了两个向量的和. 至于多个向量的和则未加定义. 怎样规定多个向量的和?

- 先看四个向量的情形. 设 $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\delta} \in V$, 由向量加法的定义, $\boldsymbol{\alpha}+\boldsymbol{\beta}$ 和 $\boldsymbol{\gamma}+\boldsymbol{\delta}$ 有意义, 从而 $(\boldsymbol{\alpha}+\boldsymbol{\beta})+(\boldsymbol{\gamma}+\boldsymbol{\delta})$ 有意义; 又 $\boldsymbol{\beta}+\boldsymbol{\gamma}$ 有意义, 因此, $\boldsymbol{\alpha}+(\boldsymbol{\beta}+\boldsymbol{\gamma})$ 有意义, 所以 $(\boldsymbol{\alpha}+(\boldsymbol{\beta}+\boldsymbol{\gamma}))+\boldsymbol{\delta}$ 也有意义. 于是得到向量 $(\boldsymbol{\alpha}+\boldsymbol{\beta})+(\boldsymbol{\gamma}+\boldsymbol{\delta})$ 与 $(\boldsymbol{\alpha}+(\boldsymbol{\beta}+\boldsymbol{\gamma}))+\boldsymbol{\delta}$.
- 当然, 还可以按照其他结合方式. 先求出向量 $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\delta}$ 中某两个向量的和, 再求出它和其他向量的和, 最后求出这四个向量的和.
- 问题是, 随着结合方式不同, 这四个向量的和是否相同? 性质 1.1 断言, 不论这四个向量的结合方式以及向量的先后次序, 所得到的和总是相等的. 这样就可以把这个和规定为这四个向量的和.

证明

(1) 首先归纳定义向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_k \in V$ 的标准和 $\boldsymbol{\alpha}_1 \oplus \boldsymbol{\alpha}_2 \oplus \cdots \oplus \boldsymbol{\alpha}_k$ 如下: 当 $k=2$ 时, 定义 $\boldsymbol{\alpha}_1 \oplus \boldsymbol{\alpha}_2 = \boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2$. 假设 $\boldsymbol{\alpha}_1 \oplus \boldsymbol{\alpha}_2 \oplus \cdots \oplus \boldsymbol{\alpha}_{k-1}$ 已经定义, 则定义

$$\boldsymbol{\alpha}_1 \oplus \boldsymbol{\alpha}_2 \oplus \cdots \oplus \boldsymbol{\alpha}_k = (\boldsymbol{\alpha}_1 \oplus \boldsymbol{\alpha}_2 \oplus \cdots \oplus \boldsymbol{\alpha}_{k-1}) + \boldsymbol{\alpha}_k.$$

(2) 我们声明

$$(\boldsymbol{\alpha}_1 \oplus \boldsymbol{\alpha}_2 \oplus \cdots \oplus \boldsymbol{\alpha}_\ell) + (\boldsymbol{\alpha}_{\ell+1} \oplus \cdots \oplus \boldsymbol{\alpha}_{\ell+k}) = \boldsymbol{\alpha}_1 \oplus \boldsymbol{\alpha}_2 \oplus \cdots \oplus \boldsymbol{\alpha}_{k+\ell}. \tag{1.1}$$

其中 k 与 ℓ 是正整数. 对 k 用归纳法. 当 $k=1$ 时, 由标准和的定义,

$$(\boldsymbol{\alpha}_1\oplus\boldsymbol{\alpha}_2\oplus\cdots\oplus\boldsymbol{\alpha}_{\ell-1}+\boldsymbol{\alpha}_\ell)=\boldsymbol{\alpha}_1\oplus\boldsymbol{\alpha}_2\oplus\cdots\oplus\boldsymbol{\alpha}_{\ell+1}.$$

因此, 当 $k=1$ 时式 (1.1) 成立. 假设 (1.1) 对 $k-1$ 成立. 则由标准和的定义以及结合律有

$$\begin{aligned}&(\boldsymbol{\alpha}_1\oplus\cdots\oplus\boldsymbol{\alpha}_\ell)+(\boldsymbol{\alpha}_{\ell+1}\oplus\cdots\oplus\boldsymbol{\alpha}_{\ell+k})\\=&(\boldsymbol{\alpha}_1\oplus\cdots\oplus\boldsymbol{\alpha}_\ell)+((\boldsymbol{\alpha}_{\ell+1}\oplus\cdots\oplus\boldsymbol{\alpha}_{\ell+k-1})+\boldsymbol{\alpha}_{\ell+k})\\=&((\boldsymbol{\alpha}_1\oplus\cdots\oplus\boldsymbol{\alpha}_\ell)+(\boldsymbol{\alpha}_{\ell+1}\oplus\cdots\oplus\boldsymbol{\alpha}_{\ell+k-1}))+\boldsymbol{\alpha}_{\ell+k}.\end{aligned}$$

由归纳假设,

$$(\boldsymbol{\alpha}_1\oplus\cdots\oplus\boldsymbol{\alpha}_\ell)+(\boldsymbol{\alpha}_{\ell+1}\oplus\cdots\oplus\boldsymbol{\alpha}_{\ell+k-1})=\boldsymbol{\alpha}_1\oplus\cdots\oplus\boldsymbol{\alpha}_\ell\oplus\boldsymbol{\alpha}_{\ell+1}\oplus\cdots\oplus\boldsymbol{\alpha}_{\ell+k-1}.$$

于是

$$(\boldsymbol{\alpha}_1\oplus\boldsymbol{\alpha}_2\oplus\cdots\oplus\boldsymbol{\alpha}_\ell)+(\boldsymbol{\alpha}_{\ell+1}\oplus\cdots\oplus\boldsymbol{\alpha}_{l+k})=(\boldsymbol{\alpha}_1\oplus\boldsymbol{\alpha}_2\oplus\cdots\oplus\boldsymbol{\alpha}_{\ell+k+1})+\boldsymbol{\alpha}_{l+k}$$

再由标准和的定义,

$$(\boldsymbol{\alpha}_1\oplus\boldsymbol{\alpha}_2\oplus\cdots\oplus\boldsymbol{\alpha}_\ell)+(\boldsymbol{\alpha}_{\ell+1}\oplus\cdots\oplus\boldsymbol{\alpha}_{\ell+k})=\boldsymbol{\alpha}_1\oplus\boldsymbol{\alpha}_2\oplus\cdots\oplus\boldsymbol{\alpha}_{\ell+k}$$

(3) 接下来说明, 改变向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 的结合方式, 所得到的和等于这 n 个向量的标准和. 事实上, 假设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n$ 按一种结合方式求得个和; 那么最后一步这个和必定是某个向量 $\boldsymbol{\beta}$ 与另外一个向量 $\boldsymbol{\gamma}$ 的和 $\boldsymbol{\beta}+\boldsymbol{\gamma}$. 其中 $\boldsymbol{\beta}$ 是前 k 个向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_k$ 按某种结合方式得到, 而 $\boldsymbol{\gamma}$ 是向量 $\boldsymbol{\alpha}_{k+1},\ldots,\boldsymbol{\alpha}_n$ 按某种结合方式得到的和. 对向量个数 n 用归纳法, 因为 $k<n$, 那么根据归纳假设有

$$\boldsymbol{\beta}=\boldsymbol{\alpha}_1\oplus\boldsymbol{\alpha}_2\oplus\cdots\oplus\boldsymbol{\alpha}_k,\quad \boldsymbol{\gamma}=\boldsymbol{\alpha}_{k+1}\oplus\cdots\oplus\boldsymbol{\alpha}_n.$$

因此由(1.1) 可得

$$\boldsymbol{\beta}+\boldsymbol{\gamma}=(\boldsymbol{\alpha}_1\oplus\boldsymbol{\alpha}_2\oplus\cdots\oplus\boldsymbol{\alpha}_k)+(\boldsymbol{\alpha}_{k+1}\oplus\cdots\oplus\boldsymbol{\alpha}_n)=\boldsymbol{\alpha}_1\oplus\boldsymbol{\alpha}_2\oplus\cdots\oplus\boldsymbol{\alpha}_n.$$

这就证明, 对向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n$, 不论按何种结合方式, 得到的和都等于标准和 $\boldsymbol{\alpha}_1\oplus\boldsymbol{\alpha}_2\oplus\cdots\oplus\boldsymbol{\alpha}_n$.

(4) 最后一步说明，改变顺序后不改变标准和. 我们要说明

$$\boldsymbol{\alpha}_1\oplus\boldsymbol{\alpha}_2\oplus\cdots\oplus\boldsymbol{\alpha}_n=\boldsymbol{\alpha}_{i_1}\oplus\boldsymbol{\alpha}_{i_2}\oplus\cdots\oplus\boldsymbol{\alpha}_{i_n}.$$

仍用归纳法. $n=2$ 显然成立。假定 $\leq n-1$ 都成立. 如果 $i_n=n$, 那么根据归纳假设

$$\boldsymbol{\alpha}_{i_1}\oplus\boldsymbol{\alpha}_{i_2}\oplus\cdots\oplus\boldsymbol{\alpha}_{i_{n-1}}=\boldsymbol{\alpha}_1\oplus\boldsymbol{\alpha}_2\oplus\cdots\oplus\boldsymbol{\alpha}_{n-1}.$$

从而根据标准和的定义

$$\begin{aligned}&\boldsymbol{\alpha}_{i_1}\oplus\boldsymbol{\alpha}_{i_2}\oplus\cdots\oplus\boldsymbol{\alpha}_{i_{n-1}}\oplus\boldsymbol{\alpha}_{i_n}=(\boldsymbol{\alpha}_{i_1}\oplus\boldsymbol{\alpha}_{i_2}\oplus\cdots\oplus\boldsymbol{\alpha}_{i_{n-1}})+\boldsymbol{\alpha}_n\\=&(\boldsymbol{\alpha}_1\oplus\boldsymbol{\alpha}_2\oplus\cdots\oplus\boldsymbol{\alpha}_{n-1})+\boldsymbol{\alpha}_n\\=&\boldsymbol{\alpha}_1\oplus\boldsymbol{\alpha}_2\oplus\cdots\oplus\boldsymbol{\alpha}_n\end{aligned}$$

如果 $i_n \neq n$, 那么存在 $1 \leq j < n$ 使得 $i_n = j$. 利用 (1.1) 可得

$$
\begin{aligned}
&\boldsymbol{\alpha}_{i_1} \oplus \boldsymbol{\alpha}_{i_2} \oplus \boldsymbol{\alpha}_{i_{j-1}} \oplus \boldsymbol{\alpha}_n \oplus \boldsymbol{\alpha}_{i_{j+1}} \cdots \oplus \boldsymbol{\alpha}_{i_n} \\
\overset{(1.1)}{=}&(\boldsymbol{\alpha}_{i_1} \oplus \boldsymbol{\alpha}_{i_2} \oplus \boldsymbol{\alpha}_{i_{j-1}}) + (\boldsymbol{\alpha}_n \oplus \boldsymbol{\alpha}_{i_{j+1}} \cdots \oplus \boldsymbol{\alpha}_{i_n}) \\
\overset{(1.1)}{=}&(\boldsymbol{\alpha}_{i_1} \oplus \boldsymbol{\alpha}_{i_2} \oplus \boldsymbol{\alpha}_{i_{j-1}}) + (\boldsymbol{\alpha}_n + (\boldsymbol{\alpha}_{i_{j+1}} \cdots \oplus \boldsymbol{\alpha}_{i_n})) \\
=&(\boldsymbol{\alpha}_{i_1} \oplus \boldsymbol{\alpha}_{i_2} \oplus \boldsymbol{\alpha}_{i_{j-1}}) + ((\boldsymbol{\alpha}_{i_{j+1}} \cdots \oplus \boldsymbol{\alpha}_{i_n}) + \boldsymbol{\alpha}_n) \quad (\text{交换律}) \\
\overset{(1.1)}{=}&(\boldsymbol{\alpha}_{i_1} \oplus \boldsymbol{\alpha}_{i_2} \oplus \boldsymbol{\alpha}_{i_{j-1}}) + (\boldsymbol{\alpha}_{i_{j+1}} \cdots \oplus \boldsymbol{\alpha}_{i_n} \oplus \boldsymbol{\alpha}_n) \quad (\text{或者根据标准和的定义}) \\
\overset{(1.1)}{=}&\boldsymbol{\alpha}_{i_1} \oplus \boldsymbol{\alpha}_{i_2} \oplus \boldsymbol{\alpha}_{i_{j-1}} \oplus \boldsymbol{\alpha}_{i_{j+1}} \cdots \oplus \boldsymbol{\alpha}_{i_n} \oplus \boldsymbol{\alpha}_n
\end{aligned}
$$

这又归到了 $i_n = n$ 的情形了.

命题 1.2 (零元素和负元素唯一)

在一个向量空间 V 里, 零向量是唯一的; 对于 V 中每一向量 $\boldsymbol{\alpha}, \boldsymbol{\alpha}$ 的负向量是由 $\boldsymbol{\alpha}$ 唯一确定的.

- 设 o 和 o' 都是 V 的零向点. 那么对于 V 中任意向量 $\boldsymbol{\alpha}$ 都有
$$o + \boldsymbol{\alpha} = \boldsymbol{\alpha}, \boldsymbol{\alpha} + \boldsymbol{o}' = \boldsymbol{\alpha}.$$
于是
$$o = o + o' = o'$$
- 现在设 $\boldsymbol{\alpha}'$ 和 $\boldsymbol{\alpha}''$ 都是 $\boldsymbol{\alpha}$ 的负向量. 那么 $\boldsymbol{\alpha}' + \boldsymbol{\alpha} = o, \boldsymbol{\alpha}'' + \boldsymbol{\alpha} = o$. 于是
$$\boldsymbol{\alpha}' = \boldsymbol{\alpha}' + o = \boldsymbol{\alpha}' + (\boldsymbol{\alpha} + \boldsymbol{\alpha}'') = (\boldsymbol{\alpha}' + \boldsymbol{\alpha}) + \boldsymbol{\alpha}'' = o + \boldsymbol{\alpha}'' = \boldsymbol{\alpha}''$$
- 通常记向量 $\boldsymbol{\alpha}$ 的唯一的负元素为 $-\boldsymbol{\alpha}$.
- 记 $\boldsymbol{\beta} + (-\boldsymbol{\alpha})$ 为 $\boldsymbol{\beta} - \boldsymbol{\alpha}$, 这个称为向量 $\boldsymbol{\beta}$ 与向量 $\boldsymbol{\alpha}$ 的差。于是我们便定义了向量的减法运算.

命题 1.3

设 $\boldsymbol{\alpha}, \boldsymbol{\beta}$ 是域 F 上向量空间 V 中任意两个向量, 则

(i) $0\boldsymbol{\alpha} = o, ko = o; k$ 为任意实数;

(ii) 如 $k\boldsymbol{\alpha} = o$, 那么 $k = 0$ 或者 $\boldsymbol{\alpha} = o$;

(iii) $k(-\boldsymbol{\alpha}) = (-k)\boldsymbol{\alpha} = -(k\boldsymbol{\alpha})$

(iv) $(-1)\boldsymbol{\alpha} = -\boldsymbol{\alpha}$;

(i)

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

同理可证 $ko = o$. 事实上,

$$ko = ko + o = ko + (ko + (-ko)) = (ko + ko) + (-ko) = k(o + o) + (-ko) = ko + (-ko) = o.$$

(iii) 由 (i), 我们有

$$k\boldsymbol{\alpha}+k(-\boldsymbol{\alpha})=k(\boldsymbol{\alpha}+(-\boldsymbol{\alpha}))=ko=o.$$

于是 $k(-\boldsymbol{\alpha})$ 是 $k\boldsymbol{\alpha}$ 的负向量. 所以 $k(-\boldsymbol{\alpha})=-(k\boldsymbol{\alpha})$. 同样因为

$$k\boldsymbol{\alpha}+(-k)\boldsymbol{\alpha}=(k-k)\boldsymbol{\alpha}=0\boldsymbol{\alpha}=o.$$

于是 $(-k)\boldsymbol{\alpha}$ 是 $k\boldsymbol{\alpha}$ 的负向量. 所以 $(-k)\boldsymbol{\alpha}=-(k\boldsymbol{\alpha})$.

(ii) 设 $k\boldsymbol{\alpha}=0$ 但 $k\neq 0$. 那么

$$\boldsymbol{\alpha}=1\boldsymbol{\alpha}=\left(\frac{1}{k}k\right)\boldsymbol{\alpha}=\frac{1}{k}(k\boldsymbol{\alpha})=\frac{1}{k}o=o.$$

最后一个等号是因为 (i). 所以 (ii) 成立.

(iv) 根据 (iii), 令 $k=1$ 再结合向量空间的定义 $1\boldsymbol{\alpha}=\boldsymbol{\alpha}$ 便得到.

引理 1.1 (消去律)

如果 $\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}\in V$ 满足 $\boldsymbol{\alpha}+\boldsymbol{\gamma}=\boldsymbol{\beta}+\boldsymbol{\gamma}$, 则 $\boldsymbol{\alpha}=\boldsymbol{\beta}$. ♡

我们约定:

- 设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 是 $\mathbb{F}$ 上向量空间 V 的 n 个向量. 我们把它们排成一行, 写成一个以向量为元素的 $1\times n$ 矩阵

$$(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)$$

- 对于任意向量 $\boldsymbol{\alpha}$ 和 $\mathbb{F}$ 中任意数 a, 约定 $\boldsymbol{\alpha}a=a\boldsymbol{\alpha}$.

定义 1.2 (由向量组成的矩阵与数量矩阵的乘法)

设 $\boldsymbol{A}=(a_{ij})$ 是数域 $\mathbb{F}$ 上一个 $n\times m$ 矩阵. 我们定义

$$(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)\boldsymbol{A}=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m),$$

这里

$$\boldsymbol{\beta}_j=\sum_{i=1}^n a_{ij}\boldsymbol{\alpha}_i=a_{1j}\boldsymbol{\alpha}_1+\cdots+a_{nj}\boldsymbol{\alpha}_n,\quad 1\leqslant j\leqslant m.$$

也就是说, 按照数域 $\mathbb{F}$ 上矩阵的乘法来定义 $(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n)$ 右乘以 $\boldsymbol{A}$. 设 $\boldsymbol{A}$ 是 $\mathbb{F}$ 上一个 $n\times m$ 矩阵, $\boldsymbol{B}$ 是 $\mathbb{F}$ 上一个 $m\times p$ 矩阵. 根据标量与向量的乘法所满足的算律, 容易证明,

$$(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)(\boldsymbol{AB})=((\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)\boldsymbol{A})\boldsymbol{B}$$

♣

练习 1.1 以下的集合 V 关于所规定的运算是否成为线性空间?

(1) 取 V 为所有实数对 (x_1,x_2) 的集合; 数域 $\mathbb{F}$ 为实数域 $\mathbb{R}$, 向量的加法规定为:

$$(x_1,x_2)+(y_1,y_2)=(x_1+y_1,x_2+y_2+x_1y_1),\quad\forall(x_1,x_2),(y_1,y_2)\in V$$

纯量与向量的乘法规定为:

$$\lambda(x_1,x_2)=\left(\lambda x_1,\lambda x_2+\frac{\lambda(\lambda-1)}{2}x_1^2\right),\quad\forall\lambda\in\mathbb{F},(x_1,x_2)\in V,$$

(2) 取 V 为所有实数对 (x_1, x_2) 的集合; 数域 $\mathbb{F}$ 为实数域 $\mathbb{R}$; 向量的加法规定为:

$$(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2), \quad \forall (x_1, x_2), (y_1, y_2) \in V,$$

纯量与向量的乘法规定为:

$$\lambda (x_1, x_2) = (x_1, x_2), \quad \forall \lambda \in F, (x_1, x_2) \in V.$$

(3) 取 V 为所有满足 $f(x^2) = f(x)^2$ 的实函数集合; $\mathbb{F}$ 为实数域; 向量的加法规定为: 对 $f, g \in V, (f+g)(x) = f(x) + g(x)$; 纯量与向量的乘法规定为: 对 $\lambda \in \mathbb{F}, f \in V, (\lambda f)(x) = \lambda f(x)$.

(4) 取 V 为所有满足 $f(-1) = 0$ 的实函数集合; 数域 $\mathbb{F}$ 为实数域 $\mathbb{R}$; 向量的加法规定为函数的加法; 纯量与向量的乘法规定为实数与函数的乘法.

解

(1) 首先对于任意 $\boldsymbol{\alpha}, \boldsymbol{\beta} \in V, V$ 中有唯一的向量 $\boldsymbol{\alpha} + \boldsymbol{\beta}$ 与之对应, 并且对任意纯量 $\lambda \in \mathbb{F}$, 向量 $\boldsymbol{\alpha} \in V, V$ 中有唯一的向量 $\lambda\boldsymbol{\alpha}$ 与之对应.

1. 加法结合律:

$$\begin{aligned}(\boldsymbol{\alpha} + \boldsymbol{\beta}) + \boldsymbol{\gamma} &= ((x_1, x_2) + (y_1, y_2)) + (z_1, z_2) \\ &= (x_1 + y_1, x_2 + y_2 + x_1 y_1) + (z_1, z_2) \\ &= (x_1 + y_1 + z_1, x_2 + y_2 + z_2 + x_1 y_1 + x_1 z_1 + y_1 z_1).\end{aligned}$$

而

$$\begin{aligned}\boldsymbol{\alpha} + (\boldsymbol{\beta} + \boldsymbol{\gamma}) &= (x_1, x_2) + ((y_1, y_2) + (z_1, z_2)) \\ &= (x_1, x_2) + (y_1 + z_1, y_2 + z_2 + y_1 z_1) \\ &= (x_1 + y_1 + z_1, x_2 + y_2 + z_2 + x_1 y_1 + x_1 z_1 + y_1 z_1).\end{aligned}$$

两者相等.

2. 加法交换律

$$\begin{aligned}\boldsymbol{\alpha} + \boldsymbol{\beta} &= (x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2 + x_1 y_1) \\ \boldsymbol{\beta} + \boldsymbol{\alpha} &= (y_1, y_2) + (x_1, x_2) = (y_1 + x_1, y_2 + x_2 + y_1 x_1).\end{aligned}$$

两者相等.

3. 具有零向量. 取 $0 = (0, 0)$, 我们有

$$\boldsymbol{\alpha} + 0 = (x_1, x_2) + (0, 0) = (x_1, x_2).$$

4. 具有负向量. 对 $\alpha = (x_1, x_2)$, 我们取 $-\boldsymbol{\alpha} = (-x_1, -x_2 + x_1^2)$, 则有

$$\boldsymbol{\alpha} + (-\boldsymbol{\alpha}) = (x_1, x_2) + \left(-x_1, -x_2 + x_1^2\right) = (0, 0).$$

5. 一方面,

$$\begin{aligned}\lambda(\mu\boldsymbol{\alpha}) = \lambda(\mu(x_1, x_2)) &= \lambda\left(\mu x_1, \mu x_2 + \frac{\mu(\mu-1)}{2} x_1^2\right) \\ &= \left(\lambda\mu x_1, \lambda\mu x_2 + \frac{\lambda\mu(\mu-1)}{2} x_1^2 + \frac{\lambda(\lambda-1)\mu^2}{2} x_1^2\right) \\ &= \left(\lambda\mu x_1, \lambda\mu x_2 + \frac{\lambda\mu(\lambda\mu-1)}{2} x_1^2\right).\end{aligned}$$

另一方面，

$$(\lambda\mu)\boldsymbol{\alpha} = (\lambda\mu)(x_1, x_2) = \left(\lambda\mu x_1, \lambda\mu x_2 + \frac{\lambda\mu(\lambda\mu - 1)}{2}x_1^2\right).$$

两者相等.

6.

$$1 \cdot \boldsymbol{\alpha} = 1 \cdot (x_1, x_2) = \left(1 \cdot x_1, 1 \cdot x_2 + \frac{1 \times (1-1)}{2}x_1^2\right) = (x_1, x_2).$$

7. 乘法对向量加法的分配律.

$$\begin{aligned}
\lambda(\boldsymbol{\alpha} + \boldsymbol{\beta}) &= \lambda\left((x_1, x_2) + (y_1, y_2)\right) = \lambda\left(x_1 + y_1, x_2 + y_2 + x_1y_1\right) \\
&= \left(\lambda\left(x_1 + y_1\right), \lambda\left(x_2 + y_2 + x_1y_1\right) + \frac{\lambda(\lambda-1)}{2}\left(x_1 + y_1\right)^2\right) \\
&= \left(\lambda x_1 + \lambda y_1, \lambda x_2 + \lambda y_2 + \lambda x_1y_1 + \frac{\lambda(\lambda-1)}{2}\left(x_1 + y_1\right)^2\right) \\
&= \left(\lambda x_1 + \lambda y_1, \frac{\lambda(\lambda-1)}{2}x_1^2 + \frac{\lambda(\lambda-1)}{2}y_1^2 + \lambda^2 x_1y_1 + \lambda x_2 + \lambda y_2\right).
\end{aligned}$$

而

$$\begin{aligned}
\lambda\boldsymbol{\alpha} + \lambda\boldsymbol{\beta} &= \lambda\left(x_1, x_2\right) + \lambda\left(y_1, y_2\right) \\
&= \left(\lambda x_1, \lambda x_2 + \frac{\lambda(\lambda-1)}{2}x_1^2\right) + \left(\lambda y_1, \lambda y_2 + \frac{\lambda(\lambda-1)}{2}y_1^2\right) \\
&= \left(\lambda x_1 + \lambda y_1, \lambda x_2 + \frac{\lambda(\lambda-1)}{2}x_1^2 + \lambda y_2 + \frac{\lambda(\lambda-1)}{2}y_1^2 + \lambda^2 x_1y_1\right) \\
&= \left(\lambda x_1 + \lambda y_1, \frac{\lambda(\lambda-1)}{2}x_1^2 + \frac{\lambda(\lambda-1)}{2}y_1^2 + \lambda^2 x_1y_1 + \lambda x_2 + \lambda y_2\right).
\end{aligned}$$

8. 乘法对纯量加法的分配律.

$$(\lambda+\mu)\boldsymbol{\alpha} = (\lambda+\mu)(x_1, x_2) = \left((\lambda+\mu)x_1, (\lambda+\mu)x_2 + \frac{(\lambda+\mu)(\lambda+\mu-1)}{2}x_1^2\right).$$

而

$$\begin{aligned}
&\lambda\boldsymbol{\alpha} + \mu\boldsymbol{\alpha} = \lambda\left(x_1, x_2\right) + \mu\left(x_1, x_2\right) \\
=& \left(\lambda x_1, \lambda x_2 + \frac{\lambda(\lambda-1)}{2}x_1^2\right) + \left(\mu x_1, \mu x_2 + \frac{\mu(\mu-1)}{2}x_1^2\right) \\
=& \left(\lambda x_1 + \mu x_1, \lambda x_2 + \frac{\lambda(\lambda-1)}{2}x_1^2 + \mu x_2 + \frac{\mu(\mu-1)}{2}x_1^2 + \lambda\mu x_1^2\right) \\
=& \left((\lambda+\mu)x_1, (\lambda+\mu)x_2 + \frac{(\lambda+\mu)(\lambda+\mu-1)}{2}x_1^2\right).
\end{aligned}$$

两者相等. 所以此集合 V 关于所规定的运算能成为线性空间.

(2) 由于

$$\begin{aligned}
&(\lambda+\mu)\boldsymbol{\alpha} = (\lambda+\mu)\left(x_1, x_2\right) = \left(x_1, x_2\right) \\
&\lambda\boldsymbol{\alpha} + \mu\boldsymbol{\alpha} = \lambda\left(x_1, x_2\right) + \mu\left(x_1, x_2\right) = \left(x_1, x_2\right) + \left(x_1, x_2\right) = \left(2x_1, 2x_2\right).
\end{aligned}$$

两者并不相等, 所以此集合 V 关于所规定的运算不能成为线性空间.

(3) 首先 $1, x \in V$. 取 $\alpha = 1, \beta = x$, 则 $\alpha + \beta = 1 + x \notin V$, 故此集合 V 关于所规定的运算不

能成为线性空间.

(4) 若 $f,g\in V$, 即 $f(-1)=g(-1)=0$, 则 $(f+g)(-1)=f(-1)+g(-1)=0, (\lambda f)(-1)=\lambda f(-1)=0$, 所以 $f+g,\lambda f\in V$. 我们可以验证八条公理对此集合 V 成立, 所以此集合 V 关于所规定的运算能成为线性空间.

✍ **练习 1.2** 对集合 $\mathbb{R}_+$, 仍取数域 $\mathbb{F}$ 为实数域 $\mathbb{R}$, 在集合 $\mathbb{R}_+$ 中规定向量加法为通常实数的乘法, 即对任意 $\alpha,\beta\in\mathbb{R}_+$, 规定向量 α 与 β 的和 $\alpha\oplus\beta=\alpha\beta$. 并规定纯量 $\lambda\in\mathbb{R}$ 与向量 $\alpha\in\mathbb{R}_+$ 的乘积 $\lambda\otimes\alpha=\alpha^\lambda$. 证明 V 是 $\mathbb{R}$ 上的向量空间。

解

1. 对于加法，显然，封闭性，结合律，交换律成立. 存在加法单位元 $(1,1,\ldots,1)$ 有

$$\begin{aligned}(1,1,\ldots,1)\oplus(a_1,a_2,\ldots,a_n)&=(a_1,a_2,\ldots,a_n)\oplus(1,1,\ldots,1)\\&=(a_1,a_2,\ldots,a_n)\end{aligned}$$

由于为正实数向量，则对于 $(a_1,\ldots,a_n)$，存在唯一的逆元 $\left(\frac{1}{a_1},\ldots,\frac{1}{a_n}\right)$，使得

$$(a_1,\ldots,a_n)\oplus\left(\frac{1}{a_1},\ldots,\frac{1}{a_n}\right)=(1,\ldots,1).$$

2. 对于数乘（纯量乘法），显然有封闭性成立。有

(a).

$$\begin{aligned}&\lambda\otimes(\mu\otimes(a_1,\ldots,a_n))\\=&\lambda\otimes(a_1^\mu,\ldots,a_n^\mu)\\=&((a_1^\mu)^\lambda,\ldots,(a_n^\mu)^\lambda)\\=&(a_1^{\lambda\mu},\ldots,a_n^{\lambda\mu})\end{aligned}$$

因此 $\lambda\otimes(\mu\otimes(a_1,\ldots,a_n))=(\lambda\mu)\otimes(a_1,\ldots,a_n)$ 成立.

(b).

$$\begin{aligned}&(\lambda+\mu)\otimes(a_1,\ldots,a_n)\\=&(a_1^{\mu+\lambda},\ldots,a_n^{\mu+\lambda})\\=&(a_1^\lambda a_1^\mu,\ldots,a_n^\lambda a_n^\mu)\\=&(a_1^\lambda,\ldots,a_n^\lambda)\oplus(a_1^\mu,\ldots,a_n^\mu),\end{aligned}$$

因此

$$(\lambda+\mu)\otimes(a_1,\ldots,a_n)=\lambda\otimes(a_1,\ldots,a_n)\oplus\mu\otimes(a_1,\ldots,a_n).$$

(c).

$$\begin{aligned}&\lambda\otimes((a_1,\ldots,a_n)\oplus(b_1,\ldots,b_n))\\=&\lambda\otimes(a_1b_1,\ldots,a_nb_n)=(a_1^\lambda b_1^\lambda,\ldots,a_n^\lambda b_n^\lambda)\\=&(a_1^\lambda,\ldots,a_n^\lambda)\oplus(b_1^\lambda,\ldots,b_n^\lambda)\end{aligned}$$

因此

$$\lambda \otimes ((a_1, \dots, a_n) + (b_1, \dots, b_n)) = \lambda \otimes (a_1, \dots, a_n) + \lambda \otimes (b_1, \dots, b_n).$$

综上有 $\mathbb{R}_+^n$ 对所定义的加法和数乘构成实数域线性空间.

八个条件你要一一验证啊，对你理解向量空间的定义很有帮助。

1.2 2023 年 3 月 7 日第一周星期四作业-子空间

不要害怕数域 $\mathbb{F}$, 你可以永远把它想象成实数域 $\mathbb{R}$ 或者复数域 $\mathbb{C}$.

设 V 是数域 $\mathbb{F}$ 上一个向量空间. W 是 V 的一个非空子集. 对于 W 中任意两个向量 $\boldsymbol{\alpha}, \boldsymbol{\beta}$, 它们的和 $\boldsymbol{\alpha}+\boldsymbol{\beta}$ 是 V 中一个向量. 一般说来, $\boldsymbol{\alpha}+\boldsymbol{\beta}$ 不一定在 W 内.

定义 1.3 (加法封闭)

如果对任意的 $\boldsymbol{\alpha}, \boldsymbol{\beta} \in W \Rightarrow \boldsymbol{\alpha}+\boldsymbol{\beta} \in W$, 那么就说, W 对于 V 的加法是封闭的. ♣

定义 1.4 (乘法封闭)

如果对于 W 中任意向量 $\boldsymbol{\alpha}$ 和数域 $\mathbb{F}$ 中任意数 $a, a\boldsymbol{\alpha} \in W$ ，那么就说, W 对于标量与向量的乘法是封闭的. ♣

定理 1.1 (加法和数乘封闭形成子空间)

设 W 是数域 $\mathbb{F}$ 上向量空间 V 的一个非空子集. 如果 W 对于 V 的加法以及标量与向量的乘法是封闭的，那么 W 本身也作成 $\mathbb{F}$ 上一个向量空间. ♡

定义 1.5 (子空间)

令 W 是数域 $\mathbb{F}$ 上向量空间 V 的一个非空子集. 如果 W 对于 V 的加法以及标量与向量的乘法来说是封闭的，那么就称 W 是 V 的一个子空间. ♣

注

检验 W 是子空间的步骤:

(1) 验证 $W \subseteq V$

(2) 验证 $0 \in W$（当然第一步已经验证，这里要强调一下）。

(3) 验证加法封闭

(4) 验证纯量乘法封闭

例题 1.5 给定 $a, b \in \mathbb{R}$, 令

$$L = \left\{(x, y) \in \mathbb{R}^2 : ax + by = 0\right\}.$$

则 L 是 $\mathbb{R}^2$ 的子空间。

例题 1.6 给定 $a, b \in \mathbb{R}$, 令

$$L = \left\{(x, y) \in \mathbb{R}^2 : ax + by = 1\right\}.$$

L 是 $\mathbb{R}^2$ 的子空间吗?

例题 1.7 子空间的例子

- 向量空间 V 总是它自身的一个子空间.
- 单独一个零向量所成的集合 $\{0\}$ 显然对于 V 的加法和标量与向量的乘法是封闭的，因而也是 V 的一个子空间，称为零空间.
- 一个向量空间 V 本身和零空间叫作 V 的平凡子空间. V 的非平凡子空间叫作 V 的真子空间.
- $\mathbb{F}[x]$ 中次数不超过一个给定的整数 n 的多项式全体连同零多项式一起作成 $\mathbb{F}[x]$ 的一个子空间.
- 闭区间 $[a, b]$ 上一切可微分函数作成 $C[a, b]$ 的一个子空间.

定理 1.2 (子空间的等价刻画)

数域 $\mathbb{F}$ 上向量空间 V 的一个非空子集 W 是 V 的一个子空间, 当且仅当对于任意 $a, b \in \mathbb{F}$ 和任意 $\boldsymbol{\alpha}, \boldsymbol{\beta} \in W$ ，都有 $a\boldsymbol{\alpha} + b\boldsymbol{\beta} \in W$.

引理 1.2 (子空间的交也是子空间)

设 W_1, W_2 是向量空间 V 的两个子空间. 那么它们的交 $W_1 \cap W_2$ 也是 V 的一个子空间. 从而任意子空间的交 $\bigcap_{i \in I} W_i$ 也是子空间.

注 一般来说，子空间的并不是子空间.

引理 1.3 (子空间的和也是子空间)

设 W_1, W_2 是向量空间 V 的两个子空间. 定义

$$\widetilde{V} := \{\boldsymbol{\alpha} + \boldsymbol{\beta} : \boldsymbol{\alpha} \in W_1, \quad \boldsymbol{\beta} \in W_2\}. \tag{1.2}$$

则 $\widetilde{V}$ 也是 V 的子空间, 称为 W_1 和 W_2 的和, 记为 $W_1 + W_2$. 这也可以推广到任意 n 个子空间的和.

命题 1.4

The union of two subspaces of V is a subspace of V if and only if one of the subspaces is contained in the other.

证明

- Let U_1, U_2 be subspaces of V.
- Suppose that one of the subspaces is contained in the other. Then either $U_1 \cup U_2 = U_1$ or $U_1 \cup U_2 = U_2$, and in both cases $U_1 \cup U_2$ is a subspace of V.
- Conversely, suppose by way of contradiction that $U_1 \cup U_2$ is a subspace of V, but neither subspace is contained in the other.
- That is, the sets $U_1 \setminus U_2$ and $U_2 \setminus U_1$ are both nonempty.
- Let $x \in U_1 \setminus U_2$ and $y \in U_2 \setminus U_1$. We claim $x + y \notin U_1$ and $x + y \notin U_2$, so that $U_1 \cup U_2$ is not

closed under addition, a contradiction.

- To see this, suppose $x+y\in U_1$. Then $(x+y)-x\in U_1$ by closure of addition in U_1, but this is absurd since this implies $y\in U_1$, contrary to assumption.
- Similarly, suppose $x+y\in U_2$. Then $(x+y)-y\in U_2$, which is also absurd since this implies $x\in U_2$.
- Therefore $U_1\cup U_2$ is not closed under addition, producing a contradiction as claimed. Thus we must have one of the subspaces contained in the other, as desired.

例题 **1.8** 设 V 是一个向量空间, 且 $V\neq\{0\}$, 证明: V 不可能表成它的两个真子空间的并集.

证明 In fact, if $V=W_1\cup W_2$ with W_1 and W_2 both being the non-trivial subspaces of V. Then $W_1\subseteq W_2$ or $W_2\subseteq W_1$ according to Proposition 1.4. Hence

$$V=W_1\cup W_2=W_1 \text{ or } W_2,$$

a contradiction.

不要害怕数域 $\mathbb{F}$, 你可以永远把它想象成实数域 $\mathbb{R}$ 或者复数域 $\mathbb{C}$.

练习 **1.3** 证明在向量空间的定义 1.1 中, 存在逆元素的条件 (4) 可以替换为

$$0\boldsymbol{\alpha}=o \text{ 对任意的 } \quad \boldsymbol{\alpha}\in V.$$

这里左边的 0 是数域 $\mathbb{F}$ 中的 0, 而右边的 o 指的是空间 V 中的加法零元素（第 (3) 条）.

解 首先假定任意的 $v\in V$ 都有一个逆元素，由于

$$0v+0v=(0+0)v=0v,$$

两边都加上 $0v$ 的逆元素后就要可得到 $0v=0$. 反之，假定 $0v=0$ 对任意的 $v\in V$ 都成立。则

$$v+(-1)v=(1+(-1))v=0v=0,$$

于是 v 有一个逆元素 $(-1)v$.

练习 **1.4** 对下面的每一个 $\mathbb{F}^3$ 的子空间[2], 判断其是否是 $\mathbb{F}^3$ 的子空间:

1. $\{(x_1,x_2,x_3)\in\mathbb{F}^3 \mid x_1+2x_2+3x_3=0\}$
2. $\{(x_1,x_2,x_3)\in\mathbb{F}^3 \mid x_1+2x_2+3x_3=4\}$
3. $\{(x_1,x_2,x_3)\in\mathbb{F}^3 \mid x_1x_2x_3=0\}$
4. $\{(x_1,x_2,x_3)\in\mathbb{F}^3 \mid x_1=5x_3\}$

解

1. 记 S 为对应的集合. 这里的 S 是子空间。这是因为首先

$$0+2\cdot 0+3\cdot 0=0,$$

从而 $0\in S$. 假定 $x=(x_1,x_2,x_3)\in S$ 和 $y=(y_1,y_2,y_3)\in S$. 则

$$x_1+2x_2+3x_3=0 \quad \text{并且} \quad y_1+2y_2+3y_3=0,$$

[2]回忆 $\mathbb{F}^3$ 的定义

于是

$$(x_1+2x_2+3x_3)+(y_1+2y_2+3y_3)=(x_1+y_1)+2(x_2+y_2)+3(x_3+y_3)$$
$$=0,$$

从而 $x+y\in S$。也就是说 S 关于加法封闭. 取 $a\in\mathbb{F}$, 则

$$a(x_1+2x_2+3x_3)=ax_1+2(ax_2)+3(ax_3)=0,$$

从而 $ax\in S$。于是 S 关于纯量乘法封闭。综上 S 是 $\mathbb{F}^3$ 的子空间.

2. 记 S 为对应的集合. 则 S 不是子空间，这是因为

$$0+2\cdot 0+3\cdot 0=0,$$

于是 S 不包含加法零元素 $(0,0,0)$。

3. 记 S 为对应的集合. 则 S 不是子空间，这是因为：记 $x=(1,0,0)$ 与 $y=(0,1,1)$. 则 $x,y\in S$, 但是 $(x_1+y_1)(x_2+y_2)(x_3+y_3)=1\cdot 1\cdot 1\neq 0$. 从而 $x+y=(1,1,1)\notin S$。那么 S 对加法不封闭.

4. 记 S 为对应的集合. 这里的 S 是子空间。这是因为首先 $0=5\cdot 0$, 于是 $0\in S$. 假定 $x=(x_1,x_2,x_3)\in S$ 和 $y=(y_1,y_2,y_3)\in S$. 则有

$$x_1=5x_3 \quad \text{以及} \quad y_1=5y_3,$$

从而

$$(x_1+y_1)=5x_3+5y_3=5(x_3+y_3),$$

进而得到 $x+y\in S$。从而 S 对加法封闭. 取 $a\in\mathbb{F}$, 则

$$a(x_1)=a(5x_3)$$

因此

$$(ax_1)=5(ax_3),$$

从而 $ax\in S$. 于是 S 关于纯量乘法封闭。综上 S 是 $\mathbb{F}^3$ 的子空间.

✍ **练习 1.5** 举一个例子：找到 $\mathbb{R}^2$ 的一个子集 U，使得 U 在加法下封闭，并且对取负元素封闭 (意思是对任意的 $u\in U$，一定有 $-u\in U$), 但是 U 又不是 $\mathbb{R}^2$ 的子空间.

解 考虑 $U=\mathbb{Z}\times\mathbb{Z}$. 这里 $\mathbb{Z}$ 代表所有的整数。则对任意的 $(a,b)\in\mathbb{Z}\times\mathbb{Z}$, 有 $(-a,-b)\in\mathbb{Z}\times\mathbb{Z}$, 因此在取加法逆元的运算下封闭。同样, 对任意的 $(c,d)\in\mathbb{Z}\times\mathbb{Z}$, 有 $(a,b)+(c,d)=(a+c,b+d)\in\mathbb{Z}\times\mathbb{Z}$, 因此在加法运算下封闭. 但是 $\mathbb{Z}\times\mathbb{Z}$ 不是 $\mathbb{R}^2$ 的子空间, 因为它关于纯量乘法不封闭. 例如, $\frac{1}{2}(1,1)=\left(\frac{1}{2},\frac{1}{2}\right)\notin\mathbb{Z}\times\mathbb{Z}$.

✍ **练习 1.6** 证明, 在数域 $\mathbb{F}$ 上向量间 V 里, 以下算律成立:

(i) $a(\boldsymbol{\alpha}-\boldsymbol{\beta})=a\boldsymbol{\alpha}-a\boldsymbol{\beta}$;

(ii) $(a-b)\boldsymbol{\alpha}=a\boldsymbol{\alpha}-b\boldsymbol{\alpha}$,

这里 $a,b\in\mathbb{F},\boldsymbol{\alpha},\boldsymbol{\beta}\in V$.

解

(i) 根据分配律

$$a(\boldsymbol{\alpha}-\boldsymbol{\beta})=a\boldsymbol{\alpha}+a(-\boldsymbol{\beta}).$$

又因为根据命题 1.3 有 $a(-\boldsymbol{\beta})=-a\boldsymbol{\beta}$. 从而 $a(\boldsymbol{\alpha}-\boldsymbol{\beta})=a\boldsymbol{\alpha}-a\boldsymbol{\beta}$

(ii) 根据分配律

$$(a-b)\boldsymbol{\alpha}=a\boldsymbol{\alpha}+(-b)\boldsymbol{\alpha}.$$

又因为根据命题 1.3的第三条有 $(-b)\boldsymbol{\alpha}=-b\boldsymbol{\alpha}$. 从而 $(a-b)\boldsymbol{\alpha}=a\boldsymbol{\alpha}-b\boldsymbol{\alpha}$.

✍ **练习 1.7** 证明：数域 $\mathbb{F}$ 上一个向量空间如果含有一个非零向量, 那么它一定含有无限多个向量.

解 取 $\boldsymbol{\alpha}\neq 0$ 为非零向量。我们说明任意的 $a\neq b\in\mathbb{F}$, 都有 $a\alpha\neq b\alpha$. 这是因为如果 $a\boldsymbol{\alpha}=b\boldsymbol{\alpha}$, 则两边加上 $-(b\boldsymbol{\alpha})=(-b)\boldsymbol{\alpha}$ (命题 1.3 的第三条) 再结合分配律可得

$$(a-b)\boldsymbol{\alpha}=a\alpha-b\alpha=0.$$

因为 $a\neq b,\boldsymbol{\alpha}\neq 0$. 这与命题 1.3 的第一条矛盾。于是

$$\{a\boldsymbol{\alpha}:a\in\mathbb{F}\}\subset V.$$

而 $\mathbb{F}$ 有无穷个元素，从而 V 一定含有无限多个向量.

✍ **练习 1.8** 证明: 对于任意正整数 n 和任意向量 $\boldsymbol{\alpha}$, 都有

$$n\boldsymbol{\alpha}=\overbrace{\boldsymbol{\alpha}+\cdots+\boldsymbol{\alpha}}^{n\text{个}}.$$

你要理解左右两边的是什么意思啊。

解 利用归纳法。显然 $n=1$ 和 $n=2$ 成立。假设对 $n-1$ 成立，那么

$$(n-1)\boldsymbol{\alpha}=\overbrace{\boldsymbol{\alpha}+\cdots+\boldsymbol{\alpha}}^{n-1\text{个}}.$$

于是

$$n\boldsymbol{\alpha}=((n-1)+1)\boldsymbol{\alpha}=(n-1)\boldsymbol{\alpha}+\boldsymbol{\alpha}=\overbrace{\boldsymbol{\alpha}+\cdots+\boldsymbol{\alpha}}^{n-1\text{个}}+\boldsymbol{\alpha}=\overbrace{\boldsymbol{\alpha}+\cdots+\boldsymbol{\alpha}}^{n\text{个}}.$$

1.3 2024 年 3 月 12 日第二周星期二作业-线性相关性

定义 1.6 (线性组合-linear combination)

设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 是向量空间 V 的 r 个向量，$a_1,a_2,\cdots,a_r$ 是数域 $\mathbb{F}$ 中任意 r 个数。我们把和

$$a_1\boldsymbol{\alpha}_1+a_2\boldsymbol{\alpha}_2+\cdots+a_r\boldsymbol{\alpha}_r$$

称为向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 的一个线性组合[a]。

[a] A linear combination of a list $v_1,\ldots,v_m$ of vectors in V is a vector of the form

$$a_1v_1+\cdots+a_mv_m,$$

where $a_1, \ldots, a_m \in \mathbf{F}$.

定义 1.7 (线性表示-linear representation)

如果 V 中某一向量 $\boldsymbol{\alpha}$ 可以表成向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r$ 的线性组合，我们也说 $\boldsymbol{\alpha}$ 可以由 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r$ 线性表示。

例题 1.9 在 $\mathbf{R}^3$ 里，取

$$\boldsymbol{\alpha}_1 = (1, -1, 0), \boldsymbol{\alpha}_2 = (0, 2, 1), \boldsymbol{\alpha}_3 = (1, -1, 2).$$

那么

$$2\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_2 + 3\boldsymbol{\alpha}_3 = 2(1, -1, 0) - (0, 2, 1) + 3(1, -1, 2) = (5, -7, 5)$$

所以向量 $(5, -7, 5)$ 可以由 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$ 线性表示。

定义 1.8

设 S 是数域 $\mathbb{F}$ 上的线性空间 V 的非空向量集合. V 中所有包含向量集合 S 的子空间的交称为由向量集合 S 生成的子空间, 记为 $V(S)$.

特别, 如果 $S = \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_k\}$, 则由 S 生成的子空间记为 $V(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_k)$, 并且称为由向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_k$ 生成的子空间. 事实上, $V(S)$ 是所有 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_k$ 的线性组合形成的空间. 我们也记[3]

$$V(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_k) = \mathcal{L}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_k).$$

向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 叫作这个子空间的一组生成元.

定义 1.9 (线性相关)

设 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r$ 是向量空间 V 的 r 个向量。如果存在 $\mathbb{F}$ 中不全为零的数 $a_1, a_2, \cdots, a_r$，使得

$$a_1\boldsymbol{\alpha}_1 + a_2\boldsymbol{\alpha}_2 + \cdots + a_r\boldsymbol{\alpha}_r = 0,$$

那么就说 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r$ 线性相关[a]。否则称 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r$ 线性无关。

[a]Let V be a vector space and let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ be a set of vectors in V. Then $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ is linearly independent if the only scalars $c_1, c_2, \ldots, c_p$ that satisfy the equation

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p = \mathbf{0}$$

are the trivial scalars $c_1 = c_2 = \cdots = c_p = 0$. If the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ is not linearly independent then we say that it is linearly dependent.

如果向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r$ 中有一个是零向量, 那么 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r$ 一定线性相关. 事实上, 例如，设 $\boldsymbol{\alpha}_1 = \mathbf{0}$. 那么

$$1\boldsymbol{\alpha}_1 + 0\boldsymbol{\alpha}_2 + \cdots + 0\boldsymbol{\alpha}_r = 0,$$

[3]张禾瑞的书

其中 $\boldsymbol{\alpha}_1$ 的系数不等于零.

例题 **1.10** Show that the following set of 2×2 matrices is linearly dependent:

$$\left\{\mathbf{A}_1=\begin{bmatrix}1&2\\0&-1\end{bmatrix},\mathbf{A}_2=\begin{bmatrix}-1&3\\1&0\end{bmatrix},\quad \mathbf{A}_3=\begin{bmatrix}5&0\\-2&-3\end{bmatrix}\right\}.$$

- It is clear that $\mathbf{A}_1$ and $\mathbf{A}_2$ are linearly independent, i.e., $\mathbf{A}_1$ cannot be written as a scalar multiple of $\mathbf{A}_2$, and vice-versa.
- Since the $(2,1)$ entry of $\mathbf{A}_1$ is zero, the only way to get the -2 in the $(2,1)$ entry of $\mathbf{A}_3$ is to multiply $\mathbf{A}_2$ by -2 .
- Similary, since the $(2,2)$ entry of $\mathbf{A}_2$ is zero, the only way to get the -3 in the $(2,2)$ entry of $\mathbf{A}_3$ is to multiply $\mathbf{A}_1$ by 3. Hence, we suspect that $3\mathbf{A}_1-2\mathbf{A}_2=\mathbf{A}_3$. Verify:
$$3\mathbf{A}_1-2\mathbf{A}_2=\begin{bmatrix}3&6\\0&-3\end{bmatrix}-\begin{bmatrix}-2&6\\2&0\end{bmatrix}=\begin{bmatrix}5&0\\-2&-3\end{bmatrix}=\mathbf{A}_3$$
- Therefore, $3\mathbf{A}_1-2\mathbf{A}_2-\mathbf{A}_3=\mathbf{0}$ and thus we have found scalars c_1,c_2,c_3 not all zero such that $c_1\mathbf{A}_1+c_2\mathbf{A}_2+c_3\mathbf{A}_3=\mathbf{0}$.

例题 **1.11**

- 令 $\mathbb{F}$ 是任意一个数域. $\mathbb{R}^3$ 中向量
$$\boldsymbol{\alpha}_1=(1,2,3),\boldsymbol{\alpha}_2=(2,4,6),\boldsymbol{\alpha}_3=(3,5,-4)$$
线性相关, 因为我们有
$$2\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2+0\boldsymbol{\alpha}_3=0.$$
- 向量
$$\boldsymbol{\beta}_1=(1,0,0),\boldsymbol{\beta}_2=(1,1,0),\boldsymbol{\beta}_3=(1,1,1)$$
线性无关.

例题 **1.12** 设 $V=M_{m\times n}(\mathbb{F})$ 为所有的 $m\times n$ 阶矩阵的全体。则
$$\{E_{ij}\mid 1\le i\le m,1\le j\le n\}$$
线性无关。而
$$\{E_{ij},A\mid 1\le i\le m,1\le j\le n\}$$
(A 是 V 中任一元素) 线性相关。这里的 E_{ij} 指的是第 (i,j) 个位置是 1 其余位置是 0 的矩阵.

例题 **1.13** 判断 $\mathbb{R}^3$ 的向量
$$\boldsymbol{\alpha}_1=(1,-2,3),\boldsymbol{\alpha}_2=(2,1,0),\boldsymbol{\alpha}_3=(1,-7,9)$$
是否线性相关.

例题 **1.14** 复向量空间 $\mathbf{C}^1$ 中任意两个向量都是线性相关的.

设 $\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathbf{C}^1$. 如果 $\boldsymbol{\alpha}=\boldsymbol{\beta}=\mathbf{0}$, 则取 $\lambda=\mu=1\in\mathbf{C}$. 于是 $\lambda\boldsymbol{\alpha}+\mu\boldsymbol{\beta}=\mathbf{0}$; 如果 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 不全为零, 则取 $\lambda=\boldsymbol{\beta}\in\mathbf{C},\mu=-\boldsymbol{\alpha}\in\mathbf{C}$. 显然, λ,μ 不全为零, 并且 $\lambda\boldsymbol{\alpha}+\mu\boldsymbol{\beta}=\mathbf{0}$. 因此, $\mathbf{C}^1$ 中任意两个向量都线性相关.

前面说的是有限个向量的线性相关性，那么任意个向量呢？

定义 1.10 (任意多个向量的线性相关性)

设 V 是数域 $\mathbb{F}$ 上的线性空间，$S \subseteq V$。如果存在向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n \in S$ 以及不全为零的纯量 $\lambda_1, \lambda_2, \cdots, \lambda_n \in \mathbb{F}$ 使得

$$\lambda_1 \boldsymbol{\alpha}_1 + \lambda_2 \boldsymbol{\alpha}_2 + \cdots + \lambda_n \alpha_n = 0$$

则称向量集合 S 是线性相关的。不是线性相关的向量集合 S 称为 线性无关的。♣

上面的定义 1.10 是定义 1.9 的一般化，因为取 $S = \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r\}$ 上面的定义 1.10 就与定义 1.9一致。

✍ **练习 1.9(选做)** 如果向量组 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 线性无关，而向量组 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n, \boldsymbol{\alpha}\}$ 线性相关，证明：$\boldsymbol{\alpha}$ 一定可以被 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_n$ 线性表示并且表示方法唯一（不算顺序）。

证明 因为 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n, \boldsymbol{\alpha}\}$ 线性相关，于是存在不全为 0 的数 $x_1, x_2, \cdots, x_{n+1}$ 使得

$$x_1 \boldsymbol{\alpha}_1 + x_2 \boldsymbol{\alpha}_2 + \cdots + x_n \boldsymbol{\alpha}_n + x_{n+1} \boldsymbol{\alpha} = 0.$$

如果 $x_{n+1} \neq 0$，则

$$x_1 \boldsymbol{\alpha}_1 + x_2 \boldsymbol{\alpha}_2 + \cdots + x_n \boldsymbol{\alpha}_n = 0.$$

但是 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 线性无关。于是

$$x_1 = x_2 = \cdots = x_n = 0 = x_{n+1}.$$

矛盾！从而 $x_{n+1} \neq 0$, 于是

$$\boldsymbol{\alpha} = \frac{-x_1}{x_{n+1}} \boldsymbol{\alpha}_1 + \cdots + \frac{x_n}{x_{n+1}} \boldsymbol{\alpha}_n.$$

从而 $\boldsymbol{\alpha}$ 一定可以被 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_n$ 线性表示。如果有两种表示

$$\boldsymbol{\alpha} = x_1 \boldsymbol{\alpha}_1 + \cdots + x_n \boldsymbol{\alpha}_n = y_1 \boldsymbol{\alpha}_1 + \cdots + y_n \boldsymbol{\alpha}_n.$$

则

$$(x_1 - y_1) \boldsymbol{\alpha}_1 + \cdots + (x_n - y_n) \boldsymbol{\alpha}_n = 0.$$

结合 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 线性无关可得

$$0 = x_1 - y_1 = x_2 - y_2 = \cdots = x_n - y_n.$$

于是表示方法唯一。

✍ **练习 1.10(选做)** 证明：若 $\boldsymbol{\alpha}$ 可以被 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_n$ 线性表示并且表示方法唯一（不算顺序），则向量组 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 线性无关。

解 如何向量组 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 线性相关，则存在不全为 0 的数 $x_1, x_2, \cdots, x_n$ 使得

$$x_1 \boldsymbol{\alpha}_1 + \cdots + x_n \boldsymbol{\alpha}_n = 0.$$

因为 $\boldsymbol{\alpha}$ 可以被 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_n$ 线性表示，则有 $y_1, y_2, \cdots, y_n$ 使得

$$\boldsymbol{\alpha} = y_1 \boldsymbol{\alpha}_1 + \cdots + y_n \boldsymbol{\alpha}_n$$

同时

$$\boldsymbol{\alpha}=\boldsymbol{\alpha}+0=y_1\boldsymbol{\alpha}_1+\cdots+y_n\boldsymbol{\alpha}_n+x_1\boldsymbol{\alpha}_1+\cdots+x_n\boldsymbol{\alpha}_n=(x_1+y_1)\boldsymbol{\alpha}_1+\cdots+(x_n+y_n)\boldsymbol{\alpha}_n.$$

因为 $x_1,x_2,\cdots,x_n$ 不全为 0. 所以至少有一个 $x_k+y_k\neq y_k$。这样 $\boldsymbol{\alpha}$ 的表示方法不唯一，矛盾。

练习 1.11 证明: 在实函数空间中, 向量组 $\{1,\cos^2(t),\cos(2t)\}$ 线性相关, 而向量组 $\{1,\sin(t),\cos(t)\}$ 线性无关。

解 因为

$$\cos^2(t)=\frac{1}{2}+\frac{1}{2}\cos(2t)=\frac{1}{2}1+\frac{1}{2}\cos(2t).$$

于是 $\cos^2(t)$ 可以被 1 和 $\cos(2t)$ 的线性表示。从而三者线性相关。

如果数 x,y,z 满足

$$x+y\sin(t)+z\cos(t)=0,\quad \forall t\in\mathbb{R}.$$

取 $t=0$ 可得 $x+z=0$, 取 $t=\frac{\pi}{2}$ 可得 $x+y=0$. 取 $t=\pi$ 可得 $x-z=0$. 于是只能 $x=y=z=0$。

练习 1.12 下列论断哪些是对的, 哪些是错的. 如果是对的, 证明; 如果是错的, 举出反例:

(i) 如果当 $a_1=a_2=\cdots=a_r=0$ 时, $a_1\boldsymbol{\alpha}_1+a_2\boldsymbol{\alpha}_2+\cdots+a_r\boldsymbol{\alpha}_r=0$, 那么 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性无关;

(ii) 如果 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性无关, 而 $\boldsymbol{\alpha}_{r+1}$ 不能由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性表示, 那么

$$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r,\boldsymbol{\alpha}_{r+1}$$

线性无关;

(iii) 如果 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性无关，那么其中每一个向量都不是其余向量的线性组合;

(iv) 如果 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性相关，那么其中每一个向量都是其余向量的线性组合.

解

(i) 不对。取 $\boldsymbol{\alpha}_1\neq 0,\quad \boldsymbol{\alpha}_2=-\boldsymbol{\alpha}_1$, 其他任意就满足条件。

(ii) 对。因为如果

$$x_1\boldsymbol{\alpha}_1+\cdots+x_r\boldsymbol{\alpha}_r+x_{r+1}\boldsymbol{\alpha}_{1+r}=0.$$

那么 $x_{r+1}=0$, 否则 $\boldsymbol{\alpha}_{r+1}$ 就被 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性表示。这样

$$x_1\boldsymbol{\alpha}_1+\cdots+x_r\boldsymbol{\alpha}_r=0.$$

又因为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性无关。于是

$$x_1=x_2=\cdots=x_r=0.$$

这样

$$x_1=x_2=\cdots=x_r=x_{1+r}=0.$$

所以

$$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r,\boldsymbol{\alpha}_{r+1}$$

线性无关。

(iii) 对。因为如果存在 $1 \leq k \leq r$ 和 $x_1, x_2, \cdots, x_{k-1}, x_{k+1}, \cdots, x_r \in \mathbb{F}$ 使得

$$\boldsymbol{\alpha}_k = x_1\boldsymbol{\alpha}_1 + x_2\boldsymbol{\alpha}_2 + \cdots + x_{k-1}\boldsymbol{\alpha}_{k-1} + x_{k+1}\boldsymbol{\alpha}_{k+1} + \cdots + x_r\boldsymbol{\alpha}_r.$$

则

$$x_1\boldsymbol{\alpha}_1 + x_2\boldsymbol{\alpha}_2 + \cdots + x_{k-1}\boldsymbol{\alpha}_{k-1} + (-1)\boldsymbol{\alpha}_k + x_{k+1}\boldsymbol{\alpha}_{k+1} + \cdots + x_r\boldsymbol{\alpha}_r = 0.$$

于是 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r$ 线性相关，矛盾！

(iv) 不对。比如

$$\boldsymbol{\alpha} = (1, 0), \quad \boldsymbol{\beta} = (-1, 0), \quad \boldsymbol{\gamma} = (0, 1) \in \mathbb{R}^2.$$

显然

$$\boldsymbol{\alpha} + \boldsymbol{\beta} + 0\boldsymbol{\gamma} = 0.$$

但是 $\boldsymbol{\gamma}$ 不可能写成 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$ 的线性组合。

练习 1.13(选做) 设 $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}$ 线性无关. 证明 $\boldsymbol{\alpha}+\boldsymbol{\beta}, \boldsymbol{\beta}+\boldsymbol{\gamma}, \boldsymbol{\gamma}+\boldsymbol{\alpha}$ 也线性无关.

解 如果

$$x(\boldsymbol{\alpha}+\boldsymbol{\beta}) + y(\boldsymbol{\beta}+\boldsymbol{\gamma}) + z(\boldsymbol{\gamma}+\boldsymbol{\alpha}) = 0.$$

则整合系数后得到

$$(x+z)\boldsymbol{\alpha} + (x+y)\boldsymbol{\beta} + (y+z)\boldsymbol{\gamma} = 0.$$

而 $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}$ 线性无关，那么

$$\begin{cases} x+y=0 \\ y+z=0 \\ x+z=0. \end{cases}$$

这个方程组只有 $x = y = z0$ 一个解。于是 $\boldsymbol{\alpha}+\boldsymbol{\beta}, \boldsymbol{\beta}+\boldsymbol{\gamma}, \boldsymbol{\gamma}+\boldsymbol{\alpha}$ 也线性无关。

练习 1.14 设向量组 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r\}\,(r \geqslant 2)$ 线性无关. 任取 $k_1, k_2, \cdots, k_{r-1} \in \mathbb{F}$. 证明, 向量组

$$\boldsymbol{\beta}_1 = \boldsymbol{\alpha}_1 + k_1\boldsymbol{\alpha}_r, \boldsymbol{\beta}_2 = \boldsymbol{\alpha}_2 + k_2\boldsymbol{\alpha}_r, \cdots, \boldsymbol{\beta}_{r-1} = \boldsymbol{\alpha}_{r-1} + k_{r-1}\boldsymbol{\alpha}_r, \boldsymbol{\alpha}_r$$

线性无关.

解 如果

$$x_1\boldsymbol{\beta}_1 + \cdots + x_{r-1}\boldsymbol{\beta}_{r-1} + x_r\boldsymbol{\alpha}_r = 0.$$

则有

$$x_1\boldsymbol{\alpha}_1 + \cdots + x_{r-1}\boldsymbol{\alpha}_{r-1} + (x_1k_1 + x_2k_2 + \cdots + x_{r-1}k_{r-1} + x_r)\boldsymbol{\alpha}_r = 0.$$

因为 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r\}\,(r \geqslant 2)$ 线性无关，于是

$$x_1 = x_2 = \cdots = x_{r-1} = x_1k_1 + x_2k_2 + \cdots + x_{r-1}k_{r-1} + x_r = 0.$$

这个方程组只有零解

$$x_1 = x_2 = \cdots = x_r = 0.$$

于是向量组

$$\boldsymbol{\beta}_1 = \boldsymbol{\alpha}_1 + k_1\boldsymbol{\alpha}_r, \boldsymbol{\beta}_2 = \boldsymbol{\alpha}_2 + k_2\boldsymbol{\alpha}_r, \cdots, \boldsymbol{\beta}_{r-1} = \boldsymbol{\alpha}_{r-1} + k_{r-1}\boldsymbol{\alpha}_r, \boldsymbol{\alpha}_r$$

线性无关.

练习 1.15 令 $\boldsymbol{\alpha}_i = (a_{i1}, a_{i2}, \cdots, a_{in}) \in \mathbb{R}^n, i = 1, 2, \cdots, n$. 证明 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 线性相关当且仅当行列式

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = 0$$

解 如果

$$x_1\boldsymbol{\alpha}_1 + \cdots + x_n\boldsymbol{\alpha}_n = 0.$$

则比较分量可得

$$\sum_{i=1}^{n} a_{ij}x_i = 0, \quad j = 1, 2, 3, 4, \cdots, n.$$

也即

$$(x_1, x_2, \cdots, x_n)\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} = 0.$$

根据齐次线性方程组有唯一解的条件知道，上面的方程组有唯一 0 解当且仅当行列式

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \neq 0$$

练习 1.16 设纯量 λ 满足下列条件之一, 求 λ :

(1) (选做) 向量 $(1+\lambda, 1-\lambda), (1-\lambda, 1+\lambda) \in \mathbb{C}^2$ 线性相关,

(2) 向量 $(\lambda, 1, 0), (1, \lambda, 1), (0, 1, \lambda) \in \mathbb{R}^3$ 线性相关.

解

(1) 首先，我们来解决第一个问题。向量 $(1+\lambda, 1-\lambda)$ 和 $(1-\lambda, 1+\lambda)$ 线性相关意味着存在不全为零的常数 c_1 和 c_2 使得:

$$c_1(1+\lambda, 1-\lambda) + c_2(1-\lambda, 1+\lambda) = 0$$

这等价于以下方程组:

$$\begin{cases} c_1(1+\lambda)+c_2(1-\lambda) & =0 \\ c_1(1-\lambda)+c_2(1+\lambda) & =0 \end{cases}$$

解这个方程组可以得到 c_1 和 c_2 的值。为了使得这个方程组有非平凡解，系数行列式应该为 0。计算系数行列式得到：

$$\begin{vmatrix} 1+\lambda & 1-\lambda \\ 1-\lambda & 1+\lambda \end{vmatrix} = (1+\lambda)(1+\lambda)-(1-\lambda)(1-\lambda)=4\lambda$$

因此，系数行列式等于 0 当且仅当 $\lambda=0$。所以对于第一个条件，满足 $\lambda=0$。

(2) 向量 $(\lambda,1,0),(1,\lambda,1),(0,1,\lambda)$ 线性相关意味着存在不全为零的常数 c_1,c_2,c_3 使得：

$$c_1(\lambda,1,0)+c_2(1,\lambda,1)+c_3(0,1,\lambda)=0$$

即，

$$\begin{cases} c_1\lambda+c_2+0 & =0 \\ c_1+c_2\lambda+c_3 & =0 \\ c_2+c_3\lambda+0 & =0 \end{cases}$$

解这个方程组也可以得到 c_1,c_2,c_3 的值。为了使得方程组有非平凡解，系数行列式应该为 0。计算系数行列式：

$$\begin{vmatrix} \lambda & 1 & 0 \\ 1 & \lambda & 1 \\ 0 & 1 & \lambda \end{vmatrix} = \lambda(\lambda^2-1)-(1)(\lambda)=\lambda^3-\lambda-\lambda=\lambda^3-2\lambda$$

所以系数行列式等于 0 当且仅当 $\lambda=0$ 或者 $\lambda=\pm\sqrt{2}$。综上所述，满足条件的 λ 是 $\lambda=0$ 或者 $\lambda=\pm\sqrt{2}$。

1.4 2024 年 3 月 14 日第二周星期四作业-极大线性无关组

今天是国际数学节，Happy π day!

例题 1.15 设 C_2 是区间 $[0,1]$ 上所有连续实函数的集合, 它关于函数的加法, 实数与函数的乘法成为实线性空间. 设 C_2 中向量集合

$$S=\left\{f_i : f_i(x)=x^i, i=0,1,2,\cdots\right\}.$$

证明: 向量集合 S 线性无关.

用反证法. 设向量集合 S 线性相关, 则存在 k 个向量 $f_{i_1}(x), f_{i_2}(x), \cdots, f_{i_k}(x)\in S$, 以及 k

个不全为零的纯量 $\lambda_1, \lambda_2, \cdots, \lambda_k \in \mathbb{R}$, 使得

$$\lambda_1 f_{i_1}(x) + \lambda_2 f_{i_2}(x) + \cdots + \lambda_k f_{i_k}(x) = 0,$$

即

$$\lambda_1 x^{i_1} + \lambda_2 x^{i_2} + \cdots + \lambda_k x^{i_k} = 0.$$

这表明, 上式左端的多项式为零多项式. 因此, 它的系数 $\lambda_1, \lambda_2, \cdots, \lambda_k$ 全为零. 矛盾. 这就证明, 向量集合 S 是线性无关的.

例题 1.16 $1, x, x^2, \ldots, x^n$ 在 $\mathbb{F}[x]$ 是线性无关的。又若 $f(x) \in \mathbb{F}[x]$，且 $\deg f(x) \leq n$。则 $1, x, x^2, \ldots, x^n, f(x)$ 是线性相关的。

事实上, 由

$$\sum_{i=0}^{n} a_i x^i = 0$$

知 $a_0 = a_1 = \cdots = a_n = 0$ 故 $1, x, x^2, \ldots, x^n$ 线性无关. 又 $\deg f(x) \leq n$. 因而有 $f(x) = \sum_{i=0}^{n} a_i x^i$. 于是

$$a_0 \cdot 1 + a_1 x + \cdots + a_n x^n + (-1) f(x) = 0,$$

$a_0, a_1, \ldots, a_n, -1$ 不全为零. 故 $1, x, x^2, \ldots, x^n, f(x)$ 线性相关.

命题 1.5

(1) 含有零向量的向量集合一定是线性相关的。

(2) 含有线性相关向量子集的向量集合 S 一定是线性相关的。

(3) 线性无关向量集合 S 的任何子集合 S_1 都是线性无关的。

(4) 向量集合 S 线性无关的充分必要条件是，S 的每一个有限子集 (即由有限个向量构成的子集) 都是线性无关的，也就是说，对于 S 的任意一组向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_k$，由

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

一定能得到 $\lambda_1 = \lambda_2 \cdots = \lambda_k = 0$，其中纯量 $\lambda_1, \lambda_2, \cdots, \lambda_k \in \mathbb{F}$。

(3) 是因为: 如果子集合 S_1 不是线性无关的, 则 S_1 线性相关, 由 (2) 可得向量集合 S 线性相关的, 矛盾. (4) 的必要性根据 (3) 可得; 充分性可用反证法证明.

命题 1.6

向量组 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r\}$ 中每一个向量 $\boldsymbol{\alpha}_i$ 都可以由这一组向量线性表示。

命题 1.7

如果向量 $\boldsymbol{\gamma}$ 可以由 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_r$ 线性表示，而每一 $\boldsymbol{\beta}_i$ 又都可以由 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_s$ 线性表示，那么 $\boldsymbol{\gamma}$ 可以由 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_s$ 线性表示。

命题 1.8

设向量组 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r\}$ 线性无关，而 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r,\boldsymbol{\beta}\}$ 线性相关。那么 $\boldsymbol{\beta}$ 一定可以由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性表示。

定理 1.3

向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r(r\geqslant 2)$ 线性相关，当且仅当其中某一个向量是其余向量的线性组合。

定理 1.4

非零向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_k$ 线性相关的充分必要条件是，存在某个向量 $\boldsymbol{\alpha}_m,2\leqslant m\leqslant k$，使得向量 $\boldsymbol{\alpha}_m$ 是它前面 $m-1$ 个向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{m-1}$ 的线性组合。

证明 设非零向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_k$ 线性相关, 则存在不全为零的纯量 $\lambda_1,\lambda_2,\cdots,\lambda_k\in\mathbb{F}$, 使得

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

设 λ_m 是纯量 $\lambda_1,\lambda_2,\cdots,\lambda_k$ 中由后往前数第一个不为零的纯量, 即

$$\lambda_m\neq 0,\lambda_{m+1}=\cdots\lambda_k=0.$$

由于纯量 $\lambda_1,\lambda_2,\cdots,\lambda_k$ 不全为零, 所以 λ_m 是存在的, 又由于向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_k$ 都是非零的, 因此, $2\leqslant m\leqslant k$. 于是,

$$\lambda_1\boldsymbol{\alpha}_1+\lambda_2\boldsymbol{\alpha}_2+\cdots+\lambda_m\boldsymbol{\alpha}_m=\mathbf{0}.$$

所以

$$\boldsymbol{\alpha}_m=\left(-\frac{\lambda_1}{\lambda_m}\right)\boldsymbol{\alpha}_1+\left(-\frac{\lambda_2}{\lambda_m}\right)\boldsymbol{\alpha}_2+\cdots+\left(-\frac{\lambda_{m-1}}{\lambda_m}\right)\boldsymbol{\alpha}_{m-1}.$$

即向量 $\boldsymbol{\alpha}_m$ 是向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{m-1}$ 的线性组合.

反之, 设向量 $\boldsymbol{\alpha}_m$ 是向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{m-1}$ 的线性组合, 则存在纯量 $\lambda_1,\lambda_2,\cdots,\lambda_{m-1}\in\mathbb{F}$, 使得

$$\boldsymbol{\alpha}_m=\lambda_1\boldsymbol{\alpha}_1+\lambda_2\boldsymbol{\alpha}_2+\cdots+\lambda_{m-1}\boldsymbol{\alpha}_{m-1}.$$

于是

$$\lambda_1\boldsymbol{\alpha}_1+\lambda_2\boldsymbol{\alpha}_2+\cdots+\lambda_{m-1}\boldsymbol{\alpha}_{m-1}+(-1)\boldsymbol{\alpha}_m=\mathbf{0}.$$

这表明, 向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_k$ 线性相关. 证毕.

定义 1.11

设

$$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r\}$$

和

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

是向量空间 V 的两个向量组。如果每一 $\boldsymbol{\alpha}_i$ 都可以由 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$ 线性表示，而每一

$\boldsymbol{\beta}_j$ 也可以由 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r$ 线性表示，那么就说这两个向量组等价。一般来说，这里的两个向量组或者说向量集合不一定有限。

引理 1.4 (容易验证)

向量集合之间的等价关系满足以下性质：

- (自反性) 对任意向量集合 $S \subseteq V$，向量集合 S 和自身等价；
- (对称性) 设向量集合 $S, T \subseteq V$，如果 S 与 T 等价，则 T 与 S 等价；
- (传递性) 设向量集合 $S, T, W \subseteq V$，如果 S 与 T 等价，T 与 W 等价，则 S 与 W 等价。

定理 1.5 (替换定理)

设向量组

$$\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r\}$$

线性无关，并且每一 $\boldsymbol{\alpha}_i$ 都可以由向量组

$$\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s\} \tag{1.3}$$

线性表示。那么 $r \leqslant s$，并且必要时可以对 (1.3) 中向量重新编号，使得用 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r$ 替换 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_r$ 后，所得的向量组

$$\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r, \boldsymbol{\beta}_{r+1}, \cdots, \boldsymbol{\beta}_s\}$$

与(1.3)等价。

证明

(1) 首先我们有存在不全为零的 $\lambda_1, \cdots, \lambda_s$ 使得

$$\boldsymbol{\alpha}_1 = \lambda_1 \boldsymbol{\beta}_1 + \cdots + \lambda_s \boldsymbol{\beta}_s$$

通过调整顺序，我们不妨假定 $\lambda_1 \neq 0$。从而有

$$\boldsymbol{\beta}_1 = \frac{1}{\lambda_1}\boldsymbol{\alpha}_1 + \frac{-\lambda_2}{\lambda_1}\boldsymbol{\beta}_2 + \cdots + \frac{-\lambda_s}{\lambda_1}\boldsymbol{\beta}_s.$$

(2) 我们得到

$$\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s\} \quad \text{等价于} \quad \{\boldsymbol{\alpha}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s\}$$

(3) 我们有存在不全为零的 $\lambda_1, \cdots, \lambda_s$ 使得

$$\boldsymbol{\alpha}_2 = \lambda_1 \boldsymbol{\alpha}_1 + \lambda_2 \boldsymbol{\beta}_2 + \cdots + \lambda_s \boldsymbol{\beta}_s.$$

如果

$$\lambda_2 = \lambda_3 = \cdots = \lambda_s = 0,$$

则 $\lambda_1 \neq 0$ 并且 $\boldsymbol{\alpha}_2 = \lambda_1 \boldsymbol{\alpha}_1$，这与 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$ 线性无关矛盾。从而

$$\lambda_2, \lambda_3, \cdots, \lambda_s$$

一定含有非零的数。不妨设 $\lambda_2 \neq 0$。于是

$$\boldsymbol{\beta}_2 = \frac{1}{\lambda_2}\boldsymbol{\alpha}_2 + \frac{-\lambda_1}{\lambda_2}\boldsymbol{\alpha}_1 + \frac{-\lambda_3}{\lambda_2}\boldsymbol{\beta}_3 + \cdots + \frac{-\lambda_s}{\lambda_2}\boldsymbol{\beta}_s.$$

(4) 我们得到

$$\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s\} \quad \text{等价于} \quad \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\beta}_3, \cdots, \boldsymbol{\beta}_s\}$$

(5) 一直进行下去，我们每次都拿新的 $\boldsymbol{\alpha}$ 来带入，这样必定有 $t \leq s$。否则，如果 $s < t$, 那么我们会得到

$$\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s\} \quad \text{等价于} \quad \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_s\}$$

这与 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_{1+s}\}$ 线性无关矛盾!

(6) 根据前面的过程，以及 $r \leq s$, 我们进行 r 步以后就得到

$$\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s\} \quad \text{等价于} \quad \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r, \boldsymbol{\beta}_{1+r}, \cdots, \boldsymbol{\beta}_s\}$$

推论 1.1

设向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_s$ 与 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_t$ 等价, 并且都线性无关, 则 $s = t$.

✍ **练习 1.17** 下面五道题你随便选做两道

(1) 判断 $\mathbf{R}^3$ 中向量 $(1,1,0),(0,1,1),(1,0,-1)$ 的线性相关性;

(2) 判断 $\mathbf{R}^3$ 中向量 $(1,-3,1),(-1,2,-2),(1,1,3)$ 的线性相关性;

(3) 判断 $\mathbf{R}[x]_3$ (次数不超过 3 的多项式) 中 $p_1(x) = 1 + x$, $p_2(x) = 1 - x$, $p_3(x) = x + x^2$ 的线性相关性;

(4) 判断连续函数全体构成的线性空间中 1, $\sin^2 x$, $\cos^2 x$ 的线性相关性;

(5) 判断连续函数全体构成的线性空间中 1, 2^x, 2^{-x} 的线性相关性.

解

1. 根据定义，应求解方程

$$\lambda_1(1,1,0) + \lambda_2(0,1,1) + \lambda_3(1,0,-1) = 0,$$

即

$$\begin{cases} \lambda_1 + \lambda_3 = 0 \\ \lambda_1 + \lambda_2 = 0 \\ \lambda_2 - \lambda_3 = 0 \end{cases}$$

解得基础解系 $k(1,-1,-1)^{\mathrm{T}}$，所以存在非零解，向量组线性相关.

2. 求解方程

$$\lambda_1(1,-3,1) + \lambda_2(-1,2,-2) + \lambda_3(1,1,3) = 0,$$

即

$$\begin{cases}\lambda_1-\lambda_2+\lambda_3=0\\-3\lambda_1+2\lambda_2+\lambda_3=0\\\lambda_1-2\lambda_2+3\lambda_3=0\end{cases}$$

解得 $\lambda_1=\lambda_2=\lambda_3=0$，此向量组线性无关.

3. 求解方程

$$\lambda_1p_1(x)+\lambda_2p_2(x)+\lambda_3p_3(x)=0$$
$$(\lambda_1+\lambda_2)+(\lambda_1-\lambda_2+\lambda_3)x+\lambda_3x^2=0$$

所以需要求解方程组

$$\begin{cases}\lambda_1+\lambda_2=0\\\lambda_1-\lambda_2+\lambda_3=0\\\lambda_3=0\end{cases}$$

解得 $\lambda_1=\lambda_2=\lambda_3=0$，此向量组线性无关.

4. 易知 $-1+\sin^2x+\cos^2x=0$，对应系数为 $-1,1,1$，不全为零，所以此向量组线性相关.

5. 求解方程

$$\lambda_1+\lambda_2 2^{-x}+\lambda_3 2^x=0$$

很明显会发现仅凭此方程是难以求解的，方程数目不足. 注意到此方程应该对于任意的 x 均成立，所以取 $x=0,x=1,x=-1$，得到方程组

$$\begin{cases}\lambda_1+\lambda_2+\lambda_3=0\\\lambda_1+\dfrac{1}{2}\lambda_2+2\lambda_3=0\\\lambda_1+2\lambda_2+\dfrac{1}{2}\lambda_3=0\end{cases}$$

解得 $\lambda_1=\lambda_2=\lambda_3=0$，此向量组线性无关.

注意上述 (3) 到 (5) 题为不能代入特殊的 x 值来说明，例如 (3) 令 $x=0$ 得到线性相关的做法是错误的，因为 (3) 中线性空间就是多项式构成的线性空间，其中的元素就是多项式，不能代入值. 注意 (5) 是特殊题型，需要构造更多的方程来求解这一问题.

✍ **练习 1.18** 求向量 $\boldsymbol{\alpha}_1=(4,-1,3,-2),\boldsymbol{\alpha}_2=(8,-2,6,-4),\boldsymbol{\alpha}_3=(3,-1,4,-2),\boldsymbol{\alpha}_4=(6,-2,8,-4)$ 的所有极大线性无关向量组.

解 注意到

$$\boldsymbol{\alpha}_2=2\boldsymbol{\alpha}_1,\quad \boldsymbol{\alpha}_4=2\boldsymbol{\alpha}_3.$$

而显然 $\boldsymbol{\alpha}_3$ 和 $\boldsymbol{\alpha}_1$ 不能互相表示。于是可得极大线性无关组为

$$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_3\},\quad \{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_4\},\quad \{\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3\},\quad \{\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_4\}.$$

✍ **练习 1.19** 设向量 $\boldsymbol{\beta}$ 可以由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性表示, 但不能由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{r-1}$ 线性表示. 证

明, 向量组 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_{r-1}, \boldsymbol{\alpha}_r\}$ 与向量组 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_{r-1}, \boldsymbol{\beta}\}$ 等价.

解 记

$$S = \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r\}, \quad T = \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_{r-1}, \boldsymbol{\beta}\}$$

因为 $\boldsymbol{\beta}$ 可以由 S 中的向量线性表示，那么 T 中所有的向量可被 S 中的向量线性表示。假定

$$\boldsymbol{\beta} = x_1\boldsymbol{\alpha}_1 + x_2\boldsymbol{\alpha}_2 + \cdots + x_r\boldsymbol{\alpha}_r, \quad x_1, x_2, \cdots, x_r \in \mathbb{F}.$$

因为 β 不能由 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_{r-1}$ 线性表示, 于是 $x_r \neq 0$。从而我们可以除掉 x_r 得到

$$\boldsymbol{\alpha}_r = \frac{1}{x_r} + \frac{-x_1}{x_r}\boldsymbol{\alpha}_1 + \cdots + \frac{-x_{r-1}}{x_r}\boldsymbol{\alpha}_{r-1}.$$

也即 $\boldsymbol{\alpha}_r$ 可以被 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_{r-1}, \boldsymbol{\beta}$ 线性表示。于是 S 中的向量可以被 T 中向量线性表示。于是二者等价。

练习 1.20 设在向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r$ 中, $\boldsymbol{\alpha}_1 \neq 0$ 并且每一 $\boldsymbol{\alpha}_i$ 都不能表成它的前 $i-1$ 个向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_{i-1}$ 的线性组合. 证明 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r$ 线性无关.

是不是可以利用我们前面讲的一个结论？反证法。

解 参见定理 1.4.

练习 1.21 设向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r$ 线性无关, 而 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r, \boldsymbol{\beta}, \boldsymbol{\gamma}$ 线性相关. 证明, 或者 β 与 γ 中至少有一个可以由 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r$ 线性表示, 或者向量组

$$\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r, \boldsymbol{\beta}\}$$

与

$$\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r, \boldsymbol{\gamma}\}$$

等价.

解 首先，根据题目给出的条件，我们知道向量组 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r, \boldsymbol{\beta}, \boldsymbol{\gamma}\}$ 线性相关。那么存在不全为零的标量 $c_1, c_2, \cdots, c_r, d, e$ 使得:

$$c_1\boldsymbol{\alpha}_1 + c_2\boldsymbol{\alpha}_2 + \cdots + c_r\boldsymbol{\alpha}_r + d\boldsymbol{\beta} + e\boldsymbol{\gamma} = \mathbf{0}$$

由于 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r\}$ 是线性无关的，那么 d, e 至少有一个不为 0。我们考虑两种情况:

(1) 如果 $d = 0$ 或 $e = 0$，假设 $d = 0, e \neq 0$，那么我们有:

$$c_1\boldsymbol{\alpha}_1 + c_2\boldsymbol{\alpha}_2 + \cdots + c_r\boldsymbol{\alpha}_r + e\boldsymbol{\gamma} = \mathbf{0}$$

由于 $e \neq 0$，我们可以得出 $\boldsymbol{\gamma}$ 可以由 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r$ 线性表示。同理，如果 $e = 0, d \neq 0$，则 $\boldsymbol{\beta}$ 可以由 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r$ 线性表示。

(2) 如果 $d \neq 0$ 且 $e \neq 0$，那么我们将向量组中的每个向量都除以 d，得到:

$$c_1\left(\frac{\boldsymbol{\alpha}_1}{d}\right) + c_2\left(\frac{\boldsymbol{\alpha}_2}{d}\right) + \cdots + c_r\left(\frac{\boldsymbol{\alpha}_r}{d}\right) + \boldsymbol{\beta} + \left(\frac{e}{d}\right)\boldsymbol{\gamma} = \mathbf{0}$$

由于向量组的等价性不受标量乘法的影响，我们可以得出

$$\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r, \boldsymbol{\beta}\}$$

和

$$\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r, \boldsymbol{\gamma}\}$$

是等价的。

综上所述，要么 $\boldsymbol{\beta}$ 与 $\boldsymbol{\gamma}$ 中至少有一个可以由 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r$ 线性表示，要么向量组

$$\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r, \boldsymbol{\beta}\}$$

和

$$\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r, \boldsymbol{\gamma}\}$$

等价。

练习 1.22 设向量组 $S = \{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_s\}$ 线性无关, 并且可以由向量组 $T = \{\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_t\}$ 线性表出. 求证:

(1) 向量组 T 与 $S \cup T$ 等价.

(2) 将 S 扩充为 $S \cup T$ 的一个极大线性无关组 $T_1 = \left\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_s, \boldsymbol{\beta}_{i_{s+1}}, \ldots, \boldsymbol{\beta}_{i_{s+k}}\right\}$, 则 T_1 与 T 等价, 且 $s + k \leq t$.

(3) 利用 (1) 和 (2) 来证明替换定理: 可以用向量 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_s$ 替换向量 $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_t$ 中某 s 个向量 $\boldsymbol{\beta}_{i_1}, \ldots, \boldsymbol{\beta}_{i_s}$, 使得到的向量组 $\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_s, \boldsymbol{\beta}_{i_{s+1}}, \ldots, \boldsymbol{\beta}_{i_t}\}$ 与 $\{\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_t\}$ 等价.

解

1. 首先 T 中的每一个向量自然可以被 $S \cup T$ 中的向量线性表示；而因为 S 中向量可以被 T 中向量线性表示，从而 $S \cup T$ 中的向量也可以被 T 中的向量线性表示。于是根据定义 $S \cup T$ 与 T 等价。
2. 为了证明这个，我们利用引理 1.5 可得 $S \cup T$ 的任意两个极大线性无关组包含相同的向量个数。根据 (1) 知 T 的极大线性无关组也是 $S \cup T$ 的极大线性无关组组，从而含有的向量个数不会超过 T 含有的向量个数，从而有 $s + k \leq t$.
3. 首先第二问可得 $t - s \geq k$, 于是

$$T = \{\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_t\}$$

中我们在选取

$$\boldsymbol{\beta}_{i_{s+1}}, \ldots, \boldsymbol{\beta}_{i_{s+k}}$$

后剩余的向量个数为 $t - k \geq s$。那么我们从这 $t - k$ 个向量中随便拿出来 s 个用 S 这个向量组替代。记新得到的向量组为

$$H = \left\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_s, \boldsymbol{\beta}_{i_{s+1}}, \ldots, \boldsymbol{\beta}_{i_{s+k}}, \cdots, \boldsymbol{\beta}_{i_t}\right\}.$$

则 $T_1 \subseteq H \subseteq S \cup T$, 因为 T_1 为 $S \cup T$ 的一个极大线性无关组，从而 T_1 为 H 的一个极大线性无关组。从而 H 和 $S \cup T$ 等价。而 (1) 又说 $S \cup T$ 与 T 等价，从而根据传递性 H 与 T 等价。

引理 1.5

如果

$$\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_m\}$$

和

$$\{\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_n\}$$

是两个等价的线性无关组，则一定有 $m = n$.

证明 利用反证法。不妨假定 $n = m + p$, 其中 $p > 0$ 是一个正整数。由于两个向量组等价，则存在 $m \times n = m \times (m+p)$ 阶的矩阵

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{C} & \boldsymbol{D} \end{pmatrix}, \quad \boldsymbol{C} \in M_{m\times m}(\mathbb{F}), \quad \boldsymbol{D} \in M_{m\times p}(\mathbb{F})$$

和一个 $n \times m = (n+p) \times m$ 阶的矩阵

$$\boldsymbol{B} = \begin{pmatrix} \boldsymbol{E} \\ \boldsymbol{F} \end{pmatrix}, \quad \boldsymbol{E} \in M_{m\times m}(\mathbb{F}), \quad \boldsymbol{F} \in M_{p\times m}(\mathbb{F})$$

使得

$$(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m)\boldsymbol{A}, \quad (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m) = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n)\boldsymbol{B}$$

于是我们有

$$(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m)\boldsymbol{AB}, \quad (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m)\boldsymbol{BA}.$$

因为这两组向量都是线性无关的，因此能被它们自身的表达方式唯一。于是有

$$\boldsymbol{AB} = \mathbf{I}_m, \quad \boldsymbol{BA} = \mathbf{I}_n.$$

这里的 $\mathbf{I}_m$ 和 $\mathbf{I}_n$ 分别是 m 阶和 n 阶的恒等矩阵。也即我们有

$$\boldsymbol{AB} = \boldsymbol{CE} + \boldsymbol{DF} = \mathbf{I}_m, \quad \boldsymbol{BA} = \begin{pmatrix} \boldsymbol{EC} & \boldsymbol{ED} \\ \boldsymbol{FC} & \boldsymbol{FD} \end{pmatrix} = \mathbf{I}_n$$

于是我们有

$$\begin{cases} \boldsymbol{EC} = \mathbf{I}_m \\ \boldsymbol{FD} = \mathbf{I}_p \\ \boldsymbol{ED} = 0 \\ \boldsymbol{FC} = 0 \\ \boldsymbol{CE} + \boldsymbol{DF} = \mathbf{I}_m \end{cases} \tag{1.4}$$

由 (1.4) 第一个式子可得 $\boldsymbol{E}, \boldsymbol{C}$ 都是 $m \times m$ 阶可逆矩阵，从而 $\boldsymbol{CE} = \mathbf{I}_m$，带入到 (1.4) 最后一个式子可得

$$\boldsymbol{DF} = 0.$$

两边乘以 $\boldsymbol{D}$ 再结合 (1.4) 第二个式子 $\boldsymbol{FD} = \mathbf{I}_p$ 可得

$$0 = \boldsymbol{DFD} = \boldsymbol{D}\mathbf{I}_p = \boldsymbol{D}.$$

于是有

$$\boldsymbol{D} = \mathbf{0}$$

与 (1.4) 第二个式子矛盾!

1.5 2024 年 3 月 19 日第三周星期二作业-矩阵的秩

定义 1.12 (极大线性无关组)

设向量集合 $S=\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s\}\subseteq V$。如果 S 中的向量 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t$ 线性无关,并且对任意 $\boldsymbol{\beta}\in S$,向量 $\boldsymbol{\beta},\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t$ 线性相关,则称向量 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t$ 是向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 的一个极大线性无关向量组。♣

定理 1.6

向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 的任意一个极大线性无关向量组都与向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 等价,而且向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 的任意两个极大线性无关向量组所含向量的个数相同。♡

定义 1.13 (向量组的秩)

向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 的极大线性无关向量组所含向量的个数 r 称为向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 的秩。♣

定理 1.7 (矩阵的行秩列秩相等)

矩阵 $\boldsymbol{A}\in\mathbb{F}^{m\times n}$ 的行秩等于列秩,并等于矩阵 $\boldsymbol{A}$ 的秩。♡

设 $\operatorname{rank}\boldsymbol{A}=r$. 则矩阵 $\boldsymbol{A}$ 具有 r 阶非零子式, 设

$$\boldsymbol{A}\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ j_1 & j_2 & \cdots & j_r \end{pmatrix}\neq 0, \tag{1.5}$$

其中 $1\leqslant i_1<i_2<\cdots<i_r\leqslant m, 1\leqslant j_1<j_2<\cdots<j_r\leqslant n$. 考察矩阵 $\boldsymbol{A}$ 相应的行向量 $\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\ldots,\boldsymbol{\alpha}_{i_r}$. 设 $\lambda_1,\lambda_2,\ldots,\lambda_r\in\mathbb{F}$, 使得

$$\lambda_1\boldsymbol{\alpha}_{i_1}+\lambda_2\boldsymbol{\alpha}_{i_2}+\cdots+\lambda_r\boldsymbol{\alpha}_{i_r}=0.$$

写成分量形式, 上式即化为

$$\begin{cases} a_{i_1 1}\lambda_1+a_{i_2 1}\lambda_2+\cdots+a_{i_r 1}\lambda_r=0 \\ a_{i_1 2}\lambda_1+a_{i_2 2}\lambda_2+\cdots+a_{i_r 2}\lambda_r=0 \\ \cdots\cdots\cdots\cdots \\ a_{i_1 n}\lambda_1+a_{i_2 n}\lambda_2+\cdots+a_{i_r n}\lambda_r=0 \end{cases}$$

由于式 (1.5), 上述方程组的系数矩阵的秩为 r, 所以, $\lambda_1=\lambda_2=\cdots=\lambda_r=0$. 这表明, 行向量 $\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\ldots,\boldsymbol{\alpha}_{i_r}$ 线性无关.

其次, 设 $\boldsymbol{\alpha}_k$ 是矩阵 $\boldsymbol{A}$ 的第 k 个行向量, 并且设 $\lambda_1,\lambda_2,\ldots,\lambda_r,\mu\in\mathbb{F}$, 使得

$$\lambda_1\boldsymbol{\alpha}_{i_1}+\lambda_2\boldsymbol{\alpha}_{i_2}+\cdots+\lambda_r\boldsymbol{\alpha}_{i_r}+\mu\boldsymbol{\alpha}_k=0.$$

写成分量形式, 上式子化为

$$\begin{cases} a_{i_1 1}\lambda_1 + a_{i_2 1}\lambda_2 + \cdots + a_{i_r 1}\lambda_r + a_{k1}\mu = 0, \\ a_{i_1 2}\lambda_1 + a_{i_2 2}\lambda_2 + \cdots + a_{i_r 2}\lambda_r + a_{k2}\mu = 0, \\ \cdots\cdots\cdots\cdots \\ a_{i_1 n}\lambda_1 + a_{i_2 n}\lambda_2 + \cdots + a_{i_r n}\lambda_r + a_{kn}\mu = 0. \end{cases}$$

由于式 (1.5), 并 $\operatorname{rank}\boldsymbol{A} = r$, 因此, 当 $1 \leqslant j \leqslant n$ 时,

$$A\begin{pmatrix} i_1 & i_2 & \cdots & i_r & k \\ j_1 & j_2 & \cdots & j_r & j \end{pmatrix} = 0.$$

即上述方程组的系数矩阵的秩为 r. 因此, 上述方程组具有非零解 $\lambda_1, \lambda_2, \ldots, \lambda_r, \mu \in \mathbb{F}$. 所以。向量 $\boldsymbol{\alpha}_{i_1}, \boldsymbol{\alpha}_{i_2}, \ldots, \boldsymbol{\alpha}_{i_k}, \boldsymbol{\alpha}_k$ 线性相关.

由向量 $\boldsymbol{\alpha}_k$ 的任意性, 向量 $\boldsymbol{\alpha}_{i_1}, \boldsymbol{\alpha}_{i_2}, \ldots, \boldsymbol{\alpha}_{i_r}$ 是向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_m$ 的极大线性无关向量组. 因此矩阵 $\boldsymbol{A}$ 的行秩等于矩阵 $\boldsymbol{A}$ 的秩. 同理可证矩阵 $\boldsymbol{A}$ 的列秩也等于矩阵 $\boldsymbol{A}$ 的秩.

例题 1.17 设 $\boldsymbol{A} = (a_{ij}) \in \mathbb{F}^{m\times n}$, $\boldsymbol{B} = (b_{ij}) \in \mathbb{F}^{n\times p}$. 证明:

$$\operatorname{rank}\boldsymbol{AB} \leqslant \min\{\operatorname{rank}\boldsymbol{A}, \operatorname{rank}\boldsymbol{B}\}.$$

例题 1.18 设 $t \leqslant n$，而 t 个 n 维行向量 $\boldsymbol{\alpha}_i = (a_{i1}, a_{i2}, \ldots, a_{in})$, $i = 1, 2, \ldots, t$，满足

$$2\left|a_{ii}\right| > \sum_{k=1}^{n}\left|a_{ik}\right|.$$

证明: 向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_t$ 线性无关。

证明 我们写成矩阵形式

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{t1} & a_{t2} & \cdots & a_{tn} \end{pmatrix}.$$

我们想说 A 秩是 t，从而它的 t 行自然线性无关。由于行秩等于列秩，我们看列秩。记它的列为

$$\beta_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \cdots \\ a_{t1} \end{pmatrix}, \quad \beta_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \cdots \\ a_{t2} \end{pmatrix}, \quad \cdots, \beta_t = \begin{pmatrix} a_{1t} \\ a_{2t} \\ \cdots \\ a_{tt} \end{pmatrix}.$$

事实上，如果

$$x_1\boldsymbol{\beta}_1 + x_2\boldsymbol{\beta}_2 + \cdots + x_t\boldsymbol{\beta}_t = 0.$$

则有

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1t} \\ a_{21} & a_{22} & \cdots & a_{2t} \\ \cdots & \cdots & & \cdots \\ a_{t1} & a_{t2} & \cdots & a_{tt} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

如果 $x_1, x_2, \cdots, x_t$ 中不全为 0。如果 $x-1 = x_2 = \cdots = x_t$, 则

$$a_{11} + a_{12} + \cdots + a_{1t} = 0.$$

这不可能。我们不妨设 $|x_1|$ 严格最大。于是

$$x_1 a_{11} + x_2 a_{12} + \cdots + x_t a_{1t} = 0.$$

而

$$\begin{aligned} |x_1||a_{11}| =& |x_2 a_{12} + \cdots + x_t a_{1t}| \\ \leq& |x_2||a_{12}| + \cdots + |x_t||a_{1t}| \\ \\ <& |x_1|(|a_{12}| + |a_{13}| + \cdots + |a_{1t}|) < |a_{11}||x_1|. \end{aligned}$$

矛盾!

练习 1.23 设 $\mathbf{A} \in \mathbf{R}^{n\times n}$，且 $\mathbf{A}$ 可逆。$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_k \in \mathbf{R}^{n\times 1}$。试证 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_k$ 线性无关当且仅当 $\mathbf{A}\boldsymbol{\alpha}_1, \mathbf{A}\boldsymbol{\alpha}_2, \ldots, \mathbf{A}\boldsymbol{\alpha}_k$ 线性无关。

解 假定 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_k$ 线性无关，如果

$$x_1 \mathbf{A}\boldsymbol{\alpha}_1 + x_2 \mathbf{A}\boldsymbol{\alpha}_2 + \ldots + x_k \mathbf{A}\boldsymbol{\alpha}_k = 0.$$

左右两边都是 $\mathbb{R}^{n\times 1}$ 中的元素。两边同时乘以 $\boldsymbol{A}^{-1}$, 再利用矩阵相乘时常数可以拿到前面可得

$$\boldsymbol{A}^{-1}(x_1 \mathbf{A}\boldsymbol{\alpha}_1 + x_2 \mathbf{A}\boldsymbol{\alpha}_2 + \ldots + x_k \mathbf{A}\boldsymbol{\alpha}_k) = x_1 \boldsymbol{\alpha}_1 + x_2 \boldsymbol{\alpha}_2 + \cdots + \boldsymbol{\alpha}_k = 0.$$

利用 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_k$ 的线性无关性，便可以得到 $\mathbf{A}\boldsymbol{\alpha}_1, \mathbf{A}\boldsymbol{\alpha}_2, \ldots, \mathbf{A}\boldsymbol{\alpha}_k$ 线性无关。反过来是同样的证明。

练习 1.24 设向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_k \in \mathbb{F}^n$ 线性相关，$k \geqslant 2$。证明对任意 $\boldsymbol{\alpha}_{k+1} \in \mathbb{F}^n$ 存在不全为零的纯量 $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{F}$，使得向量

$$\boldsymbol{\alpha}_1 + \lambda_1 \boldsymbol{\alpha}_{k+1}, \quad \boldsymbol{\alpha}_2 + \lambda_2 \boldsymbol{\alpha}_{k+1}, \cdots, \boldsymbol{\alpha}_k + \lambda_k \boldsymbol{\alpha}_{k+1}$$

线性相关。

解 因为向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_k \in \mathbb{F}^n$ 线性相关，则存在不全为零的常数 $x_1, x_2, \cdots, x_k \in \mathbb{F}$ 使得

$$x_1 \boldsymbol{\alpha}_1 + x_2 \boldsymbol{\alpha}_2 + \cdots + x_k \boldsymbol{\alpha}_k = 0$$

不失一般性（否则重新排个序）我们可以假定 $x_1 \neq 0$. 我们想找不全为零的 $y_1, y_2, \cdots, y_k$ 使得

$$y_1 \boldsymbol{\alpha}_1 + y_2 \boldsymbol{\alpha}_2 + \cdots + y_k \boldsymbol{\alpha}_k + (y_1\lambda_1 + y_2\lambda_2 + \cdots + y_k\lambda_k)\boldsymbol{\alpha}_{k+1} = 0$$

选取 $y_i = x_i, i = 1, 2, \cdots, k$. 我们想要

$$x_1\lambda_1 + x_2\lambda_2 + \cdots + x_k\lambda_k = 0$$

也即

$$\lambda_1=\frac{-x_2}{x_1}\lambda_2+\cdots+\frac{-x_k}{x_1}\lambda_k$$

(1) 如果 $x_2=x_3=\cdots=x_k=0$, 则我们只需要选取

$$\lambda_1=0,\quad \lambda_2=\lambda_3=\cdots=\lambda_k=1.$$

(2) 如果存在某一个 $2\leq j\leq n$ 使得 $x_j\neq 0$, 那么我们只需要选取

$$\lambda_1=-x_j,\quad \lambda_j=x_1,\quad \lambda_2=\lambda_3=\cdots=\lambda_{j-1}=\lambda_{j+1}=\cdots=\lambda_k=0$$

就满足条件。

✍ **练习 1.25** 请求出下面矩阵的秩，行秩和列秩。通过这个过程你看一看为什么这三个秩一样。

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & a & b & \cdots & c \\ 0 & 1 & \cdots & 0 & d & e & \cdots & f \\ \vdots & \vdots & \vdots & \vdots & a & b & \cdots & c \\ 0 & 0 & \cdots & 1 & g & h & \cdots & i \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & 0 & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

其中含有 1 的那个方阵是 $r\times r$ 阶的，其他的任意。这个练习是用来帮你理解的，请认真对待。

解 请看课堂回放，这里略。

✍ **练习 1.26** 考虑一个 4×5 的矩阵

$$\boldsymbol{A}=\begin{pmatrix} a_1 & b_1 & c_1 & d_1 & e_1 \\ a_2 & b_2 & c_2 & d_2 & e_2 \\ a_3 & b_3 & c_3 & d_3 & e_3 \\ a_4 & b_4 & c_4 & d_4 & e_4 \end{pmatrix}$$

它的行向量是

$$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4\}\subseteq\mathbb{F}^5.$$

其中

$$\boldsymbol{\alpha}_1=\begin{pmatrix} a_1 \\ b_1 \\ c_1 \\ d_1 \\ e_1 \end{pmatrix},\quad \boldsymbol{\alpha}_2=\begin{pmatrix} a_2 \\ b_2 \\ c_2 \\ d_2 \\ e_2 \end{pmatrix},\quad \boldsymbol{\alpha}_3=\begin{pmatrix} a_3 \\ b_3 \\ c_3 \\ d_3 \\ e_3 \end{pmatrix},\quad \boldsymbol{\alpha}_4=\begin{pmatrix} a_4 \\ b_4 \\ c_4 \\ d_4 \\ e_4 \end{pmatrix}$$

请根据我们课堂上讲的证明：对矩阵 A 进行初等变换（一共有 6 种啊）后不改变它的行秩。这个练习是用来帮你理解的，请认真对待。

解 请看课堂回放，这里略。

✍ **练习 1.27** 给定两个 $m\times n$ 阶的实矩阵 $\boldsymbol{A},\boldsymbol{B}$，请证明

$$\operatorname{rank}(\boldsymbol{A}+\boldsymbol{B})\le\operatorname{rank}(\boldsymbol{A})+\operatorname{rank}(\boldsymbol{B}).$$

解 假定 $\boldsymbol{A}$ 的行向量是

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

$\boldsymbol{B}$ 的行向量是

$$T=\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m\}.$$

那么 $\boldsymbol{A}+\boldsymbol{B}$ 的行向量是

$$H=\{\boldsymbol{\alpha}_1+\boldsymbol{\beta}_1,\boldsymbol{\alpha}_2+\boldsymbol{\beta}_2,\cdots,\boldsymbol{\alpha}_m+\boldsymbol{\beta}_m\}$$

显然 H 中的每一个向量都可以被 $S\cup T$ 中的向量线性表示。从而 H 的极大线性无关组的向量也可被 $S\cup T$ 的极大线性无关组的向量线性表示。于是替换定理可得出

$$\operatorname{rank}(H)\le\operatorname{rank}(S\cup T).$$

又注意到 $S\cup T$ 的极大线性无关组中的向量可以被 S 中的极大线性无关组的向量并上 T 中的极大线性无关组的向量表示出来。于是还是根据替换定理有

$$\operatorname{rank}(S\cup T)\le\operatorname{rank}(S)+\operatorname{rank}(T).$$

我们收集一些显然的结论：

引理 1.6

如果向量组

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

中的向量可以被向量组

$$T=\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m\}.$$

中的向量线性表示，那么一定有

$$\operatorname{rank}(S)\le\operatorname{rank}(T).$$

♡

证明 假定

$$\tilde{S}=\{\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_k}\}$$

是 S 的极大线性无关组。假定

$$\tilde{T}=\{\boldsymbol{\beta}_{j_1},\boldsymbol{\beta}_{j_2},\cdots,\boldsymbol{\beta}_{j_l}\}$$

是 T 的极大线性无关组。则根据线性表示的传递性可得 $\tilde{S}$ 中的向量可以被 $\tilde{T}$ 中的向量线性表示。于是根据替换定理知道 $k\le l$。也就是

$$\operatorname{rank}(S)\le\operatorname{rank}(T).$$

✍ **练习 1.28** 假定 $\boldsymbol{A}$ 是 $n\times n$ 阶的矩阵，请证明：

$$\operatorname{rank}(\boldsymbol{A}^n)=\operatorname{rank}(\boldsymbol{A}^{n+1})=\operatorname{rank}(\boldsymbol{A}^{n+2})=\cdots.$$

解 如果 A 可逆，所有的幂次的秩都是 n，就不需要证明什么了. 因此假定 $\operatorname{rank}(\boldsymbol{A}) \leq n-1$. 首先根据例子 1.17 知道对任意的 $k \geq 1$ 有

$$\operatorname{rank}(\boldsymbol{A}^{k+1}) \leq \operatorname{rank}(\boldsymbol{A}^{k}).$$

于是

$$k \to \operatorname{rank}(\boldsymbol{A}^{k})$$

关于 $k \in \{1,2,3,4,\cdots\}$ 是单调减的。又因为最大的 $\operatorname{rank}(\boldsymbol{A}) \leq n-1$, 并且秩又非负。于是一定存在 $1 \leq k_0 \leq n$ 使得

$$\operatorname{rank}(\boldsymbol{A}^{k_0}) = \operatorname{rank}(\boldsymbol{A}^{k_0+1}).$$

这是因为 $\boldsymbol{A}$ 的秩最大是 $n-1$, 在

$$\{\boldsymbol{A}, \boldsymbol{A}^2, \boldsymbol{A}^3, \cdots, \boldsymbol{A}^n, \boldsymbol{A}^{n+1}\}$$

这 $n+1$ 个矩阵中，一定有两个的秩一样。再加上递减的性质，这样的 k_0 就一定存在。首先注意到 $\boldsymbol{A}^{k_0+1}$ 的行都是 $\boldsymbol{A}^{k_0}$ 的行的线性组合，这也就是说 $\boldsymbol{A}^{k_0+1}$ 的行的极大线性无关组 (记为 S_{k_0+1}) 一定可以被 $\boldsymbol{A}^{k_0}$ 的行的的极大线性无关组 (记为 S_{k_0}) 表示。但是 $\operatorname{rank}(\boldsymbol{A}^{k_0}) = \operatorname{rank}(\boldsymbol{A}^{k_0+1})$, 那么 S_{k_0+1} 含有的向量个数一定等于 S_{k_0} 的向量个数。由替换定理可得 S_{k_0} 等价于 S_{k_0+1}。换句话说 $\boldsymbol{A}^{k_0}$ 的行的极大线性无关组一定可以被 $\boldsymbol{A}^{k_0+1}$ 的行的的极大线性无关组表示。从而 $\boldsymbol{A}^{k_0}$ 的行都是 $\boldsymbol{A}^{k_0+1}$ 的行的线性组合。

注意到矩阵 $\boldsymbol{AB}$ 所得矩阵的行其实就是对 $\boldsymbol{B}$ 的行作线性组合。这样的话对 $\boldsymbol{A}^{k_0}$ 和 $\boldsymbol{A}^{k_0+1}$ 左边同时乘以 $\boldsymbol{A}$ 后新得到的矩阵的行是分别对 $\boldsymbol{A}^{k_0}$ 和 $\boldsymbol{A}^{k_0+1}$ 的行进行同样的线性组合。于是可得 $\boldsymbol{A}^{k_0+1}$ 的行都是 $\boldsymbol{A}^{k_0+2}$ 的行的线性组合。从而

$$\operatorname{rank}(\boldsymbol{A}^{k_0+2}) \geq \operatorname{rank}(\boldsymbol{A}^{k_0+1}).$$

但是

$$\operatorname{rank}(\boldsymbol{A}^{k_0+2}) \leq \operatorname{rank}(\boldsymbol{A}^{k_0+1}).$$

于是

$$\operatorname{rank}(\boldsymbol{A}^{k_0+2}) = \operatorname{rank}(\boldsymbol{A}^{k_0+1}).$$

同样的办法走下来可得

$$\operatorname{rank}(\boldsymbol{A}^{k_0+2}) = \operatorname{rank}(\boldsymbol{A}^{k_0+3}).$$

等等一直进行下去。由于 $k_0 \leq n$, 我们一定有

$$\operatorname{rank}(\boldsymbol{A}^{n}) = \operatorname{rank}(\boldsymbol{A}^{n+1}) = \operatorname{rank}(\boldsymbol{A}^{n+2}) = \cdots.$$

1.6 2024 年 3 月 21 日第三周星期四作业-基与维数

定理 1.8

设 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 是向量空间 V 的一组不全为零的向量，而 $\{\boldsymbol{\alpha}_{i_1}, \boldsymbol{\alpha}_{i_2}, \cdots, \boldsymbol{\alpha}_{i_r}\}$ 是它的一个极大无关组。那么

$$\mathcal{L}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n) = \mathcal{L}(\boldsymbol{\alpha}_{i_1}, \boldsymbol{\alpha}_{i_2}, \cdots, \boldsymbol{\alpha}_{i_r})$$

♡

定义 1.14 (基)

设 V 是数域 $\mathbb{F}$ 上一个向量空间. V 中满足下列两个条件的向量组 S 叫作 V 的一个基:

(i) 任取 S 中有限个向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 都线性无关;

(ii) V 的每一个向量都可以由 S 中有限个向量线性表示。

♣

如果线性空间 V 存在一组基 S 由有限多个基向量组成，则线性空间 V 称为有限维的. 不是有限维的线性空间称为无限维的。

定理 1.9

设数域 $\mathbb{F}$ 上的有限维线性空间 V 具有一组基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n\}$，$n \geqslant 1$. 则 V 中任意一个线性无关向量集合 S 都是有限的，并且 S 所含向量的数目不超过 n。

♡

推论 1.2

有限维线性空间 V 任意两组基所含向量的个数相同。

♡

定义 1.15 (维数)

一个向量空间 V 的基所含向量的个数叫作 V 的维数。零空间的维数定义为 0。空间 V 的维数记作 $\dim V$。

♣

定理 1.10

设 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 是向量空间 V 的一个基。那么 V 的每一个向量可以唯一地表示成基向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 的线性组合。

♡

由定理 1.9 可知

定理 1.11

n 维向量空间中任意多于 n 个向量一定线性相关。

♡

定理 1.12

设 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r$ 是 n 维向量空间 V 中一组线性无关的向量。那么总可以添加 $n-r$ 个向量 $\boldsymbol{\alpha}_{r+1}, \cdots, \boldsymbol{\alpha}_n$，使得 $\{\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_r, \boldsymbol{\alpha}_{r+1}, \cdots, \boldsymbol{\alpha}_n\}$ 作成 V 的一个基。特别，n 维向量空间中任意 n 个线性无关的向量都可以取作基。

♡

替换定理即得。

定理 1.13

设 W_1 和 W_2 都是数域 $\mathbb{F}$ 上向量空间 V 的有限维子空间. 那么 W_1+W_2 也是有限维的, 并且

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

♡

定义 1.16 (余子空间)

设 W 是向量空间 V 的一个子空间. V 的子空间 W' 叫作 W 的一个余子空间, 如果

(i) $V=W+W'$;

(ii) $W\cap W'=\{0\}$. 在这一情形, 就说 V 是子空间 W 与 W' 的直和, 并且记作 $V=W\oplus W'$.

♣

定理 1.14

设向空间 V 是子空间 W 与 W' 的直和. 那么 V 中每一向量 $\boldsymbol{\alpha}$ 可以唯一地表成

$$\boldsymbol{\alpha}=\boldsymbol{\beta}+\boldsymbol{\beta}',\quad \boldsymbol{\beta}\in W,\boldsymbol{\beta}'\in W'.$$

♡

定理 1.15

n 维向量空间 V 的任意一个子空间 W 一定有余子空间. 如果 W' 是 W 的一个余子空间, 那么

$$\dim V=\dim W+\dim W'$$

♡

关于直和的概念可以推广到多于两个子空间的情形. 设 $W_1,W_2,\cdots,W_n$ 是 V 的子空间,

$$V=W_1+W_2+\cdots+W_n;$$

$$W_i\cap\left(W_1+\cdots+W_{i-1}+W_{i+1}+\cdots+W_n\right)=\{0\},\quad i=1,\cdots,n,$$

那么就说 V 是子空间 $W_1,W_2,\cdots,W_n$ 的直和, 并且记作

$$V=W_1\oplus W_2\oplus\cdots\oplus W_n.$$

不难证明，如果 V 是子空间 $W_1,W_2,\cdots,W_n$ 的直和，那么 V 中每一向量 $\boldsymbol{\alpha}$ 可以唯一地表成

$$\boldsymbol{\alpha}=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2+\cdots+\boldsymbol{\alpha}_n$$

的形式, 这里 $\boldsymbol{\alpha}_i\in W_i,i=1,\cdots,n$. 并且, 当 V 是有限维向量空间时,

$$\dim V=\dim W_1+\cdots+\dim W_n.$$

定义 1.17 (一般的直和)

设 V_1 与 V_2 是数域 $\mathbb{F}$ 上的 n 维线性空间 V 的子空间。如果交 $V_1\cap V_2=0$, 则子空间 V_1 与 V_2 的和 V_1+V_2 称为 V_1 与 V_2 的直和, 记为 $V_1\oplus V_2$.

♣

定理 1.16

下列命题等价.

1. 和 V_1+V_2 是直和;
2. V_1+V_2 中每个向量 $\boldsymbol{\alpha}$ 都可以唯一地表为 $\boldsymbol{\alpha}=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2$, 其中 $\boldsymbol{\alpha}_1\in V_1,\boldsymbol{\alpha}_2\in V_2$;
3. 和 V_1+V_2 中零向量可以唯一地表为 $0=0_1+0_2$, 其中 0_1 和 0_2 分别是 V_1 和 V_2 的零向量 (当然它们是线性空间 V 的零向量);
4. $\dim(V_1+V_2)=\dim V_1+\dim V_2$

定义 1.18 (任意有限个子空间的直和)

设 $V_1,V_2,\ldots,V_k$ 是数域此上的 n 维线性空间 V 的子空间. 如果子空间 $V_1,V_2,\ldots,V_k$ 的和 $V_1+V_2+\cdots+V_k$ 中每个向量 $\boldsymbol{\alpha}$ 都可以唯一地表为

$$\boldsymbol{\alpha}=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2+\cdots+\boldsymbol{\alpha}_k,$$

其中 $\boldsymbol{\alpha}_i\in V_i,1\leqslant i\leqslant k$. 则和 $V_1+V_2+\cdots+V_k$ 称为子空间 $V_1,V_2,\ldots,V_k$ 的直和, 记为

$$V_1\oplus V_2\oplus\cdots\oplus V_k.$$

定理 1.17

子空间 $V_1,V_2,\ldots,V_k$ 的和 $V_1+V_2+\cdots+V_k$ 为直和的当且仅当

$$\dim(V_1+V_2+\cdots+V_k)=\dim V_1+\dim V_2+\cdots+\dim V_k.$$

定理 1.18

子空间 $V_1,V_2,\ldots,V_k$ 的和为直和当且仅当和 $V_1+V_2+\cdots+V_k$ 中零向量可以唯一地表为子空间 $V_1,V_2,\ldots,V_k$ 中零向量 (当然是零向量) 的和。

练习 1.29 如果向量空间 V 的每一个向量都可以唯一地表示成向量

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

的线性组合。证明：$\dim(V)=n$.

解 根据习题 1.10 可知（考试的时候你要证明）

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

是线性无关的，又根据题意任意的 V 的向量都能被它们表示。从而根据基的定义知道，

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

是 V 的一组基，于是 $\dim(V)=n$.

练习 1.30 如果 W 是 n 维向量空间 V 的一个子空间，并且 $0<\dim(W)<n$，证明：W 的余子空间不止有一个。

解 假定 $\dim(W)=m$, 则 $0<m<n$, 取 W 的一组基

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

根据定理 1.12, 可以将其扩充为 V 的一组基

$$\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m,\boldsymbol{\beta}_{m+1},\cdots,\boldsymbol{\beta}_n\}.$$

首先显然

$$W_1=\mathcal{L}(\boldsymbol{\beta}_{m+1},\boldsymbol{\beta}_{m+2},\cdots,\boldsymbol{\beta}_n)$$

是 W 的一个余子空间。我们考虑

$$W_2=\mathcal{L}(\boldsymbol{\beta}_1+\boldsymbol{\beta}_{m+1},\boldsymbol{\beta}_{m+2},\cdots,\boldsymbol{\beta}_n)$$

则 $\boldsymbol{\beta}_1+\boldsymbol{\beta}_{m+1}\notin W,\boldsymbol{\beta}_1+\boldsymbol{\beta}_{m+1}\notin W_1$ 且 $\boldsymbol{\beta}_{m+1}\notin W_2$。从而 $W_1\neq W_2$ 是两个不同的子空间。我们想说 W_2 也是 W 的一个余子空间。这是因为首先注意到 $\boldsymbol{\beta}_{m+1}=(\boldsymbol{\beta}_1+\boldsymbol{\beta}_{m+1})-\boldsymbol{\beta}_1$

$$\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m,\boldsymbol{\beta}_{m+1},\cdots,\boldsymbol{\beta}_n\}.$$

与

$$\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m,\boldsymbol{\beta}_1+\boldsymbol{\beta}_{m+1},\cdots,\boldsymbol{\beta}_n\}.$$

等价。从而

$$\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m,\boldsymbol{\beta}_1+\boldsymbol{\beta}_{m+1},\cdots,\boldsymbol{\beta}_n\}.$$

也是 V 的一组基。那么自然有它的前 m 个向量生成的空间 W 与后 $n-m$ 个向量生成的空间 W_2 是互为余子空间的。

✍ **练习 1.31** 证明 $\mathbb{R}^2$ 作为实数域 $\mathbb{R}$ 上的向量空间，它的维数是 2。没错这个题很简单。

解 显然

$$\{(0,1),(1,0)\}$$

是 $\mathbb{R}^2$ 的一个基。于是 $\dim(\mathbb{R}^2)=2.$

✍ **练习 1.32** 求下列子空间的维数。

1. $\mathbb{F}[x]$ 中由

$$x-1,\quad 1-x^2,\quad x^2-x$$

   生成的子空间的维数是多少?
2. 区间 $[a,b]$ 上的所有连续函数组成的实数域 $\mathbb{R}$ 上的线性空间记为 $C[a,b]$。$C[a,b]$ 的维数是多少? 它的由

$$e^x,\quad e^{2x},\quad e^{3x}$$

   生成的子空间的维数是多少?
3. 令 S 是数域 $\mathbb{F}$ 上的所有满足 $A^T=A$ 的 $n\times n$ 阶矩阵形成的向量空间，它的维数是多少? 这里的 A^T 表示的是 A 的转置。

解

1. 首先因为

$$(x-1)+(1-x^2)+(x^2-x)=0.$$

所以这三个向量线性相关。而 $1-x^2$ 不是 $x-1$ 的常数倍可得

$$x-1, 1-x^2$$

线性无关。于是 $\{x-1, 1-x^2\}$ 组成了 $\mathcal{L}(x-1, 1-x^2, x^2-x)$ 的一个基，所以这个空间的维数是 2.

2. 如果 $a=b$，那么 $C[a,b]$ 事实上就是所有的实数组成的线性空间，也就是说 $C[a,b]=\mathbb{R}$，于是 $\dim(C[a,b])=1$. 此时当然由 e^x, e^{2x}, e^{3x} 生成的空间也是 $\mathbb{R}$，所以维数也是 1.

   下面我们假定 $a<b$。注意到

$$f_n = x^n, \quad n=0,1,2,\cdots.$$

   这些多项式看成 $C[a,b]$ 中的元素，是线性无关的。由于 n 可以任意大，所以对 $C[a,b]$ 的任意一个基一定含有无限个元素，否则与定理 1.9 矛盾。因此我们有

$$\dim(C[a,b])=\infty.$$

   如果有实数 $\lambda_1, \lambda_2, \lambda_3$ 使得

$$\lambda_1 e^x + \lambda_2 e^{2x} + \lambda_3 e^{3x} = 0, \quad \forall x \in [a,b].$$

   随便取 $[a,b]$ 中的三个数 $s,t,p \in [a,b]$，我们有

$$\begin{pmatrix} e^s & e^{2s} & e^{3s} \\ e^t & e^{2t} & e^{3t} \\ e^p & e^{2p} & e^{3p} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

   而系数矩阵的行列式可以计算得到

$$\mathrm{Det}(A) = (e^s - e^p)\left(e^t - e^p\right)\left(e^t - e^s\right) e^{p+s+t} \neq 0.$$

   不等于零是因为 p,s,t 是三个不同的实数。所以上面的方程组只有零解，也就是说

$$\lambda_1 = \lambda_2 = \lambda_3 = 0.$$

   于是 e^x, e^{2x}, e^{3x} 是线性无关的，于是其生成的空间的维数是 3.

3. 如果矩阵 $\boldsymbol{A}=(a_{ij})$ 满足 $\boldsymbol{A}^T=\boldsymbol{A}$. 也就是说

$$a_{ij} = a_{ji}$$

   这样 $\boldsymbol{A}$ 完全由

$$a_{ij} : 1 \leq i \leq j \leq n$$

   确定。可直接验证

$$\{E_{ij} + E_{ji} : 1 \leq i \leq j \leq n\}$$

   是 S 的一个基，从而 $\dim(S) = \frac{n(n+1)}{2}$. 注意这里的 E_{ij} 指的是第 (i,j) 位置是 1，其他位置是 0 的矩阵。

✍ **练习 1.33** 下面的结论对吗？对的话请证明，不对的话请给一个反例：如果 U_1, U_2, W 是向量空间 V 的三个子空间，满足

$$V = U_1 \oplus W \quad 和 \quad V = U_2 \oplus W,$$

则 $U_1 = U_2$。

解 这个结论显然不对。例如我们取

$$V = \mathbb{R}^3, \quad W = \{(x, y, 0) \in \mathbb{R}^3 | x, y \in \mathbb{R}\}, \quad U_1 = \{(0, 0, x) | x \in \mathbb{R}\}, \quad U_2 = \{(x, x, 0 | x \in \mathbb{R})\}.$$

则 W 有一组基如下:

$$\{\boldsymbol{\alpha}_1 = (1, 0, 0), \quad \boldsymbol{\alpha}_2 = (0, 1, 0)\}.$$

U_1 有基 $\{\boldsymbol{\alpha}_3 = (0, 0, 1)\}$, U_2 有基 $\{\boldsymbol{\alpha}_4 = (1, 1, 0)\}$。显然 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3\}$ 和 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_4\}$ 都是 $V = \mathbb{R}^3$ 的基。从而 U_1, U_2 都是 W 的余子空间。注意到 $U_1 \neq U_2$，从而题目的结论不对。

练习 1.34 假定

$$U = \{(x, y, x + y, x - y, 2x) \in \mathbb{R}^5 \mid x, y \in \mathbb{R}\}.$$

请找一个 $\mathbb{R}^5$ 的子空间 W 使得 $\mathbb{R}^5 = U \oplus W$。

解 注意到对任意的

$$\boldsymbol{\alpha} = (x, y, x + y, x - y, 2x) \in \mathbb{R}^5.$$

我们都可以将其表示为

$$\boldsymbol{\alpha} = x(1, 0, 1, 1, 2) + y(0, 1, 1, -1, 0) = x\boldsymbol{\beta} + y\boldsymbol{\gamma},$$

这里

$$\boldsymbol{\beta} = (1, 0, 1, 1, 2), \quad \boldsymbol{\gamma} = (0, 1, 1, -1, 0).$$

显然 $\boldsymbol{\beta}$ 与 $\boldsymbol{\gamma}$ 不能互相线性表示，因此 $\{\boldsymbol{\beta}, \boldsymbol{\gamma}\}$ 是 U 的一个基。注意到 $\boldsymbol{\gamma}$ 的第一个坐标是 0, 我们选取

$$\boldsymbol{\alpha}_1 = (1, 0, 0, 0, 0).$$

它的第五个坐标不是第一个坐标的 2 倍，因此不在空间 U 中。所以 $\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\alpha}_1$ 线性无关。再看 $\boldsymbol{\gamma}$ 的最后一个坐标是 0，选取

$$\boldsymbol{\alpha}_2 = (0, 0, 0, 0, 1)$$

它的第五个坐标不是第一个坐标的 2 倍, 因此不在 U 中。如果 $\boldsymbol{\alpha}_2$ 可以被 $\{\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\alpha}_1\}$ 线性表示，那么有实数 x, y, z 使得

$$\boldsymbol{\alpha}_2 = (0, 0, 0, 0, 1) = x\boldsymbol{\beta} + y\boldsymbol{\gamma} + z\boldsymbol{\alpha}_1 = (x + z, y, x + y, x - y, 2x)$$

于是

$$\begin{cases} x + z = 0 \\ y = 0 \\ x + y = 0 \\ x - y = 0 \\ 2x = 1 \end{cases}$$

显然无解，于是 $\{\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2\}$ 是线性无关的。再看 $\boldsymbol{\alpha}$ 的第二个坐标是 0，我们选择

$$\boldsymbol{\alpha}_3=(0,1,0,0,0).$$

那么有实数 x,y,z,w 使得

$$\boldsymbol{\alpha}_3=(0,1,0,0,0)=x\boldsymbol{\beta}+y\boldsymbol{\gamma}+z\boldsymbol{\alpha}_1+w\boldsymbol{\alpha}_2=(x+z,y,x+y,x-y,2x+w)$$

于是

$$\begin{cases}x+z=0\\y=1\\x+y=0\\x-y=0\\2x+w=0\end{cases}$$

显然无解。于是我们可得到 $\{\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3\}$ 是线性无关的。这样我们可以定义

$$W=\mathcal{L}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)=\{(x,y,0,0,z):x,y,z\in\mathbb{R}\}.$$

1.7 2024 年 3 月 26 日第四周星期二作业-坐标与坐标变换

- 设 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n\}$ 是数域 $\mathbb{F}$ 上的 n 维线性空间 V 的基，由基的定义，对任意向量 $\boldsymbol{\alpha}\in V$，存在纯量 $a_1,a_2,\ldots,a_n\in\mathbb{F}$，使得 $\boldsymbol{\alpha}$ 可以表示为

  $$\boldsymbol{\alpha}=a_1\boldsymbol{\alpha}_1+a_2\boldsymbol{\alpha}_2+\cdots a_n\boldsymbol{\alpha}_n.$$

  于是，向量 $\boldsymbol{\alpha}$ 便确定数域 $\mathbb{F}$ 上的有序 n 元数组

  $$(a_1,a_2,\ldots,a_n).$$

- n 元数组 $(a_1,a_2,\ldots,a_n)$ 是由向量 $\boldsymbol{\alpha}$ 所唯一确定的. 事实上，设另有一组纯量 $b_1,b_2,\ldots,b_n$，使得

  $$\boldsymbol{\alpha}=b_1\boldsymbol{\alpha}_1+b_2\boldsymbol{\alpha}_2+\ldots+b_n\boldsymbol{\alpha}_n.$$

  则

  $$(a_1-b_1)\boldsymbol{\alpha}_1+(a_2-b_2)\boldsymbol{\alpha}_2+\ldots+(a_n-b_n)\boldsymbol{\alpha}_n=0.$$

  因为向量

  $$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n$$

  线性无关，所以

  $$a_i-b_i=0$$

  即 $a_i=b_i,1\leqslant i\leqslant n$。

定义 1.19 (在给定基下的坐标)

由向量 $\boldsymbol{\alpha}$ 所唯一确定的 n 元数组

$$(a_1, a_2, \ldots, a_n)$$

称为向量 $\boldsymbol{\alpha}$ 在基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n\}$ 下的坐标, 而 a_i 称为向量 $\boldsymbol{\alpha}$ 的第 i 个坐标分量, $i = 1, 2, \ldots, n$.

- 设向量 $\boldsymbol{\alpha}, \boldsymbol{\beta} \in V$ 在基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n\}$ 下的坐标分别为

  $$(a_1, a_2, \ldots, a_n) \quad \text{与} \left(b_1, b_2, \ldots, b_n\right).$$

  则

  $$\boldsymbol{\alpha} = a_1\boldsymbol{\alpha}_1 + a_2\boldsymbol{\alpha}_2 + \ldots + a_n\boldsymbol{\alpha}_n, \quad \boldsymbol{\beta} = b_1\boldsymbol{\alpha}_1 + b_2\boldsymbol{\alpha}_2 + \ldots + b_n\boldsymbol{\alpha}_n.$$

  因此,

  $$\boldsymbol{\alpha} + \boldsymbol{\beta} = (a_1 + b_1)\boldsymbol{\alpha}_1 + (a_2 + b_2)\boldsymbol{\alpha}_2 + \cdots + (a_n + b_n)\boldsymbol{\alpha}_n.$$

  所以向量 $\boldsymbol{\alpha} + \boldsymbol{\beta}$ 在基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n\}$ 下的坐标为

  $$(a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n).$$

- 设纯量 $\lambda \in \mathbb{F}$, 则

  $$\lambda\boldsymbol{\alpha} = \lambda a_1\boldsymbol{\alpha}_1 + \lambda a_2\boldsymbol{\alpha}_2 + \cdots + \lambda a_n\boldsymbol{\alpha}_n.$$

  因此, 向量 $\lambda\boldsymbol{\alpha}$ 在基 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 的坐标为

  $$(\lambda a_1, \lambda a_2, \ldots, \lambda a_n).$$

记

$$\mathbb{F}^n = \{(a_1, a_2, \cdots, a_n) : a_i \in \mathbb{F}, \quad i = 1, 2, 3, \cdots, n\} \tag{1.6}$$

为数域 $\mathbb{F}$ 上所有的有序 n 数组构成的向量空间. 例如 $\mathbb{R}^n, \mathbb{C}^n$ 等等.

定义 1.20

在数域 $\mathbb{F}$ 上的 n 维线性空间 V 中取定基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n\}$. 我们可以定义线性空间 V 到 $\mathbb{F}^n$ 的映射 η 如下: 对任意 $\boldsymbol{\alpha} \in V$,

$$\eta(\boldsymbol{\alpha}) = (a_1, a_2, \ldots, a_n),$$

命题 1.9

验证

- η 是一一映射, 即既是单射又是满射.
- 对任意 $\boldsymbol{\alpha}, \boldsymbol{\beta} \in V$,

  $$\eta(\boldsymbol{\alpha} + \boldsymbol{\beta}) = \eta(\boldsymbol{\alpha}) + \eta(\boldsymbol{\beta}).$$

- 对任意 $\lambda \in \mathbb{F}, \boldsymbol{\alpha} \in V$,

  $$\eta(\lambda\boldsymbol{\alpha}) = \lambda\eta(\boldsymbol{\alpha}).$$

前面讨论了数域 $\mathbb{F}$ 上的 n 维线性空间 V 的向量在 V 的一组基下的坐标. 一般地说, 同一个向量在不同基下的坐标是不同的. 问题: 同一个向量在不同基下的坐标之间有什么关系?

- 设 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n\}$ 与 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_n\}$ 分别是数域 $\mathbb{F}$ 的 n 维线性空间 V 的基
- 设向量 $\boldsymbol{\alpha}\in V$ 在这两组基的的坐标分别为

$$x=(x_1,x_2,\ldots,x_n)^{\mathrm{T}}\ \text{与} y=(y_1,y_2,\ldots,y_n)^{\mathrm{T}}.$$

- 由于 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n\}$ 是 V 的基, 因此向量 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_n\}$ 可以由向量 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n\}$ 线性表示 (并且唯一), 所以可设

$$\begin{cases}\boldsymbol{\beta}_1=b_{11}\boldsymbol{\alpha}_1+b_{21}\boldsymbol{\alpha}_2+\cdots+b_{n1}\boldsymbol{\alpha}_n,\\ \boldsymbol{\beta}_2=b_{12}\boldsymbol{\alpha}_1+b_{22}\boldsymbol{\alpha}_2+\cdots+b_{n2}\boldsymbol{\alpha}_n,\\ \cdots\cdots\cdots\\ \boldsymbol{\beta}_n=b_{1n}\boldsymbol{\alpha}_1+b_{2n}\boldsymbol{\alpha}_2+\cdots+b_{n}\boldsymbol{\alpha}_n.\end{cases}$$

- 上式可以写成矩阵形式如下:

$$(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_n)=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n)\begin{pmatrix}b_{11}&b_{12}&\cdots&b_{1n}\\ b_{21}&b_{22}&\cdots&b_{2n}\\ \vdots&\vdots&\vdots&\vdots\\ b_{n1}&b_{n2}&\ldots&b_{nn}\end{pmatrix}$$

- 记 $\boldsymbol{B}=(b_{ij})$ 于是上式子即为

$$(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_n)=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n)\boldsymbol{B}.\tag{1.7}$$

定义 1.21 (基变换公式和过渡矩阵)

式 (1.7) 称为由基 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n\}$ 到基 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_n\}$ 的基变换公式, 矩阵 $\boldsymbol{B}$ 称为由基 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n\}$ 到基 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_n\}$ 的过渡矩阵. ♣

定理 1.19

设 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n\}$ 与 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_n\}$ 分别是数域 $\mathbb{F}$ 上的 n 维线性空间 V 的基. 设方阵 $\boldsymbol{A}$ 是由基

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

到基

$$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n\}$$

的过渡矩阵. 则方阵 $\boldsymbol{A}$ 可逆, 并且由基

$$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n\}$$

到基

$$\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_n\}$$

的过渡矩阵为 $\boldsymbol{A}^{-1}$. ♡

由于方阵 $\boldsymbol{A}=(a_{ij})$ 是由基 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_n\}$ 到基 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n\}$ 的过渡矩阵, 因此,

$$(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n)=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_n)\,\boldsymbol{A}. \tag{1.8}$$

设由基 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n\}$ 到基 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_n\}$ 的过渡矩阵为 $\boldsymbol{B}=(b_{ij})$, 则

$$(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_n)=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n)\,\boldsymbol{B}. \tag{1.9}$$

由式 (1.8) 与式 (1.9), 对于 $i,j=1,2,\ldots,n$,

$$\boldsymbol{\alpha}_k=\sum_{i=1}^n a_{ik}\boldsymbol{\beta}_i,\quad \boldsymbol{\beta}_j=\sum_{k=1}^n b_{kj}\boldsymbol{\alpha}_k.$$

因此,

$$\boldsymbol{\beta}_j=\sum_{k=1}^n b_{kj}\left(\sum_{i=1}^n a_{ik}\boldsymbol{\beta}_i\right)=\sum_{i=1}^n\left(\sum_{k=1}^n a_{ik}b_{kj}\right)\boldsymbol{\beta}_i.$$

由于向量 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_n$ 线性无关, 因此,

$$\sum_{k=1}^n a_{ik}b_{kj}=\delta_{ij}$$

其中 $1\leqslant i,j\leqslant n,\delta_{ij}$ 是 Kronecker 符号. 所以 $\boldsymbol{AB}=\mathbf{I}(n)$. 从而方阵 $\boldsymbol{A}$ 可逆, 并且 $\boldsymbol{A}^{-1}=\boldsymbol{B}$.

- 记数域 $\mathbb{F}$ 上的 n 维线性空间 V 的所有基的集合为 $\mathfrak{B}$, 数域 $\mathbb{F}$ 上的所有 n 阶可逆方阵的集合为 $\mathrm{GL}(\mathbb{F})$.
- 取定 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots\boldsymbol{\beta}_n\}\in\mathfrak{B}$, 则对于任意 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n\}\in\mathfrak{B}$, 据定理 1.19, 由
$$(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n)=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_n)\,\boldsymbol{A}.$$
所确定的矩阵 $\boldsymbol{A}$ 可逆.
- 通过上式, 可以定义集合 $\mathfrak{B}$ 到 $\mathrm{GL}_n(\mathbb{F})$ 的映射 η 如下：对于任意 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n\}\in\mathfrak{B}$
$$\eta\left(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n\right)=\boldsymbol{A}.$$
- 容易看出, 如果 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n\}$ 与 $\{\widetilde{\boldsymbol{\alpha}}_1,\widetilde{\boldsymbol{\alpha}}_2,\ldots,\widetilde{\boldsymbol{\alpha}}_n\}$ 是 V 的两组不同的基, 则由基 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_n\}$ 分别到基 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n\}$ 和基 $\{\widetilde{\boldsymbol{\alpha}}_1,\widetilde{\boldsymbol{\alpha}}_2,\ldots,\widetilde{\boldsymbol{\alpha}}_n\}$ 的过渡矩阵 $\boldsymbol{A}$ 与 $\widetilde{\boldsymbol{A}}$ 也不同, 即
$$\eta\left(\widetilde{\boldsymbol{\alpha}}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n\right)\neq\eta\left(\widetilde{\boldsymbol{\alpha}}_1,\widetilde{\boldsymbol{\alpha}}_2,\ldots,\widetilde{\boldsymbol{\alpha}}_n\right).$$
因此, 映射 η 是单射.
- 设 $\boldsymbol{A}\in\mathrm{GL}_n(\mathbb{F})$, 则由 $(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_n)\,\boldsymbol{A}$ 便确定 V 的 n 个向量, 记为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n$, 即
$$(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n)=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_n)\,\boldsymbol{A}.$$
- 如果存在纯量 $\lambda_1,\lambda_2,\ldots,\lambda_n\in\mathbb{F}$, 使得
$$\lambda_1\boldsymbol{\alpha}_1+\lambda_2\boldsymbol{\alpha}_2+\cdots+\lambda_n\boldsymbol{\alpha}_n=0$$
则记
$$x=(\lambda_1,\lambda_2,\ldots,\lambda_n)^{\mathrm{T}},$$

便得到

$$(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n)\, x = 0 \Longrightarrow (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n)\, \boldsymbol{A}x = 0.$$

- 由于 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n$ 线性无关, 因此 $\boldsymbol{A}x = 0$. 由于 $\boldsymbol{A} \in \mathrm{GL}_n(\mathbb{F})$, 故 $x = 0$, 即

$$\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0.$$

- 这表明, 向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$ 线性无关. 由于 $\dim V = n$, 所以 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n\} \in \mathfrak{B}$. 这就证明了对任意 $\boldsymbol{A} \in \mathrm{GL}_n(\mathbb{F})$, 存在 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n\} \in \mathfrak{B}$, 使得

$$\eta\, (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n) = \boldsymbol{A}.$$

因此映射 η 是满射.

定理 1.20 (同一个向量在不同基下的坐标间的关系)

设 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n\}$ 与 $\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n\}$ 分别是数域 $\mathbb{F}$ 上的 n 维线性空间 V 的基, 由基

$$\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n\}$$

到基

$$\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n\}$$

的过渡矩阵为 $\boldsymbol{A}$. 设 $\boldsymbol{\alpha} \in V$ 在基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n\}$ 与 $\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n\}$ 下的坐标分别为 x 与 y, 则

$$y = \boldsymbol{A}x. \tag{1.10}$$

♡

由假设,

$$\boldsymbol{\alpha} = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n)\, x = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n)\, y,$$

并且

$$(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n) = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n)\, \boldsymbol{A}.$$

因此,

$$\boldsymbol{\alpha} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n)\, \boldsymbol{A}x = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n)\, y.$$

由于向量 $\boldsymbol{\alpha}$ 在同一组基 $\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n\}$ 下的坐标是唯一的, 所以 $y = \boldsymbol{A}x$. 上面定理中的式 (1.10) 称为同一个向量在不同基下的坐标变换公式.

我们约定的一些表示很有用: 若向量 $\boldsymbol{\alpha}$ 在基 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n$ 下的坐标为 $(a_1, a_2, \ldots, a_n)$, 则我们也可以写为

$$\boxed{\boldsymbol{\alpha} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n) \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}}$$

甚至你可以记为

$$\boldsymbol{\alpha} = \boldsymbol{\beta} X, \quad \boldsymbol{\beta} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n), \quad X = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}.$$

练习 1.35 分别求 $p(x) = a_0 + a_1 x + a_2 x^2$ 在基 $B_1 = \{1, x, x^2\}$ 和 $B_2 = \{1, x-1, (x-1)^2\}$ 下的坐标. 这里的向量空间 V 是次数不超过 2 的实系数多项式构成的空间。

解 由于 $p(x) = a_0 \times 1 + a_1 \times x + a_2 \times x^2$, 所以在 B_1 下的坐标为

$$(a_0, a_1, a_2).$$

又因为

$$\begin{aligned} p(x) =& a_0 + a_1(1 + (x-1)) + a_2[(x-1)+1]^2 \\ =& a_0 + a_1 + a_1(x-1) + a_2(x-1)^2 + 2a_2(x-1) + a_2 \\ =& a_0 + a_1 + a_2 + (a_1 + 2a_2)(x-1) + a_2(x-1)^2. \end{aligned}$$

所以 $p(x)$ 在 B_2 下的坐标是

$$(a_1 + a_1 + a_2, a_1 + 2a_2, a_2).$$

练习 1.36 设 $\mathbf{R}^+$ 是所有正实数组成的集合，加法和数乘定义如下：

$$\forall a, b \in \mathbf{R}_+, \quad k \in \mathbf{R}: \quad a \oplus b = ab, \quad k \odot a = a^k$$

则根据例子 1.4知道 $\mathbf{R}^+$ 关于这一加法和数乘构成一个实线性空间. 求 $\mathbf{R}^+$ 的一组基并确定它的维数.

解 因为任意的 $x > 0$ 都存在唯一的 $y = \ln(x)$ 使得

$$x = e^y = y \odot e$$

所以任意的 $x > 0$ 都是 e 的线性组合，从而 $\{e\}$ 是 $\mathbb{R}_+$ 的一个基，于是 $\dim(\mathbb{R}_+) = 1$.

练习 1.37(选做) 设 $\mathbf{K} \subseteq \mathbf{F} \subseteq \mathbf{E}$ 是三个数域，已知 $\mathbf{F}$ 作为 $\mathbf{K}$ 上的线性空间是 n 维的，$\mathbf{E}$ 作为 $\mathbf{F}$ 上的线性空间是 m 维的，证明：$\mathbf{E}$ 作为 $\mathbf{K}$ 上的线性空间是 mn 维的.

解 假定

$$\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\} \subseteq \mathbf{F}$$

为 $\mathbf{F}$ 作为 $\mathbf{K}$ 上的线性空间的一组基。而假定

$$\{\boldsymbol{\beta} 1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_m\} \subseteq \mathbf{E}$$

为 $\mathbf{E}$ 作为 $\mathbf{F}$ 上的线性空间的一组基。我们想说

$$\{\boldsymbol{\alpha}_i \boldsymbol{\beta}_j : i = 1, 2, \cdots, n, j = 1, 2, \cdots, m\}$$

为 $\mathbf{E}$ 作为 $\mathbf{K}$ 上的线性空间的一组基。这是因为首先任取 $x \in \mathbf{E}$，有

$$x = \sum_{j=1}^{m} s_j \boldsymbol{\beta}_j, \quad s_1, s_2, \cdots, s_m \in \mathbf{F}.$$

又因为每一个 $s_j, j = 1, 2, \cdots, m$ 可以写成

$$s_j = \sum_{i=1}^{n} a_{ij} \boldsymbol{\alpha}_i, \quad j = 1, 2, \cdots, m, a_{ij} \in \mathbf{K}.$$

从而我们有

$$x = \sum_{j=1}^{m} \sum_{i=1}^{n} a_{ij} \boldsymbol{\beta}_j \boldsymbol{\alpha}_i.$$

即作为 $\mathbf{K}$ 上的线性空间，每一个 $\mathbf{E}$ 中的元素可以表示成

$$\{\boldsymbol{\alpha}_i \boldsymbol{\beta}_j : i = 1, 2, \cdots, n, j = 1, 2, \cdots, m\}$$

的线性组合。假定存在 $\{a_{ij} : i = 1, \cdots, n; j = 1, \cdots, m\} \subseteq \mathbf{K}$ 使得

$$\sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij} \boldsymbol{\beta}_j \boldsymbol{\alpha}_i = \sum_{j=1}^{m} \boldsymbol{\beta}_j \left(\sum_{i=1}^{n} a_{ij} \boldsymbol{\alpha}_i \right) = 0.$$

注意到

$$\sum_{i=1}^{n} a_{ij} \boldsymbol{\alpha}_i \in \mathbf{F}, \quad j = 1, 2, \cdots, m.$$

而

$$\{\boldsymbol{\beta} 1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_m\} \subseteq \mathbf{E}$$

为 $\mathbf{E}$ 作为 $\mathbf{F}$ 上的线性空间的一组基。从而有

$$\sum_{i=1}^{n} a_{ij} \boldsymbol{\alpha}_i \in \mathbf{F} = 0, \quad j = 1, 2, \cdots, m.$$

又因为

$$\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\} \subseteq \mathbf{F}$$

为 $\mathbf{F}$ 作为 $\mathbf{K}$ 上的线性空间的一组基。所以

$$a_{ij} = 0, \quad i = 1, 2, \cdots, n, j = 1, 2, \cdots, m.$$

于是

$$\{\boldsymbol{\alpha}_i \boldsymbol{\beta}_j : i = 1, 2, \cdots, n, j = 1, 2, \cdots, m\}$$

为 $\mathbf{E}$ 作为 $\mathbf{K}$ 上的线性空间中的一组线性无关的向量。于是组成了 $\mathbf{E}$ 作为 $\mathbf{K}$ 上的线性空间的一组基，从而其维数为 mn.

✍ **练习 1.38(选做)** 数域 $\mathbf{F}$ 上所有 n 阶方阵组成的线性空间 $V = \mathbf{M}_n(\mathbf{F})$，$V_1$ 表示所有对称矩阵（即 $\boldsymbol{A}^T = \boldsymbol{A}$）组成的集合，$V_2$ 表示所有反对称矩阵 (即 $\boldsymbol{A}^T = -\boldsymbol{A}$) 组成的集合. 证明：$V_1, V_2$ 都是 V 的子空间，且 $V = V_1 \oplus V_2$.

解 任取 $\boldsymbol{A}, \boldsymbol{B} \in V_1$, 则有 $\boldsymbol{A}^T = \boldsymbol{A}, \boldsymbol{B}^T = \boldsymbol{B}$. 于是任意的 $\lambda, \mu \in \mathbb{F}$ 有

$$(\lambda \boldsymbol{A} + \mu \boldsymbol{B})^T = \lambda \boldsymbol{A}^T + \mu \boldsymbol{B}^T = \lambda \boldsymbol{A} + \mu \boldsymbol{B}.$$

于是 $\lambda\boldsymbol{A}+\mu\boldsymbol{B}\in V_1$，于是 V_1 是 $\mathbf{M}^n(\mathbb{F})$ 的子空间。同理 V_2 也是 $\mathbf{M}^n(\mathbb{F})$ 的子空间。假定 $\boldsymbol{A}\in V_1\cap V_2$，则

$$\boldsymbol{A}=\boldsymbol{A}^T=-\boldsymbol{A}^T.$$

从而 $\boldsymbol{A}^T=0$，于是 $\boldsymbol{A}=0$. 于是 $V_1\cap V_2=\{0\}$. 而任意 $\boldsymbol{A}\in\mathbf{M}^n(\mathbb{F})$ 可以写成

$$\boldsymbol{A}=\frac{1}{2}(\boldsymbol{A}+\boldsymbol{A}^T)+\frac{1}{2}(\boldsymbol{A}-\boldsymbol{A}^T).$$

显然

$$\boldsymbol{A}+\boldsymbol{A}^T\in V_1,\quad \boldsymbol{A}-\boldsymbol{A}^T\in V_2.$$

所以 $V=V_1\oplus V_2$.

练习 1.39(选做) 在数域 $\mathbb{F}$ 上的所有关于 $\cos x$ 的次数不超过 n 的多项式构成的线性空间中，试写出由基

$$\left\{1,\cos x,\cos^2 x,\ldots,\cos^n x\right\}$$

到基

$$\{1,\cos x,\cos 2x,\ldots,\cos nx\}$$

的过渡矩阵. 不必在意这个叫法，你就把一组基被另一组基表示的矩阵求出来就行。

解 这道题其实考察的是你如和把 $\cos(x)^n$ 用 $\{1,\cos x,\cos 2x,\ldots,\cos nx\}$ 表示出来，这不是我们这门课的重点，是你高中学习的公式。首先，我们有：

$$(\cos(x))^n=(\cos(x))^2(\cos(x))^{n-2}$$

现在，我们使用三角恒等式 $\cos^2(x)=\frac{1+\cos(2x)}{2}$ 来展开 $(\cos(x))^2$：

$$(\cos(x))^2=\frac{1+\cos(2x)}{2}$$

将这个展开式代入我们的原始表达式中，我们得到：

$$(\cos(x))^n=\left(\frac{1+\cos(2x)}{2}\right)(\cos(x))^{n-2}$$

我们可以继续将 $(\cos(x))^{n-2}$ 展开为 $(\cos(x))^2(\cos(x))^{n-4}$，然后再次使用三角恒等式展开 $(\cos(x))^2$。重复这个过程，直到我们达到 $(\cos(x))^{n-k}$ 为止。最终，我们得到：

$$(\cos(x))^n=\frac{1}{2^n}(\cos(x))^n+\frac{n}{2^{n-1}}(\cos(x))^{n-2}\cos(2x)+\frac{n(n-1)}{2^{n-2}}(\cos(x))^{n-4}\cos(2x)^2+\cdots$$

这个展开式的一般形式为：

$$(\cos(x))^n=\sum_{k=0}^{\lfloor n/2\rfloor}\frac{n(n-1)\cdots(n-2k+1)}{2^{n-k}}(\cos(x))^{n-2k}\cos(2x)^k$$

其中 $\lfloor n/2\rfloor$ 是向下取整函数，确保我们只展开到 $(\cos(x))^0$ 或 $(\cos(2x))^0$ 的地方。这道题你只需要知道这里就行，不用再往下了。或者你只需要试试 $n=3$ 的时候，这道题的做法就行。

✍ **练习 1.40**

$$\boldsymbol{\alpha}_1 = (1, 2, -1), \boldsymbol{\alpha}_2 = (0, -1, 3), \boldsymbol{\alpha}_3 = (1, -1, 0);$$
$$\boldsymbol{\beta}_1 = (2, 1, 5), \boldsymbol{\beta}_2 = (-2, 3, 1), \boldsymbol{\beta}_3 = (1, 3, 2).$$

证明 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3\}$ 和 $\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3\}$ 都是 $\mathbf{R}^3$ 的基. 求前者到后者的过渡矩阵. 不必在意这个叫法，你就把一组基被另一组基表示的矩阵求出来就行。

解 这道题是非常简单的计算，老师就不给大家写出来了。

✍ **练习 1.41** 设 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 是 $\mathbb{F}$ 上 n 维向量空间 V 的一个基. $\boldsymbol{A}$ 是 $\mathbb{F}$ 上一个 $n \times s$ 矩阵. 令

$$(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n)\, \boldsymbol{A}$$

证明 $\dim \mathcal{L}(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s) = \operatorname{rank}(\boldsymbol{A})$.

解 假定 $\operatorname{rank}(A) = r$, 则存在 $n \times n$ 阶的可逆矩阵 $\mathbf{P}$ 以及 $s \times s$ 阶的可逆矩阵 $\mathbf{Q}$ 使得

$$\boldsymbol{A} = \mathbf{P} \begin{pmatrix} \mathbf{I}_r & 0 \\ 0 & 0 \end{pmatrix} \mathbf{Q}$$

记

$$\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s), \quad \boldsymbol{\alpha} = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n).$$

则有

$$\boldsymbol{\beta} = \boldsymbol{\alpha}\boldsymbol{A} = \boldsymbol{\alpha}\mathbf{P} \begin{pmatrix} \mathbf{I}_r & 0 \\ 0 & 0 \end{pmatrix} \mathbf{Q}$$

从而

$$\boldsymbol{\beta}\mathbf{Q}^{-1} = \boldsymbol{\alpha}\mathbf{P} \begin{pmatrix} \mathbf{I}_r & 0 \\ 0 & 0 \end{pmatrix}.$$

令

$$\tilde{\boldsymbol{\beta}} = \left(\tilde{\boldsymbol{\beta}}_1, \tilde{\boldsymbol{\beta}}_2, \cdots, \tilde{\boldsymbol{\beta}}_s\right) = \boldsymbol{\beta}\mathbf{Q}^{-1} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s)\, \mathbf{Q}^{-1}.$$

以及

$$\tilde{\boldsymbol{\alpha}} = (\tilde{\boldsymbol{\alpha}}_1, \tilde{\boldsymbol{\alpha}}_2, \cdots, \tilde{\boldsymbol{\alpha}}_n) = \boldsymbol{\alpha}\mathbf{P} = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n)\, \mathbf{P}$$

则

$$\tilde{\boldsymbol{\beta}} = \tilde{\boldsymbol{\alpha}} \begin{pmatrix} \mathbf{I}_r & 0 \\ 0 & 0 \end{pmatrix} \tag{1.11}$$

因为 $\mathbf{P}$ 是可逆的，我们可得

$$(\tilde{\boldsymbol{\alpha}}_1, \tilde{\boldsymbol{\alpha}}_2, \cdots, \tilde{\boldsymbol{\alpha}}_n)$$

也是 V 的一组基，从而线性无关。又因为 $\mathbf{Q}$ 可逆，从而 $\left(\tilde{\boldsymbol{\beta}}_1, \tilde{\boldsymbol{\beta}}_2, \cdots, \tilde{\boldsymbol{\beta}}_s\right)$ 和 $(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s)$ 可以互相线性表示，也就是说二者等价，从而 $\operatorname{rank}(\boldsymbol{\beta}) = \operatorname{rank}(\tilde{\boldsymbol{\beta}})$. 于是

$$\dim \mathcal{L}(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s) = \dim \mathcal{L}\left(\tilde{\boldsymbol{\beta}}_1, \tilde{\boldsymbol{\beta}}_2, \cdots, \tilde{\boldsymbol{\beta}}_s\right)$$

由 (1.11) 可知

$$\tilde{\boldsymbol{\beta}}_i = \tilde{\boldsymbol{\alpha}}_i, \quad i = 1, 2, \cdots, r.$$

$$\tilde{\boldsymbol{\beta}}_i = 0 \quad i = 1+r, \cdots, s.$$

因为

$$(\tilde{\boldsymbol{\alpha}}_1, \tilde{\boldsymbol{\alpha}}_2, \cdots, \tilde{\boldsymbol{\alpha}}_n)$$

也是 V 的一组基，于是有

$$\{\tilde{\boldsymbol{\beta}}_i, \quad i = 1, 2, \cdots, r.\}$$

线性无关，从而

$$\dim \mathcal{L}\left(\tilde{\boldsymbol{\beta}}_1, \tilde{\boldsymbol{\beta}}_2, \cdots, \tilde{\boldsymbol{\beta}}_s\right) = r = \operatorname{rank}(\boldsymbol{A}).$$

于是

$$\dim \mathcal{L}(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s) = \dim \mathcal{L}\left(\tilde{\boldsymbol{\beta}}_1, \tilde{\boldsymbol{\beta}}_2, \cdots, \tilde{\boldsymbol{\beta}}_s\right) = \operatorname{rank}(\boldsymbol{A})$$

1.8 2024 年 3 月 28 日第四周星期四作业-线性空间的同构

定义 1.22 (线性同构)

设 V_1 与 V_2 是同一个数域 $\mathbb{F}$ 上的两个线性空间. 如果存在 V_1 到 V_2 上的一个一一对应 η, 它把 V_1 中的向量 $\boldsymbol{\alpha}$ 映为 V_2 的向量 $\eta(\boldsymbol{\alpha})$, 使得对任意 $\boldsymbol{\alpha}, \boldsymbol{\beta} \in V_1, \lambda \in \mathbb{F}$, 都有

$$\eta(\boldsymbol{\alpha}+\boldsymbol{\beta}) = \eta(\boldsymbol{\alpha}) + \eta(\boldsymbol{\beta}), \tag{1.12}$$

$$\eta(\lambda\boldsymbol{\alpha}) = \lambda\eta(\boldsymbol{\alpha}), \tag{1.13}$$

则线性空间 V_1 与 V_2 称为同构的, 而映射 η 称为 V_1 到 V_2 上的一个同构映射.

1. 满足上述定义中条件 (2.1) 的映射 η 称为保加法的;
2. 满足条件 (2.2) 的映射称为保乘法的.

因此同构映射 η 是线性空间 V_1 到 V_2 上的保加法与保乘法的一一映射 (即双射)。

命题 1.10

当且仅当 $\alpha = 0$ 时, $\eta(\alpha) = 0$.

证明 设 $\boldsymbol{\beta} \in V_1$, 因为同构映射 η 是保加法的, 因此

$$\eta(\boldsymbol{\beta}) = \eta(\boldsymbol{\beta}+0) = \eta(\boldsymbol{\beta}) + \eta(0).$$

所以 $\eta(0) = 0$. 由于同构映射 η 是双射, 所以, 如果 $\eta(\boldsymbol{\alpha}) = 0$, 则 $\boldsymbol{\alpha} = 0$.

命题 1.11

设 $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{F}, \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_k \in V$, 则

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

对 k 用归纳法. 当 $k=1$ 时, 因为同构映射 η 保乘法. 所以

$$\eta\left(\lambda_1\boldsymbol{\alpha}_1\right)=\lambda_1\eta\left(\boldsymbol{\alpha}_1\right),$$

因此结论对 $k=1$ 成立. 假设结论对 $k-1$ 成立. 由于

$$\eta\left(\lambda_1\boldsymbol{\alpha}_1+\lambda_2\boldsymbol{\alpha}_2+\cdots+\lambda_k\boldsymbol{\alpha}_k\right)=\eta\left(\left(\lambda_1\boldsymbol{\alpha}_1+\lambda_2\boldsymbol{\alpha}_2+\cdots+\lambda_{k-1}\boldsymbol{\alpha}_{k-1}\right)+\lambda_k\boldsymbol{\alpha}_k\right),$$

并且同构映射 η 保加法, 所以

$$\eta\left(\lambda_1\boldsymbol{\alpha}_1+\lambda_2\boldsymbol{\alpha}_2+\cdots+\lambda_k\boldsymbol{\alpha}_k\right)=\eta\left(\lambda_1\boldsymbol{\alpha}_1+\lambda_2\boldsymbol{\alpha}_2+\cdots+\lambda_{k-1}\boldsymbol{\alpha}_{k-1}\right)+\eta\left(\lambda_k\boldsymbol{\alpha}_k\right).$$

由归纳假设, 以及同构映射 η 保乘法, 故

$$\eta\left(\lambda_1\boldsymbol{\alpha}_1+\lambda_2\boldsymbol{\alpha}_2+\cdots+\lambda_k\boldsymbol{\alpha}_k\right)=\lambda_1\eta\left(\boldsymbol{\alpha}_1\right)+\lambda_2\eta\left(\boldsymbol{\alpha}_2\right)+\cdots+\lambda_{k-1}\eta\left(\boldsymbol{\alpha}_{k-1}\right)+\lambda_k\eta\left(\boldsymbol{\alpha}_k\right).$$

注如果线性空间 V_1 到 V_2 的映射 η 满足; 对任意 $\boldsymbol{\alpha},\boldsymbol{\beta}\in V_1,\lambda,\mu\in\mathbb{F}$,

$$\eta(\lambda\boldsymbol{\alpha}+\mu\boldsymbol{\beta})=\lambda\eta(\boldsymbol{\alpha})+\mu\eta(\boldsymbol{\beta}),$$

则映射 η 称为保线性关系的. 命题2.2 表明, 同时保加法与保乘法的映射一定保线性关系. 反之可以证明, 保线性关系的映射一定同时保加法与保乘法, 即保线性等价于同时保加法与保乘法.

命题 1.12

向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_k\in V_1$ 线性相关的充分必要条件是, 向量

$$\eta(\boldsymbol{\alpha}_1),\eta(\boldsymbol{\alpha}_2),\cdots,\eta(\boldsymbol{\alpha}_k)$$

线性相关.

1. 设向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_k$ 线性相关, 则存在不全为零的纯量 $\lambda_1,\lambda_2,\ldots,\lambda_k\in\mathbb{F}$ 使得

$$\lambda_1\boldsymbol{\alpha}_1+\lambda_2\boldsymbol{\alpha}_2+\cdots+\lambda_k\boldsymbol{\alpha}_k=0$$

由命题2.2,

$$\eta\left(\lambda_1\boldsymbol{\alpha}_1+\lambda_2\boldsymbol{\alpha}_2+\cdots+\lambda_k\boldsymbol{\alpha}_k\right)=\lambda_1\eta\left(\boldsymbol{\alpha}_1\right)+\lambda_2\eta\left(\alpha_2\right)+\cdots+\lambda_k\eta\left(\alpha_k\right)=0.$$

因此, 向量 $\eta\left(\boldsymbol{\alpha}_1\right),\eta\left(\boldsymbol{\alpha}_2\right),\ldots,\eta\left(\boldsymbol{\alpha}_k\right)$ 线性相关.

2. 反之, 设 $\eta\left(\boldsymbol{\alpha}_1\right),\eta\left(\boldsymbol{\alpha}_2\right),\ldots,\eta\left(\boldsymbol{\alpha}_k\right)$ 线性相关, 则存在不全为零的纯量 $\lambda_1,\lambda_2,\ldots,\lambda_k\in\mathbb{F}$, 使得

$$\lambda_1\eta\left(\boldsymbol{\alpha}_1\right)+\lambda_2\eta\left(\boldsymbol{\alpha}_2\right)+\cdots+\lambda_k\eta\left(\boldsymbol{\alpha}_k\right)=0.$$

由命题2.2,

$$\lambda_1\eta\left(\boldsymbol{\alpha}_1\right)+\lambda_2\eta\left(\boldsymbol{\alpha}_2\right)+\cdots+\lambda_k\eta\left(\boldsymbol{\alpha}_k\right)=\eta\left(\lambda_1\boldsymbol{\alpha}_1+\lambda_2\boldsymbol{\alpha}_2+\cdots+\lambda_k\boldsymbol{\alpha}_k\right)=0$$

由命题2.1, $\lambda_1\boldsymbol{\alpha}_1+\lambda_2\boldsymbol{\alpha}_2+\cdots+\lambda_k\boldsymbol{\alpha}_k=0$. 因此, $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_k$ 线性相关.

命题 1.13

设 η 是数域 $\mathbb{F}$ 上的有限维线性空间 V_1 到 V_2 的同构映射, 则

$$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n\}$$

是 V_1 的基的必要且充分条件为

$$\{\eta(\boldsymbol{\alpha}_1), \eta(\boldsymbol{\alpha}_2), \cdots, \eta(\boldsymbol{\alpha}_n)\}$$

是 V_2 的基, 从而 $\dim V_1 = \dim V_2$.

命题 1.14

同构是线性空间之间的一种关系，换句话说有

1. (自反性) 数域 $\mathbb{F}$ 上的线性空间 V 与 V 自身同构。
2. (对称性) 若 $\mathbb{F}$ 上的线性空间 V_1 与 V_2 同构，则 V_2 与 V_1 也同构。
3. (传递性) 设 V_1, V_2 和 V_3 是数域 $\mathbb{F}$ 上的线性空间。如果 V_1 与 V_2 同构，V_2 与 V_3 同构，则 V_1 与 V_3 同构。

注 由于数域 F 上的线性空间之间的同构关系满足自反性、对称性和传递性，所以同构关系是数域 $\mathbb{F}$ 上的线性空间之间的一种等价关系。按照同构关系可以对线性空间进行分类: 彼此同构的线性空间归在同一个类, 彼此不同构的线性空间归在不同的类。

基本的问题是：如何判定数域 $\mathbb{F}$ 上的两个线性空间 V_1 和 V_2 是否属于同类，即如何判定线性空间 V_1 和 V_2 是否同构? 在线性空间的同构类中怎样选取代表元?

定理 1.21 (线性空间的分类)

数域 $\mathbb{F}$ 上的任意一个 n 维线性空间 V 都同构于数域 $\mathbb{F}$ 上的 n 维向量空间 $\mathbb{F}^n$

练习 1.42(选做) 假定 V_1, V_2 是数域 $\mathbb{F}$ 两个有限维的向量空间，并且 $\eta : V_1 \mapsto V_2$ 是满射且满足

$$\eta(a\boldsymbol{\alpha} + b\boldsymbol{\beta}) = a\eta(\boldsymbol{\alpha}) + b\eta(\boldsymbol{\beta}), \quad \forall \boldsymbol{\alpha}, \boldsymbol{\beta} \in V_1, \quad \forall a, b \in \mathbb{F}.$$

证明:

$$\dim(V_2) \leq \dim(V_1).$$

你需要注意的是，这里的 η 我们并没有说是同构，也就是说它不一定是一一的。

解 选取 V_2 的一个基

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

因为 η 是满射，则存在

$$\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n\} \subseteq V_1$$

使得

$$\boldsymbol{\alpha}_i = \eta(\boldsymbol{\beta}_i) \quad i = 1, 2, \cdots, n.$$

如果存在 $x_1, x_2, \cdots, x_n \in \mathbb{F}$ 使得

$$x_1\boldsymbol{\beta}_1 + \boldsymbol{\beta}_2 + \cdots + x_n\boldsymbol{\beta}_n = 0.$$

那么有

$$\eta(x_1\boldsymbol{\beta}_1 + \boldsymbol{\beta}_2 + \cdots + x_n\boldsymbol{\beta}_n) = \eta(0) = 0.$$

这里 $\eta(0)=0$ 只需根据保加法就得出来了。于是根据线性性质

$$0=\eta(x_1\boldsymbol{\beta}_1+\boldsymbol{\beta}_2+\cdots+x_n\boldsymbol{\beta}_n)=x_1\eta(\boldsymbol{\beta}_1)+x_2\eta(\boldsymbol{\beta}_2)+\cdots+x_n\eta(\boldsymbol{\beta}_n)=0.$$

也即

$$x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_n\boldsymbol{\alpha}_n=0.$$

因为

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

线性无关，从而

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

于是

$$\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n\}\subseteq V_1$$

线性无关，从而 $\dim(V_1)\geq n=\dim(V_2)$.

练习 1.43 假定 V_1,V_2 是数域 $\mathbb{F}$ 两个向量空间，并且 $\eta:V_1\mapsto V_2$ 是同构。如果 $W\subseteq V_1$ 是一个子空间，证明：$\eta(W)$ 是 V_2 的一个子空间。

解 任取 $\boldsymbol{\alpha},\boldsymbol{\beta}\in\eta(W)$, 则存在 $x,y\in W$ 使得

$$\eta(x)=\boldsymbol{\alpha},\quad \eta(y)=\boldsymbol{\beta}.$$

从而对任意的 $\lambda,\mu\in\mathbb{F}$ 有

$$\lambda\boldsymbol{\alpha}+\mu\boldsymbol{\beta}=\lambda\eta(x)+\mu\eta(y)=\eta(\lambda x+\mu y).$$

因为 W 是子空间，所以 $\lambda x+\mu y\in W$, 从而 $\lambda\boldsymbol{\alpha}+\mu\boldsymbol{\beta}\in\eta(W)$. 于是 $\eta(W)$ 对加法和数乘封闭，从而是 V_2 的子空间。

练习 1.44 记 $\mathbb{F}[x]$ 为数域 $\mathbb{F}$ 上的所有多项式组成的向量空间。证明：存在 $\mathbb{F}[x]$ 的真子空间（即不等于 $\mathbb{F}[x]$）与 $\mathbb{F}[x]$ 同构。

解 令 W 是 $\mathbb{F}[x]$ 中所有的只含有偶数次幂的多项式的集合：

$$W:=\left\{a_0+a_1x^2+a_2x^4+\cdots+a_nx^{2n}:a_i\in\mathbb{F},n=0,1,\cdots\right\}$$

则显然 W 对加法和数乘封闭，从而是 $\mathbb{F}[x]$ 的子空间。令 $\eta:\mathbb{F}[x]\mapsto W$ 为

$$\eta(a_0+a_1x+a_2x^2+\cdots+a_nx^n)=a_0+a_1x^2+a_2x^4+\cdots+a_nx^{2n}.$$

则 η 满足

$$\eta(\lambda p+\mu q)=\lambda\eta(p)+\mu\eta(q),\quad \forall p,q\in\mathbb{F}[x],\quad \forall\lambda,\mu\in\mathbb{F}.$$

同时显然 η 既是单射又是满射，所以是同构。

练习 1.45(选做) 设 $\mathbf{R}^+$ 是所有正实数组成的集合，加法和数乘定义如下：

$$\forall a,b\in\mathbf{R}_+,\ \ k\in\mathbf{R}:\ \ a\oplus b=ab,\ \ k\odot a=a^k$$

则根据例子 1.4知道 $\mathbf{R}^+$ 关于这一加法和数乘构成 $\mathbb{R}$ 上的一个线性空间. 这个空间与 $\mathbb{R}$(作为 $\mathbb{R}$ 上的线性空间) 同构吗？如果同构的话，请给出一个同构映射。

解 同构。例如

$$\eta(x)=e^x,\quad x\in\mathbb{R}$$

便满足条件。既单射又满射是显然的，而线性关系来自于

$$\eta(ax+by)=e^{ax+by}=e^{ax}e^{by}=e^{ax}\oplus e^{by}=(e^x)^a\oplus(e^y)^b=(a\odot e^x)\oplus(b\odot e^y)=(a\odot\eta(x))\oplus(b\odot\eta(y)).$$

练习 1.46 设 η 是 V_1 到 V_2 的同构映射，

$$S_1=\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_m\}$$

是 V_1 的任意一组向量，

$$S_2=\{\eta(\boldsymbol{\alpha}_1),\eta(\boldsymbol{\alpha}_2),\ldots,\eta(\boldsymbol{\alpha}_m)\}$$

证明：$\operatorname{rank}(S_1)=\operatorname{rank}(S_2)$，这里 rank 表示向量组的秩（回忆定义 1.13）。这个结论也就是说同构映射保持映射前后向量组秩不变.

解 假定 $\{\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_k}\}$ 为 S_1 的极大线性无关组，则根据命题 1.8 可知，

$$\{\eta(\boldsymbol{\alpha}_{i_1}),\eta(\boldsymbol{\alpha}_{i_2}),\cdots,\eta(\boldsymbol{\alpha}_{i_k})\}$$

线性无关。从而有

$$\operatorname{rank}(S_1)\le\operatorname{rank}(S_2).$$

再利用 η^{-1} 的同构可得

$$\operatorname{rank}(S_2)\le\operatorname{rank}(S_1).$$

练习 1.47 如果数域 $\mathbb{F}$ 上两个向量空间存在一一对应，它们一定同构吗？对的话请证明，不对的话请给出反例。提示：考虑有理数域 $\mathbb{Q}$ 上的两个空间 $\mathbb{Q}$ 和 $\mathbb{Q}^2$.

解 利用对角线法则可构造 $\mathbb{Q}^2$ 到 $\mathbb{Q}$ 的一一对应。但是作为 $\mathbb{Q}$ 上的向量空间

$$\dim(\mathbb{Q}^2)=2\neq1=\dim(\mathbb{Q}).$$

所以二者一定不同构。

1.9 2024 年 4 月 2 日第五周星期二作业-线性方程的解和商空间

假定 $\boldsymbol{A}$ 是数域 $\mathbb{F}$ 上的 $m\times n$ 阶矩阵，考虑下面方程组 $\boxed{\boldsymbol{A}X=0}$ 的解。这个方程组的所有解组成了一个线性空间：

$$\boxed{\operatorname{Ker}(A):=\{X\in\mathbb{F}^n:\boldsymbol{A}X=0\}}$$

需要注意这里我们把 $\mathbb{F}^n$ 中的元素看成列向量，

$$\boxed{X=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}.}$$

同时我们还可以定义一个 $\mathbb{F}^m$ 的子空间。

$$\boxed{\mathrm{Im}(A) := \{\boldsymbol{A}X : X \in \mathbb{F}^n\}}$$

注 事实上，$\boldsymbol{A}$ 定义了一个 $\mathbb{F}^n$ 到 $\mathbb{F}^m$ 的映射 η:

$$\eta(X) := \boldsymbol{A}X.$$

一般来说，$\mathrm{Ker}(\boldsymbol{A})$ 叫做“核”，而 $\mathrm{Im}(\boldsymbol{A})$ 叫做“像”。

定理 1.22

我们有

$$\dim(\mathrm{Ker}(\boldsymbol{A})) = n - \mathrm{rank}(\boldsymbol{A}), \quad \dim(\mathrm{Im}(\boldsymbol{A})) = \mathrm{rank}(\boldsymbol{A}).$$

♡

例题 1.19 假定 $\boldsymbol{A}$ 是 $m \times n$ 阶的矩阵，而 $\boldsymbol{B}$ 是 $n \times p$ 阶的矩阵。那么有

$$\mathrm{rank}(\boldsymbol{A}) + \mathrm{rank}(\boldsymbol{B}) - n \leq \mathrm{rank}(\boldsymbol{AB}).$$

例题 1.20 If $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ are matrices in $\mathbb{R}^{n\times n}$ and $\boldsymbol{ABC} = 0$, show that

$$\mathrm{rank}(\boldsymbol{A}) + \mathrm{rank}(\boldsymbol{B}) + \mathrm{rank}(\boldsymbol{C}) \leq 2n.$$

定义 1.23 (同余)

设 W 是数域 $\mathbb{F}$ 上的有限维线性空间 V 的子空间, 向量 $\boldsymbol{\alpha}, \boldsymbol{\beta} \in V$. 如果向量 $\boldsymbol{\alpha} - \boldsymbol{\beta} \in W$, 则向量 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 称为模 W 同余, 记为 $\boldsymbol{\alpha} = \boldsymbol{\beta}(\mathrm{mod} W)$.

♣

命题 1.15

V 中向量之间的模 W 同余关系具有以下性质:

(1) (自反性) 对任意向量 $\boldsymbol{\alpha} \in V, \boldsymbol{\alpha} \equiv \boldsymbol{\alpha}(\mathrm{mod} W)$;

(2) (对称性) 设向量 $\boldsymbol{\alpha}, \boldsymbol{\beta} \in V$, 且 $\boldsymbol{\alpha} \equiv \boldsymbol{\beta}(\mathrm{mod} W)$, 则 $\boldsymbol{\beta} \equiv \boldsymbol{\alpha}(\mathrm{mod} W)$;

(3) (传递性) 设向量 $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma} \in V$, 并且 $\boldsymbol{\alpha} \equiv \boldsymbol{\beta}(\mathrm{mod} W), \boldsymbol{\beta} \equiv \boldsymbol{\gamma}(\mathrm{mod} W)$, 则 $\boldsymbol{\alpha} \equiv \boldsymbol{\gamma}(\mathrm{mod}\ W)$.

♠

1. 我们在线性空间 V 的向量之间引进了一种关系, 即同余关系.
2. 同余关系满足 自反性、对称性和传递性.
3. 因此 V 中向量便按照同余关系划分为同余类: 即在同一个同余类的向量彼此模 W 同余, 而在不同的同余类中的向量模 W 不同余.

定义 1.24 (同余类)

向量 $\boldsymbol{\alpha}$ 所在的同余类记为 $\bar{\boldsymbol{\alpha}}$. 同余类 $\bar{\boldsymbol{\alpha}}$ 由哪些向量构成? 记

$$\boldsymbol{\alpha} + W = \{\boldsymbol{\alpha} + \boldsymbol{\gamma} \mid \boldsymbol{\gamma} \in W\}.$$

♣

1. 设 $\boldsymbol{\beta} \in \boldsymbol{\alpha} + W$ ，则 $\boldsymbol{\beta} = \boldsymbol{\alpha} + \boldsymbol{\gamma}, \boldsymbol{\gamma} \in W$. 因此 $\boldsymbol{\beta} - \boldsymbol{\alpha} = \boldsymbol{\gamma} \in W$, 即 $\boldsymbol{\beta} = \boldsymbol{\alpha}(\mathrm{mod} W)$. 所以, $\boldsymbol{\beta} \in \bar{\boldsymbol{\alpha}}$.
2. 其次设 $\boldsymbol{\beta} \in \bar{\boldsymbol{\alpha}}$, 则 $\boldsymbol{\beta} = \boldsymbol{\alpha}(\mathrm{mod} W)$, 因此存在 $\boldsymbol{\gamma} \in W$, 使得 $\boldsymbol{\beta} - \boldsymbol{\alpha} = \boldsymbol{\gamma}$, 即 $\boldsymbol{\beta} = \boldsymbol{\alpha} + \boldsymbol{\gamma}$, 于

是 $\boldsymbol{\beta} \in \boldsymbol{\alpha} + W$. 所以，$\bar{\boldsymbol{\alpha}} \subseteq \boldsymbol{\alpha} + W$, 从而 $\bar{\boldsymbol{\alpha}} = \boldsymbol{\alpha} + W$.

定义 1.25 (商空间)

所有模 W 的同余类集合记为 V/W.

1. 在集合 V/W 中规定同余类的加法如下: 对任意 $\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{\beta}} \in V/W$, 令
$$\bar{\boldsymbol{\alpha}} + \bar{\boldsymbol{\beta}} = \overline{\boldsymbol{\alpha} + \boldsymbol{\beta}}.$$
2. 纯量与同余类的乘法规定为; 对任意纯量 $\lambda \in \mathbb{F}, \bar{\boldsymbol{\alpha}} \in V/W$, 令
$$\lambda \bar{\boldsymbol{\alpha}} = \overline{\lambda \boldsymbol{\alpha}}$$

♣

加法和数乘是 "well-defined"，上面规定的同余类加法与同余类代表元的选取无关:

1. 设 $\boldsymbol{\alpha}' \in \bar{\boldsymbol{\alpha}}, \boldsymbol{\beta}' \in \bar{\boldsymbol{\beta}}$. 则有 $\boldsymbol{\alpha}' \equiv \boldsymbol{\alpha}(\bmod W), \boldsymbol{\beta}' \equiv \boldsymbol{\beta}(\bmod W)$, 即 $\boldsymbol{\alpha}' - \boldsymbol{\alpha}, \boldsymbol{\beta}' - \boldsymbol{\beta} \in W$, 因此,
$$(\boldsymbol{\alpha}' + \boldsymbol{\beta}') - (\boldsymbol{\alpha} + \boldsymbol{\beta}) = (\boldsymbol{\alpha}' - \boldsymbol{\alpha}) + (\boldsymbol{\beta}' - \boldsymbol{\beta}) \in W.$$
所以 $\boldsymbol{\alpha}' + \boldsymbol{\beta}' \equiv \boldsymbol{\alpha} + \boldsymbol{\beta}(\bmod W)$.
2. 同样可以证明, 纯量与同余类的乘法与同余类的代表元选取无关.

定义 1.26

同余类集合 V/W 在上述规定的加法与乘法下是数域 $\mathbb{F}$ 上的线性空间, 其中 W 是线性空间 V/W 的零向量. 数域 $\mathbb{F}$ 上的线性空间 V/W 称为线性空间 V 关于子空间 W 的商空间. ♣

✍ **练习 1.48** 设 $\boldsymbol{A}$ 是数域 $\mathbb{F}$ 上的一个 $m \times n$ 阶的矩阵, 秩 $\boldsymbol{A} = r$, 从 $\boldsymbol{A}$ 中任取出 s 行, 作一个 s 行的矩阵 $\boldsymbol{B}$. 证明: 秩 $\operatorname{rank}\boldsymbol{B} \geqslant r + s - m$.

解 不妨设矩阵 $\boldsymbol{B}$ 就是前 s 行拿出来的（因为不然的话可以通过行变换变成前 s 行，而行变换不改变矩阵的秩）。于是
$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{B} \\ \boldsymbol{C} \end{pmatrix}$$
这里 $\boldsymbol{B}$ 是 $s \times n$ 阶的，$\boldsymbol{C}$ 是 $(m-s) \times n$ 阶的。注意到对 $X \in \mathbb{F}^n$
$$\boldsymbol{A}X = \begin{pmatrix} \boldsymbol{B}X \\ \boldsymbol{C}X \end{pmatrix} = \begin{pmatrix} \boldsymbol{B}X \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \boldsymbol{C}X \end{pmatrix} \tag{1.14}$$
假定
$$\{\boldsymbol{B}X_1, \boldsymbol{B}X_2, \cdots, \boldsymbol{B}X_k\} \subset \mathbb{F}^s$$
是 $\operatorname{Im}(\boldsymbol{B})$ 的一个基。假定
$$\{\boldsymbol{C}Y_1, \boldsymbol{C}Y_2, \cdots, \boldsymbol{C}Y_l\} \subset \mathbb{F}^{m-s}$$
是 $\operatorname{Im}(\boldsymbol{C})$ 的一个基。根据 (1.14) 可得存在常数 $x_1, x_2, \cdots, x_k, y_1, y_2, \cdots, y_l \in \mathbb{F}$ 使得
$$\boldsymbol{A}X = \begin{pmatrix} \boldsymbol{B}X \\ \boldsymbol{C}X \end{pmatrix} = \begin{pmatrix} \boldsymbol{B}X \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \boldsymbol{C}X \end{pmatrix} = \sum_{i=1}^{k} x_k \begin{pmatrix} \boldsymbol{B}X_i \\ 0 \end{pmatrix} + \sum_{j=1}^{l} y_j \begin{pmatrix} 0 \\ \boldsymbol{C}Y_j \end{pmatrix}$$

于是

$$\mathrm{Im}(\boldsymbol{A}) = \mathrm{span}\left\{\begin{pmatrix} \boldsymbol{B}X_i \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \boldsymbol{C}Y_j \end{pmatrix} | i = 1, \cdots, k, j = 1, \cdots, l\right\}$$

因为 $\mathrm{Im}(\boldsymbol{A})$ 的维数一定不超过生成它的向量的个数，我们得到

$$\dim(\mathrm{Im}(\boldsymbol{A})) \leq k + l = \dim(\mathrm{Im}(\boldsymbol{B})) + \dim(\mathrm{Im}(\boldsymbol{C})).$$

也就是说

$$\mathrm{rank}(\boldsymbol{A}) \leq \mathrm{rank}(\boldsymbol{B}) + \mathrm{rank}(\boldsymbol{C}).$$

因为 $\boldsymbol{C}$ 是 $(m-s) \times n$ 阶的矩阵，从而 $\mathrm{rank}(\boldsymbol{C}) \leq m - s$. 这就得到

$$\mathrm{rank}(\boldsymbol{A}) \leq \mathrm{rank}(\boldsymbol{B}) + m - s.$$

练习 1.49 假定 V 是数域 $\mathbb{R}$ 上的向量空间，并且 $W \subseteq V$ 是一个子空间. 证明: $x + W = y + W$ 当切仅当 $x - y \in W$. 这里你需要回忆一下定义1.24

$$x + W := \{x + y : y \in W\}.$$

解 很直接的验证，参见上面的课件的细节。这里不再重复。

练习 1.50 假定 V 是数域 $\mathbb{R}$ 上的向量空间，并且 $W \subseteq V$ 是一个子空间. 则 $x + W = W$ 当且仅当 $x \in W$.

解 直接用上面的习题结果。

如果 $x + W = W$，则 $x + w \in W$。从而存在对某一个 $w_1 \in W$ 使得 $x + w = w_1$. 于是 $x = w_1 - w \in W$.

反之，如果 $x \in W$，考虑 $p = x + w_1 \in x + W$。注意到 $x + w_1 \in W$ ，从而 $p \in W$ ，这就得到 $x + W \subseteq W$. 同样, 如果 $w \in W$，则 $w = x + w - x \in$ 和 $w - x \in W$, 因此 $w \in x + W$。我们得到 $W \subseteq x + W$. T 于是有 $x + W = W$.

练习 1.51 假定 V 是数域 $\mathbb{R}$ 上的向量空间，并且 $W \subseteq V$ 是一个子空间. 如果 $x_1 + W = x_2 + W$ 以及 $y_1 + W = y_2 + W$ 并且 $c \in \mathbb{R}$ 证明

$$x_1 + y_1 + W = x_2 + y_2 + W, \quad cx_1 + W = cx_2 + W.$$

这验证了同余类上加法和数乘的合理性。

解 参见上面的课件的细节。这里不再重复。

练习 1.52 设 $V = \mathbb{R}^3$ 以及 $W = \mathrm{span}\{(0,0,1)\}$. 求 V/W。也就是写出这个空间的元素是什么样子，并求出它的维数。

解 注意到

$$(x_1, x_2, x_3) = (y_1, y_2, y_3)(\mathrm{mod} W) \Leftrightarrow \begin{cases} x_1 = x_2 \\ y_1 = y_2 \end{cases}$$

因此有

$$\overline{(a, b, c)} = \{(a, b, z) \mid z \in \mathbb{R}\}$$

也就是说 V/W 中的点事实上可以看成 V 中的直线. 而

$$\overline{(a,b,c)}=\overline{(a,b,0)}=a\overline{(1,0,0)}+b\overline{(0,1,0)}.$$

如果

$$x\overline{(1,0,0)}+y\overline{(0,1,0)}=\overline{(0,0,0)}.$$

则可得

$$\overline{(x,y,0)}=\overline{(0,0,0)}.$$

所以

$$x=y=0.$$

也就是说

$$\{\overline{(1,0,0)},\overline{(0,1,0)}\}$$

是 V/W 的一组基。从而其维数是 2.

练习 1.53 设 $\boldsymbol{A}$ 是数域 $\boldsymbol{P}$ 上 n 阶方阵.

1. 证明

$$\operatorname{rank}\left(\boldsymbol{A}^k\right)-\operatorname{rank}\left(\boldsymbol{A}^{k+1}\right)\geqslant\operatorname{rank}\left(\boldsymbol{A}^{k+1}\right)-\operatorname{rank}\left(\boldsymbol{A}^{k+2}\right)\geqslant 0.$$

2. 若 $\operatorname{rank}\left(\boldsymbol{A}^k\right)=\operatorname{rank}\left(\boldsymbol{A}^{k+1}\right)$, 证明

$$\operatorname{rank}\left(\boldsymbol{A}^k\right)=\operatorname{rank}\left(\boldsymbol{A}^{k+s}\right),s\in\mathbf{N}\ (\text{自然数集}).$$

解

1. 事实上, 在第五周的手写笔记里我给大家写了一个东西里面包含这个证明。根据引理 1.7（$\boldsymbol{B}$ 用 $\boldsymbol{A}$ 替代, $\boldsymbol{A}$ 被 $\boldsymbol{A}^k$ 替代）我们得到

$$\dim(\operatorname{Im}(\boldsymbol{A}^{k+1}))+\dim(\operatorname{Ker}(\boldsymbol{A})\cap\operatorname{Im}(\boldsymbol{A}^k))=\dim(\operatorname{Im}(\boldsymbol{A}^k))$$

同样

$$\dim(\operatorname{Im}(\boldsymbol{A}^{k+2}))+\dim(\operatorname{Ker}(\boldsymbol{A})\cap\operatorname{Im}(\boldsymbol{A}^{k+1}))=\dim(\operatorname{Im}(\boldsymbol{A}^{k+1}))$$

显然

$$\operatorname{Im}(\boldsymbol{A}^{k+1})\subseteq\operatorname{Im}(\boldsymbol{A}^k).$$

于是

$$\dim(\operatorname{Ker}(\boldsymbol{A})\cap\operatorname{Im}(\boldsymbol{A}^{k+1}))\leq\dim(\operatorname{Ker}(\boldsymbol{A})\cap\operatorname{Im}(\boldsymbol{A}^k)).$$

得证。

2. 首先因为 $\operatorname{rank}\left(\boldsymbol{A}^k\right)=\operatorname{rank}\left(\boldsymbol{A}^{k+1}\right)$, 而

$$\dim(\operatorname{Im}(\boldsymbol{A}^k))=\operatorname{rank}(\boldsymbol{A}^k),\quad\dim(\operatorname{Im}(\boldsymbol{A}^{k+1}))=\operatorname{rank}(\boldsymbol{A}^{k+1}).$$

再结合 $\operatorname{Im}(\boldsymbol{A}^{k+1})\subseteq\operatorname{Im}(\boldsymbol{A}^k)$ 可得

$$\operatorname{Im}(\boldsymbol{A}^{k+1})=\operatorname{Im}(\boldsymbol{A}^k).$$

也就是说对任意的 $X \in \mathbb{F}^n$, 一定存在 $Y \in \mathbb{F}^n$ 使得

$$\boldsymbol{A}^k X = \boldsymbol{A}^{k+1} Y.$$

从而

$$\boldsymbol{A}^{k+1} X = \boldsymbol{A}^{k+2} Y.$$

再利用 $X \in \mathbb{F}^n$ 的任意性可得

$$\operatorname{Im}(\boldsymbol{A}^{k+1}) \subseteq \operatorname{Im}(\boldsymbol{A}^{k+2}).$$

再结合

$$\operatorname{Im}(\boldsymbol{A}^{k+2}) \subseteq \operatorname{Im}(\boldsymbol{A}^{k+1})$$

可得

$$\operatorname{Im}(\boldsymbol{A}^{k+2}) = \operatorname{Im}(\boldsymbol{A}^{k+1}).$$

于是

$$\operatorname{rank}\left(\boldsymbol{A}^{k+1}\right) = \operatorname{rank}\left(\boldsymbol{A}^{k+2}\right)$$

一直做下去便得证。

引理 1.7

假定 $\boldsymbol{A}, \boldsymbol{B}$ 是两个 n 阶方阵，则有

$$\dim(\operatorname{Im}(\boldsymbol{BA})) + \dim(\operatorname{Ker}(\boldsymbol{B}) \cap \operatorname{Im}(\boldsymbol{A})) = \dim(\operatorname{Im}(\boldsymbol{A}))$$

♡

这个引理在我们学了线性映射后很自然，这里我们给一个直接的证明。老师写的很　嗦很详细，是为了你能看明白，希望大家好好把过程过一遍，掌握。因为这个结论的证明就是我们反复利用我们学的定义。

证明 假定

$$\{\boldsymbol{BA\alpha}_1, \boldsymbol{BA\alpha}_2, \cdots \boldsymbol{BA\alpha}_k\}, \quad \boldsymbol{\alpha}_i \in \mathbb{F}^n, \quad i = 1, 2, \cdots, k$$

为 $\operatorname{Im}(\boldsymbol{BA})$ 的一组基。则有

$$\{\boldsymbol{A\alpha}_1, \boldsymbol{A\alpha}_2, \cdots, \boldsymbol{A}\alpha_k\}$$

是线性无关的。这是因为如果

$$x_1 \boldsymbol{A\alpha}_1 + x_2 \boldsymbol{A\alpha}_2 + \cdots + x_k \boldsymbol{A\alpha}_k = 0.$$

两边再乘以矩阵 $\boldsymbol{B}$ 可得

$$x_1 \boldsymbol{BA\alpha}_1 + x_2 \boldsymbol{BA\alpha}_2 + \cdots + x_k \boldsymbol{BA\alpha}_k = 0.$$

再利用 $\{\boldsymbol{BA\alpha}_1, \boldsymbol{BA\alpha}_2, \cdots \boldsymbol{BA\alpha}_k\}$ 为 $\operatorname{Im}(\boldsymbol{BA})$ 的一组基可得 $x_1 = x_2 = \cdots = x_k = 0$. 再假定

$$\{\boldsymbol{A\beta}_1, \boldsymbol{A\beta}_2, \cdots, \boldsymbol{A\beta}_l\}, \quad \boldsymbol{\beta}_j \in \mathbb{F}^n, \quad j = 1, 2, \cdots, l$$

为 $\operatorname{Im}(\boldsymbol{A}) \cap \operatorname{Ker}(\boldsymbol{B})$ 的一组基。

我们接下来证明

$$\{\boldsymbol{A}\boldsymbol{\alpha}_1, \boldsymbol{A}\boldsymbol{\alpha}_2, \cdots, \boldsymbol{A}\alpha_k, \boldsymbol{A}\boldsymbol{\beta}_1, \boldsymbol{A}\boldsymbol{\beta}_2, \cdots, \boldsymbol{A}\boldsymbol{\beta}_l\} \tag{1.15}$$

是 $\mathrm{Im}(\boldsymbol{A})$ 的一个基。

(1) 线性无关性。假定

$$x_1\boldsymbol{A}\boldsymbol{\alpha}_1 + x_2\boldsymbol{A}\boldsymbol{\alpha}_2 + \cdots + x_k\boldsymbol{A}\boldsymbol{\alpha}_k + y_1\boldsymbol{A}\boldsymbol{\beta}_1 + y_2\boldsymbol{A}\boldsymbol{\beta}_2 + \cdots + y_l\boldsymbol{A}\boldsymbol{\beta}_l = 0. \tag{1.16}$$

因为 $\boldsymbol{B}\boldsymbol{A}\boldsymbol{\beta}_j = 0, j = 1, 2, \cdots, l$, 于是上面的式子两边都乘以矩阵 $\boldsymbol{B}$ 可得

$$x_1\boldsymbol{B}\boldsymbol{A}\boldsymbol{\alpha}_1 + x_2\boldsymbol{B}\boldsymbol{A}\boldsymbol{\alpha}_2 + \cdots + x_k\boldsymbol{B}\boldsymbol{A}\boldsymbol{\alpha}_k = 0.$$

再利用 $\{\boldsymbol{B}\boldsymbol{A}\boldsymbol{\alpha}_1, \boldsymbol{B}\boldsymbol{A}\boldsymbol{\alpha}_2, \cdots \boldsymbol{B}\boldsymbol{A}\boldsymbol{\alpha}_k\}$ 为 $\mathrm{Im}(\boldsymbol{B}\boldsymbol{A})$ 的一组基可得 $x_1 = x_2 = \cdots = x_k = 0$. 于是 (1.16) 变成了

$$y_1\boldsymbol{A}\boldsymbol{\beta}_1 + y_2\boldsymbol{A}\boldsymbol{\beta}_2 + \cdots + y_l\boldsymbol{A}\boldsymbol{\beta}_l = 0.$$

但是我们已经假定

$$\{\boldsymbol{A}\boldsymbol{\beta}_1, \boldsymbol{A}\boldsymbol{\beta}_2, \cdots, \boldsymbol{A}\boldsymbol{\beta}_l\}, \quad \boldsymbol{\beta}_j \in \mathbb{F}^n, \quad j = 1, 2, \cdots, l$$

为 $\mathrm{Im}(\boldsymbol{A}) \cap \mathrm{Ker}(\boldsymbol{B})$ 的一组基从而线性无关。于是 $y_1 = y_2 = \cdots = y_l = 0$. 于是(1.15) 线性无关。

(2) 任何一个 $\mathrm{Im}(\boldsymbol{A})$ 中的向量可以被 (1.15) 线性表示。任取 $\boldsymbol{A}\boldsymbol{\alpha} \in \mathrm{Im}(\boldsymbol{A})$. 则 $\boldsymbol{B}\boldsymbol{A}\boldsymbol{\alpha} \in \mathrm{Im}(\boldsymbol{B}\boldsymbol{A})$, 从而存在 $x_1, x_2, \cdots, x_k \in \mathbb{F}$ 使得

$$\boldsymbol{B}\boldsymbol{A}\boldsymbol{\alpha} = x_1\boldsymbol{B}\boldsymbol{A}\boldsymbol{\alpha}_1 + x_2\boldsymbol{B}\boldsymbol{A}\boldsymbol{\alpha}_2 + \cdots + x_k\boldsymbol{B}\boldsymbol{A}\boldsymbol{\alpha}_k.$$

于是

$$\boldsymbol{B}(\boldsymbol{A}\boldsymbol{\alpha} - x_1\boldsymbol{A}\boldsymbol{\alpha}_1 - x_2\boldsymbol{A}\boldsymbol{\alpha}_2 - \cdots - x_k\boldsymbol{A}\boldsymbol{\alpha}_k) = 0.$$

从而

$$\boldsymbol{A}\boldsymbol{\alpha} - x_1\boldsymbol{A}\boldsymbol{\alpha}_1 - x_2\boldsymbol{A}\boldsymbol{\alpha}_2 - \cdots - x_k\boldsymbol{A}\boldsymbol{\alpha}_k \in \mathrm{Ker}(\boldsymbol{B}).$$

显然

$$\boldsymbol{A}\boldsymbol{\alpha} - x_1\boldsymbol{A}\boldsymbol{\alpha}_1 - x_2\boldsymbol{A}\boldsymbol{\alpha}_2 - \cdots - x_k\boldsymbol{A}\boldsymbol{\alpha}_k = \boldsymbol{A}(\boldsymbol{\alpha} - x_1\boldsymbol{\alpha}_1 - x_2\boldsymbol{\alpha}_2 - \cdots - x_k\boldsymbol{\alpha}_k) \in \mathrm{Im}(\boldsymbol{A}).$$

$$\boldsymbol{A}\boldsymbol{\alpha} - x_1\boldsymbol{A}\boldsymbol{\alpha}_1 - x_2\boldsymbol{A}\boldsymbol{\alpha}_2 - \cdots - x_k\boldsymbol{A}\boldsymbol{\alpha}_k \in \mathrm{Ker}(\boldsymbol{B}) \cap \mathrm{Im}(\boldsymbol{A}).$$

因为

$$\{\boldsymbol{A}\boldsymbol{\beta}_1, \boldsymbol{A}\boldsymbol{\beta}_2, \cdots, \boldsymbol{A}\boldsymbol{\beta}_l\}, \quad \boldsymbol{\beta}_j \in \mathbb{F}^n, \quad j = 1, 2, \cdots, l$$

为 $\mathrm{Im}(\boldsymbol{A}) \cap \mathrm{Ker}(\boldsymbol{B})$ 的一组基，从而存在 $y_1, y_2, \cdots, y_l \in \mathbb{F}$ 使得

$$\boldsymbol{A}\boldsymbol{\alpha} - x_1\boldsymbol{A}\boldsymbol{\alpha}_1 - x_2\boldsymbol{A}\boldsymbol{\alpha}_2 - \cdots - x_k\boldsymbol{A}\boldsymbol{\alpha}_k = y_1\boldsymbol{A}\boldsymbol{\beta}_1 + y_2\boldsymbol{A}\boldsymbol{\beta}_2 + \cdots + y_l\boldsymbol{A}\boldsymbol{\beta}_l.$$

于是

$$\boldsymbol{A}\boldsymbol{\alpha} = x_1\boldsymbol{A}\boldsymbol{\alpha}_1 + x_2\boldsymbol{A}\boldsymbol{\alpha}_2 + \cdots + x_k\boldsymbol{A}\boldsymbol{\alpha}_k + y_1\boldsymbol{A}\boldsymbol{\beta}_1 + y_2\boldsymbol{A}\boldsymbol{\beta}_2 + \cdots + y_l\boldsymbol{A}\boldsymbol{\beta}_l.$$

综上我们得到

$$\dim(\mathrm{Im}(\boldsymbol{A})) = k + l.$$

而

$$\dim(\mathrm{Im}(\boldsymbol{B}\boldsymbol{A})) = k, \quad \dim(\mathrm{Ker}(\boldsymbol{B}) \cap \mathrm{Im}(\boldsymbol{A})) = l.$$

我们完成了证明。

练习 1.54 设 $\boldsymbol{A} \in \mathbb{F}^{m\times n}, \boldsymbol{B} \in \mathbb{F}^{m\times 1}$

$$\boldsymbol{V}_1 = \mathrm{Ker}(A) = \left\{\boldsymbol{X} \in \mathbb{F}^{n\times 1} \mid \boldsymbol{A}\boldsymbol{X} = 0\right\}.$$

$\boldsymbol{S}$ 为

$$\boldsymbol{A}\boldsymbol{X} = \boldsymbol{B}$$

的解的集合. 试证: 或者 $\boldsymbol{S} = \varnothing$ 或者存在 $\boldsymbol{X}_0 \in \mathbb{F}^{n\times 1}$ 使 $\boldsymbol{S} = \boldsymbol{X}_0 + \boldsymbol{V}_1$.

解 如果 $\boldsymbol{S} = \emptyset$ 那就不需多证。下面假定 $\boldsymbol{S} \neq \emptyset$。任取 $\mathbf{X}_0 \in \boldsymbol{S}$, 则对任意的 $\mathbf{X} \in \boldsymbol{S}$ 都有

$$\boldsymbol{A}\boldsymbol{X} = \boldsymbol{B}, \quad \boldsymbol{A}\boldsymbol{X}_0 = \boldsymbol{B}.$$

于是

$$\boldsymbol{A}(\boldsymbol{X} - \boldsymbol{X}_0) = \boldsymbol{B} - \boldsymbol{B} = 0.$$

也就是

$$\boldsymbol{X} - \boldsymbol{X}_0 \in \mathbf{V}_1.$$

于是

$$\boldsymbol{X} \in \boldsymbol{X}_0 + \boldsymbol{V}_1.$$

从而

$$\boldsymbol{S} \subseteq \boldsymbol{X}_0 + \boldsymbol{V}_1.$$

反之，任取 $\boldsymbol{Y} \in \boldsymbol{X}_0 + \boldsymbol{V}_1$, 则存在 $\mathbf{X} \in \mathbf{V}_1$ 使得

$$\boldsymbol{Y} = \boldsymbol{X}_0 + \boldsymbol{X}.$$

于是

$$\boldsymbol{A}\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{X}_0 + \boldsymbol{A}\boldsymbol{X} = \boldsymbol{B} + 0 = \boldsymbol{B}.$$

从而 $\boldsymbol{Y} \in \boldsymbol{S}$. 于是

$$\boldsymbol{X}_0 + \boldsymbol{V}_1 \subseteq \boldsymbol{S}.$$

综上

$$\boldsymbol{X}_0 + \boldsymbol{V}_1 = \boldsymbol{S}.$$

第 1 章 练习

1. 假定 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3, \boldsymbol{\alpha}_4$ 生成 V. 证明

$$\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2, \boldsymbol{\alpha}_2-\boldsymbol{\alpha}_3, \boldsymbol{\alpha}_3-\boldsymbol{\alpha}_4, \boldsymbol{\alpha}_4$$

也可生成 V.

2. 求 t 使得

$$(3,1,4),(2,-3,5),(5,9,t)$$

在 $\mathbb{R}^3$ 中不是线性独立的.

3. 证明或者给出反例: 如果 $v_1, \ldots, v_m$ 和 $w_1, \ldots, w_m$ 分别是向量空间 V 中两组线性独立的向量, 则 $v_1+w_1, \ldots, v_m+w_m$ 也是线性独立的.

4. 假定 $v_1, \ldots, v_m$ 是 V 中一组线性独立的向量，取 $w \in V$. 证明：如果 $v_1+w, \ldots, v_m+w$ 线性相关, 则

$$w \in \operatorname{span}\left(v_1, \ldots, v_m\right).$$

5. 假定 $v_1, \ldots, v_m$ 是 V 中一组线性独立的向量，取 $w \in V$. 证明：$v_1, \ldots, v_m, w$ 线性独立当且仅当

$$w \notin \operatorname{span}\left(v_1, \ldots, v_m\right).$$

6. 假定 $v_1, \ldots, v_m$ 是 V 中一列线性无关的向量. 则存在 $j \in\{1,2, \ldots, m\}$ 使得如下成立:
   (a) $v_j \in \operatorname{span}\left(v_1, \ldots, v_{j-1}\right)$;
   (b)

$$\operatorname{span}\left(v_1, \ldots, v_{j-1}, v_{j+1}, \ldots v_m\right)=\operatorname{span}\left(v_1, \ldots, v_m\right).$$

解

7. (1) 若 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 线性无关，问 $\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2, \boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3, \cdots, \boldsymbol{\alpha}_{n-1}+\boldsymbol{\alpha}_n, \boldsymbol{\alpha}_n+\boldsymbol{\alpha}_1$ 是否一定线性无关? 为什么?
   (2) 若 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 线性相关，问 $\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2, \boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3, \cdots, \boldsymbol{\alpha}_{n-1}+\boldsymbol{\alpha}_n, \boldsymbol{\alpha}_n+\boldsymbol{\alpha}_1$ 是否一定线性相关? 为什么?

解 记 $\beta_i=\alpha_i+\alpha_{i+1}(1 \leq i \leq n-1), \beta_n=\alpha_n+\alpha_1$. 则 $\beta_1, \ldots, \beta_n$ 线性无关的充分必要条件为: 方程组 $\lambda_1\left(\alpha_1+\alpha_2\right)+\cdots+\lambda_n\left(\alpha_n+\alpha_1\right)=\mathbf{0}$ 只有零解. 将方程组左边整理为 $\alpha_1, \ldots, \alpha_n$ 的线性组合, 得

$$\left(\lambda_1+\lambda_2\right) \alpha_1+\left(\lambda_2+\lambda_3\right) \alpha_2+\cdots+\left(\lambda_{n-1}+\lambda_n\right) \alpha_{n-1}+\left(\lambda_n+\lambda_1\right) \alpha_n=\mathbf{0}$$

由于 $\alpha_1, \ldots, \alpha_n$ 线性无关, (1) 成立的充分必要条件是:

$$\left\{\begin{array}{c}
\lambda_1+\lambda_2=0 \\
\cdots\cdots\cdots \\
\lambda_{n-1}+\lambda_n=0 \\
\lambda_n+\lambda_1=0
\end{array}\right.$$

即

$$\lambda_1=-\lambda_2=\lambda_3=\cdots=(-1)^i \lambda_{i+1}=\cdots=(-1)^{n-1} \lambda_n=(-1)^n \lambda_1$$

当 n 为偶数时, 取 $\lambda_1=1$ 得到非零解 $\lambda_i=(-1)^{i-1}$ 即 $(\lambda_1,\dots,\lambda_n)=(1,1,-1,1,\dots,-1,1)$. 这说明 $\beta_1,\dots,\beta_n$ 线性相关.

当 n 为奇数时, $\lambda_1=(-1)^n\lambda_1=-\lambda_1$ 迫使 $\lambda_1=0$, 所有的 $\lambda_i=(-1)^{i-1}\lambda_1=0$. 方程组 (2) 只有零解, $\beta_1,\dots,\beta_n$ 线性无关. 结论是: 当 n 为奇数时一定线性无关. 当 n 是偶数时一定线性相关. （2）一定线性相关. 如果 $\alpha_1,\dots,\alpha_n$ 线性相关, 则其中某个 α_j 可以写成其余 $n-1$ 个向量 $\alpha_{i_1},\dots,\alpha_{i_{n-1}}$ 的线性组合. 代入方程 (1) 左边, 整理成 $\alpha_{i_1},\dots,\alpha_{i_{n-1}}$ 的线性组合. 方程 (1) 变成

$$\mu_1\alpha_{i_1}+\cdots+\mu_{n-1}\alpha_{i_{n-1}}=\mathbf{0}$$

其中各个 μ_i 都是 $\lambda_1,\dots,\lambda_n$ 的线性组合. $\mu_1=\cdots=\mu_{n-1}=0$ 是由 n 个未知数 $\lambda_1,\dots,\lambda_n$ 的 $n-1$ 个方程组成的齐次线性方程组, 一定有非零解 $(\lambda_1,\dots,\lambda_n)$. 这样的非零解是方程组 (3) 和 (1) 的解, 这说明 $\beta_1,\dots,\beta_n$ 线性相关.

8. 证明：实数域 $\mathbb{R}$ 作为有理数域 $\mathbb{Q}$ 上的线性空间是无限维的。
9. 证明：复数域 $\mathbb{C}$ 作为有理数域 $\mathbb{Q}$ 上的线性空间是无限维的。
10. 在实数域 $\mathbb{R}$ 上的所有二阶方阵构成的线性空间中求一组基

$$\{\boldsymbol{A}_1,\boldsymbol{A}_2,\boldsymbol{A}_3,\boldsymbol{A}_4\}$$

使得

$$\boldsymbol{A}_j^2=\boldsymbol{A}_j,\quad j=1,2,3,4.$$

第 2 章 线性映射

内容提要

2.1 2024 年 4 月 9 日第六周星期二作业-商空间-线性映射

2.1.1 商空间续

1. 设 W 是线性空间 V 的子空间, W' 为其补, 即 $W = W \oplus W'$
2. V/W 是 V 模 W 的商空间.
3. 在 W' 与 V/W 之间可以自然地规定映射 η 如下: 对任意向量 $\boldsymbol{\alpha} \in W'$, 令

$$\eta(\boldsymbol{\alpha}) = \bar{\boldsymbol{\alpha}}$$

映射 η 称为自然映射.

定理 2.1

设 W 是数域 $\mathbb{F}$ 上的 n 维线性空间 V 的子空间, W' 是 W 的补. 则子空间 W' 和 V 关于 W 的商空间 V/W 同构, 并且 W' 到 V/W 的自然映射 η 是同构映射.

(I) 设 $\bar{\boldsymbol{\alpha}} \in V/W$, 则 $\boldsymbol{\alpha} \in V = W \oplus W'$, 因此 $\boldsymbol{\alpha} = \boldsymbol{\beta} + \boldsymbol{\gamma}$, 其中 $\boldsymbol{\beta} \in W, \boldsymbol{\gamma} \in W'$. 由于 $\boldsymbol{\alpha} - \boldsymbol{\gamma} = \boldsymbol{\beta} \in W$, 所以 $\boldsymbol{\alpha} \equiv \boldsymbol{\gamma}(\bmod W)$. 因此 $\bar{\boldsymbol{\gamma}} = \bar{\boldsymbol{\alpha}}$. 即对任意 $\bar{\boldsymbol{\alpha}} \in V/W$, 总存在向量 $\boldsymbol{\gamma} \in W'$, 使得 $\bar{\boldsymbol{\gamma}} = \bar{\boldsymbol{\alpha}}$. 由于 $\eta(\boldsymbol{\gamma}) = \bar{\boldsymbol{\gamma}} = \bar{\boldsymbol{\alpha}}$, 所以, 映射 η 是满射.

(II) 设 $\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2 \in W'$, 并且 $\eta(\boldsymbol{\gamma}_1) = \eta(\boldsymbol{\gamma}_2)$, 则 $\bar{\boldsymbol{\gamma}}_1 = \bar{\boldsymbol{\gamma}}_2$. 因此, $\boldsymbol{\gamma}_1 \equiv \boldsymbol{\gamma}_2(\bmod W)$. 即 $\boldsymbol{\gamma}_1 - \boldsymbol{\gamma}_2 \in W$. 又 $\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2 \in W'$, 故 $\boldsymbol{\gamma}_1 - \boldsymbol{\gamma}_2 \in W'$. 因此, $\boldsymbol{\gamma}_1 - \boldsymbol{\gamma}_2 \in W \cap W'$. 由于 W' 是 W 的补. 所以 $W \cap W' = 0$, 由此 $\boldsymbol{\gamma}_1 - \boldsymbol{\gamma}_2 = 0$, 即 $\boldsymbol{\gamma}_1 = \boldsymbol{\gamma}_2$. 这表明, 映射 η 是单射.

(III) 最后, 由于对任意 $\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2 \in W'$,

$$\eta(\boldsymbol{\gamma}_1 + \boldsymbol{\gamma}_2) = \overline{\boldsymbol{\gamma}_1 + \boldsymbol{\gamma}_2} = \bar{\boldsymbol{\gamma}}_1 + \bar{\boldsymbol{\gamma}}_2 = \eta(\boldsymbol{\gamma}_1) + \eta(\boldsymbol{\gamma}_2),$$

所以, 映射 η 是保加法的.

(IV) 由于对任意 $\lambda \in \mathbb{F}, \boldsymbol{\gamma} \in W'$,

$$\eta(\lambda\boldsymbol{\gamma}) = \bar{\lambda}\boldsymbol{\gamma} = \lambda\bar{\boldsymbol{\gamma}} = \lambda\eta(\gamma)$$

所以, 映射 η 是保乘法的.

(V) 因此, W' 到 V/W 上的映射 η 是同构映射, 从而 W' 和 V/W 同构.

推论 2.1

设 W 是线性空间 V 的子空间, 则 V 关于 W 的商空间 V/W 的维数为

$$\dim(V/W) = \dim V - \dim W.$$

2.1.2 线性映射

定义 2.1 (什么是映射? 像, 原像)

- 集合 S 到 T 的映射 η 是指集合 S 到集合 T 的一个对应规律, 使得对任意给定的元素 $\boldsymbol{\alpha} \in S$, 按照对应规律 η, 可以确定集合 T 的唯一一个元素 $\boldsymbol{\beta}$ 与 $\boldsymbol{\alpha}$ 相对应. 集合 S 到 T 的映射 η 记为
$$\eta : S \to T.$$
- 元素 $\boldsymbol{\beta}$ 称为元素 $\boldsymbol{\alpha}$ 在映射 η 下的象, 元素 $\boldsymbol{\alpha}$ 称为元素 $\boldsymbol{\beta}$ 的原象. 当 $\boldsymbol{\alpha}$ 遍历集合 S 的元素时, 所有 $\boldsymbol{\alpha}$ 的象的集合称为集合 S 在映射 η 下的象, 记为 $\operatorname{Im}\eta$, 或者 $\eta(S)$. 即
$$\operatorname{Im}\eta = \{\eta(\boldsymbol{\alpha}) \mid \boldsymbol{\alpha} \in S\}.$$
- 元素 $\boldsymbol{\beta}$ 的所有原象的集合记为 $\eta^{-1}(\boldsymbol{\beta})$, 即
$$\eta^{-1}(\boldsymbol{\beta}) = \{\boldsymbol{\alpha} \in S \mid \eta(\boldsymbol{\alpha}) = \boldsymbol{\beta}\}.$$

注意

给定映射 $\eta : S \to T$. 如果对任意 $\alpha_1, \alpha_2 \in S, \alpha_1 \neq \alpha_2$, 均有 $\eta(\alpha_1) \neq \eta(\alpha_2)$, 则 η 称为单射. 当且仅当由 $\eta(\alpha_1) = \eta(\alpha_2)$ 可以推出 $\alpha_1 = \alpha_2$ 时, η 是单射.

注意

给定映射 $\eta : S \to T$. 如果对任意 $\beta \in T$, 存在 $\alpha \in S$, 使得 $\eta(\alpha) = \beta$, 则 η 称为满射. 容易看出, 当且仅当 $\eta(S) = T$ 时, 映射 η 是满射. 如果 η 是满射, 则说映射 η 是 S 到 T 上的; 如果 η 不是满射, 则说映射 η 是 S 到 T 内的.

注意

如果映射 $\eta : S \to T$ 既是单射, 又是满射, 则 η 称为双射, 或者 S 到 T 上的一一对应.

给定集合 U, V 和 W, 以及映射 $\eta : U \to V, \xi : V \to W$, 定义 U 到 W 的合成映射 $\xi\eta$ 如下: 设 $u \in U$, 则令

$$(\xi\eta)(u) = \xi(\eta(u)).$$

映射 $\xi\eta : U \to W$ 称为映射 η 和 ξ 的乘积.

一般来说，映射的乘积一般是不可交换的, 即等式 $\xi\eta = \eta\xi$ 一般并不成立.

定理 2.2 (映射的乘法满足结合律)

设 S, T, U 和 V 是集合, 且 $\xi: S \to T, \eta: T \to U$ 与 $\zeta: U \to V$ 是映射, 则 $\zeta(\eta\xi) = (\zeta\eta)\xi$. 换句话说, 映射的乘法满足结合律.

集合 S 到自身的双射称为集合 S 的变换.

定义 2.2

集合 S 到自身的映射 ε_S 如下: 设 $s \in S$, 则令

$$\varepsilon_S(s) = s.$$

显然, ε_S 是集合 S 的变换, 它称为集合 S 的恒等变换, 或者单位变换.

定理 2.3

对任意映射 $\eta: S \to T$, 均有 $\eta\varepsilon_S = \varepsilon_T\eta = \eta$.

定义 2.3 (逆映射)

设 $\eta: S \to T$ 是映射, 如果存在 $\xi: T \to S$, 使得 $\xi\eta = \varepsilon_S, \eta\xi = \varepsilon_T$, 则映射 $\eta: S \to T$ 称为可逆映射, 映射 $\xi: T \to S$ 称为 η 的逆映射, 并记为 η^{-1}.

定理 2.4 (可逆映射是双射)

映射 $\eta: S \to T$ 可逆的充分必要条件是映射 η 为双射.

定理 2.5

设映射 $\eta: S \to T$ 可逆, 则 η 的逆映射是唯一的.

定理 2.6

设映射 $\eta: S \to T$ 和 $\xi: T \to U$ 可逆, 则它们的乘积 $\xi\eta: S \to U$ 也可逆.

回忆什么是同构?

定义 2.4 (线性同构)

设 V_1 与 V_2 是同一个数域 $\mathbb{F}$ 上的两个线性空间. 如果存在 V_1 到 V_2 上的一个双射 η, 它把 V_1 中的向量 $\boldsymbol{\alpha}$ 映为 V_2 的向量 $\eta(\boldsymbol{\alpha})$, 使得对任意 $\boldsymbol{\alpha}, \boldsymbol{\beta} \in V_1, \lambda \in \mathbb{F}$, 都有

$$\eta(\boldsymbol{\alpha} + \boldsymbol{\beta}) = \eta(\boldsymbol{\alpha}) + \eta(\boldsymbol{\beta}), \tag{2.1}$$

$$\eta(\lambda\boldsymbol{\alpha}) = \lambda\eta(\boldsymbol{\alpha}), \tag{2.2}$$

则线性空间 V_1 与 V_2 称为同构的, 而映射 η 称为 V_1 到 V_2 上的一个同构映射.

1. 满足上述定义中条件 (2.1) 的映射 η 称为保加法的;
2. 满足条件 (2.2) 的映射称为保乘法的.

因此同构映射 η 是线性空间 V_1 到 V_2 上的保加法与保乘法的一一映射。

注意

同构要求一一的并且保线性关系

如果不要求一一，只要求保线性关系，我们称为“同态”，或者“线性映射”

定义 2.5 (线性映射，线性变换)

假定 $\eta : V_1 \mapsto V_2$ 满足

$$\eta(x\boldsymbol{\alpha} + y\boldsymbol{\beta}) = x\eta(\boldsymbol{\alpha}) + y\eta(\boldsymbol{\beta}), \quad \forall x, y \in \mathbb{F}, \quad \forall \boldsymbol{\alpha}, \boldsymbol{\beta} \in V_1,$$

则称 η 是一个线性映射，或者 η 是 V_1 到 V_2 的同态映射 或者线性同态。上面的条件很容易验证是等价于保加法和保数乘的。

例题 2.1 对于 $\mathbf{R}^2$ 的每一向量 $\xi = (x_1, x_2)$，定义

$$\eta(\xi) = (x_1, x_1 - x_2, x_1 + x_2) \in \mathbf{R}^3,$$

η 是 $\mathbf{R}^2$ 到 $\mathbf{R}^3$ 的一个线性映射。

命题 2.1

设 $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{F}, \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_k \in V$, 则

$$\eta\left(\lambda_1\boldsymbol{\alpha}_1 + \lambda_2\boldsymbol{\alpha}_2 + \cdots + \lambda_k\boldsymbol{\alpha}_k\right) = \lambda_1\eta\left(\boldsymbol{\alpha}_1\right) + \lambda_2\eta\left(\boldsymbol{\alpha}_2\right) + \cdots + \lambda_k\eta\left(\boldsymbol{\alpha}_k\right).$$

这个性质可用归纳法证明，在同构那里已经讲过。

命题 2.2

$\eta(0) = 0$，换句话说，线性映射将零向量映成零向量.

例题 2.2 设 V 是数域 $\mathbb{F}$ 上的 n 维行向量空间, 且正整数 $m < n$. 定义映射 $\mathcal{A} : \mathbb{F}^n \to \mathbb{F}^m$ 如下: 设 $x = (x_1, x_2, \ldots, x_n) \in \mathbb{F}^n$, 则令

$$\mathcal{A}(x) = (x_1, x_2, \ldots, x_m).$$

容易验证, 映射 $\mathcal{A} : \mathbb{F}^n \to \mathbb{F}^m$ 是线性的, 它称为 $\mathbb{F}^n$ 到 $\mathbb{F}^m$ 上的投影变换.

例题 2.3 令 $C[a, b]$ 是定义在 $[a, b]$ 上一切连续实函数所成的 $\mathbf{R}$ 上向量空间. 对于每一 $f(x) \in C[a, b]$, 规定

$$\eta(f(x)) = \int_a^x f(t)\mathrm{d}t.$$

$\eta(f(x))$ 仍是 $[a, b]$ 上一个连续实函数. 根据积分的基本性质, η 是 $C[a, b]$ 到自身的一个线性映射.

例题 2.4 设 V 是数域 $\mathbb{F}$ 上线性空间, U 与 W 是 V 的子空间, 且 $V = U \oplus W$. 定义映射 $\mathscr{A} : V \to U$ 如下: 设 $\boldsymbol{\alpha} \in V$, 则存在唯一一对向量 $\boldsymbol{\beta}$ 与 $\boldsymbol{\gamma}, \boldsymbol{\beta} \in U, \boldsymbol{\gamma} \in W$, 使得 $\boldsymbol{\alpha} = \boldsymbol{\beta} + \boldsymbol{\gamma}$, 令 $\mathscr{A}(\boldsymbol{\alpha}) = \boldsymbol{\beta}$. 容易验证, 映射 $\mathscr{A} : V \to U$ 是线性的, 它称为线性空间 V 到子空间 U 上的投影变换. 注意投影变换 $\mathscr{A} : V \to U$ 是满射.

例题 2.5 设 V 和 U 是数域 $\mathbb{F}$ 上线性空间. 定义映射 $\mathscr{A}: V \to U$ 如下, 设 $\alpha \in V$, 则令 $\mathscr{A}(\alpha)=0$. 显然映射 $\mathscr{A}$ 是线性的. 它称为零映射.

例题 2.6 取定 $\mathbb{F}$ 的一个 n 元数列 $(a_1, a_2, \cdots, a_n)$. 对于 $\mathbb{F}^n$ 的每一向量 $\xi=(x_1, x_2, \cdots, x_n)$, 规定

$$\eta(\xi)=a_1 x_1+a_2 x_2+\cdots+a_n x_n \in \mathbb{F}.$$

容易验证, η 是 $\mathbb{F}^n$ 到 $\mathbb{F}$ 的一个线性映射. 这个线性映射也叫作 $\mathbb{F}$ 上一个 n 元线性函数或 $\mathbb{F}^n$ 上一个线性型.

例题 2.7 设 $\mathbb{F}^{m\times n}$ 是数域 $\mathbb{F}$ 上所有 $m\times n$ 阶矩阵构成的线性空间, P 和 Q 分别是给定的 m 阶和 n 阶方阵. 定义映射 $\mathscr{A}: \mathbb{F}^{m\times n} \to \mathbb{F}^{m\times n}$ 如下: 对任意 $X \in \mathbb{F}^{m\times n}$, 令 $\mathscr{A}(X)=PXQ$. 容易验证, 映射 $\mathcal{A}$ 是线性的,

例题 2.8 设 $\mathbb{F}[x]$ 是数域 $\mathbb{F}$ 上所有关于 x 的多项式构成的线性空间. 定义映射 $\mathscr{D}: \mathbb{F}[x] \to \mathbb{F}[x]$ 如下: 设 $f(x)=a_0+a_1 x+\cdots+a_n x^n \in \mathbb{F}[x]$, 则令

$$\mathscr{D}(f)(x)=a_1+2a_2 x+\cdots+n a_n x^{n-1}.$$

映射 $\mathscr{D}$ 称为微商变换. 容易验证, 映射 $\mathscr{D}$ 是线性的.

命题 2.3 (线性相关的继承性)

假定 $\mathscr{A}: U \mapsto V$ 是一个线性映射, 设向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_k \in U$ 线性相关, 则 $\mathscr{A}(\boldsymbol{\alpha}_1) \ldots, \mathcal{A}(\boldsymbol{\alpha}_k) \in V$ 也线性相关.

定理 2.7 (线性映射的存在性)

设 U 与 V 分别是数域 $\mathbb{F}$ 上 n 维和 m 维线性空间, $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n\}$ 是 U 的基. 设 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_n$ 是 V 中任意给定的 n 个向量. 则存在唯一线性映射 $\mathscr{A}: U \to V$, 使得 $\mathscr{A}(\boldsymbol{\alpha}_j)=\boldsymbol{\beta}_j, j=1,2,\ldots,n$.

定理 2.7 表明, 如果在线性空间 U 中取定一组基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n\}$, 则线性映射 $\mathscr{A}: U \to V$ 由基向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$ 在 $\mathscr{A}$ 下的象所唯决定.

练习 2.1 令 $\xi=(x_1, x_2, x_3)$ 是 $\mathbf{R}^3$ 的任意向量. 下列映射 σ 哪些是 $\mathbf{R}^3$ 到自身的线性映射?

(i) $\sigma(\xi)=\xi+\alpha, \alpha$ 是 $\mathbf{R}^3$ 的一个固定向量;

(ii) $\sigma(\xi)=(2x_1-x_2+x_3, x_2+x_3, -x_3)$;

(iii) $\sigma(\xi)=(x_1^2, x_2^2, x_3^2)$;

(iv) $\sigma(\xi)=(\cos x_1, \sin x_2, 0)$.

解 前两个是，后两个不是。基本训练，详情略。你考试的时候不能略。

练习 2.2 设 V 是数域 $\mathbb{F}$ 上一个一维向量空间. 证明 V 到自身的一个映射 σ 是线性映射的充要条件是: 对于任意 $\xi \in V$, 都有

$$\sigma(\xi)=a\xi.$$

这里 a 是 $\mathbb{F}$ 中一个定数.

解 首先因为 V 是一维的，从而存在 $\boldsymbol{\alpha} \neq 0$ 使得

$$V = \operatorname{span}\{\boldsymbol{\alpha}\} = \{k\boldsymbol{\alpha} : k \in \mathbb{F}\}.$$

必要性。如果 σ 是线性映射，因为 $\sigma(\boldsymbol{\alpha}) \in V$，则存在 $a \in \mathbb{F}$ 使得 $\sigma(\boldsymbol{\alpha}) = a\boldsymbol{\alpha}$. 于是对 $\xi = k\boldsymbol{\alpha} \in V$ 有

$$\sigma(\xi) = \sigma(k\boldsymbol{\alpha}) = k\sigma(\boldsymbol{\alpha}) = ka\boldsymbol{\alpha} = a\xi.$$

充分性。如果 $\sigma(\xi) = a\xi$，则对 $\boldsymbol{\beta} = k\boldsymbol{\alpha}, \boldsymbol{\gamma} = l\boldsymbol{\alpha} \in V$ 以及 $x, y \in \mathbb{F}$ 有

$$\sigma(x\boldsymbol{\beta} + y\boldsymbol{\gamma}) = \sigma(xk\boldsymbol{\alpha} + yl\boldsymbol{\alpha}) = \sigma((xk + yl)\boldsymbol{\alpha}) = (xk + yl)a\boldsymbol{\alpha}.$$

而

$$x\sigma(\boldsymbol{\beta}) + y\sigma(\boldsymbol{\gamma}) = x\sigma(k\boldsymbol{\alpha}) + y\sigma(l\boldsymbol{\alpha}) = xak\boldsymbol{\alpha} + yal\boldsymbol{\alpha} = (xy + yl)a\boldsymbol{\alpha}.$$

于是

$$\sigma(x\boldsymbol{\beta} + y\boldsymbol{\gamma}) = x\sigma(\boldsymbol{\beta}) + y\sigma(\boldsymbol{\gamma}).$$

从而是线性映射。

✍ **练习 2.3** 令 $M_n(\mathbb{F})$ 表示数域 $\mathbb{F}$ 上一切 n 阶矩阵所成的向量空间. 取定 $\boldsymbol{A} \in M_n(\mathbb{F})$. 对任意 $X \in M_n(\mathbb{F})$, 定义

$$\sigma(X) = \boldsymbol{A}X - X\boldsymbol{A}.$$

(i) 证明：σ 是 $M_n(\mathbb{F})$ 是自身的线性映射;

(ii) 证明：对于任意 $X, Y \in M_n(\mathbb{F})$,

$$\sigma(XY) = \sigma(X)Y + X\sigma(Y)$$

解 很简单的验证，不需要写答案。

✍ **练习 2.4** (选做) 设 $\mathbb{F}[x]$ 是数域 $\mathbb{F}$ 上所有一元多项式 $f(x)$ 的集合. 定义映射 $\mathscr{A} : \mathbb{F}[x] \mapsto \mathbb{F}[x]$ 如下: 设 $f(x) \in \mathbb{F}[x]$, 则令

$$\mathscr{A}(f(x)) = f(x+1) - f(x).$$

证明: $\mathscr{A}$ 是线性映射.

解 很简单的验证，不需要写答案。

✍ **练习 2.5** (选做) 数域 $\mathbb{F}$ 上 k 维列向量空间记为 $\mathbb{F}^k$. 定义映射 $\mathscr{A} : \mathbb{F}^n \to \mathbb{F}^m$ 如下: 设 $\boldsymbol{\alpha} \in \mathbb{F}^n$, 则令 $\mathscr{A}(\boldsymbol{\alpha}) = \boldsymbol{A}\boldsymbol{\alpha}$, 其中 $\boldsymbol{A}$ 是 $\mathbb{F}^{m\times n}$ 中固定的矩阵. 证明: $\mathscr{A}$ 是线性映射, 并且 $\mathscr{A}$ 为零映射的充分必要条件是 $\boldsymbol{A}$ 为零矩阵.

解 首先 $\boldsymbol{A}$ 为零矩阵时自然 $\mathscr{A}(\boldsymbol{\alpha})$ 是零映射。反过来，如果 $\mathscr{A}(\boldsymbol{\alpha}) = 0$ 对任意的 $\boldsymbol{\alpha} \in \mathbb{F}^n$, 那么取

$$\boldsymbol{\alpha}_k = (0, 0, \cdots, 1, \cdots, 0)^T, \quad \text{第 } k \text{ 个位置为 } 1\text{, 其余均为 } 0.$$

则 $\mathscr{A}(\boldsymbol{\alpha}_k)$ 为矩阵 $\boldsymbol{A}$ 的第 k 列的转置。从而 $\boldsymbol{A}$ 的第 k 列为零。又因为 k 任意所以 $\boldsymbol{A}$ 是零矩阵。

✍ **练习 2.6** (选做) 设 $\mathbf{C}$ 为复数域. 求映射 $\mathscr{A} : \mathbf{C} \to \mathbf{C}$, 使得当 $\mathbf{C}$ 视为实数域上线性空间时 $\mathscr{A}$ 是

线性的, 而当 $\mathbf{C}$ 视为复数域上线性空间时 $\mathscr{A}$ 不是线性的.

解 考虑取共轭的操作

$$\mathscr{A}(z)=\bar{z}, \quad z=x+iy\in\mathbf{C}.$$

看成 $\mathbb{R}^2\mapsto\mathbb{R}^2$ 的映射

$$\mathscr{A}(x,y)=(x,-y).$$

显然对 $\lambda,\mu\in\mathbb{R}$ 有

$$\begin{aligned}&\mathscr{A}(\lambda(x_1,y_1)+\mu(x_2,y_2))=\mathscr{A}((\lambda x_1+\mu x_2,\lambda y_1+\mu y_2))\\=&(\lambda x_1+\mu x_2,-\lambda y_1-\mu y_2)=\lambda(x_1,-y_1)+\mu(x_2,-y_2)=\lambda\mathscr{A}(x_1,y_1)+\mu\mathscr{A}(x_2,y_2).\end{aligned}$$

但是

$$\mathscr{A}(iz)=\mathscr{A}(-y,x)=\overline{iz}=(-y,-x), \quad i\mathscr{A}(z)=i\bar{z}=(y,x).$$

一般来说

$$(-y,-x)\neq(y,x).$$

所以

$$\mathscr{A}(iz)\neq i\mathscr{A}(z).$$

于是当 $\mathbf{C}$ 视为复数域上线性空间时 $\mathscr{A}$ 不是线性的.

2.2 2024 年 4 月 11 日星期四第十一次课-线性映射的矩阵表示

线性映射的矩阵表示

1. 设 U 和 V 分别是数域 $\mathbb{F}$ 上 n 维和 m 维线性空间,
$$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n\}, \quad \{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_m\}$$
分别是 U 和 V 的基.
2. 假定 $\mathscr{A}:U\to V$ 是一个线性映射
3. 显然, $\mathscr{A}(\boldsymbol{\alpha}_j)\in V$. 由于 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_m\}$ 是 V 的基, 所以

$$\boxed{\begin{aligned}\mathscr{A}(\boldsymbol{\alpha}_1)&=a_{11}\boldsymbol{\beta}_1+a_{12}\boldsymbol{\beta}_2+\cdots+a_{1m}\boldsymbol{\beta}_m,\\ \mathscr{A}(\boldsymbol{\alpha}_2)&=a_{21}\boldsymbol{\beta}_1+a_{22}\boldsymbol{\beta}_2+\cdots+a_{2m}\boldsymbol{\beta}_m,\\ &\cdots\cdots\cdots\cdots\\ \mathscr{A}(\boldsymbol{\alpha}_n)&=a_{n1}\boldsymbol{\beta}_1+a_{n2}\boldsymbol{\beta}_2+\cdots+a_{nm}\boldsymbol{\beta}_m.\end{aligned}}$$

上式可以写出矩阵形式如下:

$$\boxed{\begin{aligned}\mathscr{A}\left(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n\right)&=\left(\mathscr{A}\left(\boldsymbol{\alpha}_1\right),\mathscr{A}\left(\boldsymbol{\alpha}_2\right),\ldots,\mathscr{A}\left(\boldsymbol{\alpha}_m\right)\right)\\&=\left(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_n\right)\begin{pmatrix}a_{11}&a_{21}&\cdots&a_{n1}\\a_{12}&a_{22}&\cdots&a_{n2}\\\vdots&\vdots&&\vdots\\a_{11}&a_{2m}&\cdots&a_{nm}\end{pmatrix}\end{aligned}}$$

即

$$\boxed{\mathcal{A}\left(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n\right)=\left(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_n\right)A.}$$

定义 2.6 (矩阵表示)

矩阵 A 称为线性变换 $\mathcal{A}$ 在 U 的基 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n\}$ 与 V 的基 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_m\}$ 下的矩阵.

♣

设 $\boldsymbol{\alpha}\in U$ 在基 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n\}$ 下的坐标为 $x^{\mathrm{T}}=(x_1,x_2,\ldots,x_n)$, 则

$$\boxed{\boldsymbol{\alpha}=x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_n\boldsymbol{\alpha}_n=\left(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n\right)x.}$$

因此

$$\boxed{\begin{aligned}\mathscr{A}(\boldsymbol{\alpha})&=\mathscr{A}\left(x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_n\boldsymbol{\alpha}_n\right)\\&=x_1\mathscr{A}\left(\boldsymbol{\alpha}_1\right)+x_2\mathscr{A}\left(\boldsymbol{\alpha}_2\right)+\cdots+x_n\mathscr{A}\left(\boldsymbol{\alpha}_n\right)\\&=\left(\mathscr{A}\left(\boldsymbol{\alpha}_1\right),\mathscr{A}\left(\boldsymbol{\alpha}_2\right),\ldots,\mathscr{A}\left(\boldsymbol{\alpha}_n\right)\right)x\\&=\mathcal{A}\left(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n\right)x.\end{aligned}}$$

又因为 $\mathcal{A}(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n)=\left(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_m\right)\boldsymbol{A}$. 所以

$$\boxed{\mathscr{A}(\boldsymbol{\alpha})=\left(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_m\right)\boldsymbol{A}x}$$

这表明, 向量 $\boldsymbol{\alpha}$ 在 $\mathcal{A}$ 下的像 $\mathcal{A}(\boldsymbol{\alpha})$ 在 V 的基 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_m\}$ 下的坐标 y 为

$$y=\boldsymbol{A}x.$$

1. 所有 U 到 V 的线性映射集合记为 $\boxed{\mathscr{L}(U,V)=\mathscr{L}_{m\times n}}$.
2. 数域 $\mathbb{F}$ 上所有 $m\times n$ 矩阵的集合记为 $\boxed{\mathbb{F}^{m\times n}}$.
3. 设 $\mathscr{A}\in\mathscr{L}_{m\times n}$ 在 U 的基 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n\}$ 与 V 的基 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_m\}$ 下的矩阵为 $\boldsymbol{A}$.
4. 定义 $\mathscr{L}_{m\times n}$ 到 $\mathbb{F}^{m\times n}$ 的映射 η 如下：对任意 $\mathscr{A}\in\mathscr{L}_{m\times n}$, 令
$$\boxed{\eta(\mathscr{A})=\boldsymbol{A}.}$$
5. 显然, 对任意 $\boxed{\mathscr{A}\neq\mathscr{B}\in\mathscr{L}_{mn}}$, $\mathscr{A}$ 和 $\mathscr{B}$ 在 U 的基 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n\}$ 与 V 的基 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_m\}$ 下的矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 满足 $\boldsymbol{A}\neq\boldsymbol{B}$. 因此
$$\boxed{\eta(\mathscr{A})\neq\eta(\mathscr{B})}.$$

即映射 η 是单射.

6. 另一方面, 设 $\boldsymbol{A} \in \mathbb{F}^{m\times n}$. 记

$$\boxed{(\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \ldots, \boldsymbol{\gamma}_n) = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_m)\, \boldsymbol{A}}.$$

显然, $\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \ldots, \boldsymbol{\gamma}_n \in V$. 由定理 2.7, 存在 $\mathscr{A} \in \mathscr{L}_{m\times n}$, 使得 $\mathscr{A}(\boldsymbol{\alpha}_j) = \boldsymbol{\gamma}_j$. 因此

$$\boxed{\mathscr{A}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n) = (\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \ldots, \boldsymbol{\gamma}_n) = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_m)\, \boldsymbol{A}.}$$

7. 这表明 $\mathscr{A}$ 在 U 的基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n\}$ 与 V 的基 $\{\boldsymbol{\beta}_1, \beta_2, \ldots, \boldsymbol{\beta}_m\}$ 下的矩阵为 $\boldsymbol{A}$. 所以 $\boxed{\eta(\mathscr{A}) = \boldsymbol{A}}$, 即映射 η 是满射.

8. 于是 η 是双射. 这表明, 数域 $\mathbb{F}$ 上一个 $m\times n$ 矩阵可以表示数域 $\mathbb{F}$ 上 n 维线性空间 U 到数域 $\mathbb{F}$ 上 m 维线性空间 V 的一个线性映射. 这就是给出了矩阵的一种几何意义.

设 $\mathscr{A}: U \to V$ 是线性映射。

(I) 在 U 的基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n\}$ 与 V 的基 $\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_m\}$ 下的矩阵表示为 $\boldsymbol{A}$.

(II) 在 U 的基 $\{\tilde{\boldsymbol{\alpha}}_1, \tilde{\boldsymbol{\alpha}}_2, \ldots, \tilde{\boldsymbol{\alpha}}_n\}$ 与 V 的基 $\left\{\tilde{\boldsymbol{\beta}}_1, \tilde{\boldsymbol{\beta}}_2, \ldots, \tilde{\boldsymbol{\beta}}_m\right\}$ 下的矩阵表示为 $\boldsymbol{B}$,

(III) 也就是说

$$\boxed{\begin{aligned} &\mathscr{A}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n) = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_m)\, \boldsymbol{A}, \\ &\mathscr{A}(\tilde{\boldsymbol{\alpha}}_1, \tilde{\boldsymbol{\alpha}}_2, \ldots, \tilde{\boldsymbol{\alpha}}_n) = \left(\tilde{\boldsymbol{\beta}}_1, \tilde{\boldsymbol{\beta}}_2, \cdots, \tilde{\boldsymbol{\beta}}_m\right) \boldsymbol{B}. \end{aligned}} \tag{2.3}$$

问题

矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 有什么联系?

- 设由基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n\}$ 到基 $\{\tilde{\boldsymbol{\alpha}}_1, \tilde{\boldsymbol{\alpha}}_2, \ldots, \tilde{\boldsymbol{\alpha}}_n\}$ 的过渡矩阵为 $\boldsymbol{Q}$, 也即

$$\boxed{(\tilde{\boldsymbol{\alpha}}_1, \tilde{\boldsymbol{\alpha}}_2, \ldots, \tilde{\boldsymbol{\alpha}}_n) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n)\, \mathbf{Q}} \tag{2.4}$$

- 由基 $\left\{\tilde{\boldsymbol{\beta}}_1, \tilde{\boldsymbol{\beta}}_2, \ldots, \tilde{\boldsymbol{\beta}}_m\right\}$ 到 $\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_m\}$ 的过渡矩阵为 $\boldsymbol{P}$, 即

$$\boxed{(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_m) = \left(\tilde{\boldsymbol{\beta}}_1, \tilde{\boldsymbol{\beta}}_2, \ldots, \tilde{\boldsymbol{\beta}}_m\right) \mathbf{P}.} \tag{2.5}$$

- 由于是基之间的变换, $\mathbf{Q}$ 与 $\mathbf{P}$ 分别是 n 阶与 m 阶可逆方阵.
- 记 $\mathbf{Q} = (q_{ij})$ 的第 j 列为 $\mathbf{q}_j = (q_{1j}, q_{2j}, \ldots, q_{nj})^{\mathrm{T}}$. 由式 (2.4),

$$\boxed{\tilde{\boldsymbol{\alpha}}_j = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n)\, \mathbf{q}_j.}$$

- 由于映射 $\mathscr{A}$ 是线性的, 所以

$$\boxed{\mathscr{A}(\tilde{\boldsymbol{\alpha}}_j) = (\mathscr{A}(\boldsymbol{\alpha}_1), \ldots, \mathscr{A}(\boldsymbol{\alpha}_n))\, \mathbf{q}_j.}$$

因此,

$$\boxed{\mathscr{A}(\tilde{\boldsymbol{\alpha}}_1, \tilde{\boldsymbol{\alpha}}_2, \ldots, \tilde{\boldsymbol{\alpha}}_n) = (\mathscr{A}(\boldsymbol{\alpha}_1), \mathscr{A}(\boldsymbol{\alpha}_2), \ldots, \mathscr{A}(\boldsymbol{\alpha}_n))\, \mathbf{Q}.}$$

把式 (2.3) 代入上式得到,

$$\boxed{\mathscr{A}(\tilde{\boldsymbol{\alpha}}_1, \tilde{\boldsymbol{\alpha}}_2, \ldots, \tilde{\boldsymbol{\alpha}}_n) = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_m)\, \boldsymbol{A}\mathbf{Q}.}$$

- 再把式 (2.5) 代人上式得到,

$$\boxed{\mathscr{A}\left(\widetilde{\boldsymbol{\alpha}}_1,\widetilde{\boldsymbol{\alpha}}_2,\ldots,\widetilde{\boldsymbol{\alpha}}_n\right)=\left(\widetilde{\boldsymbol{\beta}}_1,\widetilde{\boldsymbol{\beta}}_2,\ldots,\widetilde{\boldsymbol{\beta}}_m\right)\mathbf{P}\boldsymbol{A}\mathbf{Q}.}$$

- 由于线性映射 $\mathscr{A}$ 由基向量 $\widetilde{\boldsymbol{\alpha}}_1,\widetilde{\boldsymbol{\alpha}}_2,\ldots,\widetilde{\boldsymbol{\alpha}}_n$ 在 $\mathscr{A}$ 下的象所唯一确定, 因此 $\boxed{\boldsymbol{B}=\mathbf{P}\boldsymbol{A}\mathbf{Q}}$. 即线性映射 $\mathscr{A}$ 的矩阵 A 与 B 是相抵的[1]。

线性映射在不同基下矩阵表示之间的关系

- 反之, 设矩阵 A 与 B 相抵, 即设 $\boldsymbol{B}=\mathbf{P}\boldsymbol{A}\mathbf{Q}$, 其中 $\mathbf{P}$ 和 $\mathbf{Q}$ 分别是 m 阶与 n 阶可逆方阵.
- 在 U 与 V 中分别取基 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n\}$ 与基 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_m\}$.
- 由式 (2.4) 和式 (2.5) 分别确定 U 的基 $\{\widetilde{\boldsymbol{\alpha}}_1,\widetilde{\boldsymbol{\alpha}}_2,\ldots,\widetilde{\boldsymbol{\alpha}}_n\}$ 和 V 的基 $\left\{\widetilde{\boldsymbol{\beta}}_1,\widetilde{\boldsymbol{\beta}}_2,\ldots,\widetilde{\boldsymbol{\beta}}_m\right\}$.
- 由

$$\boxed{\begin{aligned}\mathscr{A}\left(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n\right)&=\left(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_m\right)\boldsymbol{A},\\ \mathscr{B}\left(\widetilde{\boldsymbol{\alpha}}_1,\widetilde{\boldsymbol{\alpha}}_2,\ldots,\widetilde{\boldsymbol{\alpha}}_n\right)&=\left(\widetilde{\boldsymbol{\beta}}_1,\widetilde{\boldsymbol{\beta}}_2,\ldots,\widetilde{\boldsymbol{\beta}}_m\right)\boldsymbol{B}\end{aligned}}$$

  可以定义 U 到 V 的线性映射 $\mathscr{A}$ 与 $\mathscr{B}$.
- 因为 $\boldsymbol{B}=\mathbf{P}\boldsymbol{A}\mathbf{Q}$, 故

$$\boxed{\mathscr{B}\left(\widetilde{\boldsymbol{\alpha}}_1,\widetilde{\boldsymbol{\alpha}}_2,\ldots,\widetilde{\boldsymbol{\alpha}}_n\right)=\left(\widetilde{\boldsymbol{\beta}}_1,\widetilde{\boldsymbol{\beta}}_2,\ldots,\widetilde{\boldsymbol{\beta}}_m\right)\mathbf{P}\boldsymbol{A}\mathbf{Q}.}$$

- 根据 (2.5) 可得

$$\boxed{\mathscr{B}\left(\widetilde{\boldsymbol{\alpha}}_1,\widetilde{\boldsymbol{\alpha}}_2,\ldots,\widetilde{\boldsymbol{\alpha}}_n\right)=\left(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_m\right)\boldsymbol{A}\mathbf{Q}.}$$

  即

$$\boxed{\mathscr{B}\left(\left(\widetilde{\boldsymbol{\alpha}}_1,\widetilde{\boldsymbol{\alpha}}_2,\ldots,\widetilde{\boldsymbol{\alpha}}_n\right)\mathbf{Q}^{-1}\right)=\left(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_m\right)\boldsymbol{A}.}$$

  由式 (2.4) 可得到

$$\boxed{\mathscr{B}\left(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n\right)=\left(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_m\right)\boldsymbol{A}.}$$

- 由定理 2.7 可得, $\mathscr{B}=\mathscr{A}$.

这表明

相抵的矩阵是同一个线性映射在不同基下的矩阵. 这就给出了矩阵相抵的几何意义.

定理 2.8 (存在一组基使得矩阵表示很简单)

设线性映射 $\mathscr{A}:U\to V$ 在 U 的基 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_n\}$ 与 V 的基 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2\ldots,\boldsymbol{\beta}_m\}$ 下的矩阵为 $\boldsymbol{A}$, 且 $\operatorname{rank}\boldsymbol{A}=r$. 则在 U 和 V 中分别存在基 $\{\widetilde{\boldsymbol{\alpha}}_1,\widetilde{\boldsymbol{\alpha}}_2,\ldots,\widetilde{\boldsymbol{\alpha}}_m\}$ 与 $\left\{\widetilde{\boldsymbol{\beta}}_1,\widetilde{\boldsymbol{\beta}}_2,\ldots,\widetilde{\boldsymbol{\beta}}_m\right\}$, 使得

$$\mathscr{A}\left(\widetilde{\boldsymbol{\alpha}}_1,\widetilde{\boldsymbol{\alpha}}_2,\ldots,\widetilde{\boldsymbol{\alpha}}_n\right)=\left(\widetilde{\boldsymbol{\beta}}_1,\widetilde{\boldsymbol{\beta}}_2,\ldots,\widetilde{\boldsymbol{\beta}}_m\right)\begin{pmatrix}I_r&0\\0&0\end{pmatrix}.$$

♡

[1]回忆矩阵的相抵：设矩阵 $\boldsymbol{A},\boldsymbol{B}\in\mathbb{F}^{m\times n}$, 如果矩阵 $\boldsymbol{A}$ 可以通过有限次初等变换变为矩阵 $\boldsymbol{B}$, 则称矩阵 $\boldsymbol{A}$ 和矩阵 $\boldsymbol{B}$ 是相抵的。这等价于存在可逆矩阵 $\mathbf{P}\in\mathbb{F}^{m\times m}$ 和可逆矩阵 $\mathbf{Q}\in\mathbb{F}^{n\times n}$ 使得 $\boldsymbol{B}=\mathbf{P}\boldsymbol{A}\mathbf{Q}$。

定义 2.7 (线性映射的和)

设 U 与 V 分别是数域 $\mathbb{F}$ 上 n 维与 m 维线性空间, $\mathscr{A}: U \to V$ 与 $\mathscr{B}: U \to V$ 是线性映射. 定义映射 $\mathscr{A}$ 与 $\mathscr{B}$ 的和 $\mathscr{A}+\mathscr{B}$ 如下, 设 $\boldsymbol{\alpha} \in U$, 则令

$$(\mathscr{A}+\mathscr{B})(\boldsymbol{\alpha}) = \mathscr{A}(\boldsymbol{\alpha}) + \mathscr{B}(\boldsymbol{\alpha}).$$

♣

定义 2.8 (数和线性映射的乘积)

设 U 与 V 分别是数域 $\mathbb{F}$ 上 n 维与 m 维线性空间, $\mathscr{A}: U \to V$ 是线性映射. 定义纯量 $\lambda \in \mathbb{F}$ 与映射 $\mathscr{A}$ 的乘积 $\lambda\mathscr{A}$ 如下: 设 $\boldsymbol{\alpha} \in U$, 则令

$$(\lambda\mathscr{A})(\boldsymbol{\alpha}) = \lambda\mathscr{A}(\boldsymbol{\alpha}).$$

♣

引理 2.1 (和与数乘还是线性映射)

我们有

1. 映射 $\mathscr{A}$ 与 $\mathscr{B}$ 的和 $\mathscr{A}+\mathscr{B}$ 是 U 到 V 的线性映射.
2. 纯量 λ 与映射 $\mathscr{A}$ 的乘积 $\lambda\mathscr{A}$ 也是 U 到 V 的线性映射.

♡

定义 2.9 (线性映射的复合 (也叫乘积))

设 $\mathscr{A}: U \to V$ 与 $\mathscr{B}: V \to W$ 是线性映射。作为映射, 可以定义映射 $\mathscr{A}$ 与 $\mathscr{B}$ 的复合（或者乘积）$\mathscr{B}\mathscr{A}: U \to W$, 即对任意 $\boldsymbol{\alpha} \in U$,

$$(\mathscr{B}\mathscr{A})(\boldsymbol{\alpha}) = \mathscr{B}(\mathscr{A}(\boldsymbol{\alpha})).$$

♣

设 $\lambda \in \mathbb{F}, \boldsymbol{\alpha}, \boldsymbol{\beta} \in U$, 则

$$(\mathscr{B}\mathscr{A})(\boldsymbol{\alpha}+\boldsymbol{\beta}) = \mathscr{B}(\mathscr{A}(\boldsymbol{\alpha}+\boldsymbol{\beta})),$$
$$(\mathscr{B}\mathscr{A})(\lambda\boldsymbol{\alpha}) = \mathscr{B}(\mathscr{A}(\lambda\boldsymbol{\alpha})).$$

由于映射 $\mathscr{A}$ 和 $\mathscr{B}$ 是线性的, 所以

$$(\mathscr{B})(\boldsymbol{\alpha}+\boldsymbol{\beta}) = \mathscr{B}(\mathscr{A}(\boldsymbol{\alpha})) + \mathscr{B}(\mathscr{A}(\boldsymbol{\beta})) = (\mathscr{B}\mathscr{A})(\boldsymbol{\alpha}) + (\mathscr{B}\mathscr{A})(\boldsymbol{\beta}),$$
$$(\mathscr{B}A)(\lambda\boldsymbol{\alpha}) = \lambda\mathscr{B}(\mathscr{A}(\boldsymbol{\alpha})) = \lambda(\mathscr{B}\mathscr{A})(\boldsymbol{\alpha}).$$

所以 $\mathscr{B}\mathscr{A}: U \to W$ 是线性映射.

注意

1. 与映射的乘积相同, 线性映射的乘积满足结合律, 但交换律并不成立. 即一般地说, 对线性映射 $\mathscr{A} \to \mathscr{B}$, 等式 $\mathscr{B}\mathscr{A} = \mathscr{A}\mathscr{B}$ 不成立.
2. 对零映射 $\mathscr{O}: U \to V$ 与线性映射 $\mathscr{A}: V \to W$, 恒有 $\mathscr{A}\mathscr{O} = \mathscr{O}$, 其中右端的 $\mathscr{O}$ 是 U 到 W 的零映射。而对线性映射 $\mathscr{A}: W \to U$, 恒有 $\mathscr{O}\mathscr{A} = \mathscr{O}$, 其中右端的 $\mathscr{O}$ 是 W 到 V 的零映射.
3. 作为集合，线性空间 U 到自身的单位映射记为 $\mathscr{E}_U$. 容易验证, $\mathscr{E}_U: U \to U$ 是线性映射. 并且对任意线性映射 $\mathscr{A}: U \to V$, 恒有 $\mathscr{E}_V\mathscr{A} = \mathscr{A} = \mathscr{A}\mathscr{E}_U$.

引理 2.2 (线性映射的运算的矩阵表示)

假定线性映射 $\mathscr{A}: U \to V$ 在 U 的基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 与 V 的基 $\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_m\}$ 下的矩阵为 $\boldsymbol{A}$, 线性映射 $\mathscr{B}: V \to W$ 在 V 的基 $\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_m\}$ 与 W 的基 $\{\gamma_1, \gamma_2, \cdots, \gamma_1\}$ 下的矩阵为 $\boldsymbol{B}$,

(1) 设线性映射 $\mathscr{A}: U \to V$ 与 $\mathscr{B}: U \to V$ 在 U 的基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 下与 V 的基 $\{\boldsymbol{\beta}_1\ ,\ \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_m\}$ 下的矩阵分别为 $\boldsymbol{A}$ 与 $\boldsymbol{B}$, 则线性映射 $\mathscr{A}+\mathscr{B}$ 在这两组基下的矩阵为 $\boldsymbol{A}+\boldsymbol{B}$.

(2) 对于 $\lambda \in \mathbb{F}$, 线性映射 $\lambda\mathscr{A}$ 在这两组基下的矩阵为 $\lambda\boldsymbol{A}$.

(3) 线性映射 $\mathscr{A}$ 与 $\mathscr{B}$ 的乘积（复合）$\mathscr{B}\mathscr{A}$ 在这两组基下的矩阵表示等于线性映射 $\mathscr{B}$ 的矩阵表示 $\boldsymbol{B}$ 与线性映射 $\mathscr{A}$ 的矩阵表示 $\boldsymbol{A}$ 的乘积.

♡

定理 2.9

设 U 与 V 分别是数域 $\mathbb{F}$ 上 n 维与 m 维线性空间。所有 U 到 V 的线性映射集合记为

$$\mathscr{L}(U,V) = \mathscr{L}_{m\times n} := \{\mathscr{A}: U \mapsto V \ \text{ 是线性映射}\}$$

则有

1. 集合 $\mathscr{L}_{m\times n}$ 在上述的加法以及纯量与映射的乘法下构成数域 $\mathbb{F}$ 上的线性空间。
2. $\mathscr{L}_{m\times n}$ 的维数为

$$\dim \mathscr{L}_{m\times n} = mn.$$

♡

练习 2.7 假定线性映射 $\mathscr{A}: \mathbb{R}^2 \mapsto \mathbb{R}$ 定义如下

$$\mathscr{A}(x,y) = x.$$

(1) 求 $\mathscr{A}$ 在 $\mathbb{R}^2$ 的基

$$\boldsymbol{\alpha}_1 = (1,2), \quad \boldsymbol{\alpha}_2 = (3,4)$$

和 $\mathbb{R}$ 的基 $\boldsymbol{\beta}_1 = 1$ 下的矩阵表示 $\boldsymbol{A}$。

(2) 求 $\mathscr{A}$ 在 $\mathbb{R}^2$ 的基

$$\tilde{\boldsymbol{\alpha}}_1 = (1,0), \quad \tilde{\boldsymbol{\alpha}}_2 = (0,2)$$

和 $\mathbb{R}$ 的基 $\tilde{\boldsymbol{\beta}}_1 = 10$ 下的矩阵表示 $\boldsymbol{B}$。

(3) 求出两组基

$$\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2\}$$

和

$$\{\tilde{\boldsymbol{\alpha}}_1, \tilde{\boldsymbol{\alpha}}_2\}$$

之间的过渡矩阵。也就是说你找一个 2×2 的矩阵 $\mathbf{P}$ 使得

$$(\tilde{\boldsymbol{\alpha}}_1, \tilde{\boldsymbol{\alpha}}_2) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2)\mathbf{P}.$$

(4) 求出两组基

$$\{\boldsymbol{\beta}_1\}$$

和

$$\{\tilde{\boldsymbol{\beta}}_1\}$$

之间的过渡矩阵。也就是说你找一个 1×1 的矩阵 $\mathbf{Q}$ 使得

$$(\tilde{\boldsymbol{\beta}}_1)=(\boldsymbol{\beta})\mathbf{Q}.$$

这事实上就是数乘。这里为了你去理解我们讲的东西。

(5) 从上面几步你就得出 $\boldsymbol{A},\boldsymbol{B},\mathbf{P},\mathbf{Q}$ 之间的关系。然后你就验证了课堂上我们讲的 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 相抵。

这道题（练习 2.7）你一定要认真做一遍，掌握了这道题，你就理解了课堂讲的东西。

解 基本训练，答案略。

练习 2.8 (选做) 在所有 2 阶实方阵构成的线性空间 $\mathbf{R}^{2\times 2}$ 中, 取

$$\boldsymbol{A}=\begin{pmatrix} a & b\\ c & d\end{pmatrix}.$$

定义映射 $\mathscr{A}:\mathbf{R}^{2\times 2}\to\mathbb{R}^{2\times 2}$ 与 $\mathscr{B}:\mathbb{R}^{2\times 2}\to\mathbb{R}^{2\times 2}$ 如下: 设 $\boldsymbol{X}\in\mathbb{R}^{2\times 2}$, 则令

$$\mathscr{A}(\boldsymbol{X})=\boldsymbol{A}\boldsymbol{X},\mathscr{B}(\boldsymbol{X})=\boldsymbol{X}\boldsymbol{A}.$$

证明: 映射 $\mathscr{A}$ 与 $\mathscr{B}$ 都是线性的, 并求出它们在基

$$\left\{\boldsymbol{\alpha}_1=\begin{pmatrix}1&0\\0&0\end{pmatrix},\boldsymbol{\alpha}_2=\begin{pmatrix}0&0\\1&0\end{pmatrix},\boldsymbol{\alpha}_3=\begin{pmatrix}0&1\\0&0\end{pmatrix},\boldsymbol{\alpha}_4=\begin{pmatrix}0&0\\0&1\end{pmatrix}\right\}$$

下的矩阵表示.

解 基本训练，答案略。

练习 2.9 (选做) 定义映射 $\operatorname{tr}:\mathbb{F}^{n\times n}\to\mathbb{F}$ 如下: 设 $\boldsymbol{A}=(\boldsymbol{a}_{ij})\in\mathbb{F}^{n\times n}$, 则令

$$\operatorname{tr}(\boldsymbol{A})=a_{11}+a_{22}+\cdots+a_{nn}.$$

证明: 映射 tr 是线性映射, 并且对任意 $\boldsymbol{A},\boldsymbol{B}\in\mathbb{F}^{n\times n}$, $\operatorname{tr}(\boldsymbol{A}\boldsymbol{B})=\operatorname{tr}(\boldsymbol{B}\boldsymbol{A})$.

解 任取矩阵 $\boldsymbol{A}=(\boldsymbol{a}_{ij}),\boldsymbol{B}=(\boldsymbol{b}_{ij})\in\mathbb{F}^{n\times n}$ 以及 $x,y\in\mathbb{F}$，则有

$$\operatorname{tr}(x\boldsymbol{A}+y\boldsymbol{B})=\sum_{i=1}^{n}(xa_{ii}+yb_{ii})=x\sum_{i=1}^{n}a_{ii}+y\sum_{i=1}^{n}b_{ii}=x\operatorname{tr}(\boldsymbol{A})+y\operatorname{tr}(\boldsymbol{B}),$$

因此 tr 是线性映射。

回忆我们上学期学的矩阵乘法，有

$$\operatorname{tr}(\boldsymbol{A}\boldsymbol{B})=\sum_{i=1}^{n}(\boldsymbol{A}\boldsymbol{B})(i,i)=\sum_{i=1}^{n}\sum_{k=1}^{n}a_{ik}b_{ki}=\sum_{k=1}^{n}\sum_{i=1}^{n}b_{ki}a_{ik}=\sum_{k=1}^{n}(\boldsymbol{B}\boldsymbol{A})(k,k)=\operatorname{tr}(\boldsymbol{B}\boldsymbol{A}).$$

这便完成了证明。

练习 2.10 (选做) 设 $\mathscr{A}:\mathbb{F}^{n\times n}\to\mathbb{F}$ 是线性映射, 并且对任意 $\boldsymbol{A},\boldsymbol{B}\in\mathbb{F}^{n\times n}$, $\mathscr{A}(\boldsymbol{A}\boldsymbol{B})=\mathscr{A}(\boldsymbol{B}\boldsymbol{A})$. 证明:

$$\mathscr{A}=\lambda\operatorname{tr},$$

其中 $\lambda\in\mathbb{F}$.

解 我们只需要看 $\mathscr{A}$ 在 $\mathbb{F}^{n\times n}$ 的基上的作用，它的一组基是

$$\{E_{ij}:1\le i,j\le n\}$$

首先对 $1\le i,j,k,l\le n$, 回忆矩阵的乘法，我们有

$$E_{ij}E_{kl}=\begin{cases}E_{il} & j=k\\ 0 & j\neq k.\end{cases}$$

于是对任意的 $1\le i,j,k\le n$ 有

$$\mathscr{A}(E_{ij})=\mathscr{A}(E_{ik}E_{kj})=\mathscr{A}(E_{kj}E_{ik})=\begin{cases}\mathscr{A}(E_{kk}) & i=j\\ \mathscr{A}(0)=0 & i\neq j\end{cases}$$

从而我们得到存在 $\lambda\in\mathbb{F}$ 使得

$$\mathscr{A}(E_{11})=\mathscr{A}(E_{22})=\cdots=\mathscr{A}(E_{nn})=\lambda$$

以及

$$\mathscr{A}(E_{ij}),\quad \forall i\neq j.$$

注意到

$$\lambda\,\mathrm{tr}(E_{ii})=\lambda,\quad i=1,2,\cdots,n$$

以及

$$\lambda\,\mathrm{tr}(E_{ij})=0,\quad \forall i\neq j.$$

从而 $\mathscr{A}$ 与 $\lambda\,\mathrm{tr}$ 在基上的取值一样。于是一定有 $\mathscr{A}=\lambda\,\mathrm{tr}$.

✍ **练习 2.11** 令 $\mathbb{F}_n[x]$ 表示一切次数不大于 n 的多项式连同零多项式所成的空间,

$$\sigma:\mathbb{F}_n[x]\mapsto\mathbb{F}_n[x]:\qquad f(x)\longmapsto f'(x).$$

求 σ 关于以下两个基的矩阵:

(1)

$$\{\boldsymbol{\alpha}_1=1,\boldsymbol{\alpha}_2=x,\boldsymbol{\alpha}_3=x^2,\cdots,\boldsymbol{\alpha}_{n+1}=x^n\}$$

(2)

$$\left\{\boldsymbol{\beta}_1=1,\boldsymbol{\beta}_2=x-c,\boldsymbol{\beta}_3=\frac{(x-c)^2}{2!},\cdots,\boldsymbol{\beta}_{n+1}=\frac{(x-c)^n}{n!}\right\},c\in\mathbb{F}.$$

注意这里的意思是求矩阵 $\boldsymbol{A}$ 和矩阵 $\boldsymbol{B}$ 使得

$$\mathscr{A}\left(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_{n+1}\right)=\left(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_{n+1}\right)\boldsymbol{A},$$
$$\mathscr{A}\left(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_{n+1}\right)=\left(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\beta}_{n+1}\right)\boldsymbol{B}$$

这里

$$\mathscr{A}\left(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_{n+1}\right)$$

的意思就是

$$\left(\mathscr{A}\boldsymbol{\alpha}_1,\mathscr{A}\boldsymbol{\alpha}_2,\ldots,\mathscr{A}\boldsymbol{\alpha}_{n+1}\right)$$

解 基本训练，答案略。

✍ **练习 2.12** 设 $\mathbb{F}$ 上三维向量空间的线性变换 $\mathscr{A}$ 关于基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3\}$ 的矩阵是

$$\boldsymbol{A} = \begin{pmatrix} 15 & -11 & 5 \\ 20 & -15 & 8 \\ 8 & -7 & 6 \end{pmatrix}$$

换句话说有

$$(\mathscr{A}\boldsymbol{\alpha}_1, \mathscr{A}\boldsymbol{\alpha}_2, \mathscr{A}\boldsymbol{\alpha}_3) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3)\,\boldsymbol{A}$$

(1) 求 $\mathscr{A}$ 关于基

$$\begin{aligned} \boldsymbol{\beta}_1 &= 2\boldsymbol{\alpha}_1 + 3\boldsymbol{\alpha}_2 + \boldsymbol{\alpha}_3, \\ \boldsymbol{\beta}_2 &= 3\boldsymbol{\alpha}_1 + 4\boldsymbol{\alpha}_2 + \boldsymbol{\alpha}_3, \\ \boldsymbol{\beta}_3 &= \boldsymbol{\alpha}_1 + 2\boldsymbol{\alpha}_2 + 2\boldsymbol{\alpha}_3 \end{aligned}$$

的矩阵. 也即是说你要求矩阵 $\boldsymbol{B}$ 使得

$$(\mathscr{A}\boldsymbol{\beta}_1, \mathscr{A}\boldsymbol{\beta}_2, \mathscr{A}\boldsymbol{\beta}_3) = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3)\,\boldsymbol{B}$$

(2) 若 $\xi = 2\boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2 - \boldsymbol{\alpha}_3$. 求 $\mathscr{A}(\xi)$ 关于基 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3$ 的坐标.

解 基本训练，答案略。

✍ **练习 2.13** 设 $\{\gamma_1, \gamma_2, \cdots, \gamma_n\}$ 是 n 维向量空间 V 的一个基.

$$\boldsymbol{\alpha}_j = \sum_{i=1}^n a_{ij}\gamma_i, \boldsymbol{\beta}_j = \sum_{i=1}^n b_{ij}\gamma_i, j = 1, 2, \cdots, n,$$

并且 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 线性无关. 又设 σ 是 V 的一个线性变换, 使得 $\sigma(\boldsymbol{\alpha}_j) = \boldsymbol{\beta}_j, j = 1, 2, \cdots, n$. 求 σ 关于基 $\gamma_1, \gamma_2, \cdots, \gamma_n$ 的矩阵. 也即说要求矩阵 $\boldsymbol{C}$ 使得

$$(\sigma(\gamma_1), \sigma(\gamma_2), \cdots, \sigma(\gamma_n)) = (\gamma_1, \gamma_2, \cdots, \gamma_n)\boldsymbol{C}.$$

解 令矩阵 $\boldsymbol{A} = (a_{ij})_{1\le i,j\le n}, \boldsymbol{B} = (b_{ij})_{1\le i,j\le n}$, 则有

$$(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n) = (\gamma_1, \gamma_2, \cdots, \gamma_n)\boldsymbol{A}.$$

和

$$(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n) = (\gamma_1, \gamma_2, \cdots, \gamma_n)\boldsymbol{B}.$$

因为 $\{\gamma_1, \gamma_2, \cdots, \gamma_n\}$ 是基从而线性无关，而 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 也线性无关，则 $\boldsymbol{A}$ 是可逆矩阵。于是

$$(\gamma_1, \gamma_2, \cdots, \gamma_n) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n)\boldsymbol{A}^{-1}.$$

于是

$$\begin{aligned} (\sigma(\gamma_1), \sigma(\gamma_2), \cdots, \sigma(\gamma_n)) &= (\sigma(\boldsymbol{\alpha}_1), \sigma(\boldsymbol{\alpha}_2), \cdots, \sigma(\boldsymbol{\alpha}_n))\,\boldsymbol{A}^{-1} \\ &= (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n)\boldsymbol{A}^{-1} \\ &= (\gamma_1, \gamma_2, \cdots, \gamma_n)\boldsymbol{B}\boldsymbol{A}^{-1} \end{aligned}$$

于是 σ 关于基 $\gamma_1, \gamma_2, \cdots, \gamma_n$ 的矩阵是 $\boldsymbol{B}\boldsymbol{A}^{-1}$.

2.3 2024 年 4 月 16 日第七周星期二第十二次课-线性映射的运算

定义 2.10 (线性映射的和)

设 U 与 V 分别是数域 $\mathbb{F}$ 上 n 维与 m 维线性空间, $\mathscr{A}:U\to V$ 与 $\mathscr{B}:U\to V$ 是线性映射. 定义映射 $\mathscr{A}$ 与 $\mathscr{B}$ 的和 $\mathscr{A}+\mathscr{B}$ 如下, 设 $\boldsymbol{\alpha}\in U$, 则令

$$(\mathscr{A}+\mathscr{B})(\boldsymbol{\alpha})=\mathscr{A}(\boldsymbol{\alpha})+\mathscr{B}(\boldsymbol{\alpha}).$$

定义 2.11 (数和线性映射的乘积)

设 U 与 V 分别是数域 $\mathbb{F}$ 上 n 维与 m 维线性空间, $\mathscr{A}:U\to V$ 是线性映射. 定义纯量 $\lambda\in\mathbb{F}$ 与映射 $\mathscr{A}$ 的乘积 $\lambda\mathscr{A}$ 如下: 设 $\boldsymbol{\alpha}\in U$, 则令

$$(\lambda\mathscr{A})(\boldsymbol{\alpha})=\lambda\mathscr{A}(\boldsymbol{\alpha}).$$

引理 2.3 (和与数乘还是线性映射)

我们有

1. 映射 $\mathscr{A}$ 与 $\mathscr{B}$ 的和 $\mathscr{A}+\mathscr{B}$ 是 U 到 V 的线性映射.
2. 纯量 λ 与映射 $\mathscr{A}$ 的乘积 $\lambda\mathscr{A}$ 也是 U 到 V 的线性映射.

定义 2.12 (线性映射的复合 (也叫乘积))

设 $\mathscr{A}:U\to V$ 与 $\mathscr{B}:V\to W$ 是线性映射。作为映射, 可以定义映射 $\mathscr{A}$ 与 $\mathscr{B}$ 的复合（或者乘积）$\mathscr{B}\mathscr{A}:U\to W$, 即对任意 $\boldsymbol{\alpha}\in U$,

$$(\mathscr{B}\mathscr{A})(\boldsymbol{\alpha})=\mathscr{B}(\mathscr{A}(\boldsymbol{\alpha})).$$

设 $\lambda\in\mathbb{F},\boldsymbol{\alpha},\boldsymbol{\beta}\in U$, 则

$$(\mathscr{B}\mathscr{A})(\boldsymbol{\alpha}+\boldsymbol{\beta})=\mathscr{B}(\mathscr{A}(\boldsymbol{\alpha}+\boldsymbol{\beta})),$$
$$(\mathscr{B}\mathscr{A})(\lambda\boldsymbol{\alpha})=\mathscr{B}(\mathscr{A}(\lambda\boldsymbol{\alpha})).$$

由于映射 $\mathscr{A}$ 和 $\mathscr{B}$ 是线性的, 所以

$$(\mathscr{B})(\boldsymbol{\alpha}+\boldsymbol{\beta})=\mathscr{B}(\mathscr{A}(\boldsymbol{\alpha}))+\mathscr{B}(\mathscr{A}(\boldsymbol{\beta}))=(\mathscr{B}\mathscr{A})(\boldsymbol{\alpha})+(\mathscr{B}\mathscr{A})(\boldsymbol{\beta}),$$
$$(\mathscr{B}A)(\lambda\boldsymbol{\alpha})=\lambda\mathscr{B}(\mathscr{A}(\boldsymbol{\alpha}))=\lambda(\mathscr{B}\mathscr{A})(\boldsymbol{\alpha}).$$

所以 $\mathscr{B}\mathscr{A}:U\to W$ 是线性映射.

注意

1. 与映射的乘积相同, 线性映射的乘积满足结合律, 但交换律并不成立. 即一般地说, 对线性映射 $\mathscr{A}\to\mathscr{B}$, 等式 $\mathscr{B}\mathscr{A}=\mathscr{A}\mathscr{B}$ 不成立.
2. 对零映射 $\mathscr{O}:U\to V$ 与线性映射 $\mathscr{A}:V\to W$, 恒有 $\mathscr{A}\mathscr{O}=\mathscr{O}$, 其中右端的 $\mathscr{O}$ 是 U 到 W 的零映射。而对线性映射 $\mathscr{A}:W\to U$, 恒有 $\mathscr{O}\mathscr{A}=\mathscr{O}$, 其中右端的 $\mathscr{O}$ 是 W 到 V 的零映射.
3. 作为集合，线性空间 U 到自身的单位映射记为 $\mathscr{E}_U$. 容易验证, $\mathscr{E}_U:U\to U$ 是线性映

射. 并且对任意线性映射 $\mathscr{A}: U \to V$, 恒有 $\mathscr{E}_V \mathscr{A} = \mathscr{A} = \mathscr{A}\mathscr{E}_U$.

引理 2.4 (线性映射的运算的矩阵表示)

假定线性映射 $\mathscr{A}: U \to V$ 在 U 的基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 与 V 的基 $\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_m\}$ 下的矩阵为 $\boldsymbol{A}$, 线性映射 $\mathscr{B}: V \to W$ 在 V 的基 $\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_m\}$ 与 W 的基 $\{\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \cdots, \boldsymbol{\gamma}_l\}$ 下的矩阵为 $\boldsymbol{B}$,

(1) 设线性映射 $\mathscr{A}: U \to V$ 与 $\mathscr{B}: U \to V$ 在 U 的基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 下与 V 的基 $\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_m\}$ 下的矩阵分别为 $\boldsymbol{A}$ 与 $\boldsymbol{B}$, 则线性映射 $\mathscr{A} + \mathscr{B}$ 在这两组基下的矩阵为 $\boldsymbol{A} + \boldsymbol{B}$.

(2) 对于 $\lambda \in \mathbb{F}$, 线性映射 $\lambda\mathscr{A}$ 在这两组基下的矩阵为 $\lambda\boldsymbol{A}$.

(3) 线性映射 $\mathscr{A}$ 与 $\mathscr{B}$ 的乘积（复合）$\mathscr{B}\mathscr{A}$ 在这两组基下的矩阵表示等于线性映射 $\mathscr{B}$ 的矩阵表示 $\boldsymbol{B}$ 与线性映射 $\mathscr{A}$ 的矩阵表示 $\boldsymbol{A}$ 的乘积.

♡

定理 2.10

设 U 与 V 分别是数域 $\mathbb{F}$ 上 n 维与 m 维线性空间。所有 U 到 V 的线性映射集合记为

$$\mathscr{L}(U, V) = \mathscr{L}_{m\times n} := \{\mathscr{A}: U \mapsto V \quad \text{是线性映射}\}$$

则有

1. 集合 $\mathscr{L}_{m\times n}$ 在上述的加法以及纯量与映射的乘法下构成数域 $\mathbb{F}$ 上的线性空间。
2. $\mathscr{L}_{m\times n}$ 的维数为

$$\dim \mathscr{L}_{m\times n} = mn.$$

♡

练习 2.14 设 U 和 V 分别是数域 $\mathbb{F}$ 上 n 维和 m 维线性空间,

$$\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n\}$$

分别是 U 的基. $\mathbf{Q}$ 是 $\mathbb{F}$ 上的 $n \times n$ 阶的矩阵。假定 $\mathscr{A}: U \to V$ 是一个线性映射. 请验证:

$$\mathcal{A}\left(\left(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n\right)\mathbf{Q}\right) = \mathcal{A}\left(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n\right)\mathbf{Q}.$$

解 简单训练，课堂已讲，只是让你熟悉记号，略。

练习 2.15 什么是矩阵的相似？请给两个相似矩阵的例子。

解 简单训练，课堂已讲，略。

练习 2.16 假定 $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{F}^{n\times n}$ 相似，证明

(1) $\boldsymbol{A}^T$ 和 $\boldsymbol{B}^T$ 也相似。

(2) $\boldsymbol{A}^2$ 和 $\boldsymbol{B}^2$ 也相似。

(3) 如果 $\boldsymbol{A}, \boldsymbol{B}$ 可逆，那么 $\boldsymbol{A}^{-1}$ 和 $\boldsymbol{B}^{-1}$ 也相似。

解 及其容易，按定义来就直接证明了，略。

练习 2.17（选做）假定 $\boldsymbol{A} \in \mathbb{F}^{n\times n}$ 是幂等矩阵，也即 $\boldsymbol{A}^2 = \boldsymbol{A}$，证明 $\boldsymbol{A}$ 相似于

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

解 定义线性变换 $\mathscr{A}:\mathbb{F}^n\mapsto\mathbb{F}^n$ 如下

$$\mathscr{A}(X)=\boldsymbol{A}X.$$

则有 $\mathscr{A}^2=\mathscr{A}$ 并且 $\mathscr{A}$ 在基

$$\left\{\boldsymbol{\alpha}_k=\underbrace{(0,0,\cdots,0,1,0,\cdots,0)}_{\text{第 }k\text{ 个位置}}:k=1,2,3,\cdots,n\right\}$$

下的矩阵表示就是 $\boldsymbol{A}$. 取 $\mathrm{Im}(\mathscr{A})$ 的一组基

$$\{\mathscr{A}(\boldsymbol{\alpha}_1),\mathscr{A}(\boldsymbol{\alpha}_2),\cdots,\mathscr{A}(\boldsymbol{\alpha}_r)\},\quad \boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r\in\mathbb{F}^n.$$

则 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r\}$ 线性无关。取 $\mathrm{Ker}(\mathscr{A})$ 的一组基

$$\{\boldsymbol{\alpha}_{1+r},\boldsymbol{\alpha}_{2+r},\cdots,\boldsymbol{\alpha}_{r+s}\}.$$

则我们可知

$$\{\mathscr{A}(\boldsymbol{\alpha}_1),\mathscr{A}(\boldsymbol{\alpha}_2),\cdots,\mathscr{A}(\boldsymbol{\alpha}_r),\boldsymbol{\alpha}_{1+r},\boldsymbol{\alpha}_{2+r},\cdots,\boldsymbol{\alpha}_{r+s}\}$$

形成 $\mathbb{F}^n$ 的基。这是因为

- 线性无关。如果

$$x_1\mathscr{A}(\boldsymbol{\alpha}_1)+x_2\mathscr{A}(\boldsymbol{\alpha}_2)+\cdots x_r\mathscr{A}(\boldsymbol{\alpha}_r)+x_{r+1}\boldsymbol{\alpha}_{r+1}+\cdots+x_{r+s}\boldsymbol{\alpha}_{r+s}=0.$$

则

$$\begin{aligned}0=&\mathscr{A}(x_1\mathscr{A}(\boldsymbol{\alpha}_1)+x_2\mathscr{A}(\boldsymbol{\alpha}_2)+\cdots x_r\mathscr{A}(\boldsymbol{\alpha}_r)+x_{r+1}\boldsymbol{\alpha}_{r+1}+\cdots+x_{r+s}\boldsymbol{\alpha}_{r+s})\\=&x_1\mathscr{A}^2(\boldsymbol{\alpha}_1)+x_2\mathscr{A}^2(\boldsymbol{\alpha}_2)+\cdots x_r\mathscr{A}^2(\boldsymbol{\alpha}_r)+x_{r+1}\mathscr{A}(\boldsymbol{\alpha}_{r+1})+\cdots+x_{r+s}\mathscr{A}(\boldsymbol{\alpha}_{r+s})\\=&x_1\mathscr{A}(\boldsymbol{\alpha}_1)+x_2\mathscr{A}(\boldsymbol{\alpha}_2)+\cdots x_r\mathscr{A}(\boldsymbol{\alpha}_r)+x_{r+1}\mathscr{A}(\boldsymbol{\alpha}_{r+1})+\cdots+x_{r+s}\mathscr{A}(\boldsymbol{\alpha}_{r+s})\\=&x_1\mathscr{A}(\boldsymbol{\alpha}_1)+x_2\mathscr{A}(\boldsymbol{\alpha}_2)+\cdots x_r\mathscr{A}(\boldsymbol{\alpha}_r)\end{aligned}$$

又因为

$$\{\mathscr{A}(\boldsymbol{\alpha}_1),\mathscr{A}(\boldsymbol{\alpha}_2),\cdots,\mathscr{A}(\boldsymbol{\alpha}_r)\}$$

线性无关，从而 $x_1=x_2=\cdots=x_r=0$. 于是

$$x_{r+1}\boldsymbol{\alpha}_{r+1}+\cdots+x_{r+s}\boldsymbol{\alpha}_{r+s}=0$$

从而 $x_{r+1}=\cdots=x_{r+s}=0$.

- 任何一个向量可被表示。任取 $\boldsymbol{\alpha}\in U$，则存在 $x_1,x_2,\cdots,x_r\in\mathbb{F}$ 使得

$$\begin{aligned}\mathscr{A}(\boldsymbol{\alpha})=&x_1\mathscr{A}(\boldsymbol{\alpha}_1)+x_2\mathscr{A}(\boldsymbol{\alpha}_2)+\cdots x_r\mathscr{A}(\boldsymbol{\alpha}_r)\\=&x_1\mathscr{A}^2(\boldsymbol{\alpha}_1)+x_2\mathscr{A}^2(\boldsymbol{\alpha}_2)+\cdots x_r\mathscr{A}^2(\boldsymbol{\alpha}_r)\\=&\mathscr{A}(x_1\mathscr{A}(\boldsymbol{\alpha}_1)+x_2\mathscr{A}(\boldsymbol{\alpha}_2)+\cdots x_r\mathscr{A}(\boldsymbol{\alpha}_r)).\end{aligned}$$

也即

$$\mathscr{A}(\boldsymbol{\alpha}-(x_1\mathscr{A}(\boldsymbol{\alpha}_1)+x_2\mathscr{A}(\boldsymbol{\alpha}_2)+\cdots x_r\mathscr{A}(\boldsymbol{\alpha}_r)))=0.$$

于是 $\boldsymbol{\alpha}-(x_1\mathscr{A}(\boldsymbol{\alpha}_1)+x_2\mathscr{A}(\boldsymbol{\alpha}_2)+\cdots x_r\mathscr{A}(\boldsymbol{\alpha}_r))\in \mathrm{Ker}(\mathscr{A})$. 于是存在 $x_{r+1},\cdots,x_{r+s}\in\mathbb{F}$ 使得

$$\boldsymbol{\alpha}-(x_1\mathscr{A}(\boldsymbol{\alpha}_1)+x_2\mathscr{A}(\boldsymbol{\alpha}_2)+\cdots x_r\mathscr{A}(\boldsymbol{\alpha}_r))=x_{r+1}\boldsymbol{\alpha}_{r+1}+\cdots+x_{r+s}\boldsymbol{\alpha}_{r+s})$$

所以

$$\boldsymbol{\alpha}=x_1\mathscr{A}(\boldsymbol{\alpha}_1)+x_2\mathscr{A}(\boldsymbol{\alpha}_2)+\cdots x_r\mathscr{A}(\boldsymbol{\alpha}_r)+x_{r+1}\boldsymbol{\alpha}_{r+1}+\cdots+x_{r+s}\boldsymbol{\alpha}_{r+s}$$

从而 $n=r+s$. 注意到 $\mathscr{A}^2(\boldsymbol{\alpha}_k)=\mathscr{A}(\boldsymbol{\alpha}_k)$, 我们得到 $\mathscr{A}$ 在这组基

$$\{\mathscr{A}(\boldsymbol{\alpha}_1),\mathscr{A}(\boldsymbol{\alpha}_2),\cdots,\mathscr{A}(\boldsymbol{\alpha}_r),\boldsymbol{\alpha}_{1+r},\boldsymbol{\alpha}_{2+r},\cdots,\boldsymbol{\alpha}_{r+s}\}$$

下的矩阵表示为

$$\begin{pmatrix}\boldsymbol{I}_r & 0\\ 0 & \boldsymbol{0}\end{pmatrix}$$

根据定理 2.20 可得 $\boldsymbol{A}$ 相似于

$$\begin{pmatrix}\boldsymbol{I}_r & 0\\ 0 & \boldsymbol{0}\end{pmatrix}.$$

✍ **练习 2.18** 假定线性映射 $\mathscr{A}:U\to V$ 在 U 的基 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$ 与 V 的基 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m\}$ 下的矩阵为 $\boldsymbol{A}$, 线性映射 $\mathscr{B}:V\to W$ 在 V 的基 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m\}$ 与 W 的基 $\{\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\cdots,\boldsymbol{\gamma}_1\}$ 下的矩阵为 $\boldsymbol{B}$, 请证明:

(1) 设线性映射 $\mathscr{A}:U\to V$ 与 $\mathscr{B}:U\to V$ 在 U 的基 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$ 下与 V 的基 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m\}$ 下的矩阵分别为 $\boldsymbol{A}$ 与 $\boldsymbol{B}$, 则线性映射 $\mathscr{A}+\mathscr{B}$ 在这两组基下的矩阵为 $\boldsymbol{A}+\boldsymbol{B}$.

(2) 对于 $\lambda\in\mathbb{F}$, 线性映射 $\lambda\mathscr{A}$ 在这两组基下的矩阵为 $\lambda\boldsymbol{A}$.

(3) 线性映射 $\mathscr{A}$ 与 $\mathscr{B}$ 的乘积（复合）$\mathscr{B}\mathscr{A}$ 在这两组基下的矩阵表示等于线性映射 $\mathscr{B}$ 的矩阵表示 $\boldsymbol{B}$ 与线性映射 $\mathscr{A}$ 的矩阵表示 $\boldsymbol{A}$ 的乘积.

解 简单训练，课堂已讲，略。

✍ **练习 2.19** 设 U 与 V 分别是数域 $\mathbb{F}$ 上 n 维与 m 维线性空间, $\mathscr{A}:U\to V$ 是线性映射.

- 定义映射 $\mathscr{A}$ 与 $\mathscr{B}$ 的和 $\mathscr{A}+\mathscr{B}$ 如下, 设 $\boldsymbol{\alpha}\in U$, 则令

$$(\mathscr{A}+\mathscr{B})(\boldsymbol{\alpha})=\mathscr{A}(\boldsymbol{\alpha})+\mathscr{B}(\boldsymbol{\alpha}).$$

- 定义纯量 $\lambda\in\mathbb{F}$ 与映射 $\mathscr{A}$ 的乘积 $\lambda\mathscr{A}$ 如下: 设 $\boldsymbol{\alpha}\in U$, 则令

$$(\lambda\mathscr{A})(\boldsymbol{\alpha})=\lambda\mathscr{A}(\boldsymbol{\alpha}).$$

请证明 $\mathscr{A}+\mathscr{B}$ 和 $\lambda\mathscr{A}$ 也是线性映射。

解 简单训练，课堂已讲，略。

2.4 2024 年 4 月 18 日第七周星期四第十三次课-像与核

2.4.1 可逆的线性映射

定义 2.13 (可逆)

设 $\mathscr{A}: U \to V$ 是线性映射. 如果作为映射, $\mathscr{A}$ 可逆, 则 $\mathscr{A}$ 称为可逆的线性映射.

定理 2.11

设线性映射 $\mathscr{A}: U \to V$ 可逆, 则 $\dim U = \dim V$.

定理 2.12

设线性映射 $\mathscr{A}: U \to V$ 可逆, 则 $\mathscr{A}$ 的逆映射是可逆线性映射.

定理 2.13

设 $\mathscr{A}: U \to V$ 与 $\mathscr{B}: V \to W$ 是可逆线性映射. 则乘积 $\mathscr{B}\mathscr{A}: U \to W$ 是可逆线性映射.

命题 2.4 (可逆线性映射 $\mathscr{A}: U \to V$ 的逆映射 $\mathscr{B}: V \to U$ 的矩阵表示)

假定映射 $\mathscr{A}$ 可逆（所以 $\dim U = \dim V$）. 设线性映射 $\mathscr{A}$ 在 U 的基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 与 V 的基 $\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n\}$ 下的矩阵为 $\boldsymbol{A}$, 即

$$\mathscr{A}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n) = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n)\boldsymbol{A}.$$

设 $\mathscr{B}: U \to V$ 在 V 的基 $\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n\}$ 与 U 的基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 下的矩阵为 $\boldsymbol{B}$, 即

$$\mathscr{B}(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n)\boldsymbol{B}.$$

因为 $\mathscr{B}\mathscr{A} = \mathscr{E}_U, \mathscr{A}\mathscr{B} = \mathscr{E}_V$, 因此

$$(\mathscr{B}\mathscr{A})(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n)\boldsymbol{B}\boldsymbol{A} = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n)\boldsymbol{I}_{(n)},$$
$$(\mathscr{A}\mathscr{B})(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n) = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n)\boldsymbol{A}\boldsymbol{B} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n)\boldsymbol{I}_n.$$

所以 $\boldsymbol{AB} = I_n = \boldsymbol{BA}$. 即可逆矩阵线性映射 $\mathscr{A}$ 的矩阵表示 $\boldsymbol{A}$ 是可逆方阵, 并且 $\mathscr{A}$ 的逆映射 $\mathscr{B}$ 的矩阵表示是方阵 $\boldsymbol{A}$ 的逆方阵. 以后记线性映射 $\mathscr{A}$ 的逆映射为 $\mathscr{A}^{-1}$.

2.4.2 线性映射的核与像

我们恒假设 U 与 V 分别是数域 $\mathbb{F}$ 上 n 维与 m 维线性空间.

定义 2.14

设 $\mathscr{A}: U \to V$ 是线性映射. 集合

$$\operatorname{Im}(\mathscr{A}) = \{\mathscr{A}(\boldsymbol{\alpha}) \in V : \boldsymbol{\alpha} \in U\}$$

称为 U 在 $\mathscr{A}$ 下的像, 或者 $\mathscr{A}$ 的值域, 也记为 $\mathscr{A}(U)$. 集合

$$\operatorname{Ker}(\mathscr{A}) = \{\boldsymbol{\alpha} \in U : \mathscr{A}(\boldsymbol{\alpha}) = \boldsymbol{0} \in V\}$$

称为 $\mathscr{A}$ 的核, 也记为 $\mathscr{A}^{-1}(\mathbf{0})$.

定理 2.14

U 在 $\mathscr{A}$ 下的像 $\mathrm{Im}(\mathscr{A})$ 是 V 的子空间, 而 $\mathscr{A}$ 的核 $\mathrm{Ker}(\mathscr{A})$ 是 U 的子空间.

定理 2.15

设 V 和 W 是数域 $\mathbb{F}$ 上的向量空间, 而 $\sigma : V \to W$ 是一个线性映射. 那么

(i) σ 是满射 $\Longleftrightarrow \mathrm{Im}(\sigma) = W$;

(ii) σ 是单射 $\Longleftrightarrow \mathrm{Ker}(\sigma) = \{0\}$.

定理 2.16

设 $\mathscr{A} : U \to V$ 是线性映射. 则 U 模 $\mathrm{Ker}(\mathscr{A})$ 的商空间 $U/\mathrm{Ker}(\mathscr{A})$ 同构于 $\mathrm{Im}(\mathscr{A})$.

定理 2.17

设 $\mathscr{A} : U \to V$ 是线性映射, 则

$$\dim U = \dim(\mathrm{Im}(\mathscr{A})) + \dim(\mathrm{Ker}(\mathscr{A})).$$

定义 2.15

设 $\mathscr{A} : U \to V$ 是线性映射. 则 $\dim(\mathrm{Im}(\mathscr{A}))$ 称为 $\mathscr{A}$ 的秩, 记为 $\rho(\mathscr{A})$, 而 $\dim(\mathrm{Ker}(\mathscr{A}))$ 称为 $\mathscr{A}$ 的零度 (nullity), 记为 $\nu(\mathscr{A})$.

定理 2.18

线性映射 $\mathscr{A} : U \to V$ 为单射的充分必要条件是 $\mathrm{Ker}(\mathscr{A}) = \{0\}$.

定理 2.19

线性映射 $\mathscr{A} : U \to V$ 可逆的充分必要条件是 $\dim V = \dim U$, 且 $\mathrm{Ker}(\mathscr{A}) = \{\mathbf{0}\}$.

2.4.3 线性变换

这一小节我们恒假定 V 是数域 $\mathbb{F}$ 上 n 维线性空间.

注意

集合 S 到自身的映射通常也称为集合 S 的变换. 因此线性映射 $\mathcal{A} : V \to V$ 也称为线性空间 V 的线性变换. 于是线性变换是线性映射的特例, 所以前面叙述的关于线性映射的结论对线性变换也成立.

1. 考虑线性变换 $\mathscr{A} : V \to V$ 的矩阵表示. 设 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 是 V 的基。

2. 由于 $\mathscr{A}(\boldsymbol{\alpha}_1), \mathscr{A}(\boldsymbol{\alpha}_2), \cdots, \mathscr{A}(\boldsymbol{\alpha}_n) \in V$. 因此

$$\begin{aligned}
\mathscr{A}(\boldsymbol{\alpha}_1) &= a_{11}\boldsymbol{\alpha}_1 + a_{12}\boldsymbol{\alpha}_2 + \cdots + a_{1n}\boldsymbol{\alpha}_n, \\
\mathscr{A}(\boldsymbol{\alpha}_2) &= a_{21}\boldsymbol{\alpha}_1 + a_{22}\boldsymbol{\alpha}_2 + \cdots + a_{2n}\boldsymbol{\alpha}_n, \\
&\cdots\cdots\cdots\cdots \\
\mathscr{A}(\boldsymbol{\alpha}_n) &= a_{n1}\boldsymbol{\alpha}_1 + a_{n2}\boldsymbol{\alpha}_2 + \cdots + a_{nn}\boldsymbol{\alpha}_n.
\end{aligned}$$

即

$$\mathscr{A}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n) \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ & \cdots\cdots\cdots & & \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{pmatrix}.$$

3. 记 $\boldsymbol{A} = (a_{ij}) \in \mathbb{F}^{n\times n}$. 则上式为

$$\mathscr{A}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n)\boldsymbol{A}. \tag{2.6}$$

定义 2.16 (线性变换的矩阵表示)

方阵 $\boldsymbol{A}$ 称为线性变换 $\mathscr{A}: V \to V$ 在基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 下的矩阵. ♣

回忆

- 数域 $\mathbb{F}$ 上 n 维线性空间 V 的所有线性变换集合记为 $\mathscr{L}_n(V)$
- $\mathscr{L}_n(V)$ 在线性变换的加法与纯量和线性变换的乘法下成为数域 $\mathbb{F}$ 上 n^2 维线性空间.
- 数域 $\mathbb{F}$ 上所有 n 阶方阵集合记为 $\mathbb{F}^{n\times n}$.
- 定义映射 $\eta: \mathscr{L}_n(V) \to \mathbb{F}^{n\times n}$ 如下: 设 $\mathscr{A} \in \mathscr{L}_n(V)$, 且 $\mathscr{A}$ 在 V 的基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 下的方阵为 $\boldsymbol{A}$. 则令 $\eta(\mathscr{A}) = \boldsymbol{A}$.
- 映射 $\eta: \mathscr{L}_n(V) \to \mathbb{F}^{n\times n}$ 为同构映射.
- 设 $\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n\}$ 是 V 的另一组基, 线性变换 $\mathscr{A}: V \to V$ 在基 $\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n\}$ 下的方阵为 $\boldsymbol{B}$, 即
$$\mathscr{A}(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n) = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n)\boldsymbol{B}.$$
- 设基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 到基 $\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n\}$ 的过渡矩阵为 $\boldsymbol{P}$, 即
$$(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n)\boldsymbol{P}, \tag{2.7}$$
其中 P 是数域 $\mathbb{F}$ 上 n 阶可逆方阵. 于是
$$\mathscr{A}(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n) = (\mathscr{A}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n))\boldsymbol{P}.$$
- 根据 (2.6) 有
$$\mathscr{A}(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n)\boldsymbol{A}\boldsymbol{P}.$$

- 再结合 (2.7) 可得

$$\boxed{\mathscr{A}\left(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n\right)=\left(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n\right)\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}.} \tag{2.8}$$

- 再根据线性变换在同一组基下矩阵表示的唯一性得到 $\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$.

定义 2.17 (矩阵的相似)

设 $\boldsymbol{A},\boldsymbol{B}\in\boldsymbol{F}^{n\times n}$, 如果存在 n 阶可逆方阵 $\boldsymbol{P}\in\mathbb{F}^{n\times n}$, 使得 $\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$, 则方阵 $\boldsymbol{A},\boldsymbol{B}$ 称为相似的.

定理 2.20 (不同基下的矩阵表示的相似性)

设 $\mathscr{A}:V\to V$ 是线性变换, 则 $\mathscr{A}$ 在 V 的不同基下的方阵是相似的.

(1) 反之设方阵 $\boldsymbol{A},\boldsymbol{B}\in\boldsymbol{F}^{n\times n}$ 相似, 即存在可逆方阵 $\boldsymbol{P}\in\boldsymbol{F}^{n\times n}$, 使 $\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$.

(2) 在数域 $\mathbb{F}$ 上 n 维线性空间 V 中取定基 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$. 由式(2.6), 可以确定线性变换 $\mathscr{A}:V\to V$.

(3) 由式 (2.7) 可以确定 V 的基 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n\}$.

(4) 于是由式 (2.8), 线性变换 $\mathscr{A}:V\to V$ 在基 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n\}$ 下的方阵即为 $\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$.

定理 2.21 (相似矩阵可得到同一个线性变换)

设 $\boldsymbol{A}\in\mathbb{F}^{n\times n}$, 则由式 (2.6) 可以确定一个线性变换 $\mathscr{A}:V\to V$. 而且相似的方阵是同一个线性变换在不同基下的方阵.

命题 2.5

方阵之间的相似关系具有如下的性质:

1° (自反性) 对任意 $\boldsymbol{A}\in\mathbb{F}^{n\times n}$, 方阵 A 与自身相似;

2° (对称性) 设 $\boldsymbol{A},\boldsymbol{B}\in\mathbb{F}^{n\times n}$ 相似, 则 $\boldsymbol{B},\boldsymbol{A}$ 相似;

3° (传递性) 设 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}\in\mathbb{F}^{n\times n}$, 且 $\boldsymbol{A},\boldsymbol{B}$ 相似, $\boldsymbol{B},\boldsymbol{C}$ 相似, 则 $\boldsymbol{A},\boldsymbol{C}$ 相似.

回忆:

- 可逆线性变换 $\mathscr{A}:V\to V$ 是指线性空间 V 到自身的可逆线性映射.
- 设可逆线性变换 $\mathscr{A}:V\to$ 在 V 的基 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$ 下的方阵为 $\boldsymbol{A}$, 即

$$\mathscr{A}\left(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\right)=\left(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\right)\boldsymbol{A},$$

  则方阵 $\boldsymbol{A}$ 可逆, 并且 $\mathscr{A}$ 的逆变换 $\mathscr{A}^{-1}$ 在同一组基 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$ 下的方阵为 $\boldsymbol{A}^{-1}$, 即

$$\mathscr{A}^{-1}\left(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\right)=\left(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\right)\boldsymbol{A}^{-1}.$$

由前面知线性映射 $\mathscr{A}:U\to V$ 可逆的充分必要条件是 $\mathscr{A}$ 为双射. 对线性变换, 条件可以减弱。

定理 2.22 (可逆线性变换的等价条件)
设 $\mathscr{A}: V \to V$ 是线性变换, 则下述三条件等价:
(1) 线性变换 $\mathscr{A}$ 可逆;
(2) $\mathrm{Ker}(\mathscr{A}) = \{\mathbf{0}\}$;
(3) $\mathrm{Im}(\mathscr{A}) = V$.

推论 2.2 (可逆的充要条件)
线性变换 $\mathscr{A}: V \to V$ 可逆的充分必要条件是, 线性变换 $\mathscr{A}$ 把 V 的基变为 V 的基, 即如果 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 是 V 的基, 则 $\{\mathscr{A}(\boldsymbol{\alpha}_1), \mathscr{A}(\boldsymbol{\alpha}_2), \cdots, \mathscr{A}(\boldsymbol{\alpha}_n)\}$ 是 V 的基.

练习 2.20 (选做) 设 $\mathscr{A}: U \to V$ 是线性映射, 并且 $\dim \mathrm{Im}(\mathscr{A}) = r$. 证明: 存在 U 的基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 与 V 的基 $\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_m\}$, 使得

$$\mathscr{A}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n) = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_m) \begin{pmatrix} \boldsymbol{I}_r & 0 \\ 0 & \mathbf{0} \end{pmatrix}.$$

解 取 $\mathrm{Im}(\mathscr{A})$ 的一组基

$$\{\mathscr{A}(\boldsymbol{\alpha}_1), \mathscr{A}(\boldsymbol{\alpha}_2), \cdots, \mathscr{A}(\boldsymbol{\alpha}_r)\}, \quad \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r \in U.$$

则 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r\}$ 线性无关。取 $\mathrm{Ker}(\mathscr{A})$ 的一组基

$$\{\boldsymbol{\alpha}_{1+r}\boldsymbol{\alpha}_{2+r}, \cdots, \boldsymbol{\alpha}_{r+s}\}.$$

则我们可知

$$\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r, \boldsymbol{\alpha}_{1+r}\boldsymbol{\alpha}_{2+r}, \cdots, \boldsymbol{\alpha}_{r+s}\}$$

形成 U 的基。这是因为

- 线性无关。如果

$$x_1\boldsymbol{\alpha}_1 + x_2\boldsymbol{\alpha}_2 + \cdots x_r\boldsymbol{\alpha}_r + x_{r+1}\boldsymbol{\alpha}_{r+1} + \cdots + x_{r+s}\boldsymbol{\alpha}_{r+s} = 0.$$

  则

$$\begin{aligned}
0 =& \mathscr{A}(x_1\boldsymbol{\alpha}_1 + x_2\boldsymbol{\alpha}_2 + \cdots x_r\boldsymbol{\alpha}_r + x_{r+1}\boldsymbol{\alpha}_{r+1} + \cdots + x_{r+s}\boldsymbol{\alpha}_{r+s}) \\
=& x_1\mathscr{A}(\boldsymbol{\alpha}_1) + x_2\mathscr{A}(\boldsymbol{\alpha}_2) + \cdots x_r\mathscr{A}(\boldsymbol{\alpha}_r) + x_{r+1}\boldsymbol{\alpha}_{r+1} + \cdots + x_{r+s}\mathscr{A}(\boldsymbol{\alpha}_{r+s}) \\
=& x_1\mathscr{A}(\boldsymbol{\alpha}_1) + x_2\mathscr{A}(\boldsymbol{\alpha}_2) + \cdots x_r\mathscr{A}(\boldsymbol{\alpha}_r).
\end{aligned}$$

  又因为

$$\{\mathscr{A}(\boldsymbol{\alpha}_1), \mathscr{A}(\boldsymbol{\alpha}_2), \cdots, \mathscr{A}(\boldsymbol{\alpha}_r)\}$$

  线性无关，从而 $x_1 = x_2 = \cdots = x_r = 0$. 于是

$$x_{r+1}\boldsymbol{\alpha}_{r+1} + \cdots + x_{r+s}\boldsymbol{\alpha}_{r+s} = 0$$

  从而 $x_{r+1} = \cdots = x_{r+s} = 0$.

- 任何一个向量可被表示。任取 $\boldsymbol{\alpha} \in U$，则存在 $x_1, x_2, \cdots, x_r \in \mathbb{F}$ 使得

$$\mathscr{A}(\boldsymbol{\alpha}) = x_1\mathscr{A}(\boldsymbol{\alpha}_1) + x_2\mathscr{A}(\boldsymbol{\alpha}_2) + \cdots x_r\mathscr{A}(\boldsymbol{\alpha}_r) = \mathscr{A}(x_1\boldsymbol{\alpha}_1 + x_2\boldsymbol{\alpha}_2 + \cdots x_r\boldsymbol{\alpha}_r).$$

也即

$$\mathscr{A}(\boldsymbol{\alpha} - (x_1\boldsymbol{\alpha}_1 + x_2\boldsymbol{\alpha}_2 + \cdots x_r\boldsymbol{\alpha}_r)) = 0.$$

于是 $\boldsymbol{\alpha} - (x_1\boldsymbol{\alpha}_1 + x_2\boldsymbol{\alpha}_2 + \cdots x_r\boldsymbol{\alpha}_r) \in \mathrm{Ker}(\mathscr{A})$. 于是存在 $x_{r+1}, \cdots, x_{r+s} \in \mathbb{F}$ 使得

$$\boldsymbol{\alpha} - (x_1\boldsymbol{\alpha}_1 + x_2\boldsymbol{\alpha}_2 + \cdots x_r\boldsymbol{\alpha}_r) = x_{r+1}\boldsymbol{\alpha}_{r+1} + \cdots + x_{r+s}\boldsymbol{\alpha}_{r+s})$$

所以

$$\boldsymbol{\alpha} = x_1\boldsymbol{\alpha}_1 + x_2\boldsymbol{\alpha}_2 + \cdots x_r\boldsymbol{\alpha}_r + x_{r+1}\boldsymbol{\alpha}_{r+1} + \cdots + x_{r+s}\boldsymbol{\alpha}_{r+s}$$

于是我们有 $n = r + s$. 选 U 的基

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

以及将 $\mathscr{A}(\boldsymbol{\alpha}_1), \mathscr{A}(\boldsymbol{\alpha}_2), \cdots, \mathscr{A}(\boldsymbol{\alpha}_r)$ 扩充为 V 的基

$$\{\mathscr{A}(\boldsymbol{\alpha}_1), \mathscr{A}(\boldsymbol{\alpha}_2), \cdots, \mathscr{A}(\boldsymbol{\alpha}_r), \boldsymbol{\gamma}_{r+1}, \cdots, \boldsymbol{\gamma}_m\}.$$

在这两组基下，$\mathscr{A}$ 的矩阵表示就有所求的形式。

练习 2.21 (选做) 设 $\boldsymbol{A}$ 是线性映射 $\mathscr{A}: U \to V$ 在 U 的基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 与 V 的基 $\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_m\}$ 下的矩阵. 则 $\mathrm{rank}\,\boldsymbol{A} = \dim(\mathrm{Im}(\mathscr{A}))$

解 假定 $\dim(\mathrm{Im}(\mathscr{A})) = r$，则由前一题知道存在存在 U 的一组基与 V 的一组基使得 $\mathscr{A}$ 在这些基下的矩阵表示有形式

$$\boldsymbol{B} = \begin{pmatrix} \boldsymbol{I}_r & 0 \\ 0 & \boldsymbol{0} \end{pmatrix}.$$

从而有 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相抵。换句话说存在 $m \times m$ 阶可逆矩阵 $\mathbf{P}$ 与 $n \times n$ 阶可逆矩阵 $\mathbf{Q}$ 使得

$$\boldsymbol{A} = \mathbf{P}\begin{pmatrix} \boldsymbol{I}_r & 0 \\ 0 & \boldsymbol{0} \end{pmatrix}\mathbf{Q}.$$

因为乘以可逆矩阵不改变矩阵的秩，所以

$$\mathrm{rank}(\boldsymbol{A}) = \mathrm{rank}\begin{pmatrix} \boldsymbol{I}_r & 0 \\ 0 & \boldsymbol{0} \end{pmatrix} = r.$$

定义 2.18

设 $\mathscr{A}: U \to V$ 是线性映射. 则 $\dim(\mathrm{Im}(\mathscr{A}))$ 称为 $\mathscr{A}$ 的秩, 记为 $\rho(\mathscr{A})$.

练习 2.22 设 $\mathscr{A}: V \to V$ 是数域 $\mathbb{F}$ 上 n 维线性空间 V 到自身的线性映射, 且 $\rho(\mathscr{A}^2) = \rho(\mathscr{A})$. 证明:

$$\mathrm{Im}(\mathscr{A}) \cap \mathrm{Ker}(\mathscr{A}) = \{0\}.$$

解 假定 $\boldsymbol{\alpha} \in \mathrm{Im}(\mathscr{A}) \cap \mathrm{Ker}(\mathscr{A})$，则存在 $\boldsymbol{\beta} \in V$ 使得

$$\boldsymbol{\alpha} = \mathscr{A}(\boldsymbol{\beta}).$$

并且

$$0=\mathscr{A}(\boldsymbol{\alpha})=\mathscr{A}^2(\boldsymbol{\beta}).$$

于是 $\boldsymbol{\beta}\in\mathrm{Ker}(\mathscr{A}^2)$. 因为 $\rho(\mathscr{A}^2)=\rho(\mathscr{A})$，于是

$$\dim(\mathrm{Im}(\mathscr{A}))=\dim(\mathrm{Im}(\mathscr{A}^2).)$$

于是由定理 2.17 可得

$$\dim(\mathrm{Ker}(\mathscr{A}))=n-\dim(\mathrm{Im}(\mathscr{A}))=n-\dim(\mathrm{Im}(\mathscr{A}))=\dim(\mathrm{Ker}(\mathscr{A}^2)).$$

显然 $\mathrm{Ker}(\mathscr{A})\subseteq\mathrm{Ker}(\mathscr{A}^2)$, 于是一定有

$$\mathrm{Ker}(\mathscr{A})=\mathrm{Ker}(\mathscr{A}^2).$$

而前面已经得到 $\boldsymbol{\beta}\in\mathrm{Ker}(\mathscr{A}^2)$，从而 $\boldsymbol{\beta}\in\mathrm{Ker}(\mathscr{A})$, 于是

$$\boldsymbol{\alpha}=\mathscr{A}(\boldsymbol{\beta})=0.$$

再根据 $\boldsymbol{\alpha}\in\mathrm{Im}(\mathscr{A})\cap\mathrm{Ker}(\mathscr{A})$ 的任意性就得到

$$\mathrm{Im}(\mathscr{A})\cap\mathrm{Ker}(\mathscr{A})=\{0\}.$$

✍ **练习 2.23** (选做) 设 $V_0,V_1,\cdots,V_{n+1}$ 是数域 $\mathbb{F}$ 上有限维线性空间, $V_0=V_{n+1}=\{0\}$. 设 $\mathscr{A}_i:V_i\to V_{i+1}$ 是线性映射, $i=0,1,\cdots,n$, 且 $\mathrm{Ker}\left(\mathscr{A}_{i+1}\right)=\mathrm{Im}\left(\mathscr{A}_i\right),i=0,1,\cdots,n-1$. 证明:

$$\sum_{i=1}^{n}(-1)^i\dim V_i=0$$

解 对线性变换 $\mathscr{A}_i:V_i\to V_{i+1}$ 利用定理 2.17 可得

$$\dim(V_i)=\dim\mathrm{Im}(V_i)+\dim\mathrm{Ker}(V_i)=\dim\mathrm{Ker}(V_i)+\dim\mathrm{Ker}(V_{i+1}),\quad i=0,1,n-1.$$

同时因为 $\mathscr{A}_n\equiv0$ 是零映射，从而有

$$\dim(V_n)=\dim\mathrm{Ker}(V_n).$$

于是

$$\begin{aligned}\sum_{i=1}^{n}(-1)^i\dim\mathscr{A}_i=&\sum_{i=1}^{n-1}(-1)^i(\dim\mathrm{Ker}(\mathscr{A}_i)+\dim\mathrm{Ker}(\mathscr{A}_{i+1}))+(-1)^n\dim\mathrm{Ker}(\mathscr{A}_n)\\=&\sum_{i=1}^{n-1}(-1)^i\dim\mathrm{Ker}(\mathscr{A}_i)+\sum_{i=1}^{n-1}(-1)^i\dim\mathrm{Ker}(\mathscr{A}_{i+1})+(-1)^n\dim\mathrm{Ker}(\mathscr{A}_n)\\=&\sum_{i=1}^{n-1}(-1)^i\dim\mathrm{Ker}(\mathscr{A}_i)+\sum_{i=2}^{n}(-1)^{i-1}\dim\mathrm{Ker}(\mathscr{A}_i)+(-1)^n\dim\mathrm{Ker}(\mathscr{A}_n)\\=&(-1)\dim\mathrm{Ker}\mathscr{A}_1+\sum_{i=2}^{n}((-1)^i+(-1)^{i-1})\dim\mathrm{Ker}(\mathscr{A}_i)\\=&0.\end{aligned}$$

✍ **练习 2.24** (选做) 设 $\boldsymbol{A},\boldsymbol{B}\in\mathbb{F}^{m\times n}$. 证明: $\mathrm{rank}(\boldsymbol{A}+\boldsymbol{B})=\mathrm{rank}\,\boldsymbol{A}+\mathrm{rank}\,\boldsymbol{B}$ 的充分必要条件是,

存在数域 $\boldsymbol{F}$ 上 m 阶与 n 阶可逆方阵 P 与 Q 使得

$$\boldsymbol{PAQ}=\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}, \quad \boldsymbol{PBQ}=\begin{pmatrix} 0 & 0 \\ 0 & I_{(s)} \end{pmatrix},$$

其中 $r=\operatorname{rank}\boldsymbol{A}, s=\operatorname{rank}\boldsymbol{B}$, 且 $r+s\leqslant \min\{m,n\}$.

解 定义两个线性变换 $\mathcal{A}:\mathbb{F}^n\to\mathbb{F}^m$ 和 $\mathcal{B}:\mathbb{F}^n\to\mathbb{F}^m$ 如下:

$$\mathcal{A}(X)=\boldsymbol{A}X, \quad \mathcal{B}(X)=\boldsymbol{B}X, \quad \forall X\in\mathbb{F}^n.$$

当 $\operatorname{rank}(\boldsymbol{A}+\boldsymbol{B})=\operatorname{rank}(\boldsymbol{A})+\operatorname{rank}(\boldsymbol{B})$ 时, 有

$$\dim\operatorname{Im}(\mathcal{A}+\mathcal{B})=\dim\operatorname{Im}\mathcal{A}+\dim\operatorname{Im}\mathcal{B}. \tag{2.9}$$

注意到

$$\dim(\operatorname{Im}A+\operatorname{Im}\mathcal{B})=\dim\operatorname{Im}\mathcal{A}+\dim\operatorname{Im}\mathcal{B}-\dim\left(\operatorname{Im}\mathcal{A}\cap\operatorname{Im}\mathcal{B}\right),$$

而 $\operatorname{Im}(\mathcal{A}+\mathcal{B})\subseteq\operatorname{Im}\mathcal{A}+\operatorname{Im}\mathcal{B}$。因此

$$\dim\operatorname{Im}(\mathcal{A}+\mathcal{B})\leq\dim(\operatorname{Im}A+\operatorname{Im}\mathcal{B})=\dim\operatorname{Im}\mathcal{A}+\dim\operatorname{Im}\mathcal{B}-\dim\left(\operatorname{Im}\mathcal{A}\cap\operatorname{Im}\mathcal{B}\right)$$

结合 (2.9) 可得

$$\dim\left(\operatorname{Im}\mathcal{A}\cap\operatorname{Im}\mathcal{B}\right)=0.$$

于是

$$\operatorname{Im}\mathcal{A}\cap\operatorname{Im}\mathcal{B}=\{0\}.$$

于是

$$\dim\operatorname{Im}(\mathcal{A}+\mathcal{B})=\dim(\operatorname{Im}\mathcal{A}+\operatorname{Im}\mathcal{B})=\dim\operatorname{Im}\mathcal{A}+\dim\operatorname{Im}\mathcal{B}. \tag{2.10}$$

也即

$$\operatorname{Im}(\mathcal{A}+\mathcal{B})=\operatorname{Im}\mathcal{A}+\operatorname{Im}\mathcal{B}.$$

由于

$$\dim\operatorname{Im}\mathcal{A}=\operatorname{rank}\boldsymbol{A}=r \quad \dim\operatorname{Im}\mathcal{B}=\operatorname{rank}\boldsymbol{B}=s,$$

我们有

$$r+s=\dim\operatorname{Im}(\mathcal{A}+\mathcal{B})\leq\dim\left(\mathbb{F}^m\right)=m.$$

又因为

$$\begin{aligned} n &= \dim\operatorname{Ker}\mathcal{A}+\dim\operatorname{Im}\mathcal{A} \\ &= \dim\operatorname{Ker}\mathcal{B}+\dim\operatorname{Im}\mathcal{B} \\ &= \dim\operatorname{Ker}(\mathcal{A}+\mathcal{B})+\dim\operatorname{Im}(\mathcal{A}+\mathcal{B}), \end{aligned}$$

带入 (2.10) 可得

$$n-\dim\operatorname{Ker}(\mathcal{A}+\mathcal{B})=n-\dim\operatorname{Ker}\mathcal{A}+n-\dim\operatorname{Ker}\mathcal{B},$$

也即

$$n + \dim \operatorname{Ker}(\mathcal{A} + \mathcal{B}) = \dim \operatorname{Ker} \mathcal{A} + \dim \operatorname{Ker} \mathcal{B}.$$

带入对应的取值得到

$$n + \dim \operatorname{Ker}(\mathcal{A} + \mathcal{B}) = n - r + n - s,$$

于是

$$n - (r + s) = \dim \operatorname{Ker}(\mathcal{A} + \mathcal{B}) \geq 0.$$

从而

$$r + s \leq n.$$

综上有

$$r + s \leq \min\{m, n\}.$$

由于

$$\operatorname{Im}(\mathcal{A} + \mathcal{B}) = \operatorname{Im} \mathcal{A} \oplus \operatorname{Im} \mathcal{B},$$

我们可以选择 $\operatorname{Im} \mathcal{A}$ 的一组基

$$\{\mathcal{A}(\boldsymbol{\alpha}_1), \mathcal{A}(\boldsymbol{\alpha}_2), \cdots, \mathcal{A}(\boldsymbol{\alpha}_r)\}, \quad \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r \in \mathbb{F}^n,$$

和 $\operatorname{Im} \mathcal{B}$ 的一组基

$$\{\mathcal{B}(\boldsymbol{\beta}_1), \mathcal{B}(\boldsymbol{\beta}_2), \cdots, \mathcal{B}(\boldsymbol{\beta}_s)\}, \quad \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s \in \mathbb{F}^n$$

组成 $\operatorname{Im}(\mathcal{A} + \mathcal{B})$ 的一组基

$$\{\mathcal{A}(\boldsymbol{\alpha}_1), \mathcal{A}(\boldsymbol{\alpha}_2), \cdots, \mathcal{A}(\boldsymbol{\alpha}_r), \mathcal{B}(\boldsymbol{\beta}_1), \mathcal{B}(\boldsymbol{\beta}_2), \cdots, \mathcal{B}(\boldsymbol{\beta}_s)\}.$$

从而存在

$$\delta_1, \delta_2, \cdots, \delta_r, \theta_1, \theta_2, \cdots, \theta_s \in \mathbb{F}^n$$

使得

$$\begin{cases} (\mathcal{A} + \mathcal{B})(\delta_k) = \mathcal{A}(\boldsymbol{\alpha}_k) & k = 1, 2, \cdots, r \\ (\mathcal{A} + \mathcal{B})(\theta_l) = \mathcal{B}(\boldsymbol{\beta}_l) & l = 1, 2, \cdots, s. \end{cases}$$

也即

$$\begin{cases} \mathcal{A}(\delta_k) - \mathcal{A}(\boldsymbol{\alpha}_k) = -\mathcal{B}(\delta_k) \in \operatorname{Im} \mathcal{A} \cap \operatorname{Im} \mathcal{B} = \{0\} & k = 1, 2, \cdots, r \\ \mathcal{B}(\theta_l) - \mathcal{B}(\boldsymbol{\beta}_l) = -\mathcal{A}(\theta_l) \in \operatorname{Im} \mathcal{A} \cap \operatorname{Im} \mathcal{B} = \{0\} & l = 1, 2, \cdots, s. \end{cases}$$

于是

$$\begin{cases} \mathcal{A}(\delta_k) - \mathcal{A}(\boldsymbol{\alpha}_k) = -\mathcal{B}(\delta_k) = 0 & k = 1, 2, \cdots, r \\ \mathcal{B}(\theta_l) - \mathcal{B}(\boldsymbol{\beta}_l) = -\mathcal{A}(\theta_l) = 0 & l = 1, 2, \cdots, s. \end{cases}$$

也即

$$\begin{cases} \mathcal{A}(\delta_k) = \mathcal{A}(\boldsymbol{\alpha}_k) & \delta_k \in \operatorname{Ker} \mathcal{B} \quad k = 1, 2, \cdots, r \\ \mathcal{B}(\theta_l) = \mathcal{B}(\boldsymbol{\beta}_l) & \theta_l \in \operatorname{Ker} \mathcal{A} \quad l = 1, 2, \cdots, s. \end{cases} \tag{2.11}$$

再取 $\operatorname{Ker} \mathcal{A} \cap \operatorname{Ker} \mathcal{B}$ 的一组基

$$\{\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \cdots, \boldsymbol{\gamma}_p\}. \tag{2.12}$$

则我们声称

$$\{\theta_1, \theta_2, \cdots, \theta_s, \boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \cdots, \boldsymbol{\gamma}_p, \delta_1, \delta_2, \cdots, \delta_r\}$$

组成 $\mathbb{F}^n$ 的一组基。这是因为

(1) 线性无关。假定

$$\sum_{k=1}^{r} x_k \delta_k + \sum_{l=1}^{s} y_l \theta_l + \sum_{j=1}^{p} z_j \boldsymbol{\gamma}_j = 0. \tag{2.13}$$

用 $\mathcal{A}$ 作用 (2.13)，根据(2.12) 和 (2.11)，有

$$0 = \mathcal{A}\left(\sum_{k=1}^{r} x_k \delta_k + \sum_{l=1}^{s} y_l \theta_l + \sum_{j=1}^{p} z_j \boldsymbol{\gamma}_j\right) = \sum_{k=1}^{r} x_k \mathcal{A}(\delta_k) + \sum_{l=1}^{s} y_l \mathcal{A}(\theta_l) + \sum_{j=1}^{p} z_j \mathcal{A}(\boldsymbol{\gamma}_j) = \sum_{k=1}^{r} x_k \mathcal{A}(\delta_k).$$

又因为 (2.11) 和

$$\{\mathcal{A}(\boldsymbol{\alpha}_1), \mathcal{A}(\boldsymbol{\alpha}_2), \cdots, \mathcal{A}(\boldsymbol{\alpha}_r)\}, \quad \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r \in \mathbb{F}^n,$$

是 $\operatorname{Im} \mathcal{A}$ 的基，从而有

$$x_1 = x_2 = \cdots = x_r = 0.$$

用 $\mathcal{B}$ 作用 (2.13)，根据(2.12) 和 (2.11)，有

$$0 = \mathcal{B}\left(\sum_{k=1}^{r} x_k \delta_k + \sum_{l=1}^{s} y_l \theta_l + \sum_{j=1}^{p} z_j \boldsymbol{\gamma}_j\right) = \sum_{k=1}^{r} x_k \mathcal{B}(\delta_k) + \sum_{l=1}^{s} y_l \mathcal{B}(\theta_l) + \sum_{j=1}^{p} z_j \mathcal{B}(\boldsymbol{\gamma}_j) = \sum_{l=1}^{s} y_l \mathcal{B}(\theta_l).$$

又因为 (2.11) 和

$$\{\mathcal{B}(\boldsymbol{\beta}_1), \mathcal{B}(\boldsymbol{\beta}_2), \cdots, \mathcal{B}(\boldsymbol{\beta}_s)\}, \quad \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s \in \mathbb{F}^n$$

是 $\operatorname{Im} \mathcal{B}$ 的基，从而有

$$y_1 = y_2 = \cdots = y_s = 0.$$

于是根据假设 (2.13) 可得

$$\sum_{j=1}^{p} z_j \boldsymbol{\gamma}_j = 0.$$

再结合 (2.12) 可得

$$z_1 = z_2 = \cdots = z_p = 0.$$

所以

$$\{\theta_1, \theta_2, \cdots, \theta_s, \boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \cdots, \boldsymbol{\gamma}_p, \delta_1, \delta_2, \cdots, \delta_r\}$$

线性无关。

(2) 任何一个 $\mathbb{F}^n$ 中的向量可被

$$\{\theta_1,\theta_2,\cdots,\theta_s,\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\cdots,\boldsymbol{\gamma}_p,\delta_1,\delta_2,\cdots,\delta_r\}$$

线性表示。证明如下：任取 $\boldsymbol{\alpha}\in\mathbb{F}^n$，则 $\mathcal{A}(\boldsymbol{\alpha})+\mathcal{B}(\boldsymbol{\alpha})\in\mathrm{Im}\,(\mathcal{A}+\mathcal{B})$。又因为

$$\{\mathcal{A}(\boldsymbol{\alpha}_1),\mathcal{A}(\boldsymbol{\alpha}_2),\cdots,\mathcal{A}(\boldsymbol{\alpha}_r),\mathcal{B}(\boldsymbol{\beta}_1),\mathcal{B}(\boldsymbol{\beta}_2),\cdots,\mathcal{B}(\boldsymbol{\beta}_s)\}$$

是 $\mathrm{Im}\,(\mathcal{A}+\mathcal{B})$ 的一组基，我们有存在 $x_1,\cdots,x_r,y_1,\cdots,y_s\in\mathbb{F}$ 使得

$$\mathcal{A}(\boldsymbol{\alpha})+\mathcal{B}(\boldsymbol{\alpha})=\sum_{k=1}^{s}x_k\mathcal{A}(\boldsymbol{\alpha}_k)+\sum_{l=1}^{s}y_l\mathcal{B}(\boldsymbol{\beta}_l).$$

再利用 (2.11) 可得

$$\mathcal{A}(\boldsymbol{\alpha})+\mathcal{B}(\boldsymbol{\alpha})=\sum_{k=1}^{s}x_k\mathcal{A}(\delta_k)+\sum_{l=1}^{s}y_l\mathcal{B}(\theta_l).$$

于是

$$\mathcal{A}\left(\boldsymbol{\alpha}-\sum_{k=1}^{r}x_k\delta_k\right)=-\mathcal{B}\left(\boldsymbol{\alpha}-\sum_{l=1}^{s}y_l\theta_l\right)\in\mathrm{Im}\,\mathcal{A}\cap\mathrm{Im}\,\mathcal{B}=\{0\}.$$

于是

$$\boldsymbol{\alpha}-\sum_{k=1}^{r}x_k\delta_k\in\mathrm{Ker}\,\mathcal{A},\quad\boldsymbol{\alpha}-\sum_{l=1}^{s}y_l\theta_l\in\mathrm{Ker}\,\mathcal{B}.$$

于是存在 $\boldsymbol{\beta}\in\mathrm{Ker}\,\mathcal{A}$ 和 $\boldsymbol{\gamma}\in\mathrm{Ker}\,\mathcal{B}$ 使得

$$\boldsymbol{\alpha}-\sum_{k=1}^{r}x_k\delta_k=\boldsymbol{\beta},\quad\boldsymbol{\alpha}-\sum_{l=1}^{s}y_l\theta_l=\boldsymbol{\gamma}.$$

从而

$$\boldsymbol{\alpha}=\sum_{k=1}^{r}x_k\delta_k+\boldsymbol{\beta}=\sum_{l=1}^{s}y_l\theta_l+\boldsymbol{\gamma}.$$

于是根据 (2.11) 和 $\boldsymbol{\beta}\in\mathrm{Ker}\,\mathcal{A}$ 可得

$$\mathcal{A}\left(\sum_{l=1}^{s}y_l\theta_l+\boldsymbol{\gamma}\right)=\mathcal{A}(\boldsymbol{\gamma})=\mathcal{A}\left(\sum_{k=1}^{r}x_k\delta_k+\boldsymbol{\beta}\right)=\sum_{k=1}^{r}x_k\mathcal{A}(\delta_k).$$

也即

$$\mathcal{A}\left(\boldsymbol{\gamma}-\sum_{k=1}^{r}x_k\delta_k\right)=0.$$

于是

$$\boldsymbol{\gamma}-\sum_{k=1}^{r}x_k\delta_k\in\mathrm{Ker}\,\mathcal{A}.$$

又因为 $\boldsymbol{\gamma}\in\mathrm{Ker}\,\mathcal{B}$ 和 (2.11) (推出 $\delta_k\in\mathrm{Ker}\,\mathcal{B}$)，从而

$$\boldsymbol{\gamma}-\sum_{k=1}^{r}x_k\delta_k\in\mathrm{Ker}\,\mathcal{B}.$$

于是

$$\boldsymbol{\gamma}-\sum_{k=1}^{r}x_k\delta_k\in\mathrm{Ker}\,\mathcal{A}\cap\mathrm{Ker}\,\mathcal{B}.$$

根据 (2.12)

$$\{\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \cdots, \boldsymbol{\gamma}_p\}.$$

为 $\operatorname{Ker} \mathcal{A} \cap \operatorname{Ker} \mathcal{B}$ 的基，于是存在 $z_1, z_2, \cdots, z_p \in \mathbb{F}$ 使得

$$\boldsymbol{\gamma} - \sum_{k=1}^{r} x_k \delta_k = \sum_{j=1}^{p} z_j \boldsymbol{\gamma}_j.$$

从而

$$\boldsymbol{\alpha} = \sum_{l=1}^{s} y_l \theta_l + \boldsymbol{\gamma} = \sum_{l=1}^{s} y_l \theta_l + \sum_{k=1}^{r} x_k \delta_k + \sum_{j=1}^{p} z_j \boldsymbol{\gamma}_j.$$

综上则我们声称

$$\{\theta_1, \theta_2, \cdots, \theta_s, \boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \cdots, \boldsymbol{\gamma}_p, \delta_1, \delta_2, \cdots, \delta_r\}$$

组成 $\mathbb{F}^n$ 的一组基。我们将是 $\operatorname{Im}(\mathcal{A}+\mathcal{B})$ 的一组基

$$\{\mathcal{A}(\boldsymbol{\alpha}_1), \mathcal{A}(\boldsymbol{\alpha}_2), \cdots, \mathcal{A}(\boldsymbol{\alpha}_r), \mathcal{B}(\boldsymbol{\beta}_1), \mathcal{B}(\boldsymbol{\beta}_2), \cdots, \mathcal{B}(\boldsymbol{\beta}_s)\}$$

扩充为 $\mathbb{F}^m$ 的一组基

$$\{\mathcal{A}(\boldsymbol{\alpha}_1), \mathcal{A}(\boldsymbol{\alpha}_2), \cdots, \mathcal{A}(\boldsymbol{\alpha}_r), \mathcal{B}(\boldsymbol{\beta}_1), \mathcal{B}(\boldsymbol{\beta}_2), \cdots, \mathcal{B}(\boldsymbol{\beta}_s), \eta_1, \eta_2, \cdots \eta_{m-r-s}.\}$$

注意到(2.11)

$$\mathcal{A}(\boldsymbol{\alpha}_k) = \mathcal{A}(\delta_k), k = 1, \cdots, r, \quad \mathcal{B}(\beta_l) = \mathcal{B}(\theta_l), l = 1, \cdots, s.$$

从而 $\mathcal{A}$ 在 $\mathbb{F}^n$ 的基

$$\{\delta_1, \cdots, \delta_r, \boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \cdots, \boldsymbol{\gamma}_p, \theta_1, \theta_2, \cdots, \theta_s, \}$$

和 $\mathbb{F}^m$ 的基

$$\{\mathcal{A}(\delta_1), \mathcal{A}(\delta_2), \cdots, \mathcal{A}(\delta_r), \eta_1, \eta_2, \cdots \eta_{m-r-s}, \mathcal{B}(\theta_1), \mathcal{B}(\theta_2), \cdots, \mathcal{B}(\theta_s)\}$$

下的矩阵表示为

$$\begin{pmatrix} I_{(r)} & 0 \\ 0 & 0 \end{pmatrix}$$

同理 $\mathcal{B}$ 在 $\mathbb{F}^n$ 的基

$$\{\delta_1, \cdots, \delta_r, \boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \cdots, \boldsymbol{\gamma}_p, \theta_1, \theta_2, \cdots, \theta_s, \}$$

和 $\mathbb{F}^m$ 的基

$$\{\mathcal{A}(\delta_1), \mathcal{A}(\delta_2), \cdots, \mathcal{A}(\delta_r), \eta_1, \eta_2, \cdots \eta_{m-r-s}, \mathcal{B}(\theta_1), \mathcal{B}(\theta_2), \cdots, \mathcal{B}(\theta_s)\}$$

下的矩阵表示为

$$\begin{pmatrix} 0 & 0 \\ 0 & I_{(s)} \end{pmatrix}$$

利用线性映射在不同基下的矩阵表示是相抵的便得到了我们要的证明。

练习 2.25 设 σ 是数域 $\mathbb{F}$ 上 n 维向量空间 V 的一个线性变换. 证明, 总可以如此选取 V 的两个基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 和 $\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n\}$, 使得对于 V 的任意向量 ξ 来说，如果 $\xi = \sum_{i=1}^{n} x_i \boldsymbol{\alpha}_i$,

则 $\sigma(\xi)=\sum_{i=1}^{r}x_i\boldsymbol{\beta}_i$, 这里 $0\leqslant r\leqslant n$ 是一个定数。你可以用习题 2.20 来证明。

解 习题 2.20 的特例。

✍ **练习 2.26** 设 $\boldsymbol{A},\boldsymbol{B}$ 是 n 阶矩阵, 且 $\boldsymbol{A}$ 可逆. 证明: $\boldsymbol{AB}$ 与 $\boldsymbol{BA}$ 相似.

解 注意到

$$\boldsymbol{AB}=\boldsymbol{A}(\boldsymbol{BA})\boldsymbol{A}^{-1}.$$

✍ **练习 2.27** 设 $\dim V=n$. 假设 $\mathcal{A}:V\mapsto V$ 是一个线性变换。证明下面条件是等价的.

(1) $\mathcal{A}=k\,\mathrm{id}$. 这里 id 是恒等映射。

(2) $\mathcal{A}\mathcal{B}=\mathcal{B}\mathcal{A}$, 对所有的线性变换 $\mathcal{B}:V\mapsto V$ 都成立.

(3) $\mathcal{A}$ 在任何基下的矩阵都相同.

解

- $(1)\Rightarrow(2)$ 是显然的。
- $(2)\Rightarrow(3)$。任取两组基

$$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\},\quad 和 \quad \{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n\}$$

假定 $\mathcal{A}$ 在这两组基下的矩阵表示分别为 $\boldsymbol{A}$ 和 $\widetilde{\boldsymbol{A}}$, 也即

$$\mathcal{A}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)\boldsymbol{A},\quad \mathcal{A}(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n)=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n)\widetilde{\boldsymbol{A}}.$$

现在任取 $1\leq i_0,j_0\leq n$, 定义一个新的线性变换 $\mathcal{B}:V\mapsto V$ 如下

$$\begin{cases}\mathcal{B}(\boldsymbol{\alpha}_i)=0 & i=1,2,\cdots,i_0-1,i_0+1,\cdots,n\\ \mathcal{B}(\boldsymbol{\alpha}_{i_0})=\boldsymbol{\beta}_{j_0}\end{cases}$$

于是

$$\begin{aligned}\mathcal{B}\mathcal{A}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)&=\mathcal{B}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)\boldsymbol{A}\\&=(\mathcal{B}(\boldsymbol{\alpha}_1),\mathcal{B}(\boldsymbol{\alpha}_2),\cdots,\mathcal{B}(\boldsymbol{\alpha}_n))\boldsymbol{A}\\&=\underbrace{(0,0,\cdots,\boldsymbol{\beta}_{j_0},0\cdots,0)}_{\text{在第 } i_0 \text{ 个位置}}\boldsymbol{A}\\&=\left(a_{i_01}\boldsymbol{\beta}_{j_0},a_{i_02}\boldsymbol{\beta}_{j_0},\cdots,a_{i_0n}\boldsymbol{\beta}_{j_0}\right).\end{aligned}$$

又因为

$$\begin{aligned}\mathcal{A}\mathcal{B}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)&=\mathcal{A}(\mathcal{B}(\boldsymbol{\alpha}_1),\mathcal{B}(\boldsymbol{\alpha}_2),\cdots,\mathcal{B}(\boldsymbol{\alpha}_n))\\&=\mathcal{A}\underbrace{(0,0,\cdots,\boldsymbol{\beta}_{j_0},0\cdots,0)}_{\text{在第 } i_0 \text{ 个位置}}\\&=\underbrace{(0,0,\cdots,\mathcal{A}(\boldsymbol{\beta}_{j_0}),0\cdots,0)}_{\text{在第 } i_0 \text{ 个位置}}\end{aligned}$$

因为 $\mathcal{A}\mathcal{B}=\mathcal{B}\mathcal{A}$, 于是我们有

$$\underbrace{(0,0,\cdots,\mathcal{A}(\boldsymbol{\beta}_{j_0}),0\cdots,0)}_{\text{在第 } i_0 \text{ 个位置}}=\left(a_{i_01}\boldsymbol{\beta}_{j_0},a_{i_02}\boldsymbol{\beta}_{j_0},\cdots,a_{i_0n}\boldsymbol{\beta}_{j_0}\right).$$

从而

$$\begin{cases} a_{i_0 j}\boldsymbol{\beta}_{j_0} = 0 & j \neq i_0 \\ \mathcal{A}(\boldsymbol{\beta}_{j_0}) = a_{i_0 i_0}\boldsymbol{\beta}_{j_0} \end{cases}$$

由于 $\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n\}$ 是基从而不为 0, 于是

$$\begin{cases} a_{i_0 j} = 0 & j \neq i_0 \\ a_{i_0 i_0}\boldsymbol{\beta}_{j_0} = \mathcal{A}(\boldsymbol{\beta}_{j_0}) = \sum_{k=1}^n b_{k j_0}\boldsymbol{\beta}_k. \end{cases}$$

再利用 $\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n\}$ 的线性无关性可得

$$\begin{cases} a_{i_0 j} = 0 & j \neq i_0 \\ b_{k j_0} = 0 & k \neq j_0 \\ b_{j_0 j_0} = a_{i_0 i_0} \end{cases}$$

由 i_0, j_0 的任意性可得 $\boldsymbol{A} = \boldsymbol{B} = \lambda I_n$.

- $(3) \Rightarrow (1)$。任取一组基

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

假定 $\mathcal{A}$ 在这组基下的矩阵都是 $\boldsymbol{A}$。也即

$$\mathcal{A}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n)\boldsymbol{A}.$$

任取 $1 \leq i_0 \neq j_0 \leq n$, 构造一组新的基 $\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n\}$ 如下

$$\begin{cases} \boldsymbol{\beta}_i = \boldsymbol{\alpha}_i & i \neq i_0, j_0 \\ \boldsymbol{\beta}_{i_0} = \boldsymbol{\alpha}_{j_0} \\ \boldsymbol{\beta}_{j_0} = -\boldsymbol{\alpha}_{i_0} \end{cases}$$

则

$$\begin{aligned} \mathcal{A}(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n) &= \mathcal{A}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_{i_0-1}, \boldsymbol{\alpha}_{j_0}, \boldsymbol{\alpha}_{i_0+1}, \cdots, \boldsymbol{\alpha}_{j_0-1}, -\boldsymbol{\alpha}_{i_0}, \boldsymbol{\alpha}_{j_0+1}, \cdots \boldsymbol{\alpha}_n) \\ &= (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_{i_0-1}, \boldsymbol{\alpha}_{j_0}, \boldsymbol{\alpha}_{i_0+1}, \cdots, \boldsymbol{\alpha}_{j_0-1}, -\boldsymbol{\alpha}_{i_0}, \boldsymbol{\alpha}_{j_0+1}, \cdots \boldsymbol{\alpha}_n)\boldsymbol{B} \end{aligned}$$

这里 $\boldsymbol{B}$ 是将 $\boldsymbol{A}$ 的第 i_0 行第 i_0 列乘以 -1 后再将 i_0 行和 j_0 行互换，再将 i_0 列和 j_0 列互换得到的矩阵。根据假设 $\boldsymbol{B} = \boldsymbol{A}$, 也就是说将 $\boldsymbol{A}$ 的第 i_0 行第 i_0 列乘以 -1 后，再将 i_0 行和 j_0 行互换，再将 i_0 列和 j_0 列互换后保持不变。这就得到

$$\begin{cases} -a_{i_0 k} = a_{j_0 k} & k \neq i_0, j_0 \\ -a_{k i_0} = a_{k j_0} & k \neq i_0, j_0 \\ a_{i_0 i_0} = a_{j_0 j_0} \\ a_{i_0 j_0} = -a_{j_0 i_0} \end{cases} \tag{2.14}$$

再构造一组新的基 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n\}$ 如下

$$\begin{cases}\boldsymbol{\beta}_i=\boldsymbol{\alpha}_i & i\neq i_0,j_0\\ \boldsymbol{\beta}_{i_0}=\boldsymbol{\alpha}_{j_0}\\ \boldsymbol{\beta}_{j_0}=\boldsymbol{\alpha}_{i_0}\end{cases}\tag{2.15}$$

则

$$\begin{aligned}&\mathcal{A}(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n)=\mathcal{A}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{i_0-1},\boldsymbol{\alpha}_{j_0},\boldsymbol{\alpha}_{i_0+1},\cdots,\boldsymbol{\alpha}_{j_0-1},\boldsymbol{\alpha}_{i_0},\boldsymbol{\alpha}_{j_0+1},\cdots\boldsymbol{\alpha}_n)\\ =&(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{i_0-1},\boldsymbol{\alpha}_{j_0},\boldsymbol{\alpha}_{i_0+1},\cdots,\boldsymbol{\alpha}_{j_0-1},\boldsymbol{\alpha}_{i_0},\boldsymbol{\alpha}_{j_0+1},\cdots\boldsymbol{\alpha}_n)\boldsymbol{B}\end{aligned}$$

这里 $\boldsymbol{B}$ 是将 $\boldsymbol{A}$ 的 i_0 行和 j_0 行互换, 再将 i_0 列和 j_0 列互换得到的矩阵。根据假设 $\boldsymbol{B}=\boldsymbol{A}$, 也就是说将 $\boldsymbol{A}$ 的 i_0 行和 j_0 行互换，再将 i_0 列和 j_0 列互换后保持不变。这就得到

$$\begin{cases}a_{i_0k}=a_{j_0k} & k\neq i_0,j_0\\ a_{ki_0}=a_{kj_0} & k\neq i_0,j_0\\ a_{i_0i_0}=a_{j_0j_0}\\ a_{i_0j_0}=a_{j_0i_0}\end{cases}\tag{2.16}$$

结合(2.14) 和 (2.16) 可得

$$\begin{cases}a_{i_0k}=a_{j_0k}=0 & k\neq i_0,j_0\\ a_{ki_0}=a_{kj_0}=0 & k\neq i_0,j_0\\ a_{i_0i_0}=a_{j_0j_0}\\ a_{i_0j_0}=a_{j_0i_0}=0.\end{cases}$$

根据 $i_0\neq j_0$ 的任意性可得 $\boldsymbol{A}=\lambda I_n$, 从而 $\mathcal{A}=\lambda\,\mathrm{id}$.

- $(3)\Rightarrow(1)$ 还可以更简单地选其他的基。我们详细如下说。任取一组基

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

假定 $\mathcal{A}$ 在这组基下的矩阵都是 $\boldsymbol{A}$。也即

$$\mathcal{A}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)\boldsymbol{A}.$$

任取 $1\leq i_0\leq n$, 构造一组新的基 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n\}$ 如下

$$\begin{cases}\boldsymbol{\beta}_i=\boldsymbol{\alpha}_i & i\neq i_0\\ \boldsymbol{\beta}_{i_0}=-\boldsymbol{\alpha}_{i_0}\end{cases}$$

则

$$\begin{aligned}\mathcal{A}(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n)=&\mathcal{A}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{i_0-1},-\boldsymbol{\alpha}_{i_0},\boldsymbol{\alpha}_{i_0+1},\cdots,\boldsymbol{\alpha}_n)\\ =&(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{i_0-1},-\boldsymbol{\alpha}_{i_0},\boldsymbol{\alpha}_{i_0+1},\cdots,\boldsymbol{\alpha}_n)\boldsymbol{B}\end{aligned}$$

这里 $\boldsymbol{B}$ 是将 $\boldsymbol{A}$ 的第 i_0 行乘以 -1 和 i_0 列乘以 -1 得到的矩阵。根据假设 $\boldsymbol{A}=\boldsymbol{B}$, 于是

有

$$\begin{cases} a_{i_0k} = -a_{i_0k} & k \neq i_0 \\ a_{ki_0} = -a_{ki_0} & k \neq i_0 \end{cases}$$

从而

$$\begin{cases} a_{i_0k} = 0 & k \neq i_0 \\ a_{ki_0} = 0 & k \neq i_0 \end{cases}$$

于是 $\boldsymbol{A}$ 是对角矩阵 $\boldsymbol{A} = \operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$，也即

$$\mathcal{A}(\boldsymbol{\alpha}_k) = \lambda_k \boldsymbol{\alpha}_k, \quad k = 1, 2, 3, \cdot, n.$$

交换基的顺序

$$\mathcal{A}(\boldsymbol{\alpha}_{i_1}, \boldsymbol{\alpha}_{i_2}, \cdots, \boldsymbol{\alpha}_{i_n}) = (\boldsymbol{\alpha}_{i_1}, \boldsymbol{\alpha}_{i_2}, \cdots, \boldsymbol{\alpha}_{i_n}) \operatorname{diag}(\lambda_{i_1}, \lambda_{i_2}, \cdots, \lambda_{i_n}).$$

于是有

$$\operatorname{diag}(\lambda_{i_1}, \lambda_{i_2}, \cdots, \lambda_{i_n}) = \operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$$

对任意的 $\{1, 2, \cdots, n\}$ 的排列都对。于是有

$$\lambda_1 = \lambda_2 = \cdots = \lambda_n.$$

2.5 2024 年 4 月 23 日第八周星期二第十四次课-不变子空间

给定线性变换 $\mathscr{A} : V \to V$.

练习 2.28 如果 U 是 V 的子空间, 则 $\mathscr{A}(U) = \{\mathscr{A}(\boldsymbol{\alpha}) : \boldsymbol{\alpha} \in U\}$ 是 V 的子空间.

定义 2.19

记 V 的所有子空间集合为 S.

1. 定义映射 $\eta_{\mathscr{A}} : S \to S$ 如下: 设 $U \in S$, 则令 $\eta_{\mathscr{A}}(U) = \mathscr{A}(U)$.
2. 映射 $\eta_{\mathscr{A}} : S \to S$ 称为由线性变换 $\mathscr{A}$ 诱导的映射。

定义 2.20 (不变子空间)

设 $\mathscr{A} : V \to V$ 是线性变换, U 是 V 的子空间. 如果对任意 $\boldsymbol{\alpha} \in U$, 即有 $\mathscr{A}(\boldsymbol{\alpha}) \in U$, 即 $\mathscr{A}(U) \subseteq U$, 则 U 称为 $\mathscr{A}$ 线性变换 $\mathscr{A}$ 的不变子空间.

例题 2.9 两个平凡的不变子空间

- 对线性空间 V 自身, $\mathscr{A}(V) \subseteq V$. 因此 V 是线性变换 $\mathscr{A}$ 的不变子空间.
- 对 V 的零子空间 $\{0\}$, $\mathscr{A}(0) = 0$, 因此零子空间 $\{0\}$ 是线性变换 $\mathscr{A}$ 的不变子空间.
- 它们称为 $\mathscr{A}$ 的平凡不变子空间, 除 V 本身与零子空间外, 其他不变子空间称为 $\mathscr{A}$ 的非平凡不变子空间.

例题 2.10 另外两个不变子空间

- 由于 $\mathscr{A}(V)\subseteq V$, 故 $\mathscr{A}(\mathscr{A}(V))\subseteq\mathscr{A}(V)$, 因此线性变换 $\mathscr{A}$ 的像 $\mathscr{A}(V)$ 是 $\mathscr{A}$ 的不变子空间;
- 设 $\boldsymbol{\alpha}\in\operatorname{Ker}(\mathscr{A})$, 则 $\mathscr{A}(\boldsymbol{\alpha})=0$, 因此 $\mathscr{A}(\mathscr{A}(\boldsymbol{\alpha}))=0$, 所以 $\mathscr{A}(\boldsymbol{\alpha})\in\operatorname{Ker}(\mathscr{A})$. 因此线性变换 $\mathscr{A}$ 的核 $\operatorname{Ker}(\mathscr{A})$ 也是 $\mathscr{A}$ 的不变子空间.

定理 2.23 (有限和与任意交也是不变子空间)

线性变换 $\mathscr{A}$ 的有限多个不变子空间之和是 $\mathscr{A}$ 的不变子空间; $\mathscr{A}$ 的任意多个不变子空间之交是 $\mathscr{A}$ 的不变子空间.

定理 2.24

设 $\mathscr{A}:V\to V$ 是线性变换, U 是 $\mathscr{A}$ 的不变子空间. 设 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_k\}$ 是 U 的基, 则 $\mathscr{A}$ 在 V 的基 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_k,\boldsymbol{\alpha}_{k+1},\cdots,\boldsymbol{\alpha}_n\}$ 下的方阵是准上三角的.

因为 U 是 $\mathscr{A}$ 的不变子空间, 所以对 $1\leqslant j\leqslant k$, $\mathscr{A}(\boldsymbol{\alpha}_j)\in U$, 由于 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$ 是 U 的基, 因此

$$
\begin{gathered}
\mathscr{A}(\boldsymbol{\alpha}_1)=a_{11}\boldsymbol{\alpha}_1+a_{12}\boldsymbol{\alpha}_2+\cdots+a_{1k}\boldsymbol{\alpha}_k,\\
\mathscr{A}(\boldsymbol{\alpha}_2)=a_{21}\boldsymbol{\alpha}_1+a_{22}\boldsymbol{\alpha}_2+\cdots+a_{2k}\boldsymbol{\alpha}_k,\\
\cdots\cdots\cdots\cdots\\
\mathscr{A}(\boldsymbol{\alpha}_k)=a_{k1}\boldsymbol{\alpha}_1+a_{k2}\boldsymbol{\alpha}_2+\cdots+a_{kk}\boldsymbol{\alpha}_k.
\end{gathered}
$$

而当 $k+1\leqslant j\leqslant n$ 时, 由于 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_k,\boldsymbol{\alpha}_{k+1},\cdots,\boldsymbol{\alpha}_n\}$ 为 V 的基, 故

$$
\begin{aligned}
&\mathscr{A}(\boldsymbol{\alpha}_{k+1})=a_{k+1,1}\boldsymbol{\alpha}_1+a_{k+1,2}\boldsymbol{\alpha}_2+\cdots+a_{k+1,k}\boldsymbol{\alpha}_k+a_{k+1,k+1}\boldsymbol{\alpha}_{k+1}+\cdots+a_{k+1,n}\boldsymbol{\alpha}_n,\\
&\mathscr{A}(\boldsymbol{\alpha}_n)=a_{n1}\boldsymbol{\alpha}_1+a_{n2}\boldsymbol{\alpha}_2+\cdots+a_{nk}\boldsymbol{\alpha}_k+a_{n,k+1}\boldsymbol{\alpha}_{k+1}+\cdots+a_{nn}\boldsymbol{\alpha}_n.
\end{aligned}
$$

因此

$$
\mathscr{A}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)\begin{pmatrix}
a_{11} & \cdots & a_{k1} & a_{k+1,1} & \cdots & a_{n1}\\
\vdots & & \vdots & \vdots & & \vdots\\
a_{1k} & \cdots & a_{kk} & a_{k+1,k} & \cdots & a_{nk}\\
0 & \cdots & 0 & a_{k+1,k+1} & \cdots & a_{n,k+1}\\
\vdots & & \vdots & \vdots & & \vdots\\
0 & \cdots & 0 & a_{k+1,n} & \cdots & a_{nn}
\end{pmatrix}. \tag{2.17}
$$

定理 2.25

设线性变换 $\mathscr{A}:V\to V$ 在 V 的基 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$ 下的方阵为准上三角的, 即式(2.17)成立. 则由向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_k$ 生成的子空间 U 是 $\mathscr{A}$ 的不变子空间.

定义 2.21 (线性变换在不变子空间的限制)

设 U 是线性变换 $\mathscr{A}$ 的不变子空间. 定义映射 $\mathscr{A}|_U:U\to U$ 如下: 设 $\boldsymbol{\alpha}\in U$, 则令 $\mathscr{A}|_U(\boldsymbol{\alpha})=\mathscr{A}(\boldsymbol{\alpha})$. 映射 $\mathscr{A}|_U$ 称为线性变换 $\mathscr{A}$ 在不变子空间 U 上的限制.

- 由于 U 是线性变换 $\mathscr{A}$ 的不变子空间, 因此对 $\boldsymbol{\alpha}\in U$, $\mathscr{A}(\boldsymbol{\alpha})\in U$, 所以映射 $\mathscr{A}|_U$ 是有意

义的.

- 由于 $\mathscr{A}$ 是线性变换的, 因此 $\mathscr{A}|_U$ 也是线性变换.
- 定理 2.24说明, 如果 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_k\}$ 是 $\mathscr{A}$ 的不变子空间 U 的基, 且 $\mathscr{A}$ 在 V 的基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_k, \boldsymbol{\alpha}_{k+1}, \cdots, \boldsymbol{\alpha}_n\}$ 下的方阵为准上三角的, 即

$$\mathscr{A}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n) \begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \mathbf{0} & \boldsymbol{A}_{22} \end{pmatrix}, \tag{2.18}$$

其中 $\boldsymbol{A}_{11}$ 是 k 阶方阵, 则线性变换 $A|_U$ 在 U 的基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_k\}$ 下的方阵即为 $\boldsymbol{A}_{11}$. (2.18) 中 $n-k$ 阶方阵 A_{22} 的意义: 设 U 是线性变换 $\mathscr{A}: V \to V$ 的不变子空间, V/U 是 V 模 U 的商空间, V 中向量 $\boldsymbol{\alpha}$ 所属的模 U 同余类记为 $\tilde{\boldsymbol{\alpha}}$. 定义映射 $\mathscr{A}|_{(V/U)}: V/U \to V/U$ 如下: 设 $\tilde{\boldsymbol{\alpha}} \in V/U$, 则令 $\mathscr{A}|_{(V/U)}(\tilde{\boldsymbol{\alpha}}) = \widetilde{\mathscr{A}(\boldsymbol{\alpha})}$. 映射 $\mathscr{A}|_{(V/U)}$ 称为由线性变换 $\mathscr{A}$ 所诱导的商变换.

注意

如果向量 $\boldsymbol{\beta} \in \tilde{\boldsymbol{\alpha}}$, 则存在向量 $\boldsymbol{\gamma} \in U$, 使得 $\boldsymbol{\beta} = \boldsymbol{\alpha} + \boldsymbol{\gamma}$. 因此 $\mathscr{A}(\boldsymbol{\beta}) = \mathscr{A}(\boldsymbol{\alpha}) + \mathscr{A}(\boldsymbol{\gamma})$. 由于 U 是 $\mathscr{A}$ 的不变子空间, 把以 $\mathscr{A}(\boldsymbol{\gamma}) \in U$. 这表明, $\mathscr{A}(\boldsymbol{\beta}) \equiv \mathscr{A}(\boldsymbol{\alpha})(\bmod U)$, 也即 $\widetilde{\mathscr{A}(\boldsymbol{\beta})} = \widetilde{\mathscr{A}(\boldsymbol{\alpha})}$. 因此映射 $\mathscr{A}|_{V/U}$ 的定义不依赖于同余类 $\tilde{\boldsymbol{\alpha}}$ 的代表元的选取, 也就是说, 映射 $\mathscr{A}|_{(V/U)}$ 是有确切定义的.

定理 2.26

假定 U 是不变子空间 (也即 (2.18)成立). 则商变换 $\mathscr{A}|_{(V/U)}$ 是线性变换, 并且在基

$$\{\tilde{\boldsymbol{\alpha}}_{k+1}, \tilde{\boldsymbol{\alpha}}_{k+2}, \cdots, \tilde{\boldsymbol{\alpha}}_n\}$$

下的方阵为 $\boldsymbol{A}_{22}$. ♡

定理 2.27 (不变子空间分解)

设 U 与 W 是线性变换 $\mathscr{A}: V \to V$ 的不变子空间并且 $U \oplus W = V$, 且

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

与

$$\{\boldsymbol{\alpha}_{k+1}, \boldsymbol{\alpha}_{k+2}, \cdots, \boldsymbol{\alpha}_n\}$$

分别是 U 与 W 的基, 则 $\mathscr{A}$ 在 V 的基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_{k+1}, \boldsymbol{\alpha}_{k+2}, \cdots, \boldsymbol{\alpha}_n\}$ 下的方阵是准对角的, 即有

$$\begin{aligned} &\mathscr{A}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_k, \boldsymbol{\alpha}_{k+1}, \boldsymbol{\alpha}_{k+2}, \cdots, \boldsymbol{\alpha}_n) \\ =& (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_k, \boldsymbol{\alpha}_{k+1}, \boldsymbol{\alpha}_{k+2}, \cdots, \boldsymbol{\alpha}_n) \begin{pmatrix} \boldsymbol{A}_{11} & 0 \\ 0 & \boldsymbol{A}_{22} \end{pmatrix}, \end{aligned} \tag{2.19}$$

其中 $\boldsymbol{A}_{11}$ 与 $\boldsymbol{A}_{22}$ 分别是 k 阶与 $n-k$ 阶方阵. ♡

定理 2.28 (不变子空间分解)

设 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_k, \boldsymbol{\alpha}_{k+1}, \boldsymbol{\alpha}_{k+2}, \cdots, \boldsymbol{\alpha}_n\}$ 是 V 的基, 且 $\mathscr{A}$ 在此基下的方阵是准对角的, 即式 (2.19) 成立, 则 V 中分别由 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_k\}$ 与 $\{\boldsymbol{\alpha}_{k+1}, \boldsymbol{\alpha}_{k+2}, \cdots, \boldsymbol{\alpha}_n\}$ 生成的子空间 U 与 W 是 $\mathscr{A}$ 的不变子空间, 并且

$$\left.\mathscr{A}\right|_U (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_k) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_k) \boldsymbol{A}_{11},$$

$$\left.\mathscr{A}\right|_W (\boldsymbol{\alpha}_{k+1}, \boldsymbol{\alpha}_{k+2}, \cdots, \boldsymbol{\alpha}_n) = (\boldsymbol{\alpha}_{k+1}, \boldsymbol{\alpha}_{k+2}, \cdots, \boldsymbol{\alpha}_n) \boldsymbol{A}_{22}.$$

♡

在定理 2.27 的条件下, 线性变换 $\mathscr{A}$ 记为 $\mathscr{A} = \left.\mathscr{A}\right|_U \oplus \left.\mathscr{A}\right|_W$, 并称线性变换 $\mathscr{A}$ 分解为 $\mathscr{A}_U$ 与 $\mathscr{A}_W$ 的直和.

✍ **练习 2.29** 设 σ 是有限维向量空间 V 的一个线性变换. 而 W 是 σ 的一个不变子空间, 证明, 如果 σ 有逆变换, 那么 W 也在 σ^{-1} 之下不变.

解 如果 σ 可逆，则 σ 即是单射又是满射，从而是同构。因为 W 是不变子空间，那么 σ 可以限制在 W 上成为从 W 到 W 的线性映射，记为 $\sigma|_W$。同时有 $\mathrm{Ker}(\sigma|W) = \{0\}$。于是根据定理 2.22 可知

$$\sigma|_W : W \to W$$

是同构。而 σ 在 V 也是同构。那么任意的 $\alpha \in W$, 存在唯一的 $\beta \in V$ 使得 $\sigma(\beta) = \alpha$. 但 $\sigma|_W$ 是同构，则一定有 $\beta = \sigma^{-1}(\alpha) \in W$. 所以 $\sigma^{-1}(W) = W$。

✍ **练习 2.30** 设 σ, τ 是向量空间 V 的线性变换, 且 $\sigma\tau = \tau\sigma$. 证明 $\mathrm{Im}(\sigma)$ 和 $\mathrm{Ker}(\sigma)$ 都在 τ 之下不变.

解 任取 $\beta = \sigma(\alpha) \in \mathrm{Im}(\sigma)$, 则 $\tau(\beta) = \tau(\sigma(\alpha)) = \sigma(\tau(\alpha)) \in \mathrm{Im}(\sigma)$，从而 $\mathrm{Im}(\sigma)$ 在 τ 之下不变.

同理，如果取 $\alpha \in \mathrm{Ker}(\sigma)$，则 $\sigma(\alpha) = 0$, 则 $\sigma(\tau(\alpha)) = \tau(\sigma(\alpha)) = \tau(0) = 0$，从而 $\tau(\alpha) \in \mathrm{Ker}(\sigma)$。于是 $\mathrm{Ker}(\sigma)$ 在 τ 之下不变

✍ **练习 2.31** 令 σ 是数域 $\mathbb{F}$ 上向量空间 V 的一个线性变换, 并且满足条件 $\sigma^2 = \sigma$. 证明:

(i) $\mathrm{Ker}(\sigma) = \{\xi - \sigma(\xi) \mid \xi \in V\}$;

(ii) $V = \mathrm{Ker}(\sigma) \oplus \mathrm{Im}(\sigma)$;

(iii) 如果 τ 是 V 的一个线性变换, 那么 $\mathrm{Ker}(\sigma)$ 和 $\mathrm{Im}(\sigma)$ 都在 τ 之下不变的充要条件是 $\sigma\tau = \tau\sigma$.

解

(1) 首先，任取 $\xi \in V$, 则有

$$\sigma(\xi - \sigma(\xi)) = \sigma(\xi) - \sigma^2(\xi) = 0.$$

于是

$$\{\xi - \sigma(\xi) \mid \xi \in V\} \subseteq \mathrm{Ker}(\sigma).$$

反之，如果 $\sigma(\alpha) = 0$，则

$$\alpha = \alpha - 0 = \alpha - \sigma(\alpha) \in \{\xi - \sigma(\xi) \mid \xi \in V\}.$$

所以

$$\mathrm{Ker}(\sigma) \subseteq \{\xi - \sigma(\xi) \mid \xi \in V\}.$$

综上有

$$\mathrm{Ker}(\sigma) = \{\xi - \sigma(\xi) \mid \xi \in V\}.$$

(2) 首先若 $\alpha \in \mathrm{Ker}(\sigma) \cap \mathrm{Im}(\sigma)$, 则存在 $\beta \in V$ 使得 $\alpha = \sigma(\beta)$ 并且 $\sigma(\alpha) = 0$。于是

$$0 = \sigma(\alpha) = \sigma^2(\beta) = \sigma(\beta) = \alpha.$$

也即

$$\mathrm{Ker}(\sigma) \cap \mathrm{Im}(\sigma) = \{0\}.$$

另外任取 $\alpha \in V$, 则根据第一问有

$$\alpha = (\alpha - \sigma(\alpha)) + \sigma(\alpha) \in \mathrm{Ker}(\sigma) + \mathrm{Im}(\sigma).$$

综上

$$V = \mathrm{Ker}(\sigma) \oplus \mathrm{Im}(\sigma).$$

(3) $\Rightarrow$。如果 $\mathrm{Ker}(\sigma)$ 和 $\mathrm{Im}(\sigma)$ 都在 τ 之下不变。那么任取 $\alpha \in V$, 则 α 可唯一写成

$$\alpha = \beta + \gamma, \quad \beta \in \mathrm{Ker}(\sigma), \quad \gamma \in \mathrm{Im}(\sigma).$$

于是存在 $x \in V$ 使得 $\gamma = \sigma(x)$。因为 $\mathrm{Ker}(\sigma)$ 和 $\mathrm{Im}(\sigma)$ 都在 τ 之下不变。则

$$\tau(\beta) \in \mathrm{Ker}(\sigma), \quad \tau(\gamma) = \tau(\sigma(x)) \in \mathrm{Im}(\sigma).$$

从而 $\sigma(\tau(\beta)) = 0$, 以及存在 $\delta \in V$ 使得

$$\tau(\gamma) = \tau(\sigma(x)) = \sigma(\delta).$$

于是

$$\tau(\sigma(\alpha)) = \tau(\sigma(\beta)) + \tau(\sigma(\gamma)) = 0 + \tau(\sigma(\gamma)) = \tau(\sigma^2(x)) = \tau(\sigma(x)) = \tau(\gamma).$$

以及

$$\sigma(\tau(\alpha)) = \sigma(\tau(\beta)) + \sigma(\tau(\gamma)) = 0 + \sigma(\tau(\gamma)) = \sigma^2(\delta) = \sigma(\delta).$$

因为

$$\tau(\gamma) = \sigma(\delta).$$

所以

$$\tau(\sigma(\alpha)) = \sigma(\tau(\alpha)).$$

$\Leftarrow$。参看练习 2.30。

✍ **练习 2.32** (选做) 令 S 是数域 $\mathbb{F}$ 上有限维向量空间 V 的一些线性变换所成的集合. V 的一个子空间 W 如果在 S 中每一线性变换之下不变，那么就说 W 是 S 的一个不变子空间. 如果 S 在 V 中没有非平凡的不变子空间, 则是不可约的. 设 S 不可约, 而 φ 是 V 的一个线性变换, 它与 S 中每一线性变换可交换. 证明 φ 或者是零变换, 或者是可逆变换. [提示: 令 $W = \mathrm{Ker}(\varphi)$, 证

明 W 是 S 的一个不变子空间.]

解 令 $W=\operatorname{Ker}(\varphi)$。任取 $f\in S$, 则 $f\varphi=\varphi f$，那么根据习题 2.30 有，W 是 f 的不变子空间。由于 S 是任意的，那么 W 是 S 的不变子空间。但是 S 是不可约的，从而 S 只能有平凡的不变子空间，也就是说

$$W=\{0\}$$

或者

$$W=V.$$

如果 $W=V$, 则 φ 是零映射。如果 $W=0$, 则 φ 是单射，从而根据定理 2.22 知道 φ 可逆。

练习 2.33 设 V 是实数域 $\mathbf{R}$ 上 n 维线性空间. 又 $\mathcal{A}$ 是 V 上的线性变换. 证明 $\mathcal{A}$ 必有 1 维或 2 维不变子空间.

解 这道题放在这里不好，因为需要用到特征值和特征向量的东西。老师还没想到其他的方法。设 $d(\lambda)$ 为 $\mathcal{A}$ 的最小多项式. 如果 $d(\lambda)$ 在 $\mathbb{R}$ 中有根 $\lambda_0\in\mathbb{R}$, 则 λ_0 是 $\mathcal{A}$ 的特征值。设 $\boldsymbol{\alpha}_0$ 是对应的特征向量, 则

$$\mathcal{L}\left(\boldsymbol{\alpha}_0\right)=\left\{k\alpha_0:k\in\mathbb{R}\right\}$$

是 $\mathcal{A}$ 的 1 维不变子空间.

否则 $\mathcal{A}$ 无实特征根, 则有二次不可约多项式 $d_1(\lambda)=\lambda^2+p\lambda+q$, 使得 $d_1(\lambda)\mid d(\lambda)$. 令

$$d_2(\lambda)=\frac{d(\lambda)}{d_1(\lambda)},$$

则因为 $d(\lambda)$ 是最小多项式，那么一定有 $d_2(\mathcal{A})\neq 0$. 于是有 $\boldsymbol{\alpha}\in V$, 使得 $\boldsymbol{\beta}=d_2(\mathcal{A})\boldsymbol{\alpha}\neq 0$.

我们声明 $\boldsymbol{\beta},\mathcal{A}\boldsymbol{\beta}$ 线性无关。这是因为如果存在不全为零的 $x,y\in\mathbb{R}$ 使得

$$x\beta+y\mathcal{A}(\beta)=0.$$

首先 $y\neq 0$，否则 $\beta=0$ 矛盾！于是 $\mathcal{A}(\beta)=\frac{-x}{y}\beta$，于是 $\mathcal{A}$ 有实特征值，与假设矛盾！所以 $\boldsymbol{\beta},\mathcal{A}\boldsymbol{\beta}$ 线性无关。同时有

$$\left(\mathcal{A}^2+p\mathcal{A}+q\operatorname{id}\right)\boldsymbol{\beta}=d_1(\mathcal{A})d_2(\mathcal{A})=d(\mathcal{A})\boldsymbol{\alpha}=0.$$

于是 $\mathcal{A}^2\boldsymbol{\beta}=-q\boldsymbol{\beta}-p\mathcal{A}\boldsymbol{\beta}$. 因此 $\mathcal{L}(\boldsymbol{\beta},\mathcal{A}\boldsymbol{\beta})$ 是 $\mathcal{A}$ 的 2 维不变子空间.

练习 2.34 （选做）设 V 是 $\mathbf{C}$ 上 n 维线性空间. $\mathcal{A}$ 是 V 上的线性变换. 证明 $\mathcal{A}$ 在某组基下的矩阵为对角矩阵的充分必要条件是对 $\mathcal{A}$ 的任一不变子空间 W 有 $\mathcal{A}$ 的不变子空间 W' 使得 $V=W\oplus W'$.

解 $\Rightarrow$。因为 $\mathcal{A}$ 在某组基下的矩阵为对角矩阵，那么我们就选取一组基

$$\{\beta_1,\beta_2,\cdots,\beta_n\}$$

使得 $\mathscr{A}$ 在这组基下的矩阵是对角矩阵 $\operatorname{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$, 于是有

$$\mathscr{A}(\beta_j)=\lambda_j\beta_j,\quad j=1,2,3\cdots,n.$$

任取 $\mathcal{A}$ 的任一不变子空间 W，选取 W 的一组基

$$\{\alpha_1, \alpha_2, \cdots, \alpha_k\}$$

根据替换定理，可以用 $\{\alpha_1, \alpha_2, \cdots, \alpha_k\}$ 替换 $\{\beta_1, \beta_2, \cdots, \beta_n\}$ 中的 k 个向量得到新的向量组与原来的向量组等价。我们不妨假设替换的是前 k 个向量。于是

$$\{\alpha_1, \alpha_2, \cdots, \alpha_k, \beta_{k+1}, \beta_{k+2}, \cdots, \beta_n\}$$

与

$$\{\beta_1, \beta_2, \cdots, \beta_n\}$$

等价。从而 $\{\alpha_1, \alpha_2, \cdots, \alpha_k, \beta_{k+1}, \beta_{k+2}, \cdots, \beta_n\}$ 是 V 的一组基。于是

$$V = \mathcal{L}(\alpha_1, \alpha_2, \cdots, \alpha_k) \oplus \mathcal{L}(\beta_{k+1}, \beta_{k+2}, \cdots, \beta_n) = W \oplus W'.$$

此时有

$$W = \mathcal{L}(\alpha_1, \alpha_2, \cdots, \alpha_k)$$

以及

$$W' = \mathcal{L}(\beta_{k+1}, \beta_{k+2}, \cdots, \beta_n).$$

显然，因为 $\mathscr{A}(\beta_s) = \lambda_s \beta_s, s = k+1, k+2, \cdots, n$, 于是 W' 是 $\mathscr{A}$ 的不变子空间。得证。

$\Leftarrow$. 设 $\mathcal{A}$ 的特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_k$. 则

$$W = \sum_{i=1}^{k} V_{\lambda_i}(\mathcal{A})$$

是 $\mathcal{A}$ 的不变子空间. 这里的 $V_{\lambda_i}(\mathcal{A})$ 是特征子空间（参见定义 2.29）。于是根据假设有 $\mathcal{A}$ 的不变子空间 W' 使得 $V = W \oplus W'$. 若 $W' \neq \{0\}$, 那么 $\mathscr{A}|_{W'}$ 也是一个 W' 到 W' 的线性变换，从而一定有特征向量和特征值。从而有 i 使得

$$W' \cap V_{\lambda_i}(\mathcal{A}) \neq \{0\}.$$

这与 $W \cap W' = \{0\}$ 矛盾, 故 $W' = \{0\}$. 因此

$$V = V_{\lambda_1}(\mathcal{A}) \oplus V_{\lambda_2}(\mathcal{A}) \oplus \cdots \oplus V_{\lambda_k}(\mathcal{A}).$$

故 $\mathcal{A}$ 在某组基下的矩阵为对角矩阵. 上面的和是直和是因为定理 2.37.

2.6 2024 年 4 月 25 日第八周星期四第十五次课-本征值和本征向量

1. 在线性变换 $\mathscr{A} : V \to V$ 的所有不变子空间中, 一维不变子空间是极为重要的.
2. 设 U 是线性变换 $\mathscr{A}$ 的一维不变子空间. 显然 U 中任意一个非零向量 $\boldsymbol{\alpha}$ 都可以构成 U 的基 $\{\boldsymbol{\alpha}\}$. 由于 U 是 $\mathscr{A}$ 的不变子空间. 因此 $\mathscr{A}(\boldsymbol{\alpha}) \in U$. 所以

$$\mathscr{A}(\boldsymbol{\alpha}) = \lambda \boldsymbol{\alpha}, \quad \lambda \in \mathbb{F}$$

这就引出下面的定义.

定义 2.22 (本征值-本征向量或者特征值-特征向量)

设 $\mathscr{A}: V \to V$ 是线性变换. 如果存在非零向量 $\alpha \in V$, 以及纯量 $\lambda \in \mathbb{F}$, 使得

$$\mathscr{A}(\boldsymbol{\alpha}) = \lambda \boldsymbol{\alpha},$$

则 λ 称为线性变换 $\mathscr{A}$ 的本征值（特征值）, $\boldsymbol{\alpha}$ 称为线性变换 $\mathscr{A}$ 的属于特征值 λ 的本征向量（特征向量）.

注 特征向量生成的一维子空间是不变子空间。

定理 2.29 (如何求特征值？)

设线性变换 $\mathscr{A}: V \to V$ 在 V 的基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 下的方阵为 $\boldsymbol{A}$, 向量 $\boldsymbol{\alpha} \in V$ 在基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 下的坐标为 $\boldsymbol{x}$. 则 $\boldsymbol{\alpha}$ 是线性变换 $\mathscr{A}$ 的属于特征值 λ 的特征向量之充分必要条件为 x 是齐次线性方程组

$$\left(\lambda I_{(n)} - \boldsymbol{A}\right) \boldsymbol{x} = 0$$

的非零解.

- 必要性。设 $\boldsymbol{\alpha}$ 是线性变换 $\mathscr{A}$ 的属于特征值 λ 的特征向量, 则 $\mathscr{A}(\boldsymbol{\alpha}) = \lambda \boldsymbol{\alpha}$。记 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^{\mathrm{T}}$, 即
$$\boldsymbol{\alpha} = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n)\, \boldsymbol{x}.$$
由于 $\mathscr{A}$ 在基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 下的方阵为 $\boldsymbol{A}$, 所以
$$\mathscr{A}(\alpha) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n)\, \boldsymbol{A}\boldsymbol{x}.$$
由于 $\mathscr{A}(\boldsymbol{\alpha}) = \lambda \boldsymbol{\alpha}$, 因此 $\boldsymbol{A}\boldsymbol{x} = \lambda \boldsymbol{x}$, 即 $\left(\lambda I_{(n)} - \boldsymbol{A}\right) \boldsymbol{x} = \mathbf{0}$. 由于 $\boldsymbol{\alpha}$ 是 $\mathscr{A}$ 的特征向量, 故 $\boldsymbol{\alpha} \neq \mathbf{0}$. 从而 $\boldsymbol{x} \neq \mathbf{0}$. 因此 $\boldsymbol{x}$ 是方程组 $\left(\lambda \boldsymbol{I}_{(n)} - \boldsymbol{A}\right) \boldsymbol{x} = \mathbf{0}$ 的非零解.
- 充分性。设 $\boldsymbol{x}$ 是方程组 $\left(\lambda I_{(n)} - \boldsymbol{A}\right) \boldsymbol{x} = \mathbf{0}$ 的非零解. 则 $A\boldsymbol{x} = \lambda \boldsymbol{x}$. 因此
$$(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n)\, \boldsymbol{A}\boldsymbol{x} = \lambda\, (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n)\, \boldsymbol{x}.$$
所以 $\mathscr{A}(\boldsymbol{\alpha}) = \lambda \boldsymbol{\alpha}$. 由于 $\boldsymbol{x}$ 是非零的, 因此 $\boldsymbol{\alpha}$ 是非零向量. 所以 $\boldsymbol{\alpha}$ 是 $\mathscr{A}$ 的属于特征值 λ 的特征向量.

注 定理 2.29 说明, 求线性变换 $\mathscr{A}$ 的属于特征值 λ 的特征向量 $\boldsymbol{\alpha}$ 的问题等价于求齐次方程组

$$(\lambda I_n - \boldsymbol{A})\, \boldsymbol{x} = \mathbf{0}$$

的非零解问题. 显然齐次方程组 $(\lambda I_n - \boldsymbol{A})\, \boldsymbol{x} = \mathbf{0}$ 有非零解的充分必要条件是 $\det(\lambda I_n - \boldsymbol{A}) = 0$.

这就引出下面的定义.

定义 2.23 (特征多项式)

设 $\boldsymbol{A} \in \mathbb{F}^{n\times n}$, 则多项式 $\varphi(\lambda) = \det(\lambda I_n - A)$ 称为方阵 A 的特征多项式. 特征多项式 $\varphi(\lambda)$ 的根称为方阵 $\boldsymbol{A}$ 的特征值. 而方程组 $\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}$ 的非零解 $\boldsymbol{x}$ 称为方阵 $\boldsymbol{A}$ 的属于特征值 λ 的特征向量. ♣

定理 2.30 (定理 2.29 可以改述为)

设线性变换 $\mathscr{A}: V \to V$ 在 V 的基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 下的方阵为 $\boldsymbol{A}$, 向量 $\boldsymbol{\alpha} \in V$ 在基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 下的坐标为 $\boldsymbol{x}$. 则

(1) 当且仅当 λ 是方阵 $\boldsymbol{A}$ 的特征值时, λ 是线性变换 $\mathscr{A}$ 的特征值;

(2) 当且仅当 $\boldsymbol{x}$ 是方阵 $\boldsymbol{A}$ 的属于特征值 λ 的特征向量时, $\boldsymbol{\alpha}$ 是线性变换 $\mathcal{A}$ 的属于特征值 λ 的特征向量. ♡

推论 2.3 (具有一维不变子空间的充分必要条件)

设数域 $\mathbb{F}$ 上 n 维线性空间 V 到自身的线性变换 $\mathscr{A}$ 在 V 的基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 下的方阵为 $\boldsymbol{A}$. 则线性变换 $\mathscr{A}$ 具有一维不变子空间 U 的充分必要条件是, 方阵 $\boldsymbol{A}$ 的特征多项式 $\varphi(\lambda)$ 在数域 $\mathbb{F}$ 中有根 λ. ♡

注[特征值不一定存在]

- 我们知道, n 次复系数多项式一定有 n 个复根. 因此, 对 n 维复线性空间 V, 线性变换 $\mathscr{A}: V \to V$ 的一维不变子空间恒存在.
- 但是, 由于 n 次实系数多项式不一定具有实根 (尽管它的复根总存在), 所以对 n 维实线性空间 V, 线性变换 $\mathscr{A}: V \to V$ 的一维不变子空间并不是总存在的.
- 所以在讨论一维不变子空间时, 一定要注意线性空间 V 的基域 $\mathbb{F}$.
- 也就是说，如果特征多项式在域 $\mathbb{F}$ 中没有根，则不存在特征值和特征向量。

特征多项式的性质. 考虑方阵 $\boldsymbol{A}$ 的特征多项式

$$\varphi(\lambda) = \det(\lambda I_n - \boldsymbol{A}).$$

- 记 n 阶方阵 $\boldsymbol{A} = (a_{ij})_{1\le i,j\le n}$。
- 对 $1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant n, 1 \leqslant j_1 < j_2 < \cdots < j_k \leqslant n$，用
$$\boldsymbol{A}\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix}$$
表示取自第 $i_1, i_2, \cdots, i_k$ 行和第 $j_1, j_2, \cdots, j_k$ 列的 k 阶子式。
- 方阵 $\boldsymbol{A}$ 的特征多项式 $\varphi(\lambda)$ 的 n 个特征值记为 $\lambda_1, \lambda_2, \cdots, \lambda_n$.

命题 2.6 (特征多项式怎么求？)

我们有

$$\begin{aligned}\varphi(\lambda) =&\lambda^n-\left(a_{11}+a_{12}+\cdots+a_{nn}\right)\lambda^{n-1}\\&+\left(\sum_{1\leqslant i_1<i_2\leqslant n}\boldsymbol{A}\begin{pmatrix}i_1 & i_2\\ i_1 & i_2\end{pmatrix}\right)\lambda^{n-2}+\cdots\\&+(-1)^k\left(\sum_{1\leqslant i_1<i_2<\cdots<i_k\leqslant n}\boldsymbol{A}\begin{pmatrix}i_1 & i_2 & \cdots & i_k\\ i_1 & i_2 & \cdots & i_k\end{pmatrix}\right)\lambda^{n-k}+\cdots\\&+(-1)^n\det\boldsymbol{A}.\end{aligned}$$

命题 2.7 (韦达定理)

设 $\sigma_1\left(x_1,x_2,\cdots,x_n\right),\sigma_2\left(x_1,x_2,\cdots,x_n\right),\cdots,\sigma_n\left(x_1,x_2,\cdots,x_n\right)$ 是关于 $x_1,x_2,\cdots,x_n$ 的基本对称多项式. 则对 $k=1,2,\cdots,n$,

$$\sigma_k\left(\lambda_1,\lambda_2,\cdots,\lambda_n\right)=\sum_{1\leqslant i_1<i_2<\cdots<i_k\leqslant n}\boldsymbol{A}\begin{pmatrix}i_1 & i_2 & \cdots & i_k\\ i_1 & i_2 & \cdots & i_k\end{pmatrix}.$$

于是我们有

$$\boxed{\operatorname{tr}(\boldsymbol{A})=\lambda_1+\lambda_2+\cdots+\lambda_n}$$

和

$$\boxed{\det(\boldsymbol{A})=\lambda_1\lambda_2\cdots\lambda_n}.$$

命题 2.8 (可逆的充要条件)

设 $\boldsymbol{A}$ 为 n 阶复方阵. 则方阵 $\boldsymbol{A}$ 可逆的充分必要条件是方阵 $\boldsymbol{A}$ 的特征值全不为零.

命题 2.9 (准上三角矩阵的特征多项式)

设方阵 $\boldsymbol{A}$ 是准上三角的, 即设

$$\boldsymbol{A}=\begin{pmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12}\\ \mathbf{0} & \boldsymbol{A}_{22}\end{pmatrix},$$

其中 $\boldsymbol{A}_{11}$ 与 $\boldsymbol{A}_{22}$ 是子方阵. 则方阵 $\boldsymbol{A}$ 的特征多项式等于 $\boldsymbol{A}_{11}$ 与 $\boldsymbol{A}_{22}$ 的特征多项式的乘积.

练习 2.35 考虑 $\mathscr{A}:\mathbb{R}^3\mapsto\mathbb{R}^3$, 定义如下

$$\mathscr{A}(x,y,z)=(x+2y+2z,2x+y+2z,2x+2y+z)$$

(1) 考虑 $\mathbb{R}^3$ 的一组基

$$\boldsymbol{\alpha}_1=(1,0,0),\quad\boldsymbol{\alpha}_2=(0,1,0),\quad\boldsymbol{\alpha}_3=(0,0,1)$$

证明：$\mathscr{A}$ 在基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3\}$ 下的矩阵表示为

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

(2) 证明：如果 $\boldsymbol{\alpha} \in \mathbb{R}^3$ 在 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3\}$ 下的的坐标是 $X = (x, y, z)^T$(看成列向量)。则 $\boldsymbol{\alpha}$ 是 $\mathscr{A}$ 的特征值为 $\lambda \in \mathbb{R}$ 的特征向量当且仅当 X 是 $(\lambda I_3 - \boldsymbol{A})X = 0$ 的非零解。

(3) 根据上一问知道, $(\lambda I_3 - \boldsymbol{A})X = 0$ 有非零解当且仅当行列式

$$\det(\lambda I_3 - \boldsymbol{A}) = \det \begin{pmatrix} \lambda - 1 & -2 & -2 \\ -2 & \lambda - 1 & -2 \\ -2 & -2 & \lambda - 1 \end{pmatrix} = 0$$

请证明这个行列式等于 $\varphi_{\boldsymbol{A}}(\lambda) = (\lambda - 5)(\lambda + 1)^2$。这个行列式是 λ 的 3 次多项式，我们把它称为矩阵

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix}$$

的特征多项式。

(4) 注意到上面的特征多项式

$$\varphi_{\boldsymbol{A}}(\lambda) = \lambda^3 + a_2\lambda^2 + a_1\lambda + a_0 = (\lambda - x_1)(\lambda - x_2)(\lambda - x_3),$$

这里 $x_1 = x_2 = -1, x + 3 = 5$ 以及 $a_2 = -3, a_1 = -9, a_0 = -5$. 请验证一下韦达定理吧。也就是说

$$\begin{cases} x_1x_2x_3 = -a_0 \\ x_1x_2 + x_2x_3 + x_2x_3 = a_1 \\ x_1 + x_2 + x_3 = -a_2 \end{cases}$$

(5) 再验证一下

$$\begin{cases} a_2 = -(a_{11} + a_{22} + a_{33}) \\ a_1 = \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} + \det \begin{pmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{pmatrix} + \det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} \\ a_0 = -\det(\boldsymbol{A}) \end{cases}$$

这你就验证了命题2.6.

(6) 请验证

$$a_0 I_3 + a_1 \boldsymbol{A} + a_2 \boldsymbol{A}^2 + \boldsymbol{A}^3 = 0$$

这道题希望你认真做一遍，老师自己编了这个题就是想让你理解这一节课讲的东西，这道题你走下来，你就理解了这节课的东西。

✍ **练习 2.36** 求下列矩阵在实数域内的特征根和相应的特征向量:

1. （选做）

$$\begin{pmatrix} 3 & -2 & 0 \\ -1 & 3 & -1 \\ -5 & 7 & -1 \end{pmatrix}$$

2. (选做)

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

3.

$$\begin{pmatrix} 3 & 6 & 6 \\ 0 & 2 & 0 \\ -3 & -12 & -6 \end{pmatrix}$$

解

1. 特征多项式为

$$\varphi_A(x) = -x^3 + 5x^2 - 8x + 4 = -(x-2)^2(x-1).$$

所以特征值为

$$\lambda_1 = 2, \quad \lambda_2 = 2, \quad \lambda_3 = 1.$$

对应的特征向量 (自己去计算) 为

$$X_1 = \begin{pmatrix} -2 \\ -1 \\ 1 \end{pmatrix}, \quad X_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad X_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

2. 特征多项式为

$$\varphi_A(x) = -x^3 + 5x^2 - 17x + 13 = -\left((x-1)\left(x^2 - 4x + 13\right)\right).$$

所以在实数域的特征值为

$$\lambda = 1.$$

对应的特征向量 (自己去计算) 为

$$X = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$$

你还可计算在复数域上的另外两个特征值为

$$\lambda_2 = 2 + 3i, \quad \lambda_3 = 2 - 3i.$$

对应的特征向量(自己去计算)为

$$X_2=\begin{pmatrix}3-3i\\5-3i\\4\end{pmatrix},\quad X_2=\begin{pmatrix}3+3i\\5+3i\\4\end{pmatrix}.$$

你能发现什么规律吗?

3. 特征多项式为

$$\varphi_A(x)=-x^3-x^2+6x=-((x-2)x(x+3)).$$

所以特征值为

$$\lambda_1=-3,\quad \lambda_2=2,\quad \lambda_3=0.$$

对应的特征向量(自己去计算)为

$$X_1=\begin{pmatrix}-1\\0\\1\end{pmatrix},\quad X_2=\begin{pmatrix}12\\-5\\3\end{pmatrix}\quad X_3=\begin{pmatrix}-2\\0\\1\end{pmatrix}.$$

✍ **练习 2.37** 证明：对角矩阵

$$A=\begin{pmatrix}a_1&&&0\\&a_2&&\\0&&\ddots&\\0&&&a_n\end{pmatrix}\text{与 }B=\begin{pmatrix}b_1&&&0\\&b_2&&\\0&&\ddots&\\0&&&b_n\end{pmatrix}$$

相似当且仅当 $b_1,b_2,\cdots,b_n$ 是 $a_1,a_2,\cdots,a_n$ 的一个排列.

解 因为相似矩阵有相同的特征值，从而如果 A 和 B 相似，则

$$\{b_1,b_2,\cdots,b_n\}=\{a_1,a_2,\cdots,a_n\}.$$

反之，如果

$$b_1=a_{i_1},b_2=a_{i_2},\cdots,b_n=a_{i_n}.$$

此时，存在一个 n 维线性空间 V 以及 V 上的线性变换 $\mathscr{A}$ 以及的一组基

$$\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$$

使得 $\mathscr{A}$ 在上面的基下的矩阵表示为 A。取同样的基

$$\beta_1=\alpha_{i_1},\beta_2=\alpha_{i_2},\cdots,\beta_n=\alpha_{i_n}.$$

那么 $\mathscr{A}$ 在新的基下的矩阵表示为 B。因为同一个线性变换在不同基下的矩阵表示相似，所以 A 和 B 相似。

2.7 2024 年 4 月 30 日第九周星期二第十六次课-特征多项式

命题 2.10 (上三角矩阵的特征值)

设 $\boldsymbol{A}$ 是上三角方阵, 则方阵 $\boldsymbol{A}$ 的对角元是方阵 $\boldsymbol{A}$ 的特征值.

命题 2.11

相似的方阵具有相同的特征多项式, 从而具有相同的特征值.

设方阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似, 则存在可逆方阵 $\boldsymbol{P}$, 使得 $\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$, 因此

$$\begin{aligned}
\det\left(\lambda \boldsymbol{I}_n-\boldsymbol{B}\right) &= \det\left(\lambda \boldsymbol{I}_n-\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}\right) \\
&= \det \boldsymbol{P}^{-1}\left(\lambda \boldsymbol{I}_n-\boldsymbol{A}\right)\boldsymbol{P} \\
&= \det \boldsymbol{P}^{-1} \det\left(\lambda \boldsymbol{I}_n-\boldsymbol{A}\right) \det \boldsymbol{P} \\
&= \det\left(\lambda \boldsymbol{I}_n-\boldsymbol{A}\right).
\end{aligned}$$

所以方阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的特征多项式相同.

注 方阵的特征值是方阵在相似下的不变量. 但应指出, 它并不是方阵在相似下的全系不变量, 即特征值相同的方阵并不一定相似. 例如方阵

$$\boldsymbol{A}=\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad 与 \quad \boldsymbol{B}=\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

的特征值相同, 但方阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 并不相似.

定理 2.31 (Cayley-Hamilton 定理)

设 n 阶方阵 $\boldsymbol{A}$ 的特征多项为 $\varphi(\lambda)=\lambda^n+a_{n-1}\lambda^{n-1}+\cdots+a_1\lambda+a_0$, 则

$$\varphi(\boldsymbol{A})=\boldsymbol{A}^n+a_{n-1}\boldsymbol{A}^{n-1}+\cdots+a_1\boldsymbol{A}+a_0\boldsymbol{I}_n=\boldsymbol{0}.$$

线性变换的特征多项式. 考虑线性变换 $\mathscr{A}: V\to V$. 回忆同一个线性变换 $\mathscr{A}$ 在不同基下的方阵是相似的. 而相似方阵的特征多项式相同.

定义 2.24

设线性变换 $\mathscr{A}: V\to V$ 的基 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$ 下的方阵为 $\boldsymbol{A}$, 则方阵 $\boldsymbol{A}$ 的特征多项式称为<u>线性变换 $\mathscr{A}$ 的特征多项式</u>.

定理 2.32 (Cayley-Hamilton 定理)

设 $\varphi(\lambda)$ 是线性变换 $\mathscr{A}$ 的特征多项式, 则 $\varphi(\mathscr{A})=0$, 其中 0 是零变换.

定义 2.25 (化零多项式)

设方阵 $\boldsymbol{A}\in\mathbb{F}^{n\times n}$, 且非零多项式 $f(\lambda)\in\mathbb{F}[\lambda]$. 如果 $f(\boldsymbol{A})=\boldsymbol{0}$, 则 $f(\lambda)$ 称为方阵 $\boldsymbol{A}$ 的一个<u>化零多项式</u>.

定义 2.26 (最小多项式)

方阵 $\boldsymbol{A}$ 的所有化零多项式中次数最小的首一多项式称为方阵 $\boldsymbol{A}$ 的最小多项式, 记为 $d_{\boldsymbol{A}}(\lambda)$ 或 $d(\lambda)$.

- 考虑正整数集合 $M = \{\deg\ f(\lambda) : f(\lambda)$ 为方阵 $\boldsymbol{A}$ 的化零多项式 $\}$.
- 由于方阵 $\boldsymbol{A}$ 的特征多项式 $\varphi(\lambda)$ 是方阵 $\boldsymbol{A}$ 的一个化零多项式, 因此集合 $M \neq \varnothing$.
- 所以集合 M 中存在最小的正整数 m, 而且 m 是方阵 $\boldsymbol{A}$ 的某个化零多项式 $g(\lambda)$ 的次数. 于是 $d(\lambda) = a^{-1}g(\lambda)$ 是方阵 $\boldsymbol{A}$ 的最小多项式, 其中 a 是多项式 $g(\lambda)$ 的首项次数.
- 这说明, 任意一个方阵 $\boldsymbol{A}$ 的最小多项式均存在.

命题 2.12 (最小多项式整除化零多项式)

方阵 $\boldsymbol{A}$ 的最小多项式 $d(\lambda)$ 整除方阵 $\boldsymbol{A}$ 的任一化零多项式 $f(\lambda)$. 特别地, $d(\lambda)$ 整除方阵 $\boldsymbol{A}$ 的特征多项式 $\varphi(\lambda)$.

命题 2.13 (最小多项式存在唯一)

方阵 $\boldsymbol{A}$ 的最小多项式 $d(\lambda)$ 存在且唯一.

命题 2.14

相似的方阵具有相同的最小多项式.

注 方阵的最小多项式并不是方阵在相似下的全系不变量, 即具有相同的最小多项式的方阵并不一定相似.

定理 2.33

纯量 $c \in \mathbb{F}$ 为方阵 $\boldsymbol{A}$ 的最小多项式 $d(\lambda)$ 的根的充分必要条件是 c 为方阵 $\boldsymbol{A}$ 的特征值.

定义 2.27

设 $\mathscr{A} : V \to V$ 是线性变换的, 且非零多项式 $f(\lambda) \in \mathbb{F}[\lambda]$. 如果 $f(\mathcal{A}) = 0$, 则 $f(\lambda)$ 称为线性变换 $\mathscr{A}$ 的化零多项式.

定义 2.28

线性变换 $\mathscr{A}$ 的所有化零多项式中次数最小的首一多项式称为线性变换 $\mathscr{A}$ 的最小多项式.

定理 2.34

纯量 $c \in \mathbb{F}$ 为线性变换 $\mathscr{A}$ 的最小多项式的根的充分必要条件是, c 是线性变换 $\mathscr{A}$ 的特征值.

例题 2.11 设 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 是 n 阶复方阵, 则方阵 $\boldsymbol{AB}$ 与 $\boldsymbol{BA}$ 的特征多项式相同.

一般地说, 方阵 $\boldsymbol{AB}$ 与 $\boldsymbol{BA}$ 的最小多项式并不一定相同. 例如, 设

$$\boldsymbol{A}=\begin{pmatrix}0&0\\1&0\end{pmatrix},\boldsymbol{B}=\begin{pmatrix}0&-1\\0&1\end{pmatrix}.$$

则

$$\boldsymbol{AB}=\boldsymbol{0},\quad \boldsymbol{BA}=\begin{pmatrix}-1&-1\\1&1\end{pmatrix}.$$

方阵 $\boldsymbol{AB}$ 的特征多项式 (也是方阵 $\boldsymbol{BA}$ 的特征多项式) 为 $\det\left(\lambda\boldsymbol{I}_{(2)}-\boldsymbol{AB}\right)=\lambda^2$. 因此方阵 $\boldsymbol{AB}$ 的最小多项式 $d_{\boldsymbol{AB}}(\lambda)\mid\lambda^2$. 所以 $d_{\boldsymbol{AB}}(\lambda)=\lambda$, 或 λ^2, 将方阵 $\boldsymbol{AB}$ 代入 λ, 即知 $d_{\boldsymbol{AB}}(\lambda)=\lambda$. 同样, 方阵 $\boldsymbol{BA}$ 的最小多项式 $d_{\boldsymbol{BA}}(\lambda)\mid\lambda^2$, 故 $d_{\boldsymbol{BA}}(\lambda)=\lambda$ 或 λ^2. 因为 $\boldsymbol{BA}\neq\boldsymbol{0}$, 故 $d_{\boldsymbol{AB}}(\lambda)=\lambda^2$. 因此方阵 $\boldsymbol{AB}$ 与 $\boldsymbol{BA}$ 的最小多项式并不相等.

✍ **练习 2.38** 设

$$A=\begin{pmatrix}a&b\\c&d\end{pmatrix}$$

是一个实矩阵且 $ad-bc=1$. 证明:

(i) 如果 $|\operatorname{tr}\boldsymbol{A}|>2$, 那么存在可逆实矩阵 $\boldsymbol{T}$, 使得

$$\boldsymbol{T}^{-1}\boldsymbol{AT}=\begin{pmatrix}\lambda&0\\0&\lambda^{-1}\end{pmatrix},$$

这里 $\lambda\in\mathbf{R}$ 且 $\lambda\neq0,1,-1$;

(ii) 如果 $|\operatorname{tr}\boldsymbol{A}|=2$ 且 $\boldsymbol{A}\neq\pm\boldsymbol{I}$, 那么存在可逆实矩阵 $\boldsymbol{T}$, 使得

$$\boldsymbol{T}^{-1}\boldsymbol{A}\mathbf{T}=\begin{pmatrix}1&1\\0&1\end{pmatrix}\text{ 或 }\begin{pmatrix}-1&1\\0&-1\end{pmatrix};$$

(iii) 如果 $|\operatorname{tr}\boldsymbol{A}|<2$, 则存在可逆实矩阵 $\boldsymbol{T}$ 及 $\theta\in\mathbb{R}$, 使得

$$\boldsymbol{T}^{-\mathbf{1}}\boldsymbol{AT}=\begin{pmatrix}\cos\theta&\sin\theta\\-\sin\theta&\cos\theta\end{pmatrix};$$

提示：在 (iii), $\boldsymbol{A}$ 有非实共轭复特征根 $\lambda,\bar{\lambda},\lambda\bar{\lambda}=1$. 将 λ 写成三角形式. 令 $\xi\in\mathbb{C}^2$ 是 $\boldsymbol{A}$ 的属于 λ 的一个特征向量, 计算 $\boldsymbol{A}(\xi+\bar{\xi})$ 和 $\boldsymbol{A}(\mathrm{i}(\boldsymbol{\xi}-\bar{\boldsymbol{\xi}}))$。

解 首先 A 的特征多项式为

$$\varphi_A(x)=(x-a)(x-d)-bc=x^2-(a+d)x+sd-bc=x^2-(a+d)x+1.$$

1. 如果 $|a+d|>2$，那么特征多项式的判别式

$$\triangle=(a+d)^2-4>0.$$

于是 A 有两个不同的特征值 $\lambda_1\neq\lambda_2$ 并且 $\lambda_1\lambda_2=1$。

$$\varphi(0)=1\neq0,\quad \varphi(\pm1)=2\pm(a+d)\neq0.$$

所以 $\lambda_1 \neq \lambda_2 \in \mathbb{R} \setminus \{0, \pm 1\}$ 。于是 A 相似于对角矩阵

$$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \frac{1}{\lambda_1} \end{pmatrix}.$$

2. 如果 $|a+d| = 2$，则特征多项式的判别式

$$\triangle = (a+d)^2 - 4 = 0.$$

于是 A 只有一个特征值

$$\lambda = \frac{(a+d)}{2}, \quad \Rightarrow |\lambda| = 1.$$

(1) 如果 $\lambda = 1$ 并且对 A 有两个线性无关的特征向量, 那么 A 可以对角化，从而相似于 I_2. 也即 $A = I_2$, 与题意矛盾!

(2) 如果 $\lambda = 1$ 并且对 A 没有两个线性无关的特征向量。此时因为 $\varphi_A(1) = 0$, 于是 $a + d = 2$ 从而

$$\varphi_A(x) = (x-1)^2.$$

假定 X 是对应于特征值 1 的特征向量，将 X 扩充成 $\mathbb{R}^2$ 的一组基 $\{X, Y\}$。则 $AY \neq Y$，否则有两个线性无关的特征向量。假定

$$AY = tX + sY, \quad t, s \in \mathbb{R}.$$

且 $t \neq 0$，否则 Y 是特征向量。则

$$A^2Y = tAX + sAY = tX + s(tX + sY) = (t + st)X + s^2Y.$$

因为特征多项式为 $(x-1)^2$，于是 $(A - I_2)^2 = 0$，于是

$$\begin{aligned} 0 = (A - I_2)(A - I_2)Y &= A^2Y - 2AY + Y = (t + st)X + s^2Y - 2tX - 2sY + Y \\ =&(st - t)X + (s^2 - 2s + 1)Y. \end{aligned}$$

所以

$$\begin{cases} st - t = 0 \\ s^2 - 2s + 1 = 0 \end{cases}$$

于是 $s = 1$。从而

$$A\frac{Y}{t} = X + \frac{Y}{t}.$$

于是 A 在 $\mathbb{R}^2$ 的基 $\{X, \frac{Y}{t}\}$ 下的矩阵表示为

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

(3) 如果 $\lambda = -1$ 并且对 A 有两个线性无关的特征向量, 那么 A 可以对角化，从而相似于 I_2. 也即 $A = -I_2$, 与题意矛盾!

(4) 如果 $\lambda = -1$ 并且对 A 没有两个线性无关的特征向量。此时因为 $\varphi_A(-1) = 0$, 于是 $a + d = -2$ 从而

$$\varphi_A(x) = (x+1)^2.$$

假定 X 是对应于特征值 -1 的特征向量，将 X 扩充成 $\mathbb{R}^2$ 的一组基 $\{X, Y\}$。则 $AY \neq -Y$，否则有两个线性无关的特征向量。假定

$$AY = tX + sY, \quad t, s \in \mathbb{R}.$$

且 $t \neq 0$，否则 Y 是特征向量。则

$$A^2Y = tAX + sAY = -tX + s(tX + sY) = (-t + st)X + s^2Y.$$

因为特征多项式为 $(x+1)^2$，于是 $(A+I_2)^2 = 0$，于是

$$0 = (A+I_2)(A+I_2)Y = A^2Y + 2AY + Y = (-t+st)X + s^2Y + 2tX + 2sY + Y$$
$$=(st+t)X + (s^2+2s+1)Y.$$

所以

$$\begin{cases} st + t = 0 \\ s^2 + 2s + 1 = 0 \end{cases}$$

于是 $s = -1$。从而

$$A\frac{Y}{t} = X - \frac{Y}{t}.$$

于是 A 在 $\mathbb{R}^2$ 的基 $\{X, \frac{Y}{t}\}$ 下的矩阵表示为

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

3. 如果 $|a+d| < 2$, 则特征多项式 $\varphi_A(x)$ 的判别式

$$\triangle = (a+d)^2 - 4 < 0.$$

于是有两个复的特征值

$$\lambda = \frac{(a+d) + i\sqrt{4-(a+d)^2}}{2} \quad \bar{\lambda} = \frac{(a+d) - i\sqrt{4-(a+d)^2}}{2}$$

显然

$$\lambda\bar{\lambda} = 1$$

于是可假定

$$\lambda = \cos(q) + i\sin(q), \quad q \in \mathbb{R}.$$

假定对应于 λ 的特征向量为

$$X = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} x_1 + iy_1 \\ x_2 + iy_2 \end{pmatrix} \in \mathbb{C}^2.$$

则 $AX = \lambda X$，取共轭有 $A\bar{X} = \bar{\lambda}\bar{X}$。所以

$$\bar{X} = \begin{pmatrix} \bar{z}_1 \\ \bar{z}_2 \end{pmatrix} = \begin{pmatrix} x_1 - iy_1 \\ x_2 - iy_2 \end{pmatrix} \in \mathbb{C}^2.$$

是对应于特征值 $\bar{\lambda}$ 的特征向量。首先如果存在不全为零的实数 $s, t \in \mathbb{R}$ 使得

$$s\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + t\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

则

$$(s - it)X + (s + it)\bar{X} = (s - it)\begin{pmatrix} x_1 + iy_1 \\ x_2 + iy_2 \end{pmatrix} + (s + it)\begin{pmatrix} x_1 - iy_1 \\ x_2 - iy_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

由于 $s + it, s - it$ 不全为零，从而 $X, \bar{X}$ 线性相关，这是不可能的，因为属于不同的特征值的特征向量线性无关（这里我们把它看成复数域上的矩阵）。从而

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

在 $\mathbb{R}^2$ 中线性无关。

因为

$$AX = \lambda X,$$

于是

$$A\begin{pmatrix} x_1 + iy_1 \\ x_2 + iy_2 \end{pmatrix} = A\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + iA\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = (\cos(q) + i\sin(q))\begin{pmatrix} x_1 + iy_1 \\ x_2 + iy_2 \end{pmatrix}$$

比较实部和虚部可得

$$\begin{cases} A\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \cos(q)x_1 - \sin(q)y_1 \\ \cos(q)x_2 - \sin(q)y_2 \end{pmatrix} = \cos(q)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \sin(q)\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \\ A\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \cos(q)y_1 + \sin(q)x_1 \\ \cos(q)y_2 + \sin(q)x_2 \end{pmatrix} = \cos(q)\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \sin(q)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \end{cases}$$

于是 A 在 $\mathbb{R}^2$ 的基

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

下的矩阵表示为

$$\begin{pmatrix} \cos(q) & \sin(q) \\ -\sin(q) & \cos(q) \end{pmatrix}.$$

✍ **练习 2.39** 设 $a, b, c \in \mathbb{C}$. 令

$$\boldsymbol{A} = \begin{pmatrix} b & c & a \\ c & a & b \\ a & b & c \end{pmatrix}, \boldsymbol{B} = \begin{pmatrix} c & a & b \\ a & b & c \\ b & c & a \end{pmatrix}, \boldsymbol{C} = \begin{pmatrix} a & b & c \\ b & c & a \\ c & a & b \end{pmatrix}$$

(i) 证明 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ 彼此相似;

(ii) 如果 $\boldsymbol{BC} = \boldsymbol{CB}$, 那么 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ 的特征根至少有两个等于零.

解

1. 假定 V 是一个三维线性空间，其上的线性变换 $\mathscr{A}$ 在一组基 $\{\alpha, \beta, \gamma\}$ 下的矩阵表示 A。

则

$$\begin{cases} \mathscr{A}(\alpha) = b\alpha + c\beta + a\gamma \\ \mathscr{A}(\beta) = c\alpha + a\beta + b\gamma \\ \mathscr{A}(\gamma) = a\alpha + b\beta + c\gamma \end{cases}$$

于是容易验证 $\mathscr{A}$ 在 $\{\beta, \gamma, \alpha\}$ 下的矩阵为 C, 在 $\{\gamma, \alpha, \beta\}$ 下的矩阵表示为 B. 因为同一个线性变换在不同基下的矩阵表示相似，所以 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ 彼此相似。

2. 我们暴力地计算 $BC = CB$ 当且仅当

$$ab + ac + bc = a^2 + b^2 + c^2.$$

当且仅当

$$(a-b)^2 + (a-c)^2 + (b-c)^2 = 0.$$

特征多项式为

$$\begin{aligned} \varphi(x) =& \det \begin{bmatrix} x-a & -b & -c \\ -b & x-c & -a \\ -c & -a & x-b \end{bmatrix} \\ =& -a^3 + a^2x + 3abc - abx - acx + ax^2 - b^3 + b^2x - bcx + bx^2 - c^3 + c^2x + cx^2 - x^3 \end{aligned}$$

首先因为

$$A \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = (a+b+c) \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

我们知道 $a+b+c$ 是一个特征根，从而 $x-(a+b+c)$ 可以整除 $\varphi(x)$。我们用长除法可以分解 $\varphi(x)$ 为

$$\varphi(x) = -\left((a+b+c-x)\left(a^2 - ab - ac + b^2 - bc + c^2 - x^2\right)\right)$$

又因为 $ab + ac + bc = a^2 + b^2 + c^2$, 所以

$$\varphi(x) = x^2(x-a-b-c).$$

所以特征根至少有两个零。

✍ **练习 2.40** 设 $\boldsymbol{A}$ 是复数域 $\mathbf{C}$ 上一个 n 阶矩阵.

(i) 证明：存在 $\mathbf{C}$ 上 n 阶可逆矩阵 $\boldsymbol{T}$ 使得

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = \begin{pmatrix} \lambda_1 & b_{12} & \cdots & b_{1n} \\ 0 & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & b_{n2} & \cdots & b_{nn} \end{pmatrix};$$

(ii) 对 n 作数学归纳法证明, 复数域 $\mathbf{C}$ 上任意一个 n 阶矩阵都与一个上三角形矩阵上三角形

矩阵

$$\begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

相似，这里主对角线以下的元素都是零.

解

(1) 我们还是把 A 看成是 $\mathbb{C}^n$ 上的线性变换 $\mathcal{A}(X)=AX$ 在标准基下的矩阵。这里的标准基指的是

$$\alpha_k = \underbrace{(0,0,\cdots,0,1,0\cdots,0)}_{\text{第 } k \text{ 个为 } 1}, \quad k=1,2,3,\cdots,n.$$

由于 A 的特征多项式在 $\mathbb{C}$ 上一定有根，于是存在特征值 β_1 以及特征向量 $\alpha_1 \in \mathbb{C}^n$ 使得 $\mathscr{A}(\beta_1)=\lambda_1\beta_1$. 那么我们将 β_1 扩充为 $\mathbb{C}^n$ 的基

$$\{\beta_1,\beta_2,\cdots,\beta_n\}.$$

那么 $\mathscr{A}$ 在基 $\{\beta_1,\beta_2,\cdots,\beta_n\}$ 下的矩阵有形式

$$\begin{pmatrix} \lambda_1 & b_{12} & \cdots & b_{1n} \\ 0 & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & b_{n2} & \cdots & b_{nn} \end{pmatrix}$$

因为同一个线性变换在不同基下的矩阵相似，我们得到 A 相似于

$$\begin{pmatrix} \lambda_1 & b_{12} & \cdots & b_{1n} \\ 0 & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & b_{n2} & \cdots & b_{nn} \end{pmatrix}.$$

(2) 用归纳法，令

$$B=\begin{pmatrix} b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \vdots \\ b_{n2} & \cdots & b_{nn} \end{pmatrix}.$$

则根据归纳假设，存在 $n-1$ 阶可逆矩阵 Q 使得

$$QBQ^{-1}$$

是上三角矩阵。现在令

$$P=\begin{pmatrix} 1 & 0 \\ 0 & Q \end{pmatrix}$$

则有

$$P\begin{pmatrix} \lambda_1 & b_{12} & \cdots & b_{1n} \\ 0 & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & b_{n2} & \cdots & b_{nn} \end{pmatrix} P^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & Q \end{pmatrix} \begin{pmatrix} \lambda_1 & X \\ 0 & B \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & Q^{-1} \end{pmatrix} = \begin{pmatrix} \lambda_1 & XQ^{-1} \\ 0 & QBQ^{-1} \end{pmatrix}.$$

其中 $X = (b_{12}, \cdots, b_{1n})$。显然

$$\begin{pmatrix} \lambda_1 & XQ^{-1} \\ 0 & QBQ^{-1} \end{pmatrix}$$

是上三角矩阵。再根据矩阵相似的传递性，知道 A 相似于上三角阵。

✍ **练习 2.41** 设 $\boldsymbol{A}$ 是复数域 $\mathbb{C}$ 上一个 n 阶矩阵, $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是 $\boldsymbol{A}$ 的全部特征根 (重根按重数计算).

(i) 如果 $f(x)$ 是 $\mathbb{C}$ 上任意一个次数大于零的多项式，那么 $f(\lambda_1), f(\lambda_2), \cdots, f(\lambda_n)$ 是 $f(\boldsymbol{A})$ 的全部特征根;

(ii) 如果 $\boldsymbol{A}$ 可逆，那么 $\lambda_i \neq 0, i = 1, 2, \cdots, n$, 并且 $\lambda_1^{-1}, \lambda_2^{-1}, \cdots, \lambda_n^{-1}$ 是 $\boldsymbol{A}^{-1}$ 的全部特征根.

解

(1) 根据上一道练习，存在 n 阶可逆矩阵 P 以及上三角矩阵

$$\Lambda = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

使得

$$A = P\Lambda P^{-1}.$$

从而有

$$f(A) = Pf(\Lambda)P^{-1}.$$

因为上三角矩阵的幂次还是上三角矩阵，上三角矩阵相加还是上三角阵，我们有 $f(\Lambda)$ 还是上三角阵并且有如下形式:

$$f(\Lambda) = \begin{pmatrix} f(\lambda_1) & * & \cdots & * \\ 0 & f(\lambda_2) & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & f(\lambda_n) \end{pmatrix}$$

从而 $f(A)$ 的特征值与上面的矩阵的特征值一样，都是 $f(\lambda_1), f(\lambda_2), \cdots, f(\lambda_n)$。

(2) 首先 A 还是相似于上三角阵

$$\Lambda=\begin{pmatrix}\lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n\end{pmatrix}.$$

即存在 n 阶可逆矩阵 P 使得 $A=P\Lambda P^{-1}$，从而 $A^{-1}=P\Lambda^{-1}P^{-1}$。因为上三角矩阵的逆还是上三角矩阵，并且对角元素变成倒数，于是 Λ^{-1} 有如下形式

$$\begin{pmatrix}\frac{1}{\lambda_1} & * & \cdots & * \\ 0 & \frac{1}{\lambda_2} & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \frac{1}{\lambda_n}\end{pmatrix}.$$

从而 Λ^{-1} 的特征根是 $\lambda_1^{-1},\lambda_2^{-1},\cdots,\lambda_n^{-1}$。由于相似矩阵的特征根一样，所以 A^{-1} 的特征根是 $\lambda_1^{-1},\lambda_2^{-1},\cdots,\lambda_n^{-1}$。

✍ **练习 2.42** 令

$$\boldsymbol{A}=\begin{pmatrix}0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & 0 & \cdots & 0\end{pmatrix}$$

是一个 n 阶矩阵.

(i) 计算 $\boldsymbol{A}^2,\boldsymbol{A}^3,\cdots,\boldsymbol{A}^{n-1}$;

(ii) 求 $\boldsymbol{A}$ 的全部特征根.

解

(1) 首先取 $\mathbb{R}^n$ 的标准基

$$\alpha_k=\underbrace{(0,0,\cdots,0,1,0\cdots,0)}_{\text{第 } k \text{ 个为 } 1},\quad k=1,2,3,\cdots,n.$$

那么 $\mathscr{A}(X)=AX$ 在上面的基下的矩阵表示就是 A。并且

$$\mathscr{A}(\alpha_k)=\alpha_{k-1},\quad k=1,2,3,\cdots,n.\quad \text{这里我们令 } \alpha_0=\alpha_n.$$

于是对 $2\le l\le n-1$ 有

$$\mathscr{A}^l(\alpha_k)=\mathscr{A}^{l-1}(\alpha_{k-1})=\cdots=\mathscr{A}(\alpha_{k-l+1})=\alpha_{k-l},\quad k=1,2,\cdots,n.$$

这里我们采取记号: 对整数 $s\in\mathbb{Z}$,

$$\alpha_s=\alpha_k,\ \text{如果 } s=nj+k\text{，也即 } s \text{ 除以 } n \text{ 的余数是 } k\text{。}$$

于是当 $l=2$ 时,

$$\mathscr{A}^2(\alpha_k)=\alpha_{k-2},\quad k=1,2,\cdots,n.$$

也即

$$\mathscr{A}^2(\alpha_1)=\alpha_{-1}=\alpha_{n-1},\mathscr{A}^2(\alpha_2)=\alpha_0=\alpha_n,\quad ,\mathscr{A}^2(\alpha_3)=\alpha_1,\cdots,\mathscr{A}^2(\alpha_n)=\alpha_{n-2}.$$

从而 $\mathscr{A}^2$ 在自然基下的矩阵为

$$A^2=\left(\begin{array}{cc|cccc} 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ \hline 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 0 & \cdots & 0 \end{array}\right)$$

于是当 $l=3$ 时，

$$\mathscr{A}^3(\alpha_k)=\alpha_{k-3},\quad k=1,2,\cdots,n.$$

也即

$$\mathscr{A}^3(\alpha_1)=\alpha_{-2}=\alpha_{n-2},\mathscr{A}^3(\alpha_2)=\alpha_{-1}=\alpha_{n-1},\mathscr{A}^3(\alpha_3)=\alpha_0=\alpha_n,$$

和

$$\mathscr{A}^3(\alpha_4)=\alpha_1,\cdots,\mathscr{A}^3(\alpha_n)=\alpha_{n-3}.$$

从而 $\mathscr{A}^2$ 在自然基下的矩阵为

$$A^3=\left(\begin{array}{ccc|cccc} 0 & 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 1 \\ \hline 1 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & 0 & \cdots & 0 \end{array}\right)=\begin{pmatrix} 0 & I_{n-3} \\ I_3 & 0 \end{pmatrix}.$$

于是对 $2\leq l\leq n-1$，有

$$A^l=\begin{pmatrix} 0 & I_{n-l} \\ I_l & 0 \end{pmatrix}.$$

当然 $l=n$ 时

$$\mathscr{A}^n(\alpha_k)=\mathscr{A}^{n-k}(\mathscr{A}^k(\alpha_k))=\mathscr{A}^{n-k}(\alpha_0)=\mathscr{A}^{n-k}(\alpha_n)=\alpha_k,\quad ,k=1,2,3\cdots,n.$$

于是

$$\mathscr{A}^n-Id=0.$$

于是 $A^n=I_n$.

(2) 因为 $\mathscr{A}^n=1$, 所以 A 的最小多项式 $d(\lambda)|(\lambda^n-1)$. 对任意的 m 次多项式

$$f(x)=\sum_{i=0}^m a_ix^i,\quad a_m\neq 0,\quad m<n.$$

则

$$f(A)=\sum_{l=0}^{m}a_lA^l=\sum_{l=0}^{m}a_l\begin{pmatrix}0 & I_{n-l}\\ I_l & 0\end{pmatrix}$$

显然 $f(A)$ 的 $(m-l+1,1)$ 位置处的元素是 $a_m\neq 0$. 所以 $f(A)\neq 0$。也就是说任意的次数小于 n 的多项式作用在 A 上都不是零。从而 A 的特征多项式和最小多项式一样，又因为 $d(\lambda)|(\lambda^n-1)$。所以特征多项式为 $\varphi(\lambda)=\lambda^n-1$。从而特征值是 n 次单位根

$$e^{i\frac{2\pi k}{n}},\quad k=0,1,,\cdots,n-1.$$

✍ **练习 2.43** 令 $a_1,a_2,\cdots,a_n$ 是任意复数, 行列式

$$D=\begin{vmatrix}a_1 & a_2 & a_3 & \cdots & a_n\\ a_n & a_1 & a_2 & \cdots & a_{n-1}\\ a_{n-1} & a_n & a_1 & \cdots & a_{n-2}\\ \vdots & \vdots & \vdots & & \vdots\\ a_2 & a_3 & a_4 & \cdots & a_1\end{vmatrix}$$

叫作一个循环行列式. 证明:

$$\det(D)=f(\omega_1)f(\omega_2)\cdots f(\omega_n),$$

这里 $f(x)=a_1+a_2x+\cdots+a_nx^{n-1}$, 而 $\omega_1,\omega_2,\cdots,\omega_n$ 是全部 n 次单位根.

解 定义矩阵 A 如上一题，也即

$$\boldsymbol{A}=\begin{pmatrix}0 & 1 & 0 & 0 & \cdots & 0\\ 0 & 0 & 1 & 0 & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots & & \vdots\\ 0 & 0 & 0 & 0 & \cdots & 1\\ 1 & 0 & 0 & 0 & \cdots & 0\end{pmatrix}$$

则有

$$D=\sum_{k=2}^{n}a_kA^{k-1}+a_1I_n=f(A).$$

因为 A 可逆并且 A 的特征值是 n 次单位根

$$e^{i\frac{2\pi k}{n}},\quad k=0,1,,\cdots,n-1.$$

从而 $f(A)$ 的特征值是

$$f(e^{i\frac{2\pi k}{n}}),\quad k=0,1,,\cdots,n-1.$$

于是 D 的行列式为

$$\det(D)=\prod_{k=0}^{n-1}f(e^{i\frac{2\pi k}{n}}).$$

✍ **练习 2.44** 设 $\boldsymbol{A},\boldsymbol{B}$ 是复数域上 n 阶矩阵. 证明, $\boldsymbol{AB}$ 与 $\boldsymbol{BA}$ 有相同的特征根, 并且对应的特征根的重数也相同.

解 根据下面的引理 2.5 即得二者的特征多项式一样。

引理 2.5

设 $A \in \mathbb{F}^{m\times n}, B \in \mathbb{F}^{n\times m}$. 则有

$$\det(I_m - AB) = \det(I_n - BA).$$

取 $m+n$ 阶方阵

$$D = \begin{pmatrix} I_m & A \\ B & I_n \end{pmatrix}.$$

则

$$\begin{pmatrix} I_m & 0 \\ -B & I_n \end{pmatrix}\begin{pmatrix} I_m & A \\ B & I_n \end{pmatrix} = \begin{pmatrix} I_m & A \\ 0 & I_n - BA \end{pmatrix}.$$

同样

$$\begin{pmatrix} I_m & -A \\ 0 & I_n \end{pmatrix}\begin{pmatrix} I_m & A \\ B & I_n \end{pmatrix} = \begin{pmatrix} I_m - AB & 0 \\ B & I_n \end{pmatrix}.$$

取上述两式的行列式，得到

$$\det(D) = \det(I_n - BA) = \det(I_m - AB).$$

2.8 2024 年 5 月 7 日第十周星期二第十七次课-最小多项式-可对角化的矩阵

定义 2.29 (特征子空间)

设 λ_0 是线性变换 $\mathscr{A}: V \to V$ 的特征值. 属于 λ_0 的所有特征向量以及零向量构成的集合记为 V_{λ_0}, 即

$$V_{\lambda_0} = \{\boldsymbol{\alpha} \in V : \mathscr{A}(\boldsymbol{\alpha}) = \lambda_0\boldsymbol{\alpha}\}.$$

集合 V_{λ_0} 称为属于特征值 λ_0 的特征子空间.

换句话说，我们有

$$V_{\lambda_0} = \mathrm{Ker}(\lambda_0 \,\mathrm{Id} - \mathscr{A}),$$

定理 2.35 (特征子空间是不变子空间)

特征子空间 V_{λ_0} 是线性变换 $\mathscr{A}$ 的不变子空间.

令 V 是数域 $\mathbb{F}$ 上一个 n 维向量空间, 而 $\mathscr{A}$ 是 V 的一个线性变换. 设 λ_0 是 $\mathscr{A}$ 的一个本征值, V_{λ_0} 是 $\mathscr{A}$ 的属于本征值 λ_0 的本征子空间. 取 V_{λ_0} 的一个基 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_s$ 并且将它扩充为 V 的基, $\mathscr{A}$ 关于这个基的矩阵有形状

$$\boldsymbol{A} = \begin{pmatrix} \lambda_0 \mathbf{I}_s & \boldsymbol{A}_1 \\ \mathbf{0} & \boldsymbol{A}_2 \end{pmatrix},$$

这里 $\boldsymbol{I}_s$ 是一个 s 阶单位矩阵. 因此, $\boldsymbol{A}$ 的特征多项式是

$$\begin{aligned}\varphi_{\boldsymbol{A}}(\lambda) &= \begin{vmatrix} (\lambda-\lambda_0)\boldsymbol{I}_s & -\boldsymbol{A}_1 \\ \boldsymbol{0} & \lambda\boldsymbol{I}_{n-s}-\boldsymbol{A}_2 \end{vmatrix} \\ &= (\lambda-\lambda_0)^s \det\left(\lambda\boldsymbol{I}_{n-s}-\boldsymbol{A}_2\right) \\ &= (\lambda-\lambda_0)^s g(\lambda).\end{aligned}$$

由此可见, λ_0 至少是 $\varphi_{\boldsymbol{A}}(\lambda)$ 的一个 s 重根.

定义 2.30 (几何重数和代数重数)

特征子空间 V_{λ_0} 的维数 m_0 称为线性变换 $\mathcal{A}$ 的特征值 λ_0 的几何重数, 而线性变换 $\mathscr{A}$ 的特征多项式 $\varphi(\lambda)$ 的根 λ_0 的重数称为特征值 λ_0 的代数重数.

定理 2.36

线性变换 $\mathscr{A}$ 的特征值 λ_0 的几何重数不超过它的代数重数.

这一节的习题全部是课堂所讲，请看回放，不写答案了。

练习 2.45 请回答下面的概念类问题。

1. 什么是线性变换的特征值和特征向量?
2. 什么是矩阵的特征值和特征向量?
3. 什么是矩阵的特征多项式?
4. 什么是矩阵的最小多项式?
5. 什么是线性变换的特征多项式?
6. 什么是线性变换的最小多项式?
7. 什么是可对角化的线性变换?
8. 什么是可对角化的矩阵?
9. 什么是特征子空间?

练习 2.46 证明：特征子空间是不变子空间.

练习 2.47 线性变换 $\mathscr{A}$ 的特征值 λ_0 的几何重数不超过它的代数重数.

练习 2.48 设 $\mathscr{A}: V \to V$ 是线性变换, 在基

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

下的矩阵表示为 $\boldsymbol{A}$, 求证：对任意的多项式 $f(\lambda)$, 线性变换 $f(\mathscr{A})$ 在基

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

下的矩阵表示为 $f(\boldsymbol{A})$.

练习 2.49 设 $\mathscr{A}: V \to V$ 是线性变换, 在基

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

下的矩阵表示为 $\boldsymbol{A}$。假定 $f(\lambda)$ 是一个多项式，证明 $f(\mathscr{A}) = 0$ 当且仅当 $f(\boldsymbol{A}) = 0$. 你要心里明白这里的 0 都代表什么。

练习 2.50 纯量 $c \in \mathbb{F}$ 为线性变换 $\mathscr{A}$ 的最小多项式的根的充分必要条件是, c 是线性变换 $\mathscr{A}$ 的特征值.

2.9 2024 年 5 月 9 日第十周星期四第十八次课-可对角化的矩阵

定理 2.37 (不同特征值对应的特征子空间的和是直和)

假定 $\lambda_1, \lambda_2, \cdots, \lambda_t$ 是线性变换 A 的互异的特征值 (不一定是全部的特征值), 线性变换 $\mathscr{A}$ 的属于特征值 $\lambda_1, \lambda_2, \cdots, \lambda_t$ 的特征子空间依次记为 $V_{\lambda_1}, V_{\lambda_2}, \cdots, V_{\lambda_t}$。则有特征子空间 $V_{\lambda_1}, V_{\lambda_2}, \cdots, V_{\lambda_t}$ 的和

$$V_{\lambda_1} + V_{\lambda_2} + \cdots + V_{\lambda_t}$$

是直和.

定义 2.31 (可对角化的线性变换)

设 $\mathscr{A}$ 是数域 $\mathbb{F}$ 上 $n(n \geqslant 1)$ 维向量空间 V 的一个线性变换. 如果存在 V 的一个基, 使得 $\mathscr{A}$ 关于这个基的矩阵是对角矩阵

$$\begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix},$$

那么就说线性变换 $\mathscr{A}$ 为可对角化的.

定理 2.38 (可对角化的充要条件)

令 $\mathscr{A}$ 是数域 $\mathbb{F}$ 上 n 维向量空间 V 的一个线性变换, $\mathscr{A}$ 可以对角化的充要的条件是

(i) $\mathscr{A}$ 的特征多项式的根都在 $\mathbb{F}$ 内;

(ii) 对于 $\mathscr{A}$ 的特征多项式的每一根 λ 的几何重数等于它的代数重数. 换句话说, 设 $\lambda_1, \lambda_2, \cdots, \lambda_t$ 是线性变换 $\mathscr{A}: \cdot V \to V$ 的全部不同的特征值, λ_j 的几何重数等于它的代数重数, $j = 1, 2, \cdots, t$.

定义 2.32 (可对角化的矩阵)

设 $\boldsymbol{A}$ 是数域 $\mathbb{F}$ 上一个 n 阶矩阵. 如果存在 $\mathbb{F}$ 上一个 n 阶可逆矩阵 T, 使得 $T^{-1}\boldsymbol{A}T$ 为对角矩阵, 那么就说矩阵 $\boldsymbol{A}$ 可以对角化.

定义 2.33 (完全特征向量组)

设 $\mathscr{A}: V \to V$ 是线性变换, $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 是 V 的基. 如果每个向量 $\boldsymbol{\alpha}_j$ 都是 $\mathscr{A}$ 的特征向量, 则 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 称为完全特征向量组.

定理 2.39

n 维向量空间 V 的一个线性变换 $\mathscr{A}$ 可以对角化的充要的条件是，V 可以分解为 n 个在 $\mathscr{A}$ 之下不变的一维子空间 $W_1, W_2, \cdots, W_n$ 的直和.

然而一维不变子空间的每一非零向量都是 $\mathscr{A}$ 的属于某一本征值的本征向量, 所以上述条件相当于说, 在 V 中存在由 $\mathscr{A}$ 的本征向量所组成的基.

定理 2.40

线性变换 $\mathscr{A}$ 可对角化的充分必要条件是存在一组完全特征向量组.

推论 2.4

令 $\mathscr{A}$ 是数域 $\mathbb{F}$ 上 n 维向量空间 V 的一个线性变换. 如果 $\mathscr{A}$ 的特征多项式 $\varphi_{\mathscr{A}}(\lambda)$ 在 $\mathbb{F}$ 内有 n 个单根, 那么存在 V 的一个基, 使 $\mathscr{A}$ 关于这个基的矩阵是对角矩阵.

推论 2.5

令 $\boldsymbol{A}$ 是数域 $\mathbb{F}$ 上一个 n 阶矩阵. 如果 $\boldsymbol{A}$ 的特征多项式 $\varphi_{\boldsymbol{A}}(\lambda)$ 在 $\mathbb{F}$ 内有 n 个单根, 那么存在一个 n 阶可逆矩阵 $\mathbf{T}$, 使

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

推论 2.5 的条件只是一个 n 阶矩阵可以对角化的充分条件, 但不是必要条件. 例如, n 阶单位矩阵 I_n 本身就是对角矩阵, 但它的特征根只是 n 重根 1 .

例题 2.12 设 $\mathscr{A}: V \to V$ 是线性变换，并且 $\dim(V) = n$。证明下面四条等价：

(1) $\mathscr{A}$ 可对角化；

(2) $\mathscr{A}$ 有 n 个线性无关的特征向量；

(3) 如果 $\lambda_1, \lambda_2, \cdots, \lambda_t$ 是 $\mathscr{A}$ 的所有的特征值，属于特征值 $\lambda_1, \lambda_2, \cdots, \lambda_t$ 的特征子空间依次记为 $V_{\lambda_1}, V_{\lambda_2}, \cdots, V_{\lambda_t}$，则

$$V = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_t}.$$

(4) $\mathscr{A}$ 的最小多项式 $d_A(\lambda)$ 是不同的一次因式的乘积，也就是说

$$d_A(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_t)$$

定理 2.41

设 V 是复数域 $\mathbf{C}$ 上 n 维线性空间. $\varphi_{\mathscr{A}}(\lambda)$ 为 V 的线性变换 $\mathscr{A}$ 的特征多项式，且有因式分解：

$$\varphi_{\mathscr{A}}(\lambda) = (\lambda - \lambda_1)^{n_1}(\lambda - \lambda_2)^{n_2}\cdots(\lambda - \lambda_s)^{n_s},$$

其中 $\lambda_i \in \mathbb{C}$ 并且当 $i \neq j$ 时，$\lambda_i \neq \lambda_j$。则 V 可分解为 $\mathscr{A}$ 的不变子空间的直和：

$$V = \mathscr{R}_{\lambda_1}(\mathscr{A}) \oplus \mathscr{R}_{\lambda_2}(\mathscr{A}) \oplus \cdots \oplus \mathscr{R}_{\lambda_s}(\mathscr{A}),$$

其中

$$\mathscr{R}_{\lambda_i}(\mathscr{A}) = \operatorname{Ker}\left(\mathscr{A} - \lambda_i \mathrm{id}\right)^{n_i} = \left\{\alpha \in V \mid \left(\mathscr{A} - \lambda_i \mathrm{id}\right)^{n_i} \alpha = 0\right\},$$

称为 $\mathscr{A}$ 的属于 λ_i 的根子空间。

推论 2.6

我们有

(1)

$$\dim(\mathscr{R}_{\lambda_i}) = n_i.$$

(2) 对任意的 $k \geq n_i$,

$$\mathscr{R}_{\lambda_i} = \left\{\alpha \in V : (\mathscr{A} - \lambda_i \mathrm{Id})^k \alpha = 0\right\}.$$

练习 2.51 设 V 是 $\mathbb{C}$ 上的 n 维向量空间，$\mathscr{A}, \mathscr{B} : V \to V$ 是两个线性变换并且 $\mathscr{A}\mathscr{B} = \mathscr{B}\mathscr{A}$。证明

(1) $\mathscr{A}$ 和 $\mathscr{B}$ 至少有一个公共的特征向量。

(2) 如果 $\mathscr{A}$ 和 $\mathscr{B}$ 均可以对角化，则存在 V 的一组基使得 $\mathscr{A}$ 和 $\mathscr{B}$ 在这组基下的矩阵表示都是对角阵。

证明

(1) 首先 $\mathbb{C}$ 上的多项式一定有根，所以 $\mathscr{A}$ 的特征多项式一定有根 $\lambda \in \mathbb{C}$，对应的特征向量记为 α。则对应的特征子空间 V_λ 有如下形式

$$V_\lambda = \{k\alpha \mid k \in \mathbb{C}\}.$$

任取 $k\alpha \in V_\lambda$, 则有

$$\mathscr{A}(\mathscr{B}(k\alpha)) = k\mathscr{B}(\mathscr{A}(\alpha)) = k\mathscr{B}(\lambda\alpha) = \lambda\mathscr{B}(k\alpha).$$

于是 $\mathscr{B}(k\alpha) \in V_\lambda$, 于是 V_λ 是 $\mathscr{B}$ 的不变子空间，那么因为在复数域上，$\mathscr{B}|_{V_\lambda}$ 一定有特征向量, 记为 β。于是 β 就是 $\mathscr{A}$ 和 $\mathscr{B}$ 的公共特征向量。

(2) 首先因为 $\mathscr{A}$ 和 $\mathscr{B}$ 都可以对角化，于是 V 就有分解

$$V = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_k} = V_{\mu_1} \oplus V_{\mu_2} \oplus \cdots \oplus V_{\mu_l}$$

其中 $\{\lambda_1, \lambda_2, \cdots, \lambda_k\}$ 是 $\mathscr{A}$ 的所有互不相同的特征值，对应的 $V_{\lambda_1}, V_{\lambda_2}, \cdots, V_{\lambda_k}$ 是相应的特征子空间;同样 $\{\mu_1, \mu_2, \cdots, \mu_l\}$ 是 $\mathscr{B}$ 的所有互不相同的特征值,对应的 $V_{\mu_1}, V_{\mu_2}, \cdots, V_{\mu_l}$ 是相应的特征子空间。于是

$$V = \oplus_{i=1}^k \oplus_{j=1}^l V_{\lambda_i} \cap V_{\mu_j}.$$

当然有的 $V_{\lambda_i} \cap V_{\mu_j} = \{0\}$. 我们记那些相交不是零的指标如下

$$S = \{(i,j) \mid V_{\lambda_i} \cap V_{\mu_j} \neq \{0\}\}$$

则

$$V = \oplus_{(i,j)\in S} V_{\lambda_i} \cap V_{\mu_j}.$$

我们选取这些子和空间的基组成 V 的基，因为这些向量都是 $\mathscr{A}$ 和 $\mathscr{B}$ 的特征向量，从而在这组基的 $\mathscr{A}$ 和 $\mathscr{B}$ 的矩阵都是对角矩阵。

练习 2.52 设 V 是 $\mathbb{C}$ 上的 n 维向量空间，$\mathscr{A}: V \to V$ 是线性变换并且有 n 个不同的特征值。证明：$\mathscr{A}$ 只有 2^n 个不变子空间。

证明 首先假定 $\mathscr{A}$ 的 n 个不同的特征值是

$$\lambda_1, \lambda_2, \cdots, \lambda_n.$$

对应的特征向量记为

$$\alpha_1, \alpha_2, \cdots, \alpha_n$$

那么对任意的 $S \subseteq \{1, 2, 3, \cdots, n\}$,

$$V_S = \mathcal{L}(\alpha_1, \alpha_2, \cdots, \alpha_n)$$

是 V 的不变子空间, 这样的不变子空间有 2^n 个。假定 W 是任意的一个不变子空间，不妨设 $W \neq \{0\}$ 且 $W \neq V$。记 $\dim(W) = k$, 那么 $\mathscr{A}$ 在 W 上的限制也是一个 W 上的线性变换，从而一定有 k 个不同的特征值以及 k 的与之对应的特征向量。于是这 k 个特征向量就张成了 W。也即一定存在子集 $S \subseteq \{1, 2, 3, \cdots, n\}$ 使得 $W = V_S$。所以总的不变子空间的个数有 2^n 个。

练习 2.53 数域 $\boldsymbol{F}$ 上 $\boldsymbol{n}$ 维向量空间 $\boldsymbol{V}$ 的一个线性变换 $\mathscr{A}$ 叫作一个对合变换，如果 $\mathscr{A}^2 = \mathrm{Id}, \mathrm{Id}$ 是恒等变换。设 $\mathscr{A}$ 是 $\boldsymbol{V}$ 的一个对合变换。证明：

(i) $\mathscr{A}$ 的本征值只能是 ± 1;

(ii) $V = V_1 \oplus V_{-1}$, 这里 V_1 是 $\mathscr{A}$ 的属于本征值 1 的本征子空间，V_{-1} 是 $\mathscr{A}$ 的属于本征值-1 的本征子空间.

解

(i) 注意到 $\mathscr{A}^2 - Id = 0$, 因此最小多项式整除 $\lambda^2 - 1$. 从而最小多项式的根只能是 -1 和 1。由于特征值是最小多项式的根，因此特征值只能是 ± 1。

(ii) 注意到

$$\alpha = \frac{\mathscr{A}(\alpha) + \alpha}{2} + \frac{-\mathscr{A}(\alpha) + \alpha}{2}$$

并且

$$\mathscr{A}\left(\frac{\mathscr{A}(\alpha) + \alpha}{2}\right) = \frac{\mathscr{A}^2(\alpha) + \mathscr{A}(\alpha)}{2} = \frac{\mathscr{A}(\alpha) + \alpha}{2}$$

以及

$$\mathscr{A}\left(\frac{-\mathscr{A}(\alpha) + \alpha}{2}\right) = \frac{-\mathscr{A}^2(\alpha) + \mathscr{A}(\alpha)}{2} = -\frac{-\mathscr{A}(\alpha) + \alpha}{2}.$$

于是

$$\frac{\mathscr{A}(\alpha) + \alpha}{2} \in V_1, \quad \frac{-\mathscr{A}(\alpha) + \alpha}{2} \in V_{-1}.$$

所以

$$V = V_1 + V_{-1}.$$

如果 $\alpha \in V_1 \cap V_{-1}$, 则 $\mathscr{A}(\alpha) = \pm\alpha$. 于是 $-\alpha = \alpha$, 则 $\alpha = 0$, 所以 $V_1 \cap V_{-1} = 0$. 从而

$$V = V_1 \oplus V_{-1}.$$

解 如果不借助最小多项式，可以这样解。如果 $\lambda \in \mathbb{F}$ 是特征值，则有 $\alpha \in V$ 是对应的特征向量。也即

$$\alpha \neq 0, \quad \mathscr{A}(\alpha) = \lambda\alpha.$$

于是

$$\mathscr{A}^2(\alpha) = \lambda^2\alpha = \alpha.$$

因为 $\alpha \neq 0$, 所以

$$\lambda^2 = 1.$$

练习 2.54 数域 $\boldsymbol{F}$ 上一个 $\boldsymbol{n}$ 阶矩阵 $\boldsymbol{A}$ 叫作一个幂等矩阵，如果 $\boldsymbol{A}^2 = \boldsymbol{A}$. 设 $\boldsymbol{A}$ 是一个幂等矩阵。证明

(i) $\boldsymbol{I}_n + \boldsymbol{A}$ 可逆，并且求 $(\boldsymbol{I}_n + \boldsymbol{A})^{-1}$;

(ii) 秩 $\boldsymbol{A}+$ 秩 $(\boldsymbol{I}_n - \boldsymbol{A}) = \boldsymbol{n}$;

解

(i) 令 $B = I_n + A$, 则

$$(B - I_n)^2 - B + I_n = 0.$$

也即

$$B^2 - 3B + 2I_n = 0.$$

从而

$$B(3I_n - B)\frac{1}{2} = I_n.$$

于是 B 可逆，并且

$$B^{-1} = \frac{1}{2}(3I_n - B) = \frac{1}{2}(2I_n - A).$$

(ii) 考虑到

$$n - \dim(\mathrm{Im}(I_n - A)) = n - \mathrm{rank}(I_n - A) = \dim(\mathrm{Ker}(I_n - A)).$$

只需要证明

$$\dim(\mathrm{Ker}(I_n - A)) = \mathrm{rank}(A).$$

因为

$$\mathrm{rank}(A) = \dim(\mathrm{Im}(A)).$$

只需要证明

$$\dim(\mathrm{Ker}(I_n - A)) = \dim(\mathrm{Im}(A)).$$

如果 $X \in \mathrm{Ker}(I_n - A)$, 则 $AX = X$, 于是 $X \in \mathrm{Im}(A)$. 如果 $X = AY \in \mathrm{Im}(A)$, 则 $AX = A^2Y = AY = X$, 从而 $X \in \mathrm{Ker}(I_n - A)$. 于是

$$\mathrm{Ker}(I_n - A) = \mathrm{Im}(A).$$

✍ **练习 2.55** 数域 $\boldsymbol{F}$ 上 $\boldsymbol{n}$ 维向量空间 $\boldsymbol{V}$ 的一个线性变换 $\mathscr{A}$ 叫作幂零的，如果存在一个正整数 m 使 $\mathscr{A}^m = 0$. 证明：

(i) $\mathscr{A}$ 是幂零变换当且仅当它的特征多项式的根都是零；

(ii) 如果一个幂零变换 $\mathscr{A}$ 可以对角化，那么 $\mathscr{A}$ 一定是零变换.

这里默认特征多项式的根都在数域 $\mathbb{F}$ 中。

解

(i) $\Rightarrow$: 如果 $\mathscr{A}^m = 0$, 则 $\mathscr{A}$ 的最小多项式 $d(\lambda)|\lambda^m$. 从而 $d(\lambda)$ 在 $\mathbb{F}$ 只有零根。而特征值是最小多项式的根（参见习题 2.50），因此 $\mathscr{A}$ 的特征值都是零。也即特征多项式的根都是零。

$\Leftarrow$: 假定 $\mathcal{A}$ 的特征多项式的根都是零。则存在 $m \geq 1$ 使得特征多项式 $\varphi_{\mathscr{A}}(\lambda) = \lambda^m$. 于是根据 Calay-Hamiton 定理知道 $\mathscr{A}^m = 0$.

(ii) 如果 $\mathscr{A}$ 可对角化，则存在一组基 $\{\alpha_1, \cdots, \alpha_n\}$ 使得

$$\mathscr{A}(\alpha_k) = \lambda_k \alpha_k, \quad k = 1, 2, \cdots, n.$$

根据 (i) 知道 $\lambda_k = 0, \quad k = 1, 2, \cdots, n$。于是 $\mathscr{A} = 0$.

解 第一问不借助最小多项式可以这样解。

$\Rightarrow$: 如果 $\mathscr{A}^m = 0$。假定 λ 是特征多项式的根，则 λ 是特征值，对应有特征向量 $\alpha \neq 0$ 且 $\mathscr{A}(\alpha) = \lambda\alpha$. 于是

$$0 = \mathscr{A}^m(\alpha) = \lambda^m \alpha.$$

因为 $\alpha \neq 0$, 于是 $\lambda^m = 0$, 从而 $\lambda = 0$。

$\Leftarrow$: 假定 $\mathcal{A}$ 的特征多项式的根都是零。则因为特征多项式是 n 次多项式，那么特征多项式一定是 $\varphi_{\mathscr{A}}(\lambda) = \lambda^n$. 于是根据 Calay-Hamiton 定理知道 $\mathscr{A}^n = 0$.

第 3 章 欧式空间

内容提要

3.1 2024 年 5 月 11 日第十周星期六第十九次课欧式空间的定义

高中学的东西

- 在空间 V_3 中建立直角坐标系后, 向量 $\boldsymbol{\alpha}$ 便唯一地确定一个坐标 (x_1, x_2, x_3). 在空间 V 中向量的长度和向量间的夹角可以用向量的纯量积表示.
- 设向量 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 的坐标分别是 (x_1, x_2, x_3) 与 (y_1, y_2, y_3), 则向量 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 的纯量积为

$$\boldsymbol{\alpha} \cdot \boldsymbol{\beta} = x_1 y_1 + x_2 y_2 + x_3 y_3.$$

- 向量 $\boldsymbol{\alpha}$ 的长为 $\|\boldsymbol{\alpha}\| = \sqrt{(\boldsymbol{\alpha}, \boldsymbol{\alpha})}$, 向量 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 的夹角 θ 的余弦为

$$\cos\theta = \frac{\boldsymbol{\alpha} \cdot \boldsymbol{\beta}}{\sqrt{(\boldsymbol{\alpha}, \boldsymbol{\alpha})(\boldsymbol{\beta}, \boldsymbol{\beta})}}.$$

由解析几何知道 V 中向量的内积具有下列性质:

1. 对任意的 $\boldsymbol{\xi}, \boldsymbol{\eta} \in V_3$, $\boldsymbol{\xi} \cdot \boldsymbol{\eta} = \boldsymbol{\eta} \cdot \boldsymbol{\xi}$
2. 对 $\boldsymbol{\xi}, \boldsymbol{\eta}, \boldsymbol{\zeta} \in V_3$，有

$$(\boldsymbol{\xi} + \boldsymbol{\eta}) \cdot \boldsymbol{\zeta} = \boldsymbol{\xi} \cdot \boldsymbol{\zeta} + \boldsymbol{\eta} \cdot \boldsymbol{\zeta}$$

3. 对任意的 $a \in \mathbb{R}$ 和 $\boldsymbol{\xi}, \boldsymbol{\eta} \in V_3$，有 $(a\boldsymbol{\xi}) \cdot \boldsymbol{\eta} = a(\boldsymbol{\xi} \cdot \boldsymbol{\eta})$
4. 如果 $\boldsymbol{\xi} \neq 0$, 则 $\boldsymbol{\xi} \cdot \boldsymbol{\xi} > 0$。

定义 3.1 (内积)

设 V 是实数域 $\mathbb{R}$ 上一个向量空间. 如果对于 V 中任意一对向量 $\boldsymbol{\xi}, \boldsymbol{\eta}$, 有一个确定的记作 $\langle \boldsymbol{\xi}, \boldsymbol{\eta} \rangle$ 的实数与它们对应, 叫作向量 $\boldsymbol{\xi}$ 与 $\boldsymbol{\eta}$ 的内积 (或标量积), 并且下列条件被满足:

(i) (对称性) $\langle \boldsymbol{\xi}, \boldsymbol{\eta} \rangle = \langle \boldsymbol{\eta}, \boldsymbol{\xi} \rangle$;

(ii) (“加法分配律”) $\langle \boldsymbol{\xi} + \boldsymbol{\eta}, \boldsymbol{\zeta} \rangle = \langle \boldsymbol{\xi}, \boldsymbol{\zeta} \rangle + \langle \boldsymbol{\eta}, \boldsymbol{\zeta} \rangle$;

(iii) (数乘“结合律”) $\langle a\boldsymbol{\xi}, \boldsymbol{\eta} \rangle = a\langle \boldsymbol{\xi}, \boldsymbol{\eta} \rangle$;

(iv) (恒正性) 当 $\boldsymbol{\xi} \neq 0$ 时, $\langle \boldsymbol{\xi}, \boldsymbol{\xi} \rangle > 0$;

这里 $\boldsymbol{\xi}, \boldsymbol{\eta}, \boldsymbol{\zeta}$ 是 V 的任意向量, a 是任意实数, 那么 V 叫作对这个内积来说的一个欧几里得 (Euclid) 空间 (简称欧氏空间).

注

(1) 这里, 我们不用 $\boldsymbol{\xi}\cdot\boldsymbol{\eta}$ 而用 $\langle\boldsymbol{\xi},\boldsymbol{\eta}\rangle$ 表示 $\boldsymbol{\xi}$ 与 $\boldsymbol{\eta}$ 的内积.

(2) 有时候，我们也用 $(\boldsymbol{\xi},\boldsymbol{\eta})$ 表示 $\boldsymbol{\xi}$ 与 $\boldsymbol{\eta}$ 的内积.

注 事实上，上面的 (ii) 和 (iii) 等价于

$$\langle a\boldsymbol{\xi}+b\boldsymbol{\eta},\boldsymbol{\zeta}\rangle=a\langle\boldsymbol{\xi},\boldsymbol{\zeta}\rangle+b\langle\boldsymbol{\eta},\boldsymbol{\zeta}\rangle,\quad \forall a,b\in\mathbb{R}.$$

这叫线性性。

例题 3.1 在 $\mathbf{R}^n$ 里, 对于任意两个向量

$$\begin{aligned}\boldsymbol{\xi}&=(x_1,x_2,\cdots,x_n),\\ \boldsymbol{\eta}&=(y_1,y_2,\cdots,y_n),\\ \langle\boldsymbol{\xi},\boldsymbol{\eta}\rangle&=x_1y_1+x_2y_2+\cdots+x_ny_n.\end{aligned}$$

例题 3.2 在 $\mathbf{R}^n$ 里, 对于任意向量

$$\begin{aligned}\boldsymbol{\xi}&=(x_1,x_2,\cdots,x_n),\\ \boldsymbol{\eta}&=(y_1,y_2,\cdots,y_n),\end{aligned}$$

规定

$$\langle\boldsymbol{\xi},\boldsymbol{\eta}\rangle=x_1y_1+2x_2y_2+\cdots+nx_ny_n$$

例题 3.3 令 $C[a,b]$ 是定义在 $[a,b]$ 上一切连续实函数所成的向量空间. 设 $f(x),g(x)\in C[a,b]$, 我们规定

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

定积分的基本性质可推出内积公理 (i)-(iv) 都被满足, 因而 $C[a,b]$ 作成一个欧氏空间.

令 $\boldsymbol{H}$ 是一切平方和收敛的实数列

$$\boldsymbol{\xi}=(x_1,x_2,\cdots),\sum_{n=1}^{\infty}x_n^2<+\infty,$$

所成的集合. 在 $\boldsymbol{H}$ 中用自然的方式定义加法和标量与向量的乘法: 设

$$\boldsymbol{\xi}=(x_1,x_2,\cdots),\boldsymbol{\eta}=(y_1,y_2,\cdots),a\in\mathbf{R}.$$

规定

$$\boldsymbol{\xi}+\boldsymbol{\eta}=(x_1+y_1,x_2+y_2,\cdots)\quad a\xi=(ax_1,ax_2,\cdots).$$

例题 3.4 向量 $\boldsymbol{\xi}=(x_1,x_2,\cdots),\boldsymbol{\eta}=(y_1,y_2,\cdots)$ 的内积由公式

$$\langle\boldsymbol{\xi},\boldsymbol{\eta}\rangle=\sum_{n=1}^{\infty}x_ny_n$$

给出. 那么 $\boldsymbol{H}$ 是一个欧氏空间. 空间 $\boldsymbol{H}$ 通常叫作希尔伯特 (Hilbert) 空间.

命题 3.1 (内积的一些简单性质)

- 设 V 是一个欧氏空间. 由 (i) 及 (iii) 得出, 对于任意 $\boldsymbol{\xi}\in V$ 都有

$$\langle 0,\boldsymbol{\xi}\rangle=\langle\boldsymbol{\xi},0\rangle=0.$$

- 反过来, 如果对任意 $\boldsymbol{\eta} \in V$, 都有 $\langle \boldsymbol{\xi}, \boldsymbol{\eta} \rangle = 0$, 那么特别地有 $\langle \boldsymbol{\xi}, \boldsymbol{\xi} \rangle = 0$. 于是由 (iv), 必须 $\boldsymbol{\xi} = 0$.
- 其次, 由 (i), (ii), (iii), 对于任意 $\boldsymbol{\xi}, \boldsymbol{\eta}, \boldsymbol{\zeta} \in V$ 和任意 $a \in \mathbf{R}$, 我们有

$$\langle \boldsymbol{\zeta}, \boldsymbol{\xi} + \boldsymbol{\eta} \rangle = \langle \boldsymbol{\zeta}, \boldsymbol{\xi} \rangle + \langle \boldsymbol{\zeta}, \boldsymbol{\eta} \rangle;$$
$$\langle \boldsymbol{\xi}, a\boldsymbol{\eta} \rangle = a\langle \boldsymbol{\xi}, \boldsymbol{\eta} \rangle.$$

- 于是, 对于任意向量 $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \cdots, \boldsymbol{\xi}_r, \boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_3 \in \boldsymbol{V}, a_1, a_2, \cdots, a_r, b_1, b_2, \cdots, b_r \in \mathbf{R}$, 有

$$\left\langle \sum_{i=1}^{r} a_i \boldsymbol{\xi}_i, \sum_{j=1}^{r} b_j \boldsymbol{\eta}_j \right\rangle = \sum_{i=1}^{r} \sum_{j=1}^{\dot{n}} a_i b_j \left\langle \boldsymbol{\xi}_i, \boldsymbol{\eta}_j \right\rangle.$$

定义 3.2 (向量长度)

设 $\boldsymbol{\xi}$ 是欧氏空间的一个向量. 非负实数 $\langle \boldsymbol{\xi}, \boldsymbol{\xi} \rangle$ 的算术根 $\sqrt{\langle \boldsymbol{\xi}, \boldsymbol{\xi} \rangle}$ 叫作 $\boldsymbol{\xi}$ 的长度. 向量 $\boldsymbol{\xi}$ 的长度用符号 $|\boldsymbol{\xi}|$ 表示:

$$|\xi| = \sqrt{\langle \xi, \xi \rangle}.$$

欧氏空间的每一向量都有一个确定的长度. 零向量的长度是零, 任意非零向量的长度是一个正数.

例题 3.5 令 $\mathbf{R}^n$ 是例 3.1 中的欧氏空间. $\mathbf{R}^n$ 的向量

$$\xi = (x_1, x_2, \cdots, x_n)$$

的长度是

$$|\boldsymbol{\xi}| = \sqrt{\langle \boldsymbol{\xi}, \boldsymbol{\xi} \rangle} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

命题 3.2

对于欧氏空间中任意向量 $\boldsymbol{\xi}$ 和任意实数 a, 有

$$|a\boldsymbol{\xi}| = \sqrt{\langle a\boldsymbol{\xi}, a\boldsymbol{\xi} \rangle} = \sqrt{a^2 \langle \boldsymbol{\xi}, \boldsymbol{\xi} \rangle} = |a||\boldsymbol{\xi}|.$$

即一个实数 a 与一个向量 $\boldsymbol{\xi}$ 的乘积的长度等于 a 的绝对值与 $\boldsymbol{\xi}$ 的长度的乘积.

定义 3.3 (单位向量)

我们把长度是 1 的向量叫作单位向量. 由性质 3.2, 如果 ξ 是一个非零向量, 那么 $\boldsymbol{\xi}/|\boldsymbol{\xi}|$ 是一个单位向量.

定理 3.1 (Cauchy-Schwarz 不等式)

在一个欧氏空间里, 对于任意向量 $\boldsymbol{\xi}, \boldsymbol{\eta}$, 有不等式

$$\langle \boldsymbol{\xi}, \boldsymbol{\eta} \rangle^2 \leqslant \langle \boldsymbol{\xi}, \boldsymbol{\xi} \rangle \langle \boldsymbol{\eta}, \boldsymbol{\eta} \rangle \tag{3.1}$$

当且仅当 $\boldsymbol{\xi}$ 与 $\boldsymbol{\eta}$ 线性相关时 (3.1)才取等号.

例题 3.6 考虑例 3.1 的欧氏空间 $\mathbf{R}^n$. 由不等式 (3.1) 推出,

$$(a_1b_1+\cdots+a_nb_n)^2 \leqslant \left(a_1^2+\cdots+a_n^2\right)\left(b_1^2+\cdots+b_n^2\right). \tag{3.2}$$

这里 $a_1,a_2,\cdots,a_n,b_1,b_2,\cdots,b_n$ 是任意实数。不等式 (3.2) 叫作柯西 (Cauchy) 不等式.

例题 3.7 考虑例 3.3 的欧氏空间 $C[a,b]$. 由不等式 (3.1) 推出, 对于定义在 $[a,b]$ 上的任意连续函数 $f(x),g(x)$, 有不等式

$$\left|\int_a^b f(x)g(x)\mathrm{d}x\right| \leqslant \sqrt{\int_a^b f^2(x)\mathrm{d}x \int_a^b g^2(x)\mathrm{d}x}. \tag{3.3}$$

不等式 (3.3) 叫作施瓦兹 (Schwarz) 不等式.

定义 3.4 (欧氏空间中两个向量的夹角)

设 $\boldsymbol{\xi}$ 和 $\boldsymbol{\eta}$ 是欧氏空间的两个非零向量. $\boldsymbol{\xi}$ 与 $\boldsymbol{\eta}$ 的夹角 θ 由以下公式定义:

$$\cos\theta = \frac{\langle\boldsymbol{\xi},\boldsymbol{\eta}\rangle}{|\boldsymbol{\xi}||\boldsymbol{\eta}|}.$$

由不等式(3.1), 我们有

$$-1 \leqslant \frac{\langle\boldsymbol{\xi},\boldsymbol{\eta}\rangle}{|\boldsymbol{\xi}||\boldsymbol{\eta}|} \leqslant 1,$$

所以这样定义夹角是合理的. ♣

这样，欧氏空间任意两个非零向量有唯一的夹角 $\theta(0\leqslant\theta\leqslant\pi)$.

定义 3.5 (正交向量)

欧氏空间的两个向量 $\boldsymbol{\xi}$ 与 $\boldsymbol{\eta}$ 说是正交的, 如果

$$\langle\boldsymbol{\xi},\boldsymbol{\eta}\rangle = 0.$$

♣

例如, 在欧氏空间 $\mathbf{R}^n$ 里, 向量

$$\varepsilon_i = (0,\cdots,0,\overset{(i)}{1},0,\cdots,0), i=1,2,\cdots,n,$$

两两正交.

定义 3.6

在一个欧氏空间里，如果向量 $\boldsymbol{\xi}$ 与向量 $\boldsymbol{\eta}_1$，$\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_{\mathrm{r}}$ 中每一个正交，那么 ξ 与 $\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_r$, 的任意一个线性组合也正交. ♣

设 $\boldsymbol{\xi},\boldsymbol{\eta}$ 是欧氏空间的任意向量. 由定理 3.1, 我们有

$$\begin{aligned}
|\boldsymbol{\xi}+\boldsymbol{\eta}|^2 &= \langle\boldsymbol{\xi}+\boldsymbol{\eta},\boldsymbol{\xi}+\boldsymbol{\eta}\rangle \\
&= \langle\boldsymbol{\xi},\boldsymbol{\xi}\rangle + 2\langle\boldsymbol{\xi},\boldsymbol{\eta}\rangle + \langle\boldsymbol{\eta},\boldsymbol{\eta}\rangle \\
&\leqslant \langle\boldsymbol{\xi},\boldsymbol{\xi}\rangle + 2|\boldsymbol{\xi}||\boldsymbol{\eta}| + \langle\boldsymbol{\eta},\boldsymbol{\eta}\rangle \\
&= |\boldsymbol{\xi}|^2 + 2|\boldsymbol{\xi}||\boldsymbol{\eta}| + |\boldsymbol{\eta}|^2 \\
&= (|\boldsymbol{\xi}|+|\boldsymbol{\eta}|)^2
\end{aligned}$$

由于 $|\boldsymbol{\xi}+\boldsymbol{\eta}|$ 和 $|\boldsymbol{\xi}|+|\boldsymbol{\eta}|$ 都是非负实数, 所以我们有

$$|\boldsymbol{\xi}+\boldsymbol{\eta}| \leq |\boldsymbol{\xi}|+|\boldsymbol{\eta}|. \tag{3.4}$$

定义 3.7 (两个向量的距离)

在一个欧氏空间里, 两个向量 $\boldsymbol{\xi}$ 与 $\boldsymbol{\eta}$ 的距离指的是 $\boldsymbol{\xi}-\boldsymbol{\eta}$ 的长度 $|\boldsymbol{\xi}-\boldsymbol{\eta}|$. 我们用符号 $d(\boldsymbol{\xi},\boldsymbol{\eta})$ 表示 $\boldsymbol{\xi}$ 与 $\boldsymbol{\eta}$ 的距离. ♣

根据内积的定义和公式 (3.4), 容易看出,

命题 3.3 (距离具有下列性质)

(1) 当 $\boldsymbol{\xi} \neq \boldsymbol{\eta}$ 时, $d(\boldsymbol{\xi},\boldsymbol{\eta})>0$;

(2) $d(\boldsymbol{\xi},\boldsymbol{\eta})=d(\boldsymbol{\eta},\boldsymbol{\xi})$;

(3) 有如下不等式

$$d(\boldsymbol{\xi},\boldsymbol{\zeta}) \leqslant d(\boldsymbol{\xi},\boldsymbol{\eta})+d(\boldsymbol{\eta},\boldsymbol{\zeta}) \tag{3.5}$$

这里 $\boldsymbol{\xi},\boldsymbol{\eta},\boldsymbol{\zeta}$ 是欧氏空间的任意向量. ♠

注 不等式 (3.5)称为三角形不等式. 在解析几何里, 这个不等式的意义就是一个三角形两边的和大于第三边.

注 如果 W 是欧氏空间 V 的一个子空间，那么对于 V 的内积来说, W 显然也作成一个欧氏空间.

练习 3.1 证明三角不等式 (3.5)。

解 由(3.4),

$$d(\boldsymbol{\xi},\boldsymbol{\eta})+d(\boldsymbol{\eta},\boldsymbol{\zeta})=|\boldsymbol{\xi}-\boldsymbol{\eta}|+|\boldsymbol{\eta}-\boldsymbol{\zeta}| \geq |\boldsymbol{\xi}-\boldsymbol{\eta}+\boldsymbol{\eta}-\boldsymbol{\zeta}|=|\boldsymbol{\xi}-\boldsymbol{\zeta}|=d(\boldsymbol{\xi},\boldsymbol{\zeta}).$$

练习 3.2 验证例子 3.3 定义的空间是欧式空间。

解 $\forall f,g,h \in C[a,b]$,

1. $\langle f,g\rangle=\int_a^b f(x)\,g(x)\,\mathrm{d}x=\langle g,f\rangle$;
2. $\langle f+g,h\rangle=\int_a^b (f(x)+g(x))\,h(x)\,\mathrm{d}x=\int_a^b f(x)\,h(x)\,dx+\int_a^b g(x)\,h(x)\,\mathrm{d}x=\langle f,g\rangle+\langle g,h\rangle$;
3. $\langle cf,g\rangle=\int_a^b cf(x)\,g(x)\,\mathrm{d}x=c\int_a^b f(x)\,g(x)\,\mathrm{d}x=c\langle f,g\rangle$;
4. 当 f 不恒为 0 时, 设 $f(t)\neq 0$, $t\in[a,b]$, 由 f 的连续性, 存在一个 t 的在 $[a,b]$ 中邻域, 不妨记为 $[s,t]$, s.t. f 在 $[s,t]$ 上恒正或恒负, 所以 $\langle f,f\rangle=\int_a^b f^2(x)\,\mathrm{d}x \geq \int_s^t f^2(x)\,\mathrm{d}x>0$.

由定义, 例子 3.3 定义的空间是欧式空间.

练习 3.3 验证欧式空间的定义中 (ii) 和 (iii) 等价于

$$\langle a\boldsymbol{\xi}+b\boldsymbol{\eta},\boldsymbol{\zeta}\rangle=a\langle\boldsymbol{\xi},\boldsymbol{\zeta}\rangle+b\langle\boldsymbol{\eta},\boldsymbol{\zeta}\rangle,\quad \forall a,b\in\mathbb{R}.$$

解

($\Rightarrow$) 由 (ii) 和 (iii), $\langle a\boldsymbol{\xi}+b\boldsymbol{\eta},\boldsymbol{\zeta}\rangle=\langle a\boldsymbol{\xi},\boldsymbol{\zeta}\rangle+\langle b\boldsymbol{\eta},\boldsymbol{\zeta}\rangle=a\langle\boldsymbol{\xi},\boldsymbol{\zeta}\rangle+b\langle\boldsymbol{\eta},\boldsymbol{\zeta}\rangle,\quad \forall a,b\in\mathbb{R}$.

($\Leftarrow$) 分别令 $a=b=1$ 和 $b=0$ 即可.

3.2 2024 年 5 月 14 日第十一周星期二第二十次课-正交基

定义 3.8 (正交组)

欧氏空间 V 的一组两两正交的非零向量叫作 V 的一个正交组. 如果一个正交组的每一个向量都是单位向量，则叫做一个规范正交组。

例题 3.8 向量

$$\alpha_1 = (0,1,0), \alpha_2 = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right), \alpha_3 = \left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\right)$$

构成 $\mathbf{R}^3$ 的一个规范正交组，因为

$$\langle \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2 \rangle = \langle \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3 \rangle = \langle \boldsymbol{\alpha}_3, \boldsymbol{\alpha}_1 \rangle = 0. \quad |\boldsymbol{\alpha}_1| = |\boldsymbol{\alpha}_2| = |\boldsymbol{\alpha}_3| = 1.$$

例题 3.9 考虑定义在闭区间 $[0, 2\pi]$ 上一切连续函数所作成的欧氏空间 $C[0, 2\pi]$. 函数组

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

构成 $C[0, 2\pi]$ 的一个正交组。

定理 3.2 (正交组线性无关)

设 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 是欧氏空间的一个正交组，那么 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 线性无关。

定义 3.9 (正交基)

设 $\boldsymbol{V}$ 是一个 $\boldsymbol{n}$ 维欧氏空间。如果 $\boldsymbol{V}$ 中 n 个向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 构成一个正交组，那么由定理 3.2，这 $\boldsymbol{n}$ 个向量构成 $\boldsymbol{V}$ 的一个基。这样的一个基叫作 $\boldsymbol{V}$ 的一个正交基。如果 $\boldsymbol{V}$ 的一个正交基还是一个规范正交组，那么就称这个基是一个规范正交基.

- 如果 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 是 $\boldsymbol{n}$ 维欧氏空间 $\boldsymbol{V}$ 的一个规范正交基
- 令 $\boldsymbol{\xi}$ 是 V 的任意一个向量，那么 $\boldsymbol{\xi}$ 可以唯一地写成

  $$\boldsymbol{\xi} = \boldsymbol{x}_1\boldsymbol{\alpha}_1 + \boldsymbol{x}_2\boldsymbol{\alpha}_2 + \cdots + \boldsymbol{x}_n\boldsymbol{\alpha}_n,$$

  这里 $x_1, x_2, \cdots, x_n$ 是 $\boldsymbol{\xi}$ 关于基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 的坐标.
- 由于 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 是规范正交基，我们有

  $$\langle \boldsymbol{\xi}, \boldsymbol{\alpha}_i \rangle = \left\langle \sum_{j=1}^{n} \boldsymbol{x}_j \boldsymbol{\alpha}_j, \boldsymbol{\alpha}_i \right\rangle = \boldsymbol{x}_i.$$

- 这就是说，向量 $\boldsymbol{\xi}$ 关于一个规范正交基的第 $\boldsymbol{i}$ 个坐标等于 $\boldsymbol{\xi}$ 与第 $\boldsymbol{i}$ 个基向量的内积。
- 令

  $$\boldsymbol{\eta} = \boldsymbol{y}_1\boldsymbol{\alpha}_1 + \boldsymbol{y}_2\boldsymbol{\alpha}_2 + \cdots + \boldsymbol{y}_n\boldsymbol{\alpha}_n,$$

  那么

  $$\langle \boldsymbol{\xi}, \boldsymbol{\eta} \rangle = \boldsymbol{x}_1\boldsymbol{y}_1 + \boldsymbol{x}_2\boldsymbol{y}_2 + \cdots + \boldsymbol{x}_n\boldsymbol{y}_n.$$

- 由此得
$$|\boldsymbol{\xi}| = \sqrt{\langle\boldsymbol{\xi},\boldsymbol{\xi}\rangle} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

-
$$d(\boldsymbol{\xi},\boldsymbol{\eta}) = |\boldsymbol{\xi}-\boldsymbol{\eta}| = \sqrt{(x_1-y_1)^2 + \cdots + (x_n-y_n)^2}.$$

定理 3.3

设 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\}$ 是欧式空间 $\boldsymbol{V}$ 的一组线性无关的向量，那么可以求出 $\boldsymbol{V}$ 的一个正交组 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m\}$，使得 $\boldsymbol{\beta}_k$ 可以由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_k$ 线性表示，$k=1,2,\cdots,m$. ♡

注这个定理实际上给出了一个方法，使得我们可以从欧氏空间的任意一组线性无关的向量出发，得出一个正交组来. 这个方法称为施密特 (*Schimidt*) 正交化方法，简称正交化方法.

现在设 $\boldsymbol{V}$ 是一个 $n(n>0)$ 维欧氏空间，令 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$ 是 $\boldsymbol{V}$ 的任意一个基. 利用正交化方法，可以得出 V 的一个正交基 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n\}$, 再令

$$\boldsymbol{\gamma}_i = \boldsymbol{\beta}_i / \left|\boldsymbol{\beta}_i\right|, \boldsymbol{i} = 1,2,\cdots,\boldsymbol{n}.$$

那么 $\{\gamma_1,\gamma_2,\cdots,\gamma_n\}$ 就是 $\boldsymbol{V}$ 的一个规范正交基。

定理 3.4 (规范正交基一定存在)

任意 $n(n>0)$ 维欧氏空间一定有正交基，因而有规范正交基. ♡

例题 3.10 在欧氏空间 $\mathbf{R}^3$ 中，对于基
$$\boldsymbol{\alpha}_1 = (1,1,1), \boldsymbol{\alpha}_2 = (0,1,2), \boldsymbol{\alpha}_3 = (2,0,3)$$
施行正交化方法，得出 $\mathbf{R}^3$ 的一个规范正交基。

推论 3.1

n 维 Euclid 空间 V 中任意一组两两正交的单位向量组
$$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\}$$
都可以扩成 V 的一组标准正交基
$$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m,\boldsymbol{\alpha}_{m+1},\cdots,\boldsymbol{\alpha}_n\}.$$ ♡

定义 3.10 (集合正交)

1. 设 $\boldsymbol{W}$ 是欧氏空间 $\boldsymbol{V}$ 的一个非空子集. 如果: $\boldsymbol{V}$ 的一个向量 $\boldsymbol{\xi}$ 与 $\boldsymbol{W}$ 的每一向量正交，那么就说 $\boldsymbol{\xi}$ 与 $\boldsymbol{W}$ 正交，并且记作 $\langle\boldsymbol{\xi},\boldsymbol{W}\rangle = 0$.
2. 令 $W^{\perp} = \{\boldsymbol{\xi}\in V \mid \langle\boldsymbol{\xi},W\rangle = 0\}$. 那么 $\mathbf{0}\in W^{\perp}$, 因而 $W^{\perp}\neq\varnothing$.
3. 其次，设 $\boldsymbol{a},\boldsymbol{b}\in\mathbb{R},\boldsymbol{\xi},\boldsymbol{\eta}\in\boldsymbol{W}^{\perp}$. 那么对于任意 $\boldsymbol{\alpha}\in W$, 我们有
$$\langle\boldsymbol{a}\boldsymbol{\xi}+\boldsymbol{b}\boldsymbol{\eta},\boldsymbol{\alpha}\rangle = \boldsymbol{a}\langle\boldsymbol{\xi},\boldsymbol{\alpha}\rangle + \boldsymbol{b}\langle\boldsymbol{\eta},\boldsymbol{\alpha}\rangle = \mathbf{0},$$
因而 $a\xi+b\eta\in W^{\perp}$. 这样，$W^{\perp}$ 是 $\boldsymbol{V}$ 的一个子空间。 ♣

定理 3.5 (正交分解)

令 $\boldsymbol{W}$ 是欧氏空间 $\boldsymbol{V}$ 的一个有限维子空间。那么 $V = W \oplus W^{\perp}$。因而 $\boldsymbol{V}$ 的每一向量 $\boldsymbol{\xi}$ 可以唯一地写成

$$\boldsymbol{\xi} = \boldsymbol{\eta} + \boldsymbol{\zeta},$$

这里 $\boldsymbol{\eta} \in \boldsymbol{W}, \langle \boldsymbol{\zeta}, \boldsymbol{W} \rangle = \mathbf{0}$. 我们称 $\boldsymbol{\eta}$ 为 $\boldsymbol{\xi}$ 在 W 的正交投影。

定理 3.6 (正交投影距离最近)

设 $\boldsymbol{W}$ 是欧氏空间 $\boldsymbol{V}$ 的一个有限维子空间，$\boldsymbol{\xi}$ 是 $\boldsymbol{V}$ 的任意向量，$\boldsymbol{\eta}$ 是 $\boldsymbol{\xi}$ 在 $\boldsymbol{W}$ 上的的正交射影。那么对于中 $\boldsymbol{W}$ 任意向量 $\boldsymbol{\eta}' \neq \boldsymbol{\eta}$, 都有 $|\boldsymbol{\xi} - \boldsymbol{\eta}| < |\boldsymbol{\xi} - \boldsymbol{\eta}'|$.

定义 3.11 (正交矩阵)

一个 $\boldsymbol{n}$ 阶实矩阵 $\boldsymbol{U}$ 叫作一个正交矩阵，如果

$$\boldsymbol{U}\boldsymbol{U}^{\mathrm{T}} = \boldsymbol{U}^{\mathrm{T}}\boldsymbol{U} = \boldsymbol{I}_n.$$

定理 3.7 (规范正交基过渡矩阵是一个正交矩阵)

$\boldsymbol{n}$ 维欧氏空间一个规范正交基到另一规范正交基的过渡矩阵是一个正交矩阵.

定义 3.12 (欧式空间的同构)

欧氏空间 $\boldsymbol{V}$ 与 $\mathbf{V}'$ 说是同构的，如果

(i) 作为实数域上向量空间，存在 $\boldsymbol{V}$ 到 $\boldsymbol{V}'$ 的一个同构映射 $\boldsymbol{f} : \boldsymbol{V} \to \boldsymbol{V}'$

(ii) 对于任意 $\boldsymbol{\xi}, \boldsymbol{\eta} \in \boldsymbol{V}$, 都有 $\langle \boldsymbol{\xi}, \boldsymbol{\eta} \rangle = \langle \boldsymbol{f}(\boldsymbol{\xi}), \boldsymbol{f}(\boldsymbol{\eta}) \rangle$.

定理 3.8

两个有限维欧氏空间同构的充要条件是它们的维数相等.

推论 3.2

任意 $\boldsymbol{n}$ 维欧氏空间都与 $\mathbf{R}^n$ 同构.

例题 3.11 任意一个 $m \times n$ 实矩阵 $\boldsymbol{A}$ 都可以表为一个正交方阵 O 和一个对角元非负的上三角方阵 T 的乘积，即 $A = OT$. 而且，当 $\boldsymbol{A}$ 为可逆实方阵时，这种表法唯一。

练习 3.4

$$\alpha_3 = (1, 2, 0, -1), \quad \alpha_4 = (1, 0, 0, 1), \quad \alpha_1 = (0, 2, 1, 0), \quad \alpha_2 = (1, -1, 0, 0),$$

是 $\mathbf{R}^4$ 的一个基. 对这个基施行正交化方法，求出 $\mathbf{R}^4$ 的一个规范正交基.

解

$$
\begin{aligned}
\beta_1 &= \alpha_1 = (0,2,1,0)\,; \\
\beta_2 &= \alpha_2 - \frac{\langle\alpha_2,\beta_1\rangle}{\langle\beta_1,\beta_1\rangle}\beta_1 = (1,-1,0,0) - \frac{-2}{5}(0,2,1,0) = \left(1,-\frac{1}{5},\frac{2}{5},0\right); \\
\beta_3 &= \alpha_3 - \frac{\langle\alpha_3,\beta_1\rangle}{\langle\beta_1,\beta_1\rangle}\beta_1 - \frac{\langle\alpha_3,\beta_2\rangle}{\langle\beta_2,\beta_2\rangle}\beta_2 \\
&= (1,2,0,-1) - \frac{4}{5}(0,2,1,0) - \frac{1}{2}\left(1,-\frac{1}{5},\frac{2}{5},0\right) = \left(\frac{1}{2},\frac{1}{2},-1,-1\right); \\
\beta_4 &= \alpha_4 - \frac{\langle\alpha_4,\beta_1\rangle}{\langle\beta_1,\beta_1\rangle}\beta_1 - \frac{\langle\alpha_4,\beta_2\rangle}{\langle\beta_2,\beta_2\rangle}\beta_2 - \frac{\langle\alpha_4,\beta_3\rangle}{\langle\beta_3,\beta_3\rangle}\beta_3 \\
&= (1,0,0,1) - \frac{5}{6}\left(1,-\frac{1}{5},\frac{2}{5},0\right) + \frac{1}{5}\left(\frac{1}{2},\frac{1}{2},-1,-1\right) \\
&= \left(\frac{4}{15},\frac{4}{15},-\frac{8}{15},\frac{4}{5}\right).
\end{aligned}
$$

再正规化, 得:

$$
\gamma_1 = \frac{1}{\sqrt{5}}(0,2,1,0)\,,\quad \gamma_2 = \frac{1}{\sqrt{30}}(5,-1,2,0)\,,
$$
$$
\gamma_3 = \frac{1}{\sqrt{10}}(1,1,-2,-2)\,,\quad \gamma_4 = \frac{1}{\sqrt{15}}(1,1,-2,3)\,.
$$

✍ **练习 3.5** 设 $\boldsymbol{\xi},\boldsymbol{\eta}$ 是一个欧氏空间里彼此正交的向量。证明：$|\boldsymbol{\xi}+\boldsymbol{\eta}|^2 = |\boldsymbol{\xi}|^2 + |\boldsymbol{\eta}|^2$.

解 $|\boldsymbol{\xi}+\boldsymbol{\eta}|^2 = \langle\boldsymbol{\xi}+\boldsymbol{\eta},\boldsymbol{\xi}+\boldsymbol{\eta}\rangle = \langle\boldsymbol{\xi},\boldsymbol{\xi}\rangle + \langle\boldsymbol{\xi},\boldsymbol{\eta}\rangle + \langle\boldsymbol{\eta},\boldsymbol{\xi}\rangle + \langle\boldsymbol{\eta},\boldsymbol{\eta}\rangle = |\boldsymbol{\xi}|^2 + |\boldsymbol{\eta}|^2$.

✍ **练习 3.6** 设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n,\boldsymbol{\beta}$ 都是一个欧氏空间的向量，且 $\boldsymbol{\beta}$ 是 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 的线性组合、证明如果 $\boldsymbol{\beta}$ 与每一个 $\boldsymbol{\alpha}_i$ 正交，$i=1,2,\cdots,n$, 那么 $\boldsymbol{\beta}=\mathbf{0}$.

解 取 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n$ 的一个极大无关组, 不妨记为 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_s$, 因此 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_s$ 无关且非零向量. 设 $\boldsymbol{\beta} = k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + ... + k_n\boldsymbol{\alpha}_s$. 由题意, 有:

$$
\langle\boldsymbol{\alpha}_i, k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + ... + k_s\boldsymbol{\alpha}_s\rangle = 0, i = 1,2,...,s,
$$

即

$$
k_i\langle\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_i\rangle = 0, i = 1,2,...,s,
$$

所以 $k_i = 0, i = 1,2,...,s$, 因此 $\boldsymbol{\beta} = 0$.

✍ **练习 3.7** 设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 是欧氏空间的 $\boldsymbol{n}$ 个向量. 行列式

$$
\boldsymbol{G}(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n) = \begin{vmatrix}
\langle\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_1\rangle & \langle\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2\rangle & \cdots & \langle\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_n\rangle \\
\langle\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_1\rangle & \langle\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_2\rangle & \cdots & \langle\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_n\rangle \\
\vdots & \vdots & & \vdots \\
\langle\boldsymbol{\alpha}_n,\boldsymbol{\alpha}_1\rangle & \langle\boldsymbol{\alpha}_n,\boldsymbol{\alpha}_2\rangle & \cdots & \langle\boldsymbol{\alpha}_n,\boldsymbol{\alpha}_n\rangle
\end{vmatrix}
$$

叫作 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n$ 的格拉姆 (Gram) 行列式. 证明 $G(\alpha_1,\cdots,\alpha_n) = 0$ 当且仅当 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n$ 线性相关.

解

$$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n\text{相关} \Leftrightarrow \text{存在不全为零的}k_1,...,k_n \text{ s.t. } k_1\boldsymbol{\alpha}_1+\cdots+k_n\boldsymbol{\alpha}_n=0$$

$$\Leftrightarrow \text{存在不全为零的}k_1,...,k_n \text{ s.t. } \langle\boldsymbol{\alpha}_i,k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+...+k_n\boldsymbol{\alpha}_n\rangle=0, i=1,2,...,n$$

$$\Leftrightarrow \begin{cases} k_1\langle\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_1\rangle+k_2\langle\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2\rangle+...k_n\langle\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_n\rangle=0\\ k_1\langle\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_1\rangle+k_2\langle\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_2\rangle+...k_n\langle\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_n\rangle=0\\ ...\\ k_1\langle\boldsymbol{\alpha}_n,\boldsymbol{\alpha}_1\rangle+k_2\langle\boldsymbol{\alpha}_n,\boldsymbol{\alpha}_2\rangle+...k_n\langle\boldsymbol{\alpha}_n,\boldsymbol{\alpha}_n\rangle=0\end{cases}\text{有非零解}$$

$$\Leftrightarrow \begin{bmatrix}\langle\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_1\rangle & \langle\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2\rangle & \cdots & \langle\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_n\rangle\\ \langle\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_1\rangle & \langle\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_2\rangle & \cdots & \langle\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_n\rangle\\ \vdots & \vdots & & \vdots\\ \langle\boldsymbol{\alpha}_n,\boldsymbol{\alpha}_1\rangle & \langle\boldsymbol{\alpha}_n,\boldsymbol{\alpha}_2\rangle & \cdots & \langle\boldsymbol{\alpha}_n,\boldsymbol{\alpha}_n\rangle\end{bmatrix}\begin{bmatrix}k_1\\k_2\\...\\k_n\end{bmatrix}=0\text{有非零解}$$

$$\Leftrightarrow \boldsymbol{G}(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n)=0.$$

练习 3.8 设 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 是欧氏空间两个线性无关的向量，满足以下条件：$\frac{2\langle\alpha,\beta\rangle}{\langle\alpha,\alpha\rangle}$ 和 $\frac{2\langle\alpha,\beta\rangle}{\langle\beta,\beta\rangle}$ 都是 $\leqslant \mathbf{0}$ 的整数. 证明：$\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 的夹角只可能是 $\frac{\pi}{2},\frac{2\pi}{3},\frac{3\pi}{4}$ 或 $\frac{5\pi}{6}$.

解 记 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 的夹角为 θ, 所以

$$\frac{2\langle\boldsymbol{\alpha},\boldsymbol{\beta}\rangle}{\langle\boldsymbol{\alpha},\boldsymbol{\alpha}\rangle}=\frac{2|\boldsymbol{\beta}|\cos\theta}{|\boldsymbol{\alpha}|},$$

$$\frac{2\langle\boldsymbol{\alpha},\boldsymbol{\beta}\rangle}{\langle\boldsymbol{\beta},\boldsymbol{\beta}\rangle}=\frac{2|\boldsymbol{\alpha}|\cos\theta}{|\boldsymbol{\beta}|},$$

记

$$n_1=\frac{2|\boldsymbol{\beta}|\cos\theta}{|\boldsymbol{\alpha}|},\ n_2=\frac{2|\boldsymbol{\alpha}|\cos\theta}{|\boldsymbol{\beta}|},$$

所以 n_1,n_2 为 ≤ 0 的整数. 首先可得 $cos\theta\le 0$, 其次

$$\cos\theta=\frac{n_1|\boldsymbol{\alpha}|}{2|\boldsymbol{\beta}|}=\frac{n_2|\boldsymbol{\beta}|}{2|\boldsymbol{\alpha}|}\in[-1,0].$$

不妨设 $|\boldsymbol{\alpha}|\ge|\boldsymbol{\beta}|$, 所以 $\frac{|\alpha|}{|\beta|}\in[1,+\infty)\Rightarrow n_1\in[-2,0]\Rightarrow n_1=-2,-1$ 或 0.

若 $n_1=-2$, 则 $\frac{|\alpha|}{|\beta|}\in[0,1]$, 又 $|\boldsymbol{\alpha}|\ge|\boldsymbol{\beta}|$，所以 $|\boldsymbol{\alpha}|=|\boldsymbol{\beta}|$，此时 $\cos\theta=-1$，$\theta=\pi$，与 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 无关矛盾;

若 $n_1=0$, 此时 $\cos\theta=0$, $\theta=\pi/2$;

若 $n_1=-2$, 则 $\frac{|\alpha|}{|\beta|}\in[0,1]\Rightarrow n_2\in[-4,0]\Rightarrow n_2=-4,-3,-2,-1$.

$n_2=-4$ 时, $\frac{-1|\alpha|}{2|\beta|}=\frac{-4|\beta|}{2|\alpha|}\Rightarrow|\boldsymbol{\alpha}|=2|\boldsymbol{\beta}|$, 此时 $\cos\theta=-1$, $\theta=\pi$, 矛盾;

$n_2=-3$ 时, $\frac{-1|\alpha|}{2|\beta|}=\frac{-3|\beta|}{2|\alpha|}\Rightarrow|\boldsymbol{\alpha}|=\sqrt{3}|\boldsymbol{\beta}|$, 此时 $\cos\theta=-\sqrt{3}/2$, $\theta=5\pi/6$;

$n_2=-2$ 时, $\frac{-1|\alpha|}{2|\beta|}=\frac{-2|\beta|}{2|\alpha|}\Rightarrow|\boldsymbol{\alpha}|=\sqrt{2}|\boldsymbol{\beta}|$, 此时 $\cos\theta=-\sqrt{2}/2$, $\theta=3\pi/4$;

$n_2=-1$ 时, $\frac{-1|\alpha|}{2|\beta|}=\frac{-1|\beta|}{2|\alpha|}\Rightarrow|\boldsymbol{\alpha}|=|\boldsymbol{\beta}|$, 此时 $\cos\theta=-1/2$, $\theta=2\pi/3$.

综上, $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 的夹角只可能是 $\frac{\pi}{2},\frac{2\pi}{3},\frac{3\pi}{4}$ 或 $\frac{5\pi}{6}$.

练习 3.9 证明：对于任意实数 $a_1, a_2, \cdots, a_n$,

$$\sum_{i=1}^{n}|a_i| \leqslant \sqrt{n\left(a_1^2+a_2^2+\cdots+a_n^2\right)}.$$

解 在 $\mathbb{R}^n$ 中，令 $\xi=(|a_1|,|a_2|,...,|a_n|), \eta=(1,1,...,1)$, 由 Cauchy-Schwarz 不等式, 有

$$\langle\xi,\eta\rangle^2 \leq \langle\xi,\xi\rangle\langle\eta,\eta\rangle,$$

即

$$\left(\sum_{i=1}^{n}|a_i|\right)^2 \leqslant \left(a_1^2+a_2^2+\cdots+a_n^2\right)(1+1+...+1),$$

因此

$$\sum_{i=1}^{n}|a_i| \leqslant \sqrt{n\left(a_1^2+a_2^2+\cdots+a_n^2\right)}.$$

练习 3.10 令 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$ 是欧氏空间 $\boldsymbol{V}$ 的一组线性无关的向量，$\{\boldsymbol{\beta}_1,\ \boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n\}$ 是由这组向量通过正交化方法所得的正交组. 证明，这两个向量组的格拉姆行列式 [习题 3.7] 相等，即 $\boldsymbol{G}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)=\boldsymbol{G}(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n)$.

解 由题意,

$$\begin{cases}\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1,\\ \boldsymbol{\beta}_2=\boldsymbol{\alpha}_2-\frac{\langle\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1\rangle}{\langle\boldsymbol{\beta}_1,\boldsymbol{\beta}_1\rangle}\boldsymbol{\beta}_1,\\ \dots\\ \boldsymbol{\beta}_n=\boldsymbol{\alpha}_n-\frac{\langle\boldsymbol{\alpha}_n,\boldsymbol{\beta}_1\rangle}{\langle\boldsymbol{\beta}_1,\boldsymbol{\beta}_1\rangle}\boldsymbol{\beta}_1-\dots-\frac{\langle\boldsymbol{\alpha}_n,\boldsymbol{\beta}_{n-1}\rangle}{\langle\boldsymbol{\beta}_{n-1},\boldsymbol{\beta}_{n-1}\rangle}\boldsymbol{\beta}_1,\end{cases}$$

所以

$$\begin{cases}\boldsymbol{\alpha}_1=\boldsymbol{\beta}_1,\\ \boldsymbol{\alpha}_2=\boldsymbol{\beta}_2-\frac{\langle\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1\rangle}{\langle\boldsymbol{\beta}_1,\boldsymbol{\beta}_1\rangle}\boldsymbol{\beta}_1,\\ \dots\\ \boldsymbol{\alpha}_n=\boldsymbol{\beta}_n-\frac{\langle\boldsymbol{\alpha}_n,\boldsymbol{\beta}_1\rangle}{\langle\boldsymbol{\beta}_1,\boldsymbol{\beta}_1\rangle}\boldsymbol{\beta}_1-\dots-\frac{\langle\boldsymbol{\alpha}_n,\boldsymbol{\beta}_{n-1}\rangle}{\langle\boldsymbol{\beta}_{n-1},\boldsymbol{\beta}_{n-1}\rangle}\boldsymbol{\beta}_1,\end{cases}$$

即

$$(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,...,\boldsymbol{\alpha}_n)=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,...,\boldsymbol{\beta}_n)\begin{pmatrix}1 & -\frac{\langle\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1\rangle}{\langle\boldsymbol{\beta}_1,\boldsymbol{\beta}_1\rangle} & \dots & -\frac{\langle\boldsymbol{\alpha}_n,\boldsymbol{\beta}_1\rangle}{\langle\boldsymbol{\beta}_1,\boldsymbol{\beta}_1\rangle}\\ 0 & 1 & \dots & -\frac{\langle\boldsymbol{\alpha}_n,\boldsymbol{\beta}_2\rangle}{\langle\boldsymbol{\beta}_2,\boldsymbol{\beta}_2\rangle}\\ \dots & \dots & & \dots\\ 0 & 0 & \dots & 1\end{pmatrix}$$

记

$$C=\begin{pmatrix}1 & -\frac{\langle\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1\rangle}{\langle\boldsymbol{\beta}_1,\boldsymbol{\beta}_1\rangle} & \dots & -\frac{\langle\boldsymbol{\alpha}_n,\boldsymbol{\beta}_1\rangle}{\langle\boldsymbol{\beta}_1,\boldsymbol{\beta}_1\rangle}\\ 0 & 1 & \dots & -\frac{\langle\boldsymbol{\alpha}_n,\boldsymbol{\beta}_2\rangle}{\langle\boldsymbol{\beta}_2,\boldsymbol{\beta}_2\rangle}\\ \dots & \dots & & \dots\\ 0 & 0 & \dots & 1\end{pmatrix},$$

有 $(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,...,\boldsymbol{\alpha}_n)=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,...,\boldsymbol{\beta}_n)C$, 且 $\det(C)=1$.

因此

$$\begin{bmatrix} \langle \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_1 \rangle & \langle \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2 \rangle & \cdots & \langle \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_n \rangle \\ \langle \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_1 \rangle & \langle \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_2 \rangle & \cdots & \langle \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_n \rangle \\ \vdots & \vdots & & \vdots \\ \langle \boldsymbol{\alpha}_n, \boldsymbol{\alpha}_1 \rangle & \langle \boldsymbol{\alpha}_n, \boldsymbol{\alpha}_2 \rangle & \cdots & \langle \boldsymbol{\alpha}_n, \boldsymbol{\alpha}_n \rangle \end{bmatrix} = \begin{bmatrix} \boldsymbol{\alpha}_1^T \boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_1^T \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_1^T \boldsymbol{\alpha}_n \\ \boldsymbol{\alpha}_2^T \boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2^T \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_2^T \boldsymbol{\alpha}_n \\ \vdots & \vdots & & \vdots \\ \boldsymbol{\alpha}_n^T \boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_n^T \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_n^T \boldsymbol{\alpha}_n \end{bmatrix}$$

$$= (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, ..., \boldsymbol{\alpha}_n)^T (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, ..., \boldsymbol{\alpha}_n)$$

$$= C^T (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, ..., \boldsymbol{\beta}_n)^T (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, ..., \boldsymbol{\beta}_n) C.$$

所以

$$\begin{aligned} G(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, ..., \boldsymbol{\alpha}_n) &= \det \begin{bmatrix} \langle \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_1 \rangle & \langle \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2 \rangle & \cdots & \langle \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_n \rangle \\ \langle \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_1 \rangle & \langle \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_2 \rangle & \cdots & \langle \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_n \rangle \\ \vdots & \vdots & & \vdots \\ \langle \boldsymbol{\alpha}_n, \boldsymbol{\alpha}_1 \rangle & \langle \boldsymbol{\alpha}_n, \boldsymbol{\alpha}_2 \rangle & \cdots & \langle \boldsymbol{\alpha}_n, \boldsymbol{\alpha}_n \rangle \end{bmatrix} \\ &= \det \left(C^T (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, ..., \boldsymbol{\beta}_n)^T (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, ..., \boldsymbol{\beta}_n) C \right) \\ &= \det \left((\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, ..., \boldsymbol{\beta}_n)^T (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, ..., \boldsymbol{\beta}_n) \right) \\ &= G(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, ..., \boldsymbol{\beta}_n). \end{aligned}$$

练习 3.11 证明实系数线性方程组

$$\sum_{j=1}^{n} a_{ij} x_j = b_i, i = 1, 2, \cdots, n,$$

有解的充要条件是向量 $\boldsymbol{\beta} = (\boldsymbol{b}_1, \boldsymbol{b}_2, \cdots, \boldsymbol{b}_n) \in \mathbf{R}^n$ 与齐次线性方程组

$$\sum_{j=1}^{n} a_{ji} x_j = 0, i = 1, 2, \cdots, n,$$

的解空间正交.

解 记 $A = (a_{ij}), x = (x_1, ..., x_n)^T, M = \{x | Ax = \boldsymbol{\beta}\}, N = \{x | A^T x = 0\}$.

($\Rightarrow$) 设 $x \in M$ 为一个解, 任取 $y \in N$, 则 $Ax = \boldsymbol{\beta}$ 且 $A^T y = 0$, 因此 $\langle \beta, y \rangle = \beta^T y = x^T A^T y = 0$.

($\Leftarrow$) 设 $\boldsymbol{\beta}$ 与 N 正交, 任取 $y \in N$ 即 $A^T y = 0$, 则有 $\boldsymbol{\beta}^T y = 0$, 因此 $A^T x = 0$ 与 $\begin{pmatrix} A^T \\ \boldsymbol{\beta}^T \end{pmatrix} x = 0$ 同解, 所以 $\operatorname{rank}(A^T) = \operatorname{rank} \begin{pmatrix} A^T \\ \boldsymbol{\beta}^T \end{pmatrix}$, 故 $\operatorname{rank}(A) = \operatorname{rank}(A, \boldsymbol{\beta})$, 方程组 $Ax = \boldsymbol{\beta}$ 有解.

3.3 2024 年 5 月 21 日第十二周星期二第二十二次课-正交变换

定义 3.13 (正交变换)

欧氏空间 V 的一个线性变换 $\boldsymbol{\sigma}$ 叫作一个正交变换，如果对于任意 $\xi \in V$ 都有 $|\boldsymbol{\sigma}(\xi)| = |\xi|$。

定理 3.9

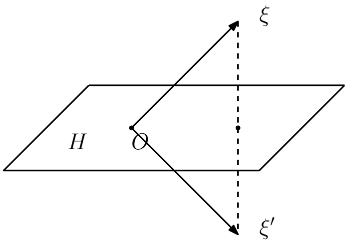
欧氏空间 $\boldsymbol{V}$ 的一个线性变换 $\boldsymbol{\sigma}$ 是正交变换的充要条件是：对于 $\boldsymbol{V}$ 中任意向量 $\boldsymbol{\xi}, \boldsymbol{\eta}$,

$$\langle \boldsymbol{\sigma}(\boldsymbol{\xi}), \boldsymbol{\sigma}(\boldsymbol{\eta}) \rangle = \langle \boldsymbol{\xi}, \boldsymbol{\eta} \rangle.$$

例题 3.12

(1) 在 $\mathbb{R}^2$ 里, 把每一向量旋转一个角 φ 的线性变换是 $\mathbb{R}^2$ 的一个正交变换.

(2) 令 H 是空间 $\mathbb{R}^3$ 里过原点的一个平面. 对于每一向量 $\boldsymbol{\xi} \in \mathbb{R}^3$, 令 ξ 对于 H 的镜面反射 ξ' 与它对应。则 $\sigma : \boldsymbol{\xi} \to \boldsymbol{\xi}'$ 是 $\mathbb{R}^3$ 的一个正交变换.



推论 3.3 (正交变换保持夹角不变)

如果 $\boldsymbol{\sigma}$ 是一个正交变换，$\boldsymbol{\theta}$ 是向量 $\boldsymbol{\xi}$ 与 $\boldsymbol{\eta}$ 的夹角，$\boldsymbol{\theta}'$ 是向量 $\boldsymbol{\sigma}(\boldsymbol{\xi})$ 与 $\boldsymbol{\sigma}(\boldsymbol{\eta})$ 的夹角，$0 \leqslant \boldsymbol{\theta} \leqslant \pi, 0 \leqslant \boldsymbol{\theta}' \leqslant \pi$. 那么

$$\boldsymbol{\theta} = \arccos \frac{\langle \boldsymbol{\xi}, \boldsymbol{\eta} \rangle}{|\boldsymbol{\xi}||\boldsymbol{\eta}|} = \arccos \frac{\langle \sigma(\boldsymbol{\xi}), \sigma(\boldsymbol{\eta}) \rangle}{|\sigma(\boldsymbol{\xi})||\sigma(\boldsymbol{\eta})|} = \boldsymbol{\theta}'$$

定理 3.10

设 $\boldsymbol{V}$ 是一个 $\boldsymbol{n}$ 维欧氏空间，$\boldsymbol{\sigma}$ 是 $\boldsymbol{V}$ 的一个线性变换。

(1) 如果 $\boldsymbol{\sigma}$ 是正交变换，那么 $\boldsymbol{\sigma}$ 把 $\boldsymbol{V}$ 的任意一个规范正交基仍旧变成 $\boldsymbol{V}$ 的一个规范正交基。

(2) 反过来，如果 $\boldsymbol{\sigma}$ 把 $\boldsymbol{V}$ 的某一规范正交基仍旧变成 $\boldsymbol{V}$ 的一个规范正交基，那么 $\boldsymbol{\sigma}$ 是 $\boldsymbol{V}$ 的一个正交变换.

定理 3.11

$\boldsymbol{n}$ 维欧氏空间 $\boldsymbol{V}$ 的一个正交变换 $\boldsymbol{\sigma}$ 关于 $\boldsymbol{V}$ 的任意规范正交基的矩阵是一个正交矩阵。反过来，如果 $\boldsymbol{V}$ 的一个线性变换关于某一规范正交基的矩阵是正交矩阵，那么 $\boldsymbol{\sigma}$ 是一个正交变换.

定理 3.12 (正交变换的几个等价条件)

设 $\mathscr{A}$ 是 n 维 Euclid 空间的线性变换. 则下述命题等价:

(1) $\mathscr{A}$ 是正交变换;

(2) $\mathscr{A}$ 是保内积的, 即对任意 $\boldsymbol{\alpha},\boldsymbol{\beta}\in V$, $(\mathscr{A}(\boldsymbol{\alpha}),\mathscr{A}(\boldsymbol{\beta}))=(\boldsymbol{\alpha},\boldsymbol{\beta})$;

(3) $\mathscr{A}$ 把 V 的标准正交基变为标准正交基, 即设 $\{\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\cdots,\boldsymbol{\xi}_n\}$ 是 V 的标准正交基, 则 $\{\mathscr{A}(\boldsymbol{\xi}_1),\mathscr{A}(\boldsymbol{\xi}_2),\cdots,\mathscr{A}(\boldsymbol{\xi}_n)\}$ 也是 V 的标准正交基;

(4) $\mathscr{A}$ 在 V 的标准正交基下的方阵是正交方阵;

♡

定理 3.13

(1) 正交方阵的行列式等于 ±1;

(2) 正交方阵的复特征值 λ_i 的模 $|\lambda_i|=1$, 实特征值 $\lambda_i=\pm1$.

♡

- 设 A 是正交方阵. 对 $A^{\mathrm{T}}A=I_n$ 两边取行列式得 $\det A^{\mathrm{T}}\det A=1$ 即 $(\det\boldsymbol{A})^2=1$, 从而 $\det\boldsymbol{A}=\pm1$.
- 设 λ_i 是正交方阵 $\boldsymbol{A}$ 的任一特征值, $0\neq\boldsymbol{X}\in\mathbb{C}^{n\times1}$ 是 $\boldsymbol{A}$ 的属于特征值 λ_i 的特征向量, 则

$$\boldsymbol{AX}=\lambda_i\boldsymbol{X}\tag{3.6}$$

- 对等式 (3.6) 两边的矩阵同时取共轭转置, 得

$$\bar{X}^{\mathrm{T}}\overline{\boldsymbol{A}}^{\mathrm{T}}=\bar{\lambda}\overline{\boldsymbol{X}}^{T}\tag{3.7}$$

- 由于 A 是实方阵, (3.6) 中的 $\bar{A}^{\mathrm{T}}=A^{\mathrm{T}}$. 将 (3.7) 两边分别左乘 (3.6) 两边, 得

$$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{AX}=\bar{\lambda}_i\lambda_i\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\tag{3.8}$$

- 由于 $\boldsymbol{A}$ 是正交方阵, $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{I}$, 又 $\bar{\lambda}_i\lambda_i=|\lambda_i|^2$.
- 设 $\boldsymbol{X}=(x_1,\cdots,x_n)^{\mathrm{T}}$, 则 $x_1,\cdots,x_n$ 是不全为 0 的复数,

$$\overline{\boldsymbol{X}}^{T}\boldsymbol{X}=(\bar{x}_1,\cdots,\bar{x}_n)\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=\bar{x}_1x_1+\cdots+x_nx_n=|x_1|^2+\cdots+|x_n|^2\neq0$$

- 因此 (3.8) 即

$$\overline{\boldsymbol{X}}^{\mathrm{T}}\boldsymbol{X}=|\lambda_i|^2\overline{\boldsymbol{X}}^{\mathrm{T}}\boldsymbol{X}\Rightarrow|\lambda_i|^2=1\Rightarrow|\lambda_i|=1.$$

  且当 λ_i 是实数时, $|\lambda_i|=1\Leftrightarrow\lambda_i=\pm1$.

命题 3.4

设 σ 是欧氏空间 V 上的正交变换。如果 1 与 -1 都是 σ 的特征值, 则特征子空间 $V_1\perp V_{-1}$. ♠

3.3.1 2 维空间的正交变换

- 设 σ 是 $\mathbb{R}^2$ 的一个正交变换.

- σ 关于 $\mathbb{R}^2$ 的一个规范正交基 $\{\gamma_1, \gamma_2\}$ 的矩阵是
$$U = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$
那么 $\boldsymbol{U}$ 是一个正交矩阵.
- 于是
$$a^2 + c^2 = 1, b^2 + d^2 = 1, ab + cd = 0. \tag{3.9}$$
- 由第一个等式, 存在一个角 α 使
$$a = \cos\alpha, \quad c = \pm\sin\alpha.$$
- 由于 $\cos\alpha = \cos(\pm\alpha), \pm\sin\alpha = \sin(\pm\alpha)$, 因此可以令
$$a = \cos\varphi, \quad c = \sin\varphi$$
这里 $\varphi = \alpha$ 或 $-\alpha$.
- 同理，由 (3.9) 的第二个等式，存在一个角 ψ 使
$$b = \cos\psi, \quad d = \sin\psi$$
- 将 a, b, c, d 代人 (3.9) 的第三个等式得
$$\cos\varphi\cos\psi + \sin\varphi\sin\psi = 0,$$
或
$$\cos(\varphi - \psi) = 0.$$
- 最后等式表明, $\varphi - \psi$ 是 $\frac{\pi}{2}$ 的一个奇数倍.
- 由此得
$$\cos\psi = \mp\sin\varphi, \sin\psi = \pm\cos\varphi.$$
- 所以
$$U = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$$
或
$$U = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix}$$
- 在前一情形, σ 是将 $\mathbb{R}^2$ 的每一向量旋转角 φ 的旋转;
- 在后一情形, σ 将 $\mathbb{R}^2$ 中以 (x, y) 为坐标的向量变成以
$$(x\cos\varphi + y\sin\varphi, x\sin\varphi - y\cos\varphi)$$
为坐标的向量. 这时 σ 是关于直线 $y = \left(\tan\frac{\varphi}{2}\right)x$ 的反射. 这怎么看呢？首先
$$U = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}$$

- 注意到
$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$
对应的是关于 x 轴的反射。
- 先旋转 $-\varphi$ 角度后再关于 x 轴反射，等价于关于直线 $y=\left(\tan\frac{\varphi}{2}\right)x$ 的反射。
- 这样, $\mathbb{R}^2$ 的正交变换或者是一个旋转, 或者是关于一条过原点的直线的反射.
- 如果是后一情形，我们可以取这条直线上一个单位向量 $\boldsymbol{\gamma}_1'$ 和垂直于这条直线的一个单位向量 γ_2' 作为 $\mathbb{R}^2$ 的一个规范正交基, 而 $\boldsymbol{\sigma}$ 关于基 $\{\boldsymbol{\gamma}_1',\boldsymbol{\gamma}_2'\}$ 的矩阵有形状
$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

3.3.2 3 维空间的正交变换

- 设 σ 是 $\mathbb{R}^3$ 的一个正交变换.
- σ 的特征多项式是一个实系数三次多项式, 因而至少有一个实根 r.
- 令 γ_1 是 σ 的属于本征值 r 的一个本征向量, 并且取 γ_1 是一个单位向量. 再添加单位向量 γ_2,γ_3 使 $\{\gamma_1,\gamma_2,\gamma_3\}$ 是 $\mathbb{R}^3$ 的一个规范正交基.
- 那么 σ 关于这个基的矩阵有形状
$$U=\begin{pmatrix} r & s & t \\ 0 & a & b \\ 0 & c & d \end{pmatrix}$$
- 由于 U 是正交矩阵, 我们有 $r^2=1, rs=rt=0$, 从而 $r=\pm1, s=t=0$. 于是
$$\boldsymbol{U}=\begin{pmatrix} \pm1 & 0 & 0 \\ 0 & a & b \\ 0 & c & d \end{pmatrix}.$$
- 由 $\boldsymbol{U}$ 的正交性推出，矩阵
$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$
是一个二阶正交矩阵.
- 由上面的讨论, 存在一个角 φ 使
$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}=\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \text{或} \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix}.$$
- 在前一情形,
$$U=\begin{pmatrix} \pm1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix}$$

- 在后一情形, 根据对 $\mathbb{R}^2$ 的正交变换的讨论, 我们可以取 $\mathbb{R}^3$ 的一个规范正交基 $\{\gamma_1, \gamma_2', \gamma_3'\}$ 使 σ 关于这个基的矩阵是

$$\boldsymbol{T}=\begin{pmatrix}\pm1&0&0\\0&1&0\\0&0&-1\end{pmatrix}.$$

- 如果在 T 中左上角的元素是 1, 那么重新排列基向量, σ 关于基 $\{\gamma_3', \gamma_2', \gamma_1\}$ 的矩阵是

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

- 如果左上角的元素是 -1 , 那么 σ 关于基 $\{\gamma_2', \gamma_3', \gamma_2\}$ 的矩阵是

$$\begin{pmatrix}1&0&0\\0&-1&0\\0&0&-1\end{pmatrix}=\begin{pmatrix}1&0&0\\0&\cos\pi&-\sin\pi\\0&\sin\pi&\cos\pi\end{pmatrix}.$$

- 这样, $\mathbb{R}^3$ 的任意正交变换 σ 关于某一规范正交基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2,\ \boldsymbol{\alpha}_3\}$ 的矩阵是下列三种类型之一:

$$\begin{pmatrix}1&0&0\\0&\cos\varphi&-\sin\varphi\\0&\sin\varphi&\cos\varphi\end{pmatrix},\begin{pmatrix}-1&0&0\\0&1&0\\0&0&1\end{pmatrix}$$

- 或

$$\begin{pmatrix}-1&0&0\\0&\cos\varphi&-\sin\varphi\\0&\sin\varphi&\cos\varphi\end{pmatrix}=\begin{pmatrix}1&0&0\\0&\cos\varphi&-\sin\varphi\\0&\sin\varphi&\cos\varphi\end{pmatrix}\begin{pmatrix}-1&0&0\\0&1&0\\0&0&1\end{pmatrix}.$$

- 在第一种情形, σ 是绕通过 $\boldsymbol{\alpha}_1$ 的直线 $\mathscr{L}(\boldsymbol{\alpha}_1)$ 的一个旋转;
- 在第二种情形, σ 是对于平面 $\mathscr{L}(\alpha_2, \alpha_3)$ 的反射;
- 第三种情形, σ 是前两种变换的合成.

例题 3.13 设 $A=(a_{ij})$ 是三阶正交矩阵, 且 $\det A=1$, 求证:

(1) $\lambda=1$ 必为 A 的特征值.

(2) 存在正交阵 T, 使 $T^TAT=\begin{pmatrix}1&0&0\\0&\cos\theta&-\sin\theta\\0&\sin\theta&\cos\theta\end{pmatrix}$;

(3) $\theta=\arccos\frac{\operatorname{tr}A-1}{2}$.

✍ **练习 3.12** 如果你不了解复数，请用高中知识证明：$\mathbb{R}^2$ 上的点 (x,y) 绕原点旋转 φ 后的坐标是

$$(x\cos\varphi-y\sin\varphi, x\sin\varphi+y\cos\varphi).$$

验证一下这是个正交变换吧。

解 记这个变换为 σ.

$$
\begin{aligned}
\langle\sigma(x,y),\sigma(z,w)\rangle=&\langle(x\cos\varphi-y\sin\varphi,x\sin\varphi+y\cos\varphi),(z\cos\varphi-w\sin\varphi,z\sin\varphi+w\cos\varphi)\rangle\\
=&(x\cos\varphi-y\sin\varphi)(z\cos\varphi-w\sin\varphi)+(x\sin\varphi+y\cos\varphi)(z\sin\varphi+w\cos\varphi)\\
=&xz\left(\cos^2\varphi+\sin^2\varphi\right)+xw(-\cos\varphi\sin\varphi+\sin\varphi\cos\varphi)\\
&+yz(-\sin\varphi\cos\varphi+\cos\varphi\sin\varphi)+yw\left(\sin^2\varphi+\cos^2\varphi\right)\\
=&xz+yw\\
=&\langle(x,y),(z,w)\rangle.
\end{aligned}
$$

因此这是一个正交变换.

✍ **练习 3.13** 设 $\boldsymbol{V}$ 是一个 $\boldsymbol{n}$ 维欧氏空间，$\boldsymbol{\sigma}$ 是 $\boldsymbol{V}$ 的一个线性变换。证明：

(1) 如果 $\boldsymbol{\sigma}$ 是正交变换，那么 $\boldsymbol{\sigma}$ 把 $\boldsymbol{V}$ 的任意一个规范正交基仍旧变成 $\boldsymbol{V}$ 的一个规范正交基。

(2) 反过来，如果 $\boldsymbol{\sigma}$ 把 $\boldsymbol{V}$ 的某一规范正交基仍旧变成 $\boldsymbol{V}$ 的一个规范正交基，那么 $\boldsymbol{\sigma}$ 是 $\boldsymbol{V}$ 的一个正交变换.

解

(1) 因为正交变换保持内积和长度, 所以正交变换后的向量仍然正交且仍然是单位向量, 因此为 $\boldsymbol{V}$ 的一个规范正交基.

(2) 设 $\alpha_1,...,\alpha_n$ 和 $\sigma(\alpha_1),...,\sigma(\alpha_n)$ 为 $\boldsymbol{V}$ 的两个规范正交基, 因此有

$$
\langle\alpha_i,\alpha_j\rangle=\langle\sigma(\alpha_i),\sigma(\alpha_j)\rangle=0,i\neq j;\ \langle\alpha_i,\alpha_i\rangle=\langle\sigma(\alpha_i),\sigma(\alpha_i)\rangle=1.
$$

$\forall\xi,\eta\in\boldsymbol{V}$, $\exists k_1,...,k_n$ 和 $s_1,...,s_n$ s.t. $\xi=k_1\alpha_1+...+k_n\alpha_n$, $\eta=s_1\alpha_1+...+s_n\alpha_n$. 因此 $\sigma(\xi)=\sigma(k_1\alpha_1+...+k_n\alpha_n)=k_1\sigma(\alpha_1)+...+k_n\sigma(\alpha_n)$, 同理 $\sigma(\eta)=s_1\sigma(\alpha_1)+...+s_n\sigma(\alpha_n)$. 所以

$$
\begin{aligned}
\langle\sigma(\xi),\sigma(\eta)\rangle=&\langle k_1\sigma(\alpha_1)+...+k_n\sigma(\alpha_n),s_1\sigma(\alpha_1)+...+s_n\sigma(\alpha_n)\rangle\\
=&\langle k_1\sigma(\alpha_1),s_1\sigma(\alpha_1)\rangle+\langle k_2\sigma(\alpha_2),s_2\sigma(\alpha_2)\rangle+...+\langle k_n\sigma(\alpha_n),s_n\sigma(\alpha_n)\rangle\\
=&k_1s_1+k_2s_2+...+k_ns_n\\
=&\langle\xi,\eta\rangle.
\end{aligned}
$$

因此 $\boldsymbol{\sigma}$ 是 $\boldsymbol{V}$ 的一个正交变换.

✍ **练习 3.14** 证明：n 维欧氏空间的两个正交变换的乘积是一个正交变换；一个正交变换的逆变换还是一个正交变换.

解 首先因为正交变换把任意一个规范正交基变成一个规范正交基, 因此正交变换一定可逆. 设 σ 和 τ 为 n 维欧氏空间 V 的两个正交变换, 因为 $\forall\xi,\eta\in V$,

$$
\langle\sigma\tau(\xi),\sigma\tau(\eta)\rangle=\langle\sigma(\tau(\xi)),\sigma(\tau(\eta))\rangle=\langle\tau(\xi),\tau(\eta)\rangle=\langle\xi,\eta\rangle,
$$
$$
\langle\xi,\eta\rangle=\langle\sigma\sigma^{-1}(\xi),\sigma\sigma^{-1}(\eta)\rangle=\langle\sigma^{-1}(\xi),\sigma^{-1}(\eta)\rangle.
$$

所以 $\sigma\tau$ 和 σ^{-1} 为正交变换.

✍ **练习 3.15** 设 σ 是 n 维欧氏空间 V 的一个正交变换。证明：如果 V 的一个子空间 W 在 σ 之下

不变，那么 W 的正交补 $W^{\perp}$ 也在 σ 之下不变.

解 设 $\varepsilon_1,...,\varepsilon_s$ 为 W 的一个基, 因为 σ 为正交变换, 所以 $\sigma(\varepsilon_1),...,\sigma(\varepsilon_s)$ 线性无关, 又 $\sigma(W)\subset W$, 所以 $\sigma(W)=W$. $\forall\alpha\in W^{\perp}$, $\forall\beta\in W$, 因此 $\sigma^{-1}(\beta)\in W$, 所以

$$\langle\sigma(\alpha),\beta\rangle=\left\langle\sigma(\alpha),\sigma\sigma^{-1}(\beta)\right\rangle=\left\langle\alpha,\sigma^{-1}(\beta)\right\rangle=0.$$

因此 $\sigma(\alpha)\in W^{\perp}$,$W^{\perp}$ 在 σ 之下不变.

练习 3.16 设 V 是一个欧氏空间，$\alpha\in V$ 是一个非零向量. 对于 $\xi\in V$ ，规定

$$\tau(\xi)=\xi-\frac{2\langle\xi,\alpha\rangle}{\langle\alpha,\alpha\rangle}\alpha.$$

证明: τ 是 V 的一个正交变换, 且 $\tau^2=\iota$, ι 是单位变换.

解 $\forall\xi,\eta\in V$,

$$\begin{aligned}\langle\tau(\xi),\tau(\eta)\rangle&=\left\langle\xi-\frac{2\langle\xi,\alpha\rangle}{\langle\alpha,\alpha\rangle}\alpha,\eta-\frac{2\langle\eta,\alpha\rangle}{\langle\alpha,\alpha\rangle}\alpha\right\rangle\\&=\langle\xi,\eta\rangle-\left\langle\xi,\frac{2\langle\eta,\alpha\rangle}{\langle\alpha,\alpha\rangle}\alpha\right\rangle-\left\langle\frac{2\langle\xi,\alpha\rangle}{\langle\alpha,\alpha\rangle}\alpha,\eta\right\rangle+\left\langle\frac{2\langle\xi,\alpha\rangle}{\langle\alpha,\alpha\rangle}\alpha,\frac{2\langle\eta,\alpha\rangle}{\langle\alpha,\alpha\rangle}\alpha\right\rangle\\&=\langle\xi,\eta\rangle-\frac{2\langle\eta,\alpha\rangle\langle\xi,\alpha\rangle}{\langle\alpha,\alpha\rangle}-\frac{2\langle\xi,\alpha\rangle\langle\alpha,\eta\rangle}{\langle\alpha,\alpha\rangle}+\frac{4\langle\xi,\alpha\rangle\langle\eta,\alpha\rangle}{\langle\alpha,\alpha\rangle}\\&=\langle\xi,\eta\rangle,\end{aligned}$$

且

$$\begin{aligned}\tau(\tau(\xi))=&\tau(\xi)-\frac{2\langle\tau(\xi),\alpha\rangle}{\langle\alpha,\alpha\rangle}\alpha\\=&\xi-\frac{2\langle\xi,\alpha\rangle}{\langle\alpha,\alpha\rangle}\alpha-\frac{2\left\langle\xi-\frac{2\langle\xi,\alpha\rangle}{\langle\alpha,\alpha\rangle}\alpha,\alpha\right\rangle}{\langle\alpha,\alpha\rangle}\alpha\\=&\xi-\frac{2\langle\xi,\alpha\rangle}{\langle\alpha,\alpha\rangle}\alpha-\frac{2\langle\xi,\alpha\rangle-\frac{4\langle\xi,\alpha\rangle}{\langle\alpha,\alpha\rangle}\langle\alpha,\alpha\rangle}{\langle\alpha,\alpha\rangle}\alpha\\=&\xi-\frac{2\langle\xi,\alpha\rangle}{\langle\alpha,\alpha\rangle}\alpha+\frac{2\langle\xi,\alpha\rangle}{\langle\alpha,\alpha\rangle}\alpha\\=&\xi.\end{aligned}$$

因此 τ 是 V 的一个正交变换, 且 τ^2 是单位变换.

3.4 2024 年 5 月 23 日第十二周星期四第二十三次课-对称变换

设 V 是一个 n 维欧氏空间, σ 是 V 的一个线性变换. 要使 V 有一个正交基，而 σ 在这个基之下的矩阵是对角形式: 问 σ 应该满足什么条件? 这相当于说，σ 满足什么条件才能使得 V 有一个由 σ 的本征向量所组成的正交基?

如果 V 的一个线性变换 σ 具有以上的性质。

- 首先, σ 的特征多项式的根必须都是实数.

- 设 $c_1, c_2, \cdots, c_n \in \mathbb{R}$ 是 σ 的全部本征值 (重根按重数计算), α_i 是属于 c_i 的一个本征向量, 且 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 两两正交.
- 不失一般性，可设 $\|\alpha_i\| = 1, i = 1, 2, \cdots$, n.
- 设
$$\xi = \sum_{i=1}^{n} x_i \alpha_i, \quad \eta = \sum_{i=1}^{n} y_i \alpha_i$$
是 V 的任意向量.
- 因为 $\sigma(\alpha_i) = c_i \alpha_i, 1 \leqslant i \leqslant n$, 所以
$$\begin{aligned}\langle\sigma(\xi), \eta\rangle &= \left\langle\sum_{i=1}^{n} x_i \sigma(\alpha_i), \sum_{j=1}^{n} y_i \alpha_i\right\rangle = \left\langle\sum_{i=1}^{n} c_i x_i \alpha_i, \sum_{j=1}^{n} y_j \alpha_j\right\rangle \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} c_i x_i y_j \langle\alpha_i, \alpha_j\rangle = \sum_{i=1}^{n} c_i x_i y_i.\end{aligned}$$
- 同样的计算, 我们有
$$\langle\boldsymbol{\xi}, \sigma(\eta)\rangle = \sum_{i=1}^{n} c_i x_i y_i.$$
- 因此，对于任意 $\xi, \eta \in V$ ，等式
$$\langle\boldsymbol{\sigma}(\boldsymbol{\xi}), \boldsymbol{\eta}\rangle = \langle\boldsymbol{\xi}, \boldsymbol{\sigma}(\boldsymbol{\eta})\rangle$$
成立。

定义 3.14 (对称变换)

设 σ 是欧氏空间 V 的一个线性变换. 如果对于 V 中任意向量 $\boldsymbol{\xi}, \boldsymbol{\eta}$, 等式
$$\langle\boldsymbol{\sigma}(\boldsymbol{\xi}), \boldsymbol{\eta}\rangle = \langle\boldsymbol{\xi}, \boldsymbol{\sigma}(\boldsymbol{\eta})\rangle$$
成立, 那么就称 σ 是一个对称变换. ♣

定理 3.14

设 σ 是 n 维欧氏空间 V 的一个对称变换. $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 是 V 的任意一个规范正交基. $A = (a_{ij})$ 是 σ 关于这个基的矩阵. 那么 $\boldsymbol{A}^T = \boldsymbol{A}$. ♡

设 $\boldsymbol{A}$ 是某一数域 $\boldsymbol{F}$ 上的 n 阶矩阵. 如果 $\boldsymbol{A}$ 与它的转置 $\boldsymbol{A}^{\mathrm{T}}$ 相等, 那么就称 A 是一个对称矩阵. 定理 3.14 是说, n 维欧氏空间的对称变换关于任意规范正交基的矩阵是一个实对称矩阵.

定理 3.15

设 σ 是 n 维欧氏空间 V 的一个线性变换. 如果 σ 关于一个规范正交基的矩阵是对称矩阵，那么 σ 是一个对称变换. ♡

定理 3.16

(1) 实对称矩阵的特征根都是实数.

(2) n 维欧氏空间的一个对称变换的属于不同本征值的本征向量彼此正交.

练习 3.17 设 σ 是 n 维欧氏空间 V 的一个线性变换. 证明: 如果 σ 满足下列三个条件中的任意两个，那么它必然满足第三个：

(i) σ 是正交变换;

(ii) σ 是对称变换;

(iii) $\sigma^2 = \imath$ 是单位变换.

解 设 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m\}$ 为 V 的一个标准正交基, σ 关于其的矩阵为 $\boldsymbol{A}$, 则 σ 是正交变换 $\Leftrightarrow$ $\boldsymbol{A}\boldsymbol{A}^T = \boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{I}$; σ 是对称变换 $\Leftrightarrow$ $\boldsymbol{A} = \boldsymbol{A}^T$; σ^2 是单位变换 $\Leftrightarrow$ $\boldsymbol{A}^2 = \boldsymbol{I}$, 因此我们只需考虑矩阵的问题, 而这是显然的.

练习 3.18 设 σ 是 n 维欧氏空间 V 的一个对称变换, 且 $\sigma^2 = \sigma$. 证明存在 V 的一个规范正交基, 使得 σ 关于这个基的矩阵有形状

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

解 由定理 3.18, 存在 V 的一个标准正交基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m\}$, 使得 σ 关于其的矩阵 $\boldsymbol{A}$ 为对角矩阵, 且由题意得 $\boldsymbol{A} = \boldsymbol{A}^T$, $\boldsymbol{A}^2 = \boldsymbol{A}$.

我们只要证明 $\boldsymbol{A}$ 的特征值只可能为 0 或 1 即可. 为此, 设 λ 为 $\boldsymbol{A}$ 的特征值, $\boldsymbol{x}$ 为对应的特征向量, 即 $\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}, \boldsymbol{x} \neq 0$. 那么 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^2\boldsymbol{x} = \boldsymbol{A}\lambda\boldsymbol{x} = \lambda\boldsymbol{A}\boldsymbol{x} = \lambda^2\boldsymbol{x}$, 所以 $(\lambda - \lambda^2)\boldsymbol{x} = 0$, 又因为 $\boldsymbol{x} \neq 0$, 所以 $\lambda - \lambda^2 = 0$, 即 $\lambda = 0$ 或 1.

练习 3.19 证明：两个对称变换的和还是一个对称变换. 两个对称变换的乘积是不是对称变换? 找出两个对称变换的乘积是对称变换的一个充要条件.

解 设 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m\}$ 为 V 的一个标准正交基, 对称变换 σ, τ 关于其的矩阵为对称矩阵 $\boldsymbol{A}, \boldsymbol{B}$, 即 $\boldsymbol{A} = \boldsymbol{A}^T, \boldsymbol{B} = \boldsymbol{B}^T$, 因此 $(\boldsymbol{A}+\boldsymbol{B})^T = \boldsymbol{A}^T + \boldsymbol{B}^T = \boldsymbol{A}+\boldsymbol{B}$, 所以 $\boldsymbol{A}+\boldsymbol{B}$ 为对称矩阵.

而 $(\boldsymbol{A}\boldsymbol{B})^T = \boldsymbol{B}^T\boldsymbol{A}^T$ 一般不等于 $\boldsymbol{A}^T\boldsymbol{B}^T$, 所以两个对称变换的乘积不一定是对称变换, 两个对称变换的乘积是对称变换的一个充要条件为 $\boldsymbol{A}\boldsymbol{B} = \boldsymbol{B}\boldsymbol{A}$.

3.5 2024 年 5 月 28 日第十三周星期二第二十四次课-对称变换续

定理 3.17

(1) 设 σ 是 n 维欧氏空间 V 的一个对称变换. 那么存在 V 的一个规范正交基，使得 σ 关于这个基的矩阵是实对角形式.

(2) 设 A 是一个 n 阶实对称矩阵. 那么存在一个 n 阶正交矩阵 $\boldsymbol{U}$, 使得 $\boldsymbol{U}^T\boldsymbol{A}\boldsymbol{U}$ 是实对角形.

- 对 n 作数学归纳法. $n=1$ 时是显然的, 因为一阶矩阵自然是对角形式.
- 设 $n>1$ 并且假设对于 $n-1$ 维欧氏空间的对称变换来说定理成立. 现在设 σ 是 n 维欧氏空间 V 的一个对称变换.
- 由定理 3.16 的 (1) 知道 σ 的本征值都是实数. 令 λ 是 σ 的一个本征值, α_1 是 V 中属于 λ 的一个本征向量, 并且可设 α_1 是单位向量:
$$\sigma\left(\alpha_1\right)=\lambda\alpha_1, \quad \left|\alpha_1\right|=1.$$
令 $W=\mathscr{L}\left(\boldsymbol{\alpha}_1\right)$ 是 $\boldsymbol{\alpha}_1$ 所生成的一维子空间. W 在 σ 之下不变.
- 有正交分解,
$$V=W\oplus W^{\perp}.$$
$W^{\perp}$ 也在 σ 之下不变. 事实上, 设 $\xi\in W^{\perp}$. 对于任意 $\eta\in W$, 我们有
$$\langle\boldsymbol{\sigma}(\boldsymbol{\xi}), \boldsymbol{\eta}\rangle=\langle\boldsymbol{\xi}, \boldsymbol{\sigma}(\boldsymbol{\eta})\rangle=0$$
所以 $\sigma(\xi)\in W^{\perp}$.
- σ 在 $W^{\perp}$ 上的限制 $\sigma|_{W^{\perp}}$ 是 $W^{\perp}$ 的一个对称变换, 并且 $\sigma|_{W^{\perp}}$ 的本征值都是 σ 的本征值.
- 因为 $\dim W^{\perp}=n-1$, 所以由归纳法的假设, 存在 $W^{\perp}$ 的一个规范正交基 $\boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$, 使 $\boldsymbol{\sigma}|_{W^{\perp}}$ 关于这个基的矩阵是实对角形式.
- 而 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 就是 V 的一个规范正交基. σ 关于这个基的矩阵是实对角形式.

如何对角化一个矩阵? 步骤如下:

- 为了求出 $\boldsymbol{U}$, 我们可以用以下的方法. 首先注意, 由于 $\boldsymbol{U}$ 是正交矩阵, 所以 $\boldsymbol{U}^{\top}=\boldsymbol{U}^{-1}$, 因此 $\boldsymbol{U}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{U}$ 与 $\boldsymbol{A}$ 相似.
- 我们可以通过求特征向量的方法求出一个可逆矩阵 T, 使得 $T^{-1}AT$ 是对角形式.
- 这样求出的矩阵 $\boldsymbol{T}$ 一般说来还不是正交矩阵.
- 然而注意到 $\boldsymbol{T}$ 的列向量都是 $\boldsymbol{A}$ 的特征向量, $\boldsymbol{A}$ 的属于不同特征根的特征向量彼此正交, 因此只要再对 T 中属于 $\boldsymbol{A}$ 的同一特征根的列向量施行正交化手续, 就得到 $\mathbf{R}^n$ 的一个规范正交组.
- 以这样得到的规范正交组作列, 就得到一个满足要求的 n 阶正交矩阵 $\boldsymbol{U}$.

再总结一下

- 第一步, 求出 A 的特征值 $\lambda_1, \lambda_2, \cdots, \lambda_k$.

- 第二步, 对 λ_i 求出齐次线性方程组

$$(\lambda_i I_n - A) X = 0$$

的基础解系[1], 并用 Schmidt 方法将其正交化, 即求出 $E_{\lambda_i}(A)$ 的标准正交基

$$\eta_{i1}, \eta_{i2}, \cdots, \eta_{in_i}$$

- 第三步, 令

$$T = (\eta_{11}, \cdots, \eta_{1n_1}, \eta_{21}, \cdots, \eta_{kn_k})$$

则 T 为正交矩阵, 且

$$T^T A T = \operatorname{diag}(\lambda_1 I_{n_1}, \lambda_2 I_{n_2}, \cdots, \lambda_k I_{n_k})$$

例题 **3.14** 设

$$A = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix}.$$

找出一个正交矩阵 $\boldsymbol{U}$ 使 $\boldsymbol{U}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{U}$ 是对角形矩阵.

- 第一步, 先求 $\boldsymbol{A}$ 的全部特征根. 我们有

$$\det(xI - A) = (x-2)^2(x-8).$$

所以 A 的特征根是 $2, 2, 8$.

- 第二步, 先对于特征根 2 , 求出齐次线性方程组

$$\begin{pmatrix} -2 & -2 & -2 \\ -2 & -2 & -2 \\ -2 & -2 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

的一个基础解系

$$\eta_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \eta_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}.$$

再把 $\{\boldsymbol{\eta}_1, \boldsymbol{\eta}_2\}$ 正交化, 得

$$\gamma_1 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \gamma_2 = \begin{pmatrix} -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{pmatrix}.$$

- 对于特征根 8 , 求出属于它的一个单位特征向量

$$\gamma_3 = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}$$

[1]一个齐次线性方程组的解空间的一个基叫作这个方程组的一个基础解系.

- 第三步, 以 $\gamma_1, \gamma_2, \gamma_3$ 为列, 作一个矩阵

$$U=\begin{pmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix}.$$

那么 $\boldsymbol{U}$ 是正交矩阵, 并且

$$\boldsymbol{U}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{U}=\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 8 \end{pmatrix}.$$

例题 3.15 求正交矩阵 T 使 T^TAT 为对角阵, 其中

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

- 由

$$|\lambda I_4 - A| = (\lambda-1)^3(\lambda+3),$$

知 A 的特征根为 1 (三重) 与 -3 .

- $\lambda=1$ 时, $(I_4-A)\,X=0$ 的基础解系为

$$\alpha_1=\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \alpha_2=\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \alpha_3=\begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

- 将基础解系正交化有

$$\beta_1=\alpha_1, \beta_2=\alpha_2-\frac{(\alpha_2,\alpha_1)}{(\alpha_1,\alpha_1)}\alpha_1=\begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 0 \\ 0 \end{pmatrix},$$

$$\beta_3=\alpha_3-\frac{(\alpha_3,\alpha_1)}{(\alpha_1,\alpha_1)}\alpha_1-\frac{(\alpha_3,\beta_2)}{(\beta_2,\beta_2)}\beta_2=\begin{pmatrix} -\frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \\ 1 \end{pmatrix}.$$

- 再单位化, 得

$$\eta_1=\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \\ 0 \end{pmatrix}, \quad \eta_2=\begin{pmatrix} \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \\ 0 \end{pmatrix}, \quad \eta_3=\begin{pmatrix} -\frac{1}{\sqrt{12}} \\ \frac{1}{\sqrt{12}} \\ \frac{1}{\sqrt{12}} \\ \frac{3}{\sqrt{12}} \end{pmatrix}.$$

- $\lambda=-3$ 时, $(-3I_4-A)\,X=0$ 的非零解为

$$\eta_4=\begin{pmatrix}\frac{1}{2}\\-\frac{1}{2}\\-\frac{1}{2}\\\frac{1}{2}\end{pmatrix}.$$

于是 $T=(\eta_1,\eta_2,\eta_3,\eta_4)$.

$$T^TAT=\mathrm{diag}(\mathbf{1},\mathbf{1},\mathbf{1},-3)$$

为 A 的对角化矩阵.

例题 3.16 请计算

$$A^{2024}=?$$

✍ **练习 3.20** 假定 V 是欧式空间，$\mathscr{A}:V\to V$ 是线性变换，

$$\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$$

是一个标准正交基，证明：$\mathscr{A}$ 是对称变换当且仅当

$$\langle\mathscr{A}(\alpha_i),\alpha_j\rangle=\langle\alpha_i,\mathscr{A}(\alpha_j)\rangle$$

解

$(\Rightarrow)$ 显然.

$(\Leftarrow)$ $\forall\xi,\eta\in V$, 设 $\xi=k_1\alpha_1+...+k_n\alpha_n,\eta=s_1\alpha_1+...+s_n\alpha_n$, 因此

$$\begin{aligned}\langle\mathscr{A}\left(\xi\right),\eta\rangle&=\langle\mathscr{A}\left(k_1\alpha_1+...+k_n\alpha_n\right),s_1\alpha_1+...+s_n\alpha_n\rangle\\&=\sum_{i,j=1}^{n}k_is_j\left\langle\mathscr{A}\left(\alpha_i\right),\alpha_j\right\rangle\\&=\sum_{i,j=1}^{n}k_is_j\left\langle\alpha_i,\mathscr{A}\left(\alpha_j\right)\right\rangle\\&=\langle k_1\alpha_1+...+k_n\alpha_n,\mathscr{A}\left(s_1\alpha_1+...+s_n\alpha_n\right)\rangle\\&=\langle\xi,\mathscr{A}\left(\eta\right)\rangle.\end{aligned}$$

✍ **练习 3.21** $\mathbb{R}^n$ 上我们给内积如下:

$$\langle X,Y\rangle=\sum_{k=1}^{n}x_ky_k.$$

其中

$$X=\begin{pmatrix}x_1\\x_2\\\cdots\\x_n\end{pmatrix},\quad Y=\begin{pmatrix}y_1\\y_2\\\cdots\\y_n\end{pmatrix}\in\mathbb{R}^n$$

证明:

$$\langle X,Y\rangle=Y^TX.$$

解 $\langle X, Y\rangle = x_1y_1 + ... + x_ny_n = Y^TX.$

✍ **练习 3.22** $\mathbb{R}^n$ 上我们给内积如下:

$$\langle X, Y\rangle = \sum_{k=1}^{n} x_k y_k.$$

其中

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$$

任意给一个 $n \times n$ 的实矩阵 A, 定义线性变换 $\mathscr{A} : \mathbb{R}^n \to \mathbb{R}^n$ 如下:

$$\mathscr{A}(X) = AX, \quad X \in \mathbb{R}^n.$$

证明: $\mathscr{A}$ 在基

$$\left\{ \alpha_1 = \begin{pmatrix} 1 \\ 0 \\ \cdots \\ 0 \end{pmatrix}, \alpha_2 = \begin{pmatrix} 0 \\ 1 \\ \cdots \\ 0 \end{pmatrix}, \cdots, \alpha_n = \begin{pmatrix} 0 \\ 0 \\ \cdots \\ 1 \end{pmatrix} \right\}$$

下的矩阵表示就是 A.

解 因为 $\mathscr{A}(\alpha_i) = A\alpha_i = A$的第一列, 所以 $\mathscr{A}(\alpha_1, ..., \alpha_n) = (\alpha_1, ..., \alpha_n) A$, 即 $\mathscr{A}$ 在基

$$\left\{ \alpha_1 = \begin{pmatrix} 1 \\ 0 \\ \cdots \\ 0 \end{pmatrix}, \alpha_2 = \begin{pmatrix} 0 \\ 1 \\ \cdots \\ 0 \end{pmatrix}, \cdots, \alpha_n = \begin{pmatrix} 0 \\ 0 \\ \cdots \\ 1 \end{pmatrix} \right\}$$

下的矩阵表示就是 A.

✍ **练习 3.23** 已知

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

求 A^{520}.

✍ **练习 3.24** 已知矩阵 $A = \begin{pmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{pmatrix}$.

(1) 求 A 的特征多项式, 并确定其是否有重根.

(2) 求一个正交矩阵 P 使得 PAP^{-1} 为对角阵.

(3) (选做) 令 V 是所有与 A 可交换的实矩阵全体, 证明 V 是一个实数域上的线性空间, 并确定 V 的维数.

解

(1) $|\lambda I-A|=(\lambda-10)(\lambda-1)^2$, 有重根 1.

(2) 对 $\lambda=10$, 解得其对应的特征向量为 $\eta_1=\begin{pmatrix}1\\2\\-2\end{pmatrix}$, 再将其单位化, 得 $\gamma_1=\begin{pmatrix}1/3\\2/3\\-2/3\end{pmatrix}$;

对 $\lambda=1$, 解得其对应的特征向量为 $\eta_2=\begin{pmatrix}2\\0\\1\end{pmatrix}$, $\eta_3=\begin{pmatrix}0\\1\\1\end{pmatrix}$, 再将其单位正交化, 得

$\gamma_2=\begin{pmatrix}2/\sqrt{5}\\0\\1/\sqrt{5}\end{pmatrix}$, $\gamma_3=\begin{pmatrix}-2/3\sqrt{5}\\\sqrt{5}/3\\4/3\sqrt{5}\end{pmatrix}$.

令 $P=\begin{pmatrix}\frac{1}{3}&\frac{2}{\sqrt{5}}&-\frac{2}{3\sqrt{5}}\\\frac{2}{3}&0&\frac{\sqrt{5}}{3}\\-\frac{2}{3}&\frac{1}{\sqrt{5}}&\frac{4}{3\sqrt{5}}\end{pmatrix}^T$, 因此 $PAP^{-1}=\mathrm{diag}(10,1,1)$.

练习 3.25 设 A,B 都是 n 阶实对称矩阵. 证明存在正交矩阵 T 使 $T^{-1}AT=B$ 的充分必要件是 A,B 的特征多项式相等（或特征多项式的根全部相同).

解 因为 A,B 都为实对称矩阵, 所以存在正交矩阵 T_1,T_2 使得

$$T_1^{-1}AT_1=\begin{pmatrix}\lambda_1&&&\\&\lambda_2&&\\&&\ddots&\\&&&\lambda_n\end{pmatrix},\ T_2^TBT_2=\begin{pmatrix}\mu_1&&&\\&\mu_2&&\\&&\ddots&\\&&&\mu_n\end{pmatrix},$$

其中 $\lambda_1,\lambda_2,\cdots,\lambda_n;\mu_1,\mu_2,\cdots,\mu_n$ 分别为 A,B 的全部特征值. 若 A,B 特征多项式相等（或特征多项式的根全部相同), 则可重新排列 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 使得 $\lambda_i=\mu_i$, $i=1,2,\cdots,n$, 于是 $T_1^{-1}AT_1=T_2^TBT_2$, 即 $(T_1T_2^{-1})^{-1}A(T_1T_2^{-1})=B$.

反之, 若存在正交矩阵 T 使 $T^{-1}AT=B$, 由相似矩阵有相同的特征多项式可得之.

练习 3.26 设 $A\in\mathbf{R}^{n\times n}$, 且 $A^T=A$. 试证

(1) 若 $A^2=A$, 则存在正交矩阵 T 使得 $T^TAT=\mathrm{diag}\,(I_r,0)$;

(2) 若 $A^2=I_n$, 则存在正交矩阵 T 使得 $T^TAT=\mathrm{diag}\,(I_r,-I_{n-r})$.

解

(1) 因为 A 为实对称矩阵, 所以存在正交矩阵 T_1 使得 $T_1^TAT_1=\mathrm{diag}(\lambda_1,...,\lambda_n)$, 因此 $T_1^TA^2T_1=\mathrm{diag}(\lambda_1^2,...,\lambda_n^2)=\mathrm{diag}(\lambda_1,...,\lambda_n)$, 所以 $\lambda_i=1$或$0, i=1,2,...,n$. 再经过重新排列即得存在正交矩阵 T 使得 $T^TAT=\mathrm{diag}(I_r,0)$;

(2) 同理, 存在正交矩阵 T_1 使得 $T_1^TAT_1=\mathrm{diag}(\lambda_1,...,\lambda_n)$, 因此 $T_1^TA^2T_1=\mathrm{diag}(\lambda_1^2,...,\lambda_n^2)=I_n$, 所以 $\lambda_i=1$或$-1, i=1,2,...,n$. 再经过重新排列即得存在正交矩阵 T 使得 $T^TAT=\mathrm{diag}(I_r,-I_{n-r})$.

第 4 章 二次型

内容提要

4.1 2024 年 5 月 30 日第十三周星期四第二十五次课-线性函数

酉空间不要求你掌握，只需了解即可。

定义 4.1

设 V 为数域 $\mathbb{F}$ 上的线性空间. V 到 $\mathbb{F}$ 上的映射 f 若满足

(1) $f(\alpha+\beta)=f(\alpha)+f(\beta), \forall \alpha, \beta \in V$,

(2) $f(k\alpha)=k\alpha, \forall \alpha \in V, k \in \mathbb{F}$,

则称 f 是 V 上的线性函数.

简言之，V 上线性函数 f 即为 V 到 $\mathbb{F}$ 的线性映射. 以 V^* 表示 V 上所有线性函数的集合.

命题 4.1

容易得到线性函数的下列性质.

(1) V 上函数 (即 V 到 $\mathbb{F}$ 的映射) f 为线性函数当且仅当

$$f(k\alpha+l\beta)=kf(\alpha)+lf(\beta), \quad \forall k, l \in \mathbb{F}, \alpha, \beta \in V.$$

(2) 若 $f \in V^*$, 则 $f(0)=0, f(-\alpha)=-f(\alpha)$.

(3) 若 $f \in V^*, \alpha_i \in V, k_i \in \mathbb{F}, 1 \leq i \leq s$ 则

$$f\left(\sum_{i=1}^{s} k_i \alpha_i\right)=\sum_{i=1}^{s} k_i f\left(\alpha_i\right).$$

命题 4.2

(1) 设 $f, g \in V^*, \varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ 为 V 的一组基. 则 $f=g$ 当且仅当

$$f\left(\varepsilon_i\right)=g\left(\varepsilon_i\right), i=1,2, \ldots, n.$$

(2) 设 $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ 为 V 的一组基. $a_1, a_2, \ldots, a_n \in \mathbb{F}$. 则存在唯一的 $f \in V^*$ 使得

$$f\left(\varepsilon_i\right)=a_i, i=1,2, \ldots, n.$$

称 $\left(f\left(\varepsilon_1\right), \ldots, f\left(\varepsilon_n\right)\right)$ 为 f 在基 $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ 下的矩阵.

定理 4.1

设 V 是数域 $\mathbb{F}$ 上的线性空间. V^* 为 V 上所有线性函数的集合. 又设 $f, g \in V^*$ 定义 f 与 g 的和 $f+g$ 如下

$$(f+g)(\alpha) = f(\alpha) + g(\alpha), \quad \forall \alpha \in V$$

又设 $c \in \mathbb{F}$, 定义 c 与 f 的积 cf 如下

$$(cf)(\alpha) = cf(\alpha), \quad \forall \alpha \in V$$

则 $f+g, cf \in V^*$; V^* 对上述加法及数量乘法构成 $\mathbb{F}$ 上线性空间; $\dim V^* = \dim V$.

定义 4.2

线性空间 V 上所有线性函数所构成的线性空间 V^* 称为 V 的对偶空间. 又设 $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ 为 V 的基. 由

$$f_i\left(\varepsilon_j\right) = \delta_{ij}, \quad 1 \leq i, j \leq n$$

所决定的 V^* 的基 $f_1, f_2, \ldots, f_n$ 称为 $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ 的对偶基.

定理 4.2

设 $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ 与 $\eta_1, \eta_2, \ldots, \eta_n$ 都是线性空间 V 的基. $f_1, f_2, \ldots, f_n$ 与 $g_1, g_2, \ldots, g_n$ 分别为它们的对偶基. 则

$$T\begin{pmatrix} \varepsilon_1 & \varepsilon_2 & \ldots & \varepsilon_n \\ \eta_1 & \eta_2 & \ldots & \eta_n \end{pmatrix}^T T\begin{pmatrix} f_1 & f_2 & \ldots & f_n \\ g_1 & g_2 & \ldots & g_n \end{pmatrix} = I_n$$

这里的

$$T\begin{pmatrix} \varepsilon_1 & \varepsilon_2 & \ldots & \varepsilon_n \\ \eta_1 & \eta_2 & \ldots & \eta_n \end{pmatrix}$$

代表的是过渡矩阵。

引理 4.1

设 V 是数域 $\mathbb{F}$ 上的线性空间. V^* 是 V 的对偶空间. 又设 $\alpha \in V$. 则由下面等式定义的 V^* 上的的数 α^{**} :

$$\alpha^{**}(f) = f(\alpha), \quad \forall f \in V^*$$

是 V^* 上的线性函数, 即 $\alpha^{**} \in (V^*)^*$.

事实上, $\forall f, g \in V^*, k \in \mathbb{F}$, 我们有

$$\begin{aligned} \alpha^{**}(f+g) &= (f+g)(\alpha) = f(\alpha) + g(\alpha) = \alpha^{**}(f) + \alpha^{**}(g), \\ \alpha^{**}(kf) &= kf(\alpha) = k\alpha^{**}(f). \end{aligned}$$

定理 4.3

设 $(V^*)^*$ 为线性空间 V 的对偶空间 V^* 的对偶空间. 对 $\alpha \in V, \alpha^{**}$ 定义如上. 则 V 到 $(V^*)^*$ 的映射 $\alpha \longrightarrow \alpha^{**}$ 是线性同构映射. ♡

于是可以将 α^{**} 与 α 等同起来, 这时 $(V^*)^*$ 等同于 V, 即 V 为 V^* 的对偶空间.

例题 4.1 设 $V=\mathbb{F}^{n\times n}$ 为所有的 $n\times n$ 阶矩阵的全体; 则 $\operatorname{tr}: V \longrightarrow \mathbb{F}$,

$$\operatorname{tr} A=\sum_{i=1}^{n} \operatorname{ent}_{ii} A$$

为 V 的线性函数.

例题 4.2 设 $V=\mathbb{F}[X]$ 是 $\mathbb{F}$ 上多项式的全体, $t\in\mathbb{F}$. 定义 V 上函数 L_t 如下:

$$L_t(f(x))=f(t), \quad \forall f(x)\in V,$$

则 L_t 是 V 的线性函数.

例题 4.3 设

$$V=\mathbb{F}[X]_n=\{f(x)\in\mathbb{F}[X] \mid \deg f(x)<n \quad \text{或} \quad f(x)=0\}$$

给定 $a_1, a_2, \ldots, a_n \in \mathbb{F}$, 互不相等.

$$p_i(x)=\prod_{j\neq i}\frac{x-a_j}{a_i-a_j}, i=1,2,\ldots,n.$$

于是

$$p_i\left(a_j\right)=\delta_{ij}, \quad 1\leq i,j\leq n.$$

若 $c_1, c_2, \ldots, c_n \in \mathbb{F}$ 使得

$$\sum_{i=1}^{n} c_i p_i(x)=0$$

则

$$\sum_{i=1}^{n} c_i p_i\left(a_j\right)=c_j=0, \quad 1\leq j\leq n.$$

故 $p_1(x), p_2(x), \ldots, p_n(x)$ 为 V 的基. 而 $L_{a_1}, L_{a_2}, \ldots, L_{a_n}$ 为其对偶基. 并且 $\forall f(x)\in V$, 有

$$f(x)=\sum_{i=1}^{n} f\left(a_i\right) p_i(x)$$

✍ **练习 4.1** 假定 V 是 n 维的欧式空间，给定 $\alpha\in V$, 定义函数

$$f_\alpha: V\to\mathbb{R}, \quad \beta\mapsto f_\alpha(\beta)=\langle\alpha,\beta\rangle.$$

(1) f_α 是一个线性函数。

(2) 假定 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 是 V 的一组规范正交基（或者说标准正交基，或者说 ONB）, 请证明

$$\{f_{\alpha_1}, f_{\alpha_2}, \cdots, f_{\alpha_n}\}$$

是 V^* 的关于 $\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$ 的对偶基 (参见定义 4.2)。

解

(1) $\forall\beta_1,\beta_2, f_\alpha(k\beta_1+\beta_2)=\langle\alpha,k\beta_1+\beta_2\rangle=k\langle\alpha,\beta_1\rangle+\langle\alpha,\beta_2\rangle=kf_\alpha(\beta_1)+f_\alpha(\beta_2)$, 因此 f_α 为线性函数.

(2) $f_{\alpha_i}(\alpha_j)=\langle\alpha_i,\alpha_j\rangle=\delta_{ij}, i=1,2,...,n$, 由定义, $\{f_{\alpha_1},f_{\alpha_2},\cdots,f_{\alpha_n}\}$ 是 V^* 的关于 $\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$ 的对偶基.

练习 4.2 设 V 是数域 $\mathbb{F}$ 上三维线性空间. $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 为 V 的一组基.

(1) 若 $f\in V^*$, 且 $f(\varepsilon_1+\varepsilon_3)=1, f(\varepsilon_2-2\varepsilon_3)=-1, f(\varepsilon_1+\varepsilon_2)=-3$. 求

$$f(x_1\varepsilon_1+x_2\varepsilon_2+x_3\varepsilon_3).$$

(2) 求 $f\in V^*$, 使得

$$f(\varepsilon_1+\varepsilon_3)=f(\varepsilon_2-3\varepsilon_3)=0, f(\varepsilon_1+\varepsilon_2)=1.$$

(3) 设 f_1,f_2,f_3 是 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 的对偶基. 试证 $\beta_1=\varepsilon_1-\varepsilon_3,\beta_2=\varepsilon_1+\varepsilon_2+\varepsilon_3,\beta_3=\varepsilon_2+\varepsilon_3$ 也是 V 的一组基, 并求它的对偶基 (用 f_1,f_2,f_3 表示).

解

(1)

$$\begin{cases} f(\varepsilon_1+\varepsilon_3)=f(\varepsilon_1)+f(\varepsilon_3)=1\\ f(\varepsilon_2-2\varepsilon_3)=f(\varepsilon_2)-2f(\varepsilon_3)=-1\\ f(\varepsilon_1+\varepsilon_2)=f(\varepsilon_1)+f(\varepsilon_2)=-3\end{cases}\Rightarrow\begin{cases} f(\varepsilon_1)=4\\ f(\varepsilon_2)=-7\\ f(\varepsilon_3)=-3\end{cases}$$

所以 $f(x_1\varepsilon_1+x_2\varepsilon_2+x_3\varepsilon_3)=x_1f(\varepsilon_1)+x_2f(\varepsilon_2)+x_3f(\varepsilon_3)=4x_1-7x_2-3x_3$.

(2)

$$\begin{cases} f(\varepsilon_1+\varepsilon_3)=f(\varepsilon_1)+f(\varepsilon_3)=0\\ f(\varepsilon_2-3\varepsilon_3)=f(\varepsilon_2)-3f(\varepsilon_3)=0\\ f(\varepsilon_1+\varepsilon_2)=f(\varepsilon_1)+f(\varepsilon_2)=1\end{cases}\Rightarrow\begin{cases} f(\varepsilon_1)=-\frac{1}{2}\\ f(\varepsilon_2)=\frac{3}{2}\\ f(\varepsilon_3)=\frac{1}{2}\end{cases}$$

因此可定义 f 为

$$\forall\xi=x_1\varepsilon_1+x_2\varepsilon_2+x_3\varepsilon_3,\ f(\xi)=-\frac{1}{2}x_1+\frac{3}{2}x_2+\frac{1}{2}x_3.$$

(3) 由 $f_i(\varepsilon_j)=\delta_{i,j}, i,j=1,2,3$ 得

$$\begin{cases} f_1(\beta_1,\beta_2,\beta_3)=(f_1(\varepsilon_1)-f_1(\varepsilon_3), f_1(\varepsilon_1)+f_1(\varepsilon_2)+f_1(\varepsilon_3), f_1(\varepsilon_2)+f_1(\varepsilon_3))=(1,1,0),\\ f_2(\beta_1,\beta_2,\beta_3)=(f_2(\varepsilon_1)-f_2(\varepsilon_3), f_2(\varepsilon_1)+f_2(\varepsilon_2)+f_2(\varepsilon_3), f_2(\varepsilon_2)+f_2(\varepsilon_3))=(0,1,1),\\ f_3(\beta_1,\beta_2,\beta_3)=(f_3(\varepsilon_1)-f_3(\varepsilon_3), f_3(\varepsilon_1)+f_3(\varepsilon_2)+f_3(\varepsilon_3), f_3(\varepsilon_2)+f_3(\varepsilon_3))=(-1,1,1).\end{cases}$$

设

$$\begin{pmatrix} g_1\\ g_2\\ g_3\end{pmatrix}=A\begin{pmatrix} f_1\\ f_2\\ f_3\end{pmatrix}.$$

所以

$$\begin{pmatrix} g_1(\beta_1,\beta_2,\beta_3) \\ g_2(\beta_1,\beta_2,\beta_3) \\ g_3(\beta_1,\beta_2,\beta_3) \end{pmatrix} = A \begin{pmatrix} f_1(\beta_1,\beta_2,\beta_3) \\ f_2(\beta_1,\beta_2,\beta_3) \\ f_3(\beta_1,\beta_2,\beta_3) \end{pmatrix} = A \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ -1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix}.$$

因此

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ -1 & 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 2 & -1 \end{pmatrix},$$

即

$$\begin{cases} g_1 = f_2 - f_3, \\ g_2 = f_1 - f_2 + f_3, \\ g_3 = -f_1 + 2f_2 - f_3. \end{cases}$$

4.2 2024 年 6 月 4 日第十四周星期二第二十六次课-双线性函数

定义 4.3 (双线性函数的定义)

数域 $\mathbb{F}$ 上线性空间 V 上的一个二元函数 f (即对 V 中任意一对元素 α,β, 根据 f 都唯一地对应于 $\mathbb{F}$ 中一个数 $f(\alpha,\beta)$) 如有下列性质:

(1) $f(k_1\alpha_1 + k_2\alpha_2, \beta) = k_1 f(\alpha_1,\beta) + k_2 f(\alpha_2,\beta)$,

(2) 对 $\forall k_1,k_2 \in \mathbb{F}, \alpha,\alpha_1,\alpha_2,\beta,\beta_1,\beta_2 \in V$, 都有

$$f(\alpha, k_1\beta_1 + k_2\beta_2) = k_1 f(\alpha,\beta_1) + k_2 f(\alpha,\beta_2),$$

则称 f 为 V 上的 双线性函数.

♣

显然, 如果 $f(\alpha,\beta)$ 为 V 的双线性函数, 则固定 α(或 β) 则 $f(\alpha,\beta)$ 是以 β (相应地以 α) 为变量的的线性函数.

例题 4.4 设 V 为 Euclid 空间, 则 V 的内积为 V 上双线性函数.

例题 4.5 设 $V = \mathbb{F}^{n\times 1}, A \in \mathbb{F}^{n\times n}$ 定义 f 如下:

$$f(X,Y) = X^T AY, \forall X,Y \in V$$

则 f 为 V 上双线性函数.

下面我们将证明，有限维线性空间上的双线性函数都可以表示为例 4.5 的形式. 为此, 先引进下面的定义.

定义 4.4 (度量矩阵)

设是 n 维线性空间 V 上的双线性函数. $\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n$ 为 V 的基. 称 n 阶方阵

$$G\left(f;\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n\right) = \begin{pmatrix} f\left(\varepsilon_1,\varepsilon_1\right) & f\left(\varepsilon_1,\varepsilon_2\right) & \cdots & f\left(\varepsilon_1,\varepsilon_n\right) \\ f\left(\varepsilon_2,\varepsilon_1\right) & f\left(\varepsilon_2,\varepsilon_2\right) & \cdots & f\left(\varepsilon_2,\varepsilon_n\right) \\ \cdots & \cdots & \cdots & \cdots \\ f\left(\varepsilon_n,\varepsilon_1\right) & f\left(\varepsilon_n,\varepsilon_2\right) & \cdots & f\left(\varepsilon_n,\varepsilon_n\right) \end{pmatrix}$$

为 f 在 $\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n$ 下的(度量) 矩阵.

命题 4.3

设 $\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n$ 与 $\eta_1,\eta_2,\ldots,\eta_n$ 都是线性空间 V 的基，并且过渡矩阵为

$$M = M\begin{pmatrix} \varepsilon_1 & \varepsilon_2 & \ldots & \varepsilon_n \\ \eta_1 & \eta_2 & \ldots & \eta_n \end{pmatrix}$$

则有

(1) 设 $\alpha,\beta\in V$, 它们在 $\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n$ 下的坐标分别为 X,Y . f 是 V 上双线性函数. 则

$$f(\alpha,\beta) = X^T G\left(f;\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n\right) Y.$$

(2) V 上双线性函数 f,g 满足 $f=g$ 的充分必要条件是

$$G\left(f;\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n\right) = G\left(g;\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n\right).$$

(3) 对任意 $A\in\mathbb{F}^{n\times n}$ 存在唯一的双线性函数 f 使得

$$A = G\left(f;\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n\right)$$

(4) f 为 V 上双线性函数, 则

$$G\left(f;\eta_1,\eta_2,\ldots,\eta_n\right) = M^T G\left(f;\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n\right) M$$

$4.3-4.5$ 你可以不做啊。$4.6-4.9$ 还是要做的。

练习 4.3 (选做) 设 $f_1,f_2,\ldots,f_s$ 是线性空间 V 上的 s 个非零线性函数. 试证存在 $\alpha\in V$, 使得 $f_i(\alpha)\neq 0, 1\leq i\leq s$. 温馨提示：你去搜一搜有限覆盖定理，可以证明一下有限覆盖定理或者直接用。这次期中考试实际上第二道证明题就是有限覆盖定理，这道题期中考试出的不好。

解 令

$$V_i = \mathrm{Ker}(f_i), \quad i=1,2,\cdots,s.$$

因为 f_i 不是零函数，从而 V_i 是 V 的真子空间，于是根据有限覆盖定理

$$\bigcup_{i=1}^{s} V_i \neq V.$$

于是一定存在 $\alpha \in V \setminus \{\bigcup_{i=1}^{s} V_i\}$, 也即

$$f_i(\alpha) \neq 0, \quad i = 1, 2, 3, \cdots, s.$$

✍ **练习 4.4** (选做) 设 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 是线性空间 V 的 s 个非零向量. 试证存在 $f \in V^*$, 使得 $f(\alpha_i) \neq 0, 1 \leq i \leq s$. 温馨提示：你可以利用前一道题来做。

解 首先根据定理 4.3, $\alpha_1^{**}, \alpha_2^{**}, \cdots, \alpha_s^{**}$ 是 V^* 上的 s 个非零线性函数。于是根据上一题，一定存在 $f \in V^*$ 使得

$$\alpha_i^{**}(f) = f(\alpha_i) \neq 0, \quad i = 1, 2, \cdots, s.$$

✍ **练习 4.5** (选做) 设 V 是数域 P 上线性空间. $f_1, f_2, \ldots, f_k \in V^*$. 试证

1) $W = \{\alpha \in V \mid f_i(\alpha) = 0, 1 \leq i \leq k\}$ 是 V 的子空间, 称为 $f_1, f_2, \ldots, f_k$ 的零化子空间, 记为

$$\operatorname{ann}\{f_1, f_2, \ldots, f_k\}$$

2) 零化子空间的维数为

$$\begin{aligned} &\dim \operatorname{ann}\{f_1, f_2, \ldots, f_k\} \\ &= \dim V - \operatorname{rank}\{f_1, f_2, \ldots, f_k\}. \end{aligned}$$

3) V 的任一子空间皆为某些线性函数的零化子空间.

解

(1) 任取 $\alpha, \beta \in W$ 和 $x, y \in P$, 则 $f_i(\alpha) = f_i(\beta) = 0$ 对 $i = 1, 2, \cdots, k$ 都成立。于是

$$f_i(x\alpha + y\beta) = xf(\alpha) + yf(\beta) = 0.$$

于是 $x\alpha + y\beta \in W$. 所以 W 是子空间。

(2) 若 $\{f_{i_1}, f_{i_2}, \cdots, f_{i_s}\}$ 为 $\{f_1, f_2, \cdots, f_k\}$ 的极大线性无关组，显然

$$f_i(\alpha) = 0, \forall i = 1, 2, \cdots, k \Leftrightarrow f_{i_t}(\alpha) = 0, t = 1, 2, \cdots, s.$$

因此我们不妨假定 $f_1, f_2, \cdots, f_k$ 是线性无关的。从而可以扩充为 V^* 的一组基

$$\{f_1, f_2, \cdots, f_k, f_{k+1}, \cdots, f_n\}.$$

根据定理 4.3 可知 W 同构于 V^{**} 的子空间

$$W^{**} = \{\alpha^{**} \in V^{**} \mid \alpha^{**}(f_i) = 0, \forall 1 \leq i \leq k\}$$

我们想说

$$\{f_{k+1}^*, f_{k+2}^*, \cdots, f_n^*\}$$

构成了 W^{**} 的一组基。这是因为，根据对偶基的定义知道任意的 $j = k+1, k+2, \cdots, n$ 有

$$f_j^*(f_i) = 0, i = 1, 2, 3, \cdots, k.$$

所以

$$\{f_{k+1}^*, f_{k+2}^*, \cdots, f_n^*\} \subset W^{**}.$$

其次任取 $h \in W^{**}$. 因为

$$\{f_1^*, f_2^*, \cdots, f_k^*, f_{k+1}^*, \cdots, f_n^*\}.$$

是 V^{**} 的基，于是我们有

$$h = x_1 f_1^* + x_2 f_2^* + \cdots + x_k f_k^* + x_{k+1} f_{k+1}^* + \cdots + x_n f_n^*.$$

于是对 $1 \leq i \leq k$,

$$0 = h(f_i) = \sum_{j=1}^{n} x_j f_j^*(f_i) = x_i.$$

所以

$$h = x_{k+1} f_{k+1}^* + \cdots + x_n f_n^* \in \mathcal{L}(f_{k+1}^*, f_{k+2}^*, \cdots, f_n^*).$$

也即是

$$W^{**} \subseteq \mathcal{L}(f_{k+1}^*, f_{k+2}^*, \cdots, f_n^*).$$

从而

$$W^{**} = \mathcal{L}(f_{k+1}^*, f_{k+2}^*, \cdots, f_n^*).$$

于是

$$\dim(W) = \dim(W^{**}) = n - k.$$

(3) 任取 V 的子空间 W, 假定 W 的基为

$$\{\alpha_1, \alpha_2, \cdots, \alpha_k\}.$$

可扩充为 V 的基

$$\{\alpha_1, \alpha_2, \cdots, \alpha_k, \alpha_{k+1}, \cdots, \alpha_n\}.$$

则对偶基的子集

$$\{\alpha_{k+1}^*, \alpha_{k+2}^*, \cdots, \alpha_n^*\}$$

的零化子空间就是 W。你直接去验证一下吧。

✍ **练习 4.6** 假定 $f: V_1 \to V_2$ 是线性映射，并且 $\mathrm{Ker}(f) = \{0\}$。如果 $\dim(V_1) = \dim(V_2)$。请证明 f 是一个同构映射。这个结论很重要，请你掌握。

解 课堂已经证明。请看笔记或者回放。

✍ **练习 4.7** 如果 $f: V \times V \to \mathbb{F}$ 是 V 上的双线性函数。固定 $\alpha \in V$, 请证明:

$$\beta \mapsto f(\alpha, \beta)$$

是 V 上的线性函数。

解 根据定义直接计算，太简单。略去。

✍ **练习 4.8** 设 $\mathbb{F}$ 为数域. $A \in \mathbb{F}^{m \times m}$ 是一个 $m \times m$ 阶的矩阵, $V = \mathbb{F}^{m \times n}$ 是所有的 $m \times n$ 阶矩阵组成的全体. 定义 V 上二元函数

$$f(X, Y) = \mathrm{tr}\left(X^T A Y\right), X, Y \in V.$$

1) 试证 $f(X,Y)$ 是 V 上双线性函数.

2) 求 f 在基

$$\{E_{ij} \mid 1 \leq i \leq m, 1 \leq j \leq n\}$$

下的度量矩阵. 这里 E_{ij} 指的是 (i,j) 位置是 1，其他位置是 0 的矩阵。

解 根据定义直接计算，太简单。略去。

练习 4.9 设 $V=\mathbb{F}^4$ 上的双线性函数 $f(X,Y)$ 为

$$f(X,Y)=3x_1y_2-5x_2y_1+x_3y_4-4x_4y_3,$$

其中 $X=(x_1,x_2,x_3,x_4)\,, Y=(y_1,y_2,y_3,y_4)$. 回忆度量矩阵的定义 4.4。

1) 求 $G(f;\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4)$, 其中

$$\begin{aligned}\varepsilon_1&=(1,-2,1,0), & \varepsilon_2&=(1,-1,1,0)\\ \varepsilon_3&=(-1,2,1,1), & \varepsilon_4&=(-1,-1,0,1)\end{aligned}$$

2) 求 $G(f;\eta_1,\eta_2,\eta_3,\eta_4)$, 其中

$$(\eta_1,\eta_2,\eta_3,\eta_4)=(\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4)\,T$$

$$T=\begin{pmatrix}1&1&1&1\\1&1&-1&-1\\1&-1&1&-1\\1&-1&-1&1\end{pmatrix}.$$

解 根据定义直接计算，太简单。略去。

4.3 2024 年 6 月 6 日第十四周星期四第二十七次课-对称双线性函数

定义 4.5

设 A,B 是数域 $\mathbb{F}$ 上两个 n 阶矩阵. 如果存在 $\mathbb{F}$ 上一个非奇异矩阵 $\boldsymbol{P}$, 使得

$$\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{B},$$

那么就称 $\boldsymbol{B}$ 与 $\boldsymbol{A}$ 合同.

命题 4.4 (矩阵的合同关系具有以下性质)

(1) 自反性：任意矩阵 A 都与自身合同，因为 $IAI=A$.

(2) 对称性：如果 $\boldsymbol{B}$ 与 $\boldsymbol{A}$ 合同, 那么 $\boldsymbol{A}$ 也与 $\boldsymbol{B}$ 合同, 因为由 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{B}$ 可以得出

$$\left(\boldsymbol{P}^{-1}\right)^{\mathrm{T}}\boldsymbol{B}\boldsymbol{P}^{-1}=\left(\boldsymbol{P}^{\mathrm{T}}\right)^{-1}\boldsymbol{B}P^{-1}=A.$$

(3) 传递性: 如果 $\boldsymbol{B}$ 与 $\boldsymbol{A}$ 合同, C 又与 $\boldsymbol{B}$ 合同, 那么 $\boldsymbol{C}$ 与 $\boldsymbol{A}$ 合同. 事实上, 由 $\boldsymbol{P}^{\mathrm{T}}A\boldsymbol{P}=$

$\boldsymbol{B}$ 和 $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{Q}=\boldsymbol{C}$ 可得

$$(\boldsymbol{P}\boldsymbol{Q})^{\mathrm{T}}A(P\boldsymbol{Q})=\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}\boldsymbol{Q}=\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{Q}=\boldsymbol{C}.$$

合同的矩阵显然有相同的秩, 并且与一个对称矩阵合同的矩阵仍是对称的.

定义 4.6

设 f 是线性空间 V 上的双线性函数. 若从

$$f(\alpha,\beta)=0,\forall\beta\in V$$

可推出 $\alpha=0$, 则称 f 是非退化的或满秩的.

定理 4.4

设 f 是线性空间 V 上的双线性函数. $\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n$ 是 V 的基. 则下面条件等价:

(1) f 是非退化的;

(2) $G\left(f;\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n\right)$ 是非退化的;

(3) 若 $f(\alpha,\beta)=0,\forall\alpha\in V$, 则 $\beta=0$.

定义 4.7

如果线性空间 V 的双线性函数 f 满足

$$f(\alpha,\beta)=f(\beta,\alpha),\quad\forall\alpha,\beta\in V,$$

则称 f 为对称双线性函数. 如果 f 满足

$$f(\alpha,\beta)=-f(\beta,\alpha),\quad\forall\alpha,\beta\in V$$

则称 f 为反对称双线性函数.

定理 4.5

设 V 是数域 $\mathbb{F}$ 上 n 维线性空间. $f(\alpha,\beta)$ 为 V 上对称双线性函数. 则存在 V 的基 $\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n$ 使得

$$G\left(f;\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n\right)$$

为对角矩阵, 即

$$G\left(f;\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n\right)=\operatorname{diag}\left(d_1,d_2,\ldots,d_n\right)$$

(1) 若 $f=0$, 则定理自然成立.

(2) 如果 $f(\alpha,\alpha)=0,\forall\alpha\in V$. 则由

$$f(\alpha,\beta)=\frac{1}{2}(f(\alpha+\beta,\alpha+\beta)-f(\alpha,\alpha)-f(\beta,\beta))=0$$

知 $f=0$.

(3) 故如果 $f\neq 0$, 则有 $\varepsilon_1\in V$, 使得 $f\left(\varepsilon_1,\varepsilon_1\right)\neq 0$.

(4) 自然 $\varepsilon_1 \neq 0$. 在 V 中取基 $\varepsilon_1, \eta_2, \ldots, \eta_n$. 令

$$\varepsilon_i' = \eta_i - \frac{f(\varepsilon_1, \eta_i)}{f(\varepsilon_1, \varepsilon_1)} \varepsilon_1, i = 2, \ldots, n.$$

于是

$$f(\varepsilon_1, \varepsilon_i') = 0, \quad i = 2, \ldots, n.$$

(5) 而且 $\varepsilon_1, \eta_2, \ldots, \eta_n$ 可被 $\varepsilon_1, \varepsilon_2', \ldots, \varepsilon_n'$ 线性表出. 故 $\varepsilon_1, \varepsilon_2', \ldots, \varepsilon_n'$ 为 V 的基. 即有

$$V = \mathcal{L}(\varepsilon_1) \oplus \mathcal{L}(\varepsilon_2', \varepsilon_3', \ldots, \varepsilon_n').$$

显然, $\forall \alpha \in L(\varepsilon_1), \beta \in L(\varepsilon_2', \varepsilon_3', \ldots, \varepsilon_n')$ 有

$$f(\alpha, \beta) = 0.$$

而且 f 在 $L(\varepsilon_2', \varepsilon_3', \ldots, \varepsilon_n')$ 上的限制仍然是对称双线性函数.

(6) 于是对 $\dim V$ 作归纳可得本定理.

推论 4.1

(1) 设 $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ 如上.

$$\alpha = \sum_{i=1}^{n} x_i \varepsilon_i, \beta = \sum_{j=1}^{n} y_j \varepsilon_j$$

则 $f(\alpha, \beta) = \sum_{i=1}^{n} d_i x_i y_i$.

(2) 若 $\mathbb{F} = \mathbb{C}$. 则可取 $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$, 使得

$$G = (f; \varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n) = \operatorname{diag}(I_k, 0)$$

(3) 若 $\mathbb{F} = \mathbb{R}$, 则可取 $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$, 使得

$$G = (f; \varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n) = \operatorname{diag}(I_p, -I_q, 0)$$

♡

✍ **练习 4.10** 设 V 是数域 $\mathbb{F}$ 上一切 $m \times n$ 矩阵所构成的向量空间. C 是一个取定的 $m \times m$ 矩阵. 定义 $f: V \times V \to \mathbb{F}$,

$$f(A, B) = \operatorname{tr}\left(A^{\mathrm{T}} C B\right).$$

证明: f 是 V 上一个双线性函数. f 是不是对称的?

解 根据定义直接计算，太简单。略去。不是对称的。

✍ **练习 4.11** 写出 $\mathbb{R}^3$ 上一切对称双线性函数.

✍ **练习 4.12** 设 V 是数域 F 上一个有限维内积空间. 配备了一个内积 f. 证明以下两条件等价：

(i) $\{\alpha \in V \mid f(\alpha, \beta) = 0 \text{ 对一切 } \beta \in V\} = \{0\}$;

(ii) f 关于 V 的任意基的格拉姆矩阵非奇异. 满足上述条件的内积叫作非退化的，这与定义 4.6 一致.

这就是定理 4.4，请你再证明一下吧。不会的话看一下课堂回放。

解 参见定理 4.4.

✍ **练习 4.13** 设 f 是数域 $\mathbb{F}$ 上有限维向空间 V 的一个非退化内积 (定义参看上一题或者我们的定

义 4.6). $\varphi: V \to \mathbb{F}$ 是 V 上一个线性函数. 证明存在唯一的向量 $\alpha \in V$, 使得对于任意 $\beta \in V$ 来说，都有 $\varphi(\beta) = f(\alpha, \beta)$.

解 首先 $\varphi \in V^*$。对任意的 $\beta \in V$, 定义函数

$$g_\beta : V \mapsto \mathbb{F}, \quad x \mapsto g_\beta(x) = f(x, \beta).$$

则任意的 $\beta \in V$, $g_\beta \in V^*$. 取 V 的一组基

$$\{\beta_1, \beta_2, \cdots, \beta_n\}.$$

因为 f 非退化，则

$$\{g_{\beta_1}, g_{\beta_2}, \cdots, g_{\beta_n}\}$$

是 V^* 的一组基。这是因为如果

$$t_1 g_{\beta_1} + t_2 g_{\beta_2} + \cdots + t_n g_{\beta_n} = 0.$$

则对任意的 $x \in V$ 都有

$$t_1 g_{\beta_1}(x) + t_2 g_{\beta_2}(x) + \cdots + t_n g_{\beta_n}(x) = 0.$$

也即

$$f(x, t_1\beta_1 + \cdots + t_n\beta_n) = 0, \quad \forall x \in V.$$

因为 f 非退化，根据定理 4.4, 则

$$t_1\beta_1 + \cdots + t_n\beta_n = 0.$$

于是

$$t_1 = t_2 = \cdots = t_n = 0.$$

所以

$$\{g_{\beta_1}, g_{\beta_2}, \cdots, g_{\beta_n}\}$$

线性无关，从而是 V^* 的一组基。于是存在唯一的 $h \in (V^*)^*$, 使得

$$h(g_{\beta_i}) = \varphi(\beta_i), \quad i = 1, 2, 3, \cdots.$$

这是因为线性函数只要在基上取值定了，这个函数也就定了。于是根据定理 4.3 知道存在唯一的 $\alpha \in V$ 使得 $(\alpha^*)^* = h$. 也即

$$(\alpha^*)^*(g_{\beta_i}) = g_{\beta_i}(\alpha) = h(g_{\beta_i}) = \varphi(\beta_i), \quad i = 1, 2, 3, \cdots.$$

也即

$$f(\alpha, \beta_i) = \varphi(\beta_i), \quad i = 1, 2, 3, \cdots n.$$

因为

$$\{\beta_1, \beta_2, \cdots, \beta_n\}.$$

是基，线性函数由基唯一确定。所以

$$f(\alpha, \beta) = \varphi(\beta), \quad \forall \beta \in V.$$

你也可以尝试用度量矩阵来证，很简单。不用写这么多，老师这里写这么多是想让你回忆前面的知识点，你也看到了，用到了前面很多知识点的。

✍ **练习 4.14** 设 f 是数域 $\mathbb{F}$ 上有限维向量空间 V 上一个非退化内积. $g: V \times V \to \mathbb{F}$ 是 $\mathbb{F}$ 上另一个内积. 证明存在 V 的一个线性变换 σ, 使得对于一切 $\alpha, \beta \in V$, 都有 $g(\alpha, \beta) = f(\sigma(\alpha), \beta)$. 证明：$g$ 是非退化的当且仅当 σ 是非奇异线性变换 (就是一一的，或者说即是单射又是满射).

解 取定 V 的一组基

$$\{\alpha_1, \alpha_2, \cdots, \alpha_n\}.$$

假定 f 在上面的基下的度量矩阵是 A, g 在上面的基下的度量矩阵是 B。因为 f 非退化，根据定理 4.4, 所以 A 可逆。则对任意的

$$\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)X, \quad \beta = (\alpha_1, \alpha_2, \cdots, \alpha_n)Y,$$

有

$$f(\alpha, \beta) = X^T AY, \quad g(\alpha, \beta) = X^T BY.$$

定义 $\sigma : V \to V$ 为

$$\sigma(\alpha) = (\alpha_1, \alpha_2, \cdots, \alpha_n)(A^T)^{-1}B^T.$$

也即 σ 在基 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 下的矩阵是 $(A^T)^{-1}B^T$. 于是

$$f(\sigma(\alpha), \beta) = ((A^T)^{-1}B^T X)^T AY = X^T BA^{-1}AY = X^T BY = g(\alpha, \beta).$$

根据定理 4.4，g 非退化当且仅当 B 可逆当且仅当 $(A^T)^{-1}B^T$ 可逆当且仅当 σ 可逆。

4.4 2024 年 6 月 11 日第十五周星期二第二十八次课-二次型

定义 4.8 (二次型的定义)

设 $\mathbb{F}$ 是一个数域, F 上 n 元二次齐次多项式

$$\begin{aligned} q(x_1, x_2, \cdots, x_n) = & a_{11}x_1^2 + a_{22}x_2^2 + \cdots + a_{nn}x_n^2 \\ & + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + \cdots \\ & + 2a_{n-1,n}x_{n-1}x_n \end{aligned} \tag{4.1}$$

叫作 $\mathbb{F}$ 上一个 ***n* 元二次型**.

♣

- $\mathbb{F}$ 上 n 元多项式总可以看成 $\mathbb{F}$ 上 n 个变量的函数.
- 二次型 (4.1) 定义了一个函数 $q : \mathbb{F}^n \to \mathbb{F}$. 所以 n 元二次型也称为 n 个变元的二次型.
- 在 (4.1) 中令 $a_{ij} = a_{ji}(1 \leqslant i, j \leqslant n)$. 因为 $x_ix_j = x_jx_i$, 所以 (4.1) 式可以写成以下形式:

$$q(x_1, x_2, \cdots, x_n) = \sum_{i=1}^n \sum_{j=1}^n a_{ij}x_ix_j, \quad 这里\ a_{ij} = a_{ji}. \tag{4.2}$$

定义 4.9 (二次型的矩阵)

令 $\boldsymbol{A}=(a_{ij})$ 是 (4.2) 式右端的系数所构成的矩阵, 称为二次型 $q(x_1,x_2,\cdots,x_n)$ 的矩阵. 因为 $a_{ij}=a_{ji}$, 所以 A 是 $\mathbb{F}$ 上一个 n 阶对称矩阵. 利用矩阵的乘法, (4.2) 式可以写成

$$q(x_1,x_2,\cdots,x_n)=(x_1,x_2,\cdots,x_n)\,A\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}. \tag{4.3}$$

二次型 (4.3) 的秩指的就是矩阵 $\boldsymbol{A}$ 的秩.

定义 4.10

如果对二次型 (4.3) 的变量施行如下的一个变换:

$$x_i=\sum_{j=1}^{n}p_{ij}y_j,\quad i=1,2,\cdots,n, \tag{4.4}$$

这里 $p_{ij}\in\mathbb{F}(1\leqslant i,j\leqslant n)$, 那么就得到一个关于 $y_1,y_2,\cdots,y_n$ 的二次型

$$q'(y_1,y_2,\cdots,y_n).$$

(4.4) 式称为变量的线性变换.

令 $\boldsymbol{P}=(p_{ij})$ 是 (4.4) 的系数所构成的矩阵, 则 (4.4) 式可以写成

$$\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=P\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix} \tag{4.5}$$

将 (4.5)代人 (4.3) 就得到

$$q'(y_1,y_2,\cdots,y_n)=(y_1,y_2,\cdots,y_n)\,\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix} \tag{4.6}$$

定义 4.11

矩阵 $\boldsymbol{P}$ 称为线性变换 (4.4) 的矩阵. 如果 $\boldsymbol{P}$ 是非奇异的, 就称 (4.4) 是一个非奇异线性变换. 因为 $\boldsymbol{A}$ 是对称矩阵, 所以 $\left(\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}\right)^{\mathrm{T}}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{P}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}$. $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}$ 也是对称矩阵.

于是就得到

定理 4.6

设 $\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j$ 是数域 $\mathbb{F}$ 上一个以 A 为矩阵的 n 元二次型. 对它的变量施行一次以 P 为矩阵的线性变换后所得到的二次型的矩阵是 $\boldsymbol{P}^T\boldsymbol{A}\boldsymbol{P}$. ♡

推论 4.2

一个二次型的秩在变量的非奇异线性变换之下保持不变. ♡

注如果不取二次型的矩阵是对称矩阵, 则推论 4.2 不成立. 例如, 根据定义, 二次型 $q(x_1,x_2) = 2x_1x_2$ 的矩阵应该是

$$A_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

如果直接取

$$A_2 = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}$$

作为该二次型的矩阵, 那么经过变量的非奇异线性变换

$$x_1 = y_1 - y_2, \quad x_2 = y_1 + y_2,$$

就得到二次型 $2y_1^2 - 2y_2^2$. 它的矩阵是 $\begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$, 秩为 2 , 而 A_2 的秩为 1 .

命题 4.5

设 q 和 q' 是数域 $\mathbb{F}$ 上两个 n 元二次型, 它们的矩阵分别为 $\boldsymbol{A}$ 和 B.

(1) 如果可以通过变量的非奇异线性变换将 q 变为 q', 则 B 与 $\boldsymbol{A}$ 合同.

(2) 反之, 设 $\boldsymbol{B}$ 与 $\boldsymbol{A}$ 合同. 于是存在 $\mathbb{F}$ 上非奇异矩阵 $\boldsymbol{P}$ 使得 $B = P^{\mathrm{T}}AP$. 通过以 P 为矩阵的非奇异线性变换就将 q 变成 q'. ♠

定义 4.12

$\mathbb{F}$ 上两个二次型叫作等价的, 如果可以通过变量的非奇异线性变换将其中一个变成另一个. ♣

于是有

定理 4.7

数域 $\mathbb{F}$ 上两个二次型等价的充要条件是它们的矩阵合同. 等价的二次型具有相同的秩. ♡

定理 4.8

设 $A=(a_{ij})$ 是数域 $\mathbb{F}$ 上一个 n 阶对称矩阵. 总存在 $\mathbb{F}$ 上一个 $\boldsymbol{n}$ 阶非奇异矩阵 $\boldsymbol{P}$, 使得

$$\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}=\begin{pmatrix} c_1 & & & 0 \\ & c_2 & & \\ & & \ddots & \\ 0 & & & c_n \end{pmatrix}.$$

即 $\mathbb{F}$ 上每一个 n 阶对称矩阵都与一个对角矩阵合同.

定理 4.9

数域 $\mathbb{F}$ 上每一个 n 元二次型 $\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j$ 可以通过变量的非奇异线性变换化为

$$c_1y_1^2+c_2y_2^2+\cdots+c_ny_n^2,$$

其中

$$c_1,c_2,\cdots,c_n\in\mathbb{F}.$$

此平方和称为 $q\left(x_1,x_2,\ldots,x_n\right)$ 的 标准形.

练习 4.15 写出二次型

$$\sum_{i=1}^{3}\sum_{j=1}^{3}|i-j|x_ix_j$$

的矩阵

解 过于简单，没什么要写的。但是你一定要会。

练习 4.16 证明一个非奇异的对称矩阵必与它的逆矩阵合同.

解 首先 A 合同与对角阵，即存在可逆矩阵 P 和对角矩阵 $\Lambda=\operatorname{diag}(\lambda_1,\cdots,\lambda_n)$ 使得 $A=P^T\Lambda P$. 由于 A 可逆那么 $\lambda_i\neq 0, i=1,2,\cdots,n$。于是

$$A^{-1}=P^{-1}\Lambda^{-1}(P^{-1})^T=P^{-1}\Lambda^{-1}(P^T)^{-1}P^T\Lambda PP^{-1}\Lambda^{-1}(P^{-1})^T=P^{-1}\Lambda^{-1}(P^T)^{-1}AP^{-1}\Lambda^{-1}(P^{-1})^T.$$

练习 4.17 令 $\boldsymbol{A}$ 是数域 $\mathbb{F}$ 上一个 n 阶反对称矩阵, 即满足条件 $\boldsymbol{A}^{\mathrm{T}}=-\boldsymbol{A}$.

(1) $\boldsymbol{A}$ 必与如下形式的一个矩阵合同:

$$\left(\begin{array}{cc|c|cc|ccc} 0 & 1 & & & & & & 0 \\ -1 & 0 & & & & & & \\ \hline & & \ddots & & & & & \\ \hline & & & 0 & 1 & & & \\ & & & -1 & 0 & & & \\ \hline & & & & & 0 & & \\ & & & & & & \ddots & \\ & & & & & & & 0 \end{array}\right)$$

(2) 反对称矩阵的秩一定是偶数;

(3) $\mathbb{F}$ 上两个 n 阶反对称矩阵合同的充要条件是它们存相同的秩.

解 参见定理 4.10.

定理 4.10

设 $f(\alpha,\beta)$ 是 n 维线性空间 V 上的反对标双线性函数. 则存在 V 的一组基 $\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n$, 使得

$$G\left(f;\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n\right)=\operatorname{diag}\left(S_2,S_2,\ldots,S_2,0,0,\ldots,0\right),$$

其中

$$S_2=\begin{pmatrix}0&1\\-1&0\end{pmatrix}$$

证明 对 $\dim V$ 作归纳证明 $\dim V=1$ 时, $f=0$. 定理自然成立.

$\dim V=2$ 时, 若 $f=0$, 定理成立. 现设 $f\neq 0$. 于是有 $\varepsilon_1,\varepsilon_2'$ 使得 $f\left(\varepsilon_1,\varepsilon_2'\right)\neq 0$. 显然 $\varepsilon_1,\varepsilon_2'$ 线性无关. 于是

$$\left\{\varepsilon_1,\ \varepsilon_2=f\left(\varepsilon_1,\varepsilon_2'\right)^{-1}\varepsilon_2'\right\}$$

是 V 的基. 而且 $G\left(f;\varepsilon_1,\varepsilon_2\right)=S_2$.

设 $\dim V<n(n\geq 2)$ 时定理成立. 现证 $\dim V=n$ 时定理成立. 若 $f=0$, 结论自然成立. 故设 $f\neq 0$. 如同对 $\dim V=2$ 时的证明, 此时有 $\varepsilon_1,\varepsilon_2\in V$ 使得 $f\left(\varepsilon_1,\varepsilon_2\right)=1,\varepsilon_1,\varepsilon_2$ 显然线性无关. 于是可将 $\varepsilon_1,\varepsilon_2$ 扩充为 V 的基 $\varepsilon_1,\varepsilon_2,\eta_3,\ldots,\eta_n$. 令

$$\varepsilon_i'=\eta_i+f\left(\eta_i,\varepsilon_1\right)\varepsilon_2-f\left(\eta_i,\varepsilon_2\right)\varepsilon_1, 3\leq i\leq n$$

于是

$$f\left(\varepsilon_1,\varepsilon_i'\right)=f\left(\varepsilon_1,\eta_i\right)+f\left(\eta_i,\varepsilon_1\right)f\left(\varepsilon_1,\varepsilon_2\right)=0$$
$$f\left(\varepsilon_2,\varepsilon_i'\right)=f\left(\varepsilon_2,\eta_i\right)-f\left(\eta_i,\varepsilon_2\right)f\left(\varepsilon_2,\varepsilon_1\right)=0$$

显然，$\varepsilon_1,\varepsilon_2,\varepsilon_3',\ldots,\varepsilon_n'$ 仍为 V 的基. 令

$$V_1=\mathcal{L}\left(\varepsilon_1,\varepsilon_2\right),\quad V_2=\mathcal{L}L\left(\varepsilon_3',\ldots,\varepsilon_n'\right)$$

则

$$V=V_1\oplus V_2$$
$$f(\alpha,\beta)=0,\quad \alpha\in V_1,\beta\in V_2.$$

由于 f 在 V_2 上的限制仍为反对称双线性函数, 又 $\dim V_2=\dim V-2$. 故由归纳假设有 V_2 的基 $\varepsilon_3,\ldots,\varepsilon_n$ 使得

$$G\left(f\mid V_2;\varepsilon_3,\ldots,\varepsilon_n\right)=\operatorname{diag}(\underbrace{S_2,\ldots,S_2}_{k-1\text{ 个}},0,\ldots,0).$$

显然, $\varepsilon_1,\varepsilon_2,\varepsilon_3,\ldots,\varepsilon_n$ 为 V 的基, 且

$$G\left(f;\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n\right)=\operatorname{diag}(\underbrace{S_2,\ldots,S_2}_{k\text{ 个}},0,\ldots,0).$$

于是定理成立.

推论 4.3

如果 n 维线性空间 V 有非退化的反对称双线性函数 f. 则 $n=2m$ 为偶数, 且存在 V 的基 $\eta_1,\eta_2,\ldots,\eta_n$ 使得

$$G\left(f;\eta_1,\eta_2,\ldots,\eta_n\right)=\begin{pmatrix}0 & I_m\\ -I_m & 0\end{pmatrix}.$$

♡

事实上, 由 f 非退化及定理 4.10 , 知有 V 的基 $\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n$ 使得

$$G\left(f;\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n\right)=\operatorname{diag}(\underbrace{S_2,S_2,\ldots,S_2}_{m\text{个}}).$$

于是 $n=2m$. 令

$$\begin{aligned}\eta_k&=\varepsilon_{2k-1}, & 1\le k\le m\\ \eta_{k+m}&=\varepsilon_{2k}, & 1\le k\le m.\end{aligned}$$

于是 $\eta_1,\ldots,\eta_m,\eta_{m+1},\ldots,\eta_n$ 为 V 的基. 且

$$G\left(f;\eta_1,\eta_2,\ldots,\eta_n\right)=\begin{pmatrix}0 & I_m\\ -I_m & 0\end{pmatrix}$$

4.5 2024 年 6 月 13 日第十五周星期四第二十九次课-二次型-续集

我们结课了啊!

定理 4.11 (求标准型的方法-配方法)

设

$$f\left(x_1,x_2,\ldots,x_n\right)=\sum_{i,j=1}^{n}a_{ij}x_ix_j, a_{ij}=a_{ji}.$$

(1) 若有 i_0 使得 $a_{i_0j}=a_{ji_0}=0, 1\le j\le n$. 则 f 是 $n-1$ 元二次型.

(2) 故可以假定, $\forall i_0$, 总有 j_0 使得 $a_{i_0j_0}\neq 0$. 分两种情形讨论.

1. 若 a_{ii} 不全为零. 不妨设 $a_{11} \neq 0$. 故

$$\begin{aligned}
&f\left(x_1, x_2, \ldots, x_n\right) \\
&= a_{11} x_1^2 + 2 x_1 \sum_{j=2}^n a_{1j} x_j + \sum_{i,j=2}^n a_{ij} x_i x_j \\
&= a_{11}\left(x_1 + \sum_{j=2}^n \frac{a_{1j}}{a_{11}} x_j\right)^2 - a_{11}\left(\sum_{j=2}^n \frac{a_{1j}}{a_{11}} x_j\right)^2 \\
&\quad + \sum_{i,j=2}^n a_{ij} x_i x_j
\end{aligned}$$

于是

$$f_1\left(x_2, \ldots, x_n\right) = -\frac{1}{a_{11}}\left(\sum_{j=2}^n a_{1j} x_j\right)^2 + \sum_{i,j=2}^n a_{ij} x_i x_j$$

为一个 $n-1$ 元二次型. 令

$$y_1 = x_1 + \sum_{j=2}^n \frac{a_{1j}}{a_{11}} x_j, \quad y_i = x_i, \quad 2 \leq i \leq n$$

或

$$x_1 = y_1 - \sum_{j=2}^n \frac{a_{1j}}{a_{11}} y_j, \quad x_i = y_i, \quad 2 \leq i \leq n.$$

显然, 这是非退化的线性替换. 此时有

$$f\left(x_1, x_2, \ldots, x_n\right) = a_{11} y_1^2 + f_1\left(y_2, \ldots, y_n\right).$$

再继续将 $f_1\left(y_2, \ldots, y_n\right)$ 化为平方和即可.

2. 若 $a_{11} = a_{22} = \cdots = a_{nn} = 0$. 则有 $a_{ij} \neq 0$. 不妨设 $a_{12} \neq 0$. 令

$$x_1 = z_1 + z_2, \quad x_2 = z_1 - z_2, \quad x_i = z_i, \quad 3 \leq i \leq n.$$

显然这是非退化线性替换. 此时有

$$\begin{aligned}
&f\left(x_1, x_2, \ldots, x_n\right) \\
&= 2 a_{12}\left(z_1 + z_2\right)\left(z_1 - z_2\right) + 2\left(a_{13}\left(z_1 + z_2\right) z_3 + \cdots\right. \\
&= 2 a_{12} z_1^2 - 2 a_{12} z_2^2 + \cdots
\end{aligned}$$

此时已变为前一种情形, 故可按前一情形做下去. ♡

例题 4.6 化二次型

$$f\left(x_1, x_2, x_3\right) = x_1^2 + 2 x_2^2 + 4 x_3^2 + 2 x_1 x_2 + 6 x_2 x_3 + 2 x_1 x_3$$

为标准形.

- 由 $a_{11}=1\neq 0$. 故为第一种情形. 因而有

$$\begin{aligned}
&f\left(x_1,x_2,x_3\right)\\
=&x_1^2+2x_1\left(x_2+x_3\right)+\left(x_2+x_3\right)^2-\left(x_2+x_3\right)^2+2x_2^2+4x_3^2+6x_2x_3\\
=&\left(x_1+x_2+x_3\right)^2+x_2^2+3x_3^2+4x_2x_3\\
=&\left(x_1+x_2+x_3\right)^2+\left(x_2+2x_3\right)^2-x_3^2.
\end{aligned}$$

- 令 $y_1=x_1+x_2+x_3, y_2=x_2+2x_3, y_3=x_3$, 或

$$\begin{cases}
x_1=y_1-y_2+y_3\\
x_2=y_2-2y_3,\\
x_3=y_3,
\end{cases}$$

则有

$$f\left(x_1,x_2,x_3\right)=y_1^2+y_2^2-y_3^2.$$

例题 4.7 将四元二次型

$$f=2x_1x_2+2x_1x_3-2x_1x_4-2x_2x_3+2x_2x_4+2x_3x_4$$

化为标准形.

- 由于 f 中不含平方项, 故属于第二种情形. 先作替换

$$x_1=y_1+y_2,\quad x_2=y_1-y_2,\quad x_3=y_3,\quad x_4=y_4.$$

则有

$$\begin{aligned}
f=&2y_1^2-2y_2^2+4y_2y_3-4y_2y_4+2y_3y_4\\
=&2y_1^2-2\left(y_2^2-2y_2\left(y_3-y_4\right)+\left(y_3-y_4\right)^2\right)+2\left(y_3-y_4\right)^2+2y_3y_4\\
=&2y_1^2-2\left(y_2-y_3+y_4\right)^2+2y_3^2+2y_4^2-2y_3y_4\\
=&2y_1^2-2\left(y_2-y_3+y_4\right)^2+2\left(y_3-\frac{1}{2}y_4\right)^2+\frac{3}{2}y_4^2
\end{aligned}$$

- 令

$$\begin{pmatrix} z_1\\ z_2\\ z_3\\ z_4\end{pmatrix}=\begin{pmatrix}1&0&0&0\\0&1&-1&1\\0&0&1&-\frac{1}{2}\\0&0&0&1\end{pmatrix}\begin{pmatrix} y_1\\ y_2\\ y_3\\ y_4\end{pmatrix}$$

则有

$$f\left(x_1,x_2,x_3,x_4\right)=2z_1^2-2z_2^2+2z_3^2+\frac{3}{2}z_4^2$$

其中 $(x_1,x_2,x_3,x_4)^T$ 为

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 & 1 \end{pmatrix}^{-1}\begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{pmatrix}$$
$$=\begin{pmatrix} 1 & 1 & 1 & -\frac{1}{2} \\ 1 & -1 & -1 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{2} \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{pmatrix}.$$

定理 4.12 (求标准型的方法-初等变换法)

(1) 由定理 4.9 知, 如果 $A^T=A$, 则有可逆矩阵 M 使得 M^TAM 为对角矩阵. 由 M 可逆, 于是 M 可分解为初等矩阵的乘积,

$$M=P_1P_2\cdots P_k$$

(2) 此时

$$M^T=P_k^T\cdots P_2^TP_1^T.$$

(3) 于是

$$M^TAM=P_k^T\left(\cdots\left(P_2^T\left(P_1^TAP_1\right)P_2\right)\cdots\right)P_k.$$

我们可以假定 M 为初等矩阵。M^TAT 与 A 的关系如下:

(a) $M=P(i(c)),\ c\neq 0,\ M^T=P(i(c))$. 此时, 我们将 A 的第 i 行乘 c, 得到 M^TA, 再将此矩阵的第 i 列乘 c, 即得 M^TAT.

(b) $M=P(i,j), M^T=P(i,j)$. 此时, 我们将 A 的第 i 行与第 j 行互换, 得 M^TA, 再将 M^TA 的第 i 列与第 j 列互换, 得 M^TAT.

(c) $M=P(i,j(k)), M^T=P(j,i(k))$. 此时, 将 A 的第 j 行加上第 i 行的 k 倍, 得 M^TA. 再将 M^TA 的第 j 列加上第 i 列的 k 倍得 M^TAT.

总之, 将 A 进行一次行变换, 而后进行一次相立的列变换, 这样一直变为对角矩阵. 我们的方法是为求出 T, 我们将 (A,I_n) 进行行变换, 而后将前面的部分进行相应的列变换. 最后得到 $\left(D,M^T\right)$, 其中 $D=\operatorname{diag}\left(d_1,d_2,\ldots,d_n\right)$ 为对角矩阵.

♡

例题 4.8

$$A=\begin{pmatrix} 0 & 0 & 0 & 3 \\ 0 & 3 & -6 & 0 \\ 0 & -6 & 12 & -4 \\ 3 & 0 & -4 & 0 \end{pmatrix}.$$

(1) 交换 A 的第一列和第二列, 第一行和第二行, 同时交换 I_4 的第一列和第二列. 这时 A 和

I_4 分别化为

$$A_1 = \begin{pmatrix} 3 & 0 & -6 & 0 \\ 0 & 0 & 0 & 3 \\ -6 & 0 & 12 & -4 \\ 0 & 3 & -4 & 0 \end{pmatrix}, \quad P_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

(2) 把 A_1 的第一列乘以 2 加到第三列, 第一行乘以 2 加到第三行, 同时把 $\boldsymbol{P}_1$ 的第一列乘以 2 加到第三列, 我们分别得到

$$\boldsymbol{A}_2 = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & -4 \\ 0 & 3 & -4 & 0 \end{pmatrix}, \quad \boldsymbol{P}_2 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 2 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

(3) 把 $\boldsymbol{A}_2$ 的第四列加到第二列, 第四行加到第二行, 同时把 $\boldsymbol{P}_2$ 的第四列加到第二列, 得

$$A_3 = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 6 & -4 & 3 \\ 0 & -4 & 0 & -4 \\ 0 & 3 & -4 & 0 \end{pmatrix}, P_3 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 2 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

(4) 以 $\frac{2}{3}$ 和 $-\frac{1}{2}$ 乘 A 的第二列依次加到第三列和第四列上，再以 $\frac{2}{3}$ 和 $-\frac{1}{2}$ 乘第二行依次加到第三行和第四行上, 同时对 P_3 的列施行同样的初等变换得

$$A_4 = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & -\frac{8}{3} & -2 \\ 0 & 0 & -2 & -\frac{3}{2} \end{pmatrix}, P_4 = \begin{pmatrix} 0 & 1 & \frac{2}{3} & -\frac{1}{2} \\ 1 & 0 & 2 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & \frac{2}{3} & \frac{1}{2} \end{pmatrix}.$$

(5) 最后, 以 $-\frac{3}{4}$ 乘 A_4 的第三列加到第四列上, 再以 $-\frac{3}{4}$ 乘第三行加到第四行上, 并且对 $\boldsymbol{P}_4$ 的列施行同样的初等变换, 我们得到

$$A_5 = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & -\frac{8}{3} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, P_5 = \begin{pmatrix} 0 & 1 & \frac{2}{3} & -1 \\ 1 & 0 & 2 & -\frac{3}{2} \\ 0 & 0 & 1 & -\frac{3}{4} \\ 0 & 1 & \frac{2}{3} & 0 \end{pmatrix}.$$

取 $M = \boldsymbol{P}_5$. 于是

$$M^T A M = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & -\frac{8}{3} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

4.5.1 实数域和复数域上的二次型

一般, 数域 $\mathbb{F}$ 上 n 元二次型 $f(x_1, x_2, \ldots, x_n)$ 的标准型

$$f(x_1, x_2, \ldots, x_n) = d_1 y_1^2 + d_2 y_2^2 + \cdots + d_n y_n^2$$

并不是唯一的. 事实上, 若 $t_1, t_2, \ldots, t_n$ 全不为零, 令 $y_i = t_i z_i$。这是非退化线性替换. 且

$$f(x_1, x_2, \ldots, x_n) = d_1 t_1^2 z_1^2 + d_2 t_2^2 z_2^2 + \cdots + d_n t_n^2 z_n^2$$

也是 f 的标准形. 因而 f 有无穷多个标准形.

但有一点可以肯定, f 的任何标准形中所含的非零项的个数是相同的. 都等于二次型的矩阵的秩。

定理 4.13

复数域 $\mathbb{C}$ 上的 n 元二次型 $f(x_1, x_2, \ldots, x_n)$ 经过非退化线性替换可化为

$$f(x_1, x_2, \ldots, x_n) = z_1^2 + z_2^2 + \cdots + z_r^2.$$

右式称为 $f(x_1, x_2, \ldots, x_n)$ 的 典范形式. $f(x_1, x_2, \ldots, x_n)$ 的典范形式是唯一的.

- $f(x_1, x_2, \ldots, x_n)$ 经过非退化线性替换可化为标准形

$$d_1 y_1^2 + d_2 y_2^2 + \cdots + d_r y_r^2, \quad d_i \neq 0, 1 \leq i \leq r.$$

- 再令

$$y_i = \begin{cases} \frac{1}{\sqrt{d_i}} z_i & 1 \leq i \leq r \\ z_i, & r+1 \leq i \leq n. \end{cases}$$

则可得

$$f(x_1, x_2, \ldots, x_n) = z_1^2 + z_2^2 + \cdots + z_r^2.$$

又因为 r 为 f 的秩, 故 f 的典范形式唯一.

定理 4.14

这个定理用矩阵语言叙述就是: 任一复 n 阶对称方阵 A 必合同于

$$\operatorname{diag}(I_r, 0), \quad r = \operatorname{rank} A.$$

或等价地叙述为, 两个复 n 阶对称方阵合同的充分必要是它们的秩相等.

定理 4.15 (“惯性定理”)

实数域 $\mathbb{R}$ 上的 n 元二次型 $f(x_1, x_2, \ldots, x_n)$ 经过非退化线性替换可化为

$$f(x_1, x_2, \ldots, x_n) = z_1^2 + z_2^2 + \cdots + z_p^2 - z_{p+1}^2 - \cdots - z_r^2.$$

左式称为 $f(x_1, x_2, \ldots, x_n)$ 的典范形式. $f(x_1, x_2, \ldots, x_n)$ 的规范形是唯一的.

- 由 $f(x_1, x_2, \ldots, x_n)$ 是 $\mathbb{R}$ 上二次型, 于是经过非退化线性替换可化为标准形

$$f(x_1, x_2, \ldots, x_n) = d_1 y_1^2 + d_2 y_2^2 + \cdots + d_p y_p^2 - d_{p+1} y_{p+1}^2 - \cdots - d_r y_r^2.$$

- 这里, $d_i > 0, 1 \leq i \leq r$, r 为 $f(x_1, x_2, \ldots, x_n)$ 的秩.
- 再令

$$\begin{cases} y_i = \frac{1}{\sqrt{d_i}} z_i, & 1 \leq i \leq r \\ y_i = z_i, & r+1 \leq i \leq n \end{cases}$$

则有

$$f(x_1, x_2, \ldots, x_n) = z_1^2 + z_2^2 + \cdots + z_p^2 - z_{p+1}^2 - \cdots - z_r^2.$$

即 f 被化为典范形式了.

- 设 f 有另一规范形

$$f(x_1, x_2, \ldots, x_n) = w_1^2 + \cdots + w_q^2 - w_{q+1}^2 - \cdots - w_r^2.$$

于是有可逆矩阵 $G \in \mathbb{R}^{n \times n}$ 使得

$$\begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = G \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}$$

- 记 $W' = (w_1, w_2, \ldots, w_n)$, $\quad Z' = (z_1, z_2, \ldots, z_n)$.
- 如果 $q < p$, 线性方程组

$$\begin{cases} (\operatorname{row}_1 G) Z = 0 \\ \cdots\cdots\cdots\cdots\cdots \\ (\operatorname{row}_q G) Z = 0 \\ z_{p+1} = 0 \\ \cdots\cdots\cdots \\ z_n = 0 \end{cases}$$

中有 $q + (n-p) = n - (p-q) < n$ 个方程, 于是有非零解

$$Z_0 = (k_1, \ldots, k_p, k_{p+1}, \ldots, k_n)' = (k_1, \ldots, k_p, 0, \ldots, 0)'.$$

于是

$$k_1^2 + k_2^2 + \cdots + k_p^2 > 0$$

另一方面有

$$W_0 = G Z_0 = (0, \ldots, 0, w_{0q+1}, \ldots, w_{0n})'$$

而

$$-w_{0q+1}^2 - w_{0q+2}^2 - \cdots - w_{0n}^2 \leq 0.$$

这就得到矛盾. 故 $q \geq p$.

- 同样 $p \geq q$. 即 $p = q$. 因而 f 的典范形式是唯一的.

定义 4.13 (惯性指数)

在实二次型 $f(x_1, x_2, \ldots, x_n)$ 的典范形式中正平方项的个数 p 称为 $f(x_1, x_2, \ldots, x_n)$ 的正惯性指数, 负平方项的个数 $r - p = q$ 称为 $f(x_1, x_2, \ldots, x_n)$ 的负惯性指数, 它们的差 $p - q = p - (r - p) = 2p - r$ 称为 $f(x_1, x_2, \ldots, x_n)$ 的符号差. ♣

4.5.2 正定二次型

不要求你掌握。

定义 4.14 (正定二次型)

$\mathbb{R}$ 上一个 n 元二次型 $q(x_1, x_2, \cdots, x_n)$ 可以看成定义在实数域上 n 个变量的实函数. 如果对于变量 $x_1, x_2, \cdots, x_n$ 的每一组不全为零的值, 函数值 $q(x_1, x_2, \cdots, x_n)$ 都是正数, 那么就称 $q(x_1, x_2, \cdots, x_n)$ 是一个正定二次型. ♣

定义 4.15 (负定二次型)

如果对于变量 $x_1, x_2, \cdots, x_n$ 的每一组不全为零的值, $q(x_1, x_2, \cdots, x_n)$ 都是负数, 就称 $q(x_1, x_2, \cdots, x_n)$ 是负定的. ♣

定理 4.16

实数域上二次型 $q(x_1, x_2, \cdots, x_n)$ 是正定的充要条件是它的秩和符号差都等于 n. $q(x_1, x_2, \cdots, x_n)$ 是负定的充要条件是它的秩等于 n, 符号差等于 $-n$. ♡

例题 4.9 将二次型 $Q(\boldsymbol{\alpha})$ 化为标准形式.

$$Q(x_1, x_2, x_3) = x_1x_2 + x_2x_3,$$

- 将二次型 $Q(\boldsymbol{\alpha})$ 写成矩阵形式:

$$Q(\boldsymbol{\alpha}) = x\boldsymbol{S}\boldsymbol{x}^{\mathrm{T}},$$

 其中

$$\boldsymbol{S} = \begin{pmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \end{pmatrix}.$$

- 则方阵 $\boldsymbol{S}$ 的特征多项式 $\varphi(\lambda) = \det\left(\lambda \boldsymbol{I}_{(3)} - \boldsymbol{S}\right) = \lambda^3 - \frac{1}{2}\lambda$.
- 所以方阵 $\boldsymbol{S}$ 的特征值为 $\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}, 0$.
- 方阵 $\boldsymbol{S}$ 的属于特征值 $\frac{\sqrt{2}}{2}$ 的单位特征向量

$$\boldsymbol{x} = (x_1, x_2, x_3)$$

满足方程 $\boldsymbol{x}\left(\frac{\sqrt{2}}{2}I_{(3)}-\boldsymbol{S}\right)=\boldsymbol{0}$, 即得齐次方程组:

$$\begin{cases}\frac{\sqrt{2}}{2}x_1-\frac{1}{2}x_2=0,\\ -\frac{1}{2}x_1+\frac{\sqrt{2}}{2}x_2-\frac{1}{2}x_3=0,\\ -\frac{\sqrt{2}}{2}x_2+\frac{\sqrt{2}}{2}x_3=0.\end{cases}$$

解得 $x_1=x_3, x_2=\sqrt{2}x_3$.

- 由于是单位向量, 所以 $x_1^2+x_2^2+x_3^2=\frac{1}{4}x_3^2=1$. 因此 $x_3=\pm\frac{1}{2}$. 取 $x_3=\frac{1}{2}$, 则

$$\boldsymbol{x}=\left(\frac{1}{2},\frac{\sqrt{2}}{2},\frac{1}{2}\right).$$

- 同时可以求得分别属于方阵 $\boldsymbol{S}$ 的特征值 $-\frac{\sqrt{2}}{2}$ 与 0 的单位特征向量为

$$\left(\frac{1}{2},-\frac{\sqrt{2}}{2},\frac{1}{2}\right),\quad\left(\frac{\sqrt{2}}{2},0,\frac{\sqrt{2}}{2}\right).$$

- 由于属于方阵 $\boldsymbol{S}$ 的不同特征值的特征向量是正交的, 所以

$$\boldsymbol{O}=\begin{pmatrix}\frac{1}{2}&\frac{\sqrt{2}}{2}&\frac{1}{2}\\ \frac{1}{2}&-\frac{\sqrt{2}}{2}&\frac{1}{2}\\ \frac{\sqrt{2}}{2}&0&\frac{\sqrt{2}}{2}\end{pmatrix},$$

且为正交方阵. 而且,

$$\boldsymbol{O}^{\mathrm{T}}\boldsymbol{S}\boldsymbol{O}=\operatorname{diag}\left(\frac{\sqrt{2}}{2},-\frac{\sqrt{2}}{2},0\right).$$

- 令 $\boldsymbol{x}=\boldsymbol{y}\boldsymbol{O}^{\mathrm{T}}$. 因此

$$\begin{aligned}Q(x_1,x_2,x_3)=\boldsymbol{x}\boldsymbol{S}\boldsymbol{x}^{\mathrm{T}}&=\boldsymbol{y}\left(\boldsymbol{O}^{\mathrm{T}}\boldsymbol{S}\boldsymbol{O}\right)\boldsymbol{y}^{\mathrm{T}}\\&=\boldsymbol{y}\operatorname{diag}\left(\frac{\sqrt{2}}{2},-\frac{\sqrt{2}}{2},0\right)\boldsymbol{y}^{\mathrm{T}}\\&=\frac{\sqrt{2}}{2}y_1^2-\frac{\sqrt{2}}{2}y_2^2.\end{aligned}$$

- 典范形式是

$$z_1^2-z_2^2.$$

✍ **练习 4.18** 对下列每一矩阵 A, 分别求一可逆矩阵 P, 使 P^TAP 为对角矩阵:

(i)

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

(ii)

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

(iii)

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

请你试试初等变换吧，还可以用正交相似化吧。你可以都试试啊。

解 过于简单，没什么要写的。但是你一定要会。

练习 **4.19** 写出二次型

$$\sum_{i=1}^{3}\sum_{j=1}^{3}|i-j|x_ix_j$$

的矩阵, 并将这个二次型化为一个与它等价的二次型, 使后者只含变量的平方项.

解 过于简单，没什么要写的。但是你一定要会。

第 5 章 期末大复习

5.1 知识框架

第六章要点

- 线性空间的定义，子空间
- 线性相关，线性表示，极大线性无关组，基，维数，维数公式，替换定理
- 矩阵的秩，坐标，直和

第七章要点

- 线性变换，线性变换的矩阵表示
- 线性变换在不同基下的矩阵是相似的，所以什么是相似；在处理矩阵相似的时候，可以转换成线性变换来考虑
- 特征值，特征向量，可对角化的矩阵

第八章要点

- 内积的定义，要会验证是否是内积
- 正交的定义
- 正交基，不同正交基之间的过渡矩阵是正交矩阵，Schmidt 正交化方法要掌握
- 正交变换，正交矩阵
- 对称变换，对称矩阵，对称矩阵的正交相似对角化要掌握

第九章要点

- 二次型的矩阵要会写
- 二次型的标准型要会求
- 二次型的典范形式要会求

5.2 一些例子

以下习题来自于全国各大高校的高等代数期末考试题，供你参考。老师不负责提供答案。

第 5 章 练习

1. 设 $A \in M_n(\mathbb{C})$, 即 A 是复数域上的 $n \times n$ 阶的矩阵，定义线性变换

$$\sigma : M_n(\mathbb{C}) \to M_n(\mathbb{C}) : \quad X \ \to AX - XA$$

证明：若 A 可对角化，则 σ 可对角化.

解 因为 A 可对角化，则存在可逆矩阵 P 以及对角矩阵 $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$ 使得

$$A = P^{-1}\Lambda P.$$

定义新的线性变换

$$\mathscr{A} : M_n(\mathbb{C}) \to M_n(\mathbb{C}) : \quad X \ \to \mathscr{A}(X) = \Lambda X - X\Lambda.$$

则有

$$\begin{aligned}
\sigma(P^{-1}XP) =& AP^{-1}XP - P^{-1}XPA \\
=& P^{-1}\Lambda PP^{-1}XP - P^{-1}XPP^{-1}\Lambda P \\
=& P^{-1}\Lambda XP - P^{-1}X\Lambda P \\
=& P^{-1}\mathscr{A}(X)P.
\end{aligned}$$

首先注意到

$$\mathscr{A}(X) = \Lambda X - X\Lambda = \Big((\lambda_i - \lambda_j)x_{ij}\Big)_{i,j=1,\cdots,n}, \qquad X = (x_{ij}) \in M_n(\mathbb{C}).$$

对 $1 \le k, l \le n$, 记 $E_{k,l}$ 为 (k,l) 位置是 1, 其他位置是 0 的矩阵。那么

$$\{E_{k,l} | 1 \le k, l \le n\}$$

是 $M_n(\mathbb{C})$ 的基。并且

$$\mathscr{A}(E_{k,l}) = (\lambda_k - \lambda_l)E_{k,l}.$$

所以 $E_{k,l}$ 是 $\mathscr{A}$ 的特征向量，对应的特征值是 $\lambda_k - \lambda_l$. 于是 $\mathscr{A}$ 可以被对角化。因为 P 可逆，从而

$$\{P^{-1}E_{k,l}P | 1 \le k, l \le n\}$$

也是 $M_n(\mathbb{C})$ 的基。在这组基下有

$$\sigma(P^{-1}E_{k,l}P) = P^{-1}\mathscr{A}(E_{k,l})P = (\lambda_k - \lambda_l)P^{-1}E_{k,l}P, \quad 1 \le k, l \le n.$$

于是

$$\{P^{-1}E_{k,l}P | 1 \le k, l \le n\}$$

是 σ 的一组特征向量，又是基，从而 σ 可对角化。

2. 已知 $A = \begin{pmatrix} 1 & -1 & 1 \\ x & 4 & y \\ -1 & -1 & 3 \end{pmatrix}$ 有 3 个线性无关的特征向量，且行列式 $\det(A) = 16$, 其中 4 是 A 的一个特征值, 求：

(1) x, y 的值.

(2) A^m, 其中 m 是正整数.

解 课堂已经讲过。

3. 设 $A,B,C\in M_n(\mathbb{R})$, 且 $AB-BA=CC^T$, A,B 均可对角化, 证明: A,B 可同时对角化.

解 记 $C=(c_{ij})_{1\le i,j\le n}$, 则

$$0=\mathrm{Trace}(AB-BA)=\mathrm{Trace}(CC^T)=\sum_{i,j=1}^{n}c_{ij}^2.$$

但 C 是实矩阵，于是 $C=0$。也即 $AB=BA$。根据练习 2.51, 可得 A,B 可同时对角化。

4. 设三阶实矩阵 A 的 3 个列向量 α,β,γ 线性无关, 二次型

$$f(x)=\left(\alpha^Tx\right)^2+\left(\beta^Tx\right)^2+\left(\gamma^Tx\right)^2.$$

其中 $x=(x_1,x_2,x_3)^T$.

(1) 求此二次型的矩阵 B;

(2) 问：此二次型是否正定？并写出此二次型的规范型;

(3) 是否存在正定矩阵 S, 使得 $B=S^3$? 并说明理由.

5. 设 A,B,C 为 n 阶复矩阵, 且 C 的特征值都是实数,

$$AB-BA=C^2.$$

证明: $C^n=O$.

解 根据练习 2.41 知道，如果 $\{\lambda_1,\lambda_2,\cdots,\lambda_n\}$ 是 C 的特征值，那么 $\{\lambda_1^2,\lambda_2^2,\cdots,\lambda_n^2\}$ 是 C^2 的特征值。从而

$$\mathrm{Trace}(C^2)=\lambda_1^2+\lambda_2^2+\cdots+\lambda_n^2.$$

因为 $\mathrm{Trace}(AB-BA)=0$，于是有

$$\lambda_1^2+\lambda_2^2+\cdots+\lambda_n^2=0.$$

但是 $\{\lambda_1,\lambda_2,\cdots,\lambda_n\}$ 是实数，于是

$$\lambda_1=\lambda_2=\cdots=\lambda_n=0.$$

从而 C 的特征多项式为 λ^n. 根据 Calay-Hamilton 定理可得 $C^n=0$。

6. 设

$$A=\begin{pmatrix} a & b & d\\ 0 & a & c\\ 0 & 0 & a\end{pmatrix}$$

求 A^n.

7. 设 V 是实数域上由矩阵 A 的全体实系数多项式组成的线性空间, 其中

$$A=\begin{pmatrix} 1 & 0 & 0\\ 0 & \omega & 0\\ 0 & 0 & \omega^2\end{pmatrix},\omega=\frac{-1+\sqrt{3}i}{2}.$$

求 V 的维数与一组基。

8. 证明矩阵 $A=\begin{pmatrix} 2 & -1 \\ 1 & 4 \end{pmatrix}$ 不能相似于对角形。

解 可计算特征多项式为

$$f(\lambda)=(\lambda-2)(\lambda-4)+1=(\lambda-3)^2.$$

从而特征根为 3, 代数重数为 2。但是

$$3I_2-A=\begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$$

的秩是 1。于是特征值 3 的特征子空间的维数为 $2-1=1$, 不等于代数重数，从而 A 不能相似于对角阵。

9. 设 $A=\begin{pmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{pmatrix}$, 求正交矩阵 T, 使得 $T^{-1}AT$ 是对角矩阵.

解 参见第 22. 题。

10. 设 α,β,γ 是有理数域上线性空间 V 中的向量, 其中 $\alpha\neq 0$, 假如存在 V 上线性变换 $\mathscr{T}$, 使得

$$\mathscr{T}\alpha=\beta, \mathscr{T}\beta=-\gamma, \mathscr{T}\gamma=\alpha-\beta.$$

证明: α,β,γ 在 V 中线性无关.

解 先说明 α,β 线性无关. 若 α,β 线性相关，因为 $\alpha\neq 0$, 则存在有理数 x 使得 $\beta=x\alpha$。于是

$$\begin{aligned}
\gamma &= -\mathscr{T}(\beta)=-\mathscr{T}(x\alpha)=-x\mathscr{T}(\alpha)=-x\beta=-x^2\alpha \\
\alpha-\beta &= \mathscr{T}(\gamma)=-x^2\mathscr{T}(\alpha)=-x^2\beta=-x^3\alpha.
\end{aligned}$$

于是

$$\alpha-x\alpha+x^3\alpha=0$$

因为 $\alpha\neq 0$，所以 $x^3-x+1=0$。注意到 $x^3-x+1=0$ 没有有理根，于是 α,β 线性无关。

下面再证 α,β,γ 线性无关，若 α,β,γ 线性相关，则

$$\gamma=x\alpha+y\beta.$$

于是

$$\alpha-\beta=\mathscr{T}(\gamma)=x\mathscr{T}(\alpha)+y\mathscr{T}(\beta)=x\beta-y\gamma=x\beta-xy\alpha-y^2\beta.$$

因此

$$(1+xy)\alpha+\left(y^2-x-1\right)\beta=0.$$

因为 α,β 线性无关，从而

$$\begin{cases} xy=-1 \\ y^2-x=1 \end{cases} \Rightarrow y^2+\frac{1}{y}=1$$

于是 $y^3-y+1=0$，但是这个方程没有有理根。从而 α,β,γ 线性无关。

另外的办法是：

解 注意到

$$\mathscr{T}(\gamma)=-\mathscr{T}^2(\beta)=-\mathscr{T}^3(\alpha)=\alpha-\beta=\alpha-\mathscr{T}(\alpha).$$

所以

$$(\mathscr{T}^3-\mathscr{T}+Id)(\alpha)=0.$$

令 $f(x)=x^3-x+1$, 则 $f(\mathscr{T})(\alpha)=0$。如果 α,β,γ 线性相关，则存在不全为零的有理数 p,q,s 使得

$$p\alpha+q\beta+s\gamma=0.$$

于是

$$p\alpha+q\mathscr{T}(\alpha)-s\mathscr{T}^2(\alpha)=0.$$

令 $g(x)=p+qx-sx^2$, 于是 $g\neq 0$ 并且 $g(\mathscr{T})(\alpha)=0$。因为 $g(x)$ 是一个次数不超过 2 的多项式，$f(x)$ 是一个三次多项式并且 f 在有理数域没有根。那么 $f(x),g(x)$ 一定是互素的。从而存在两个系数是有理数的多项式 $h(x)$ 和 $J(x)$ 使得

$$h(x)f(x)+J(x)g(x)=1.$$

于是

$$\alpha=h(\mathscr{T})f(\mathscr{T})(\alpha)+J(\mathscr{T})g(\mathscr{T})(\alpha)=0.$$

这与 $\alpha\neq 0$ 矛盾。

> 这两种方法都要用到 $f(x)=x^3-x+1$ 在有理数域中没有根。如果存在有理数 p/q, 其中 p,q 互素，使得 $f(p/q)=0$，则有
>
> $$p^3=pq^2-q^3=q^2(p-q)$$
>
> 于是 $p|(p-q)$，从而 $p|q$, 这与 p,q 互素矛盾！所以 $f(x)=0$ 没有有理根。

11. 设 V 是由 $\mathbb{R}^{3\times 3}$ 中所有斜对称矩阵组成的线性空间, 对任意的 $A,B\in V$ 定义: $(A,B)=\frac{1}{2}\operatorname{tr}\left(AB^T\right)$.

(1) 证明: $(\,,\,)$ 是 V 上的一个内积.

(2) 令 $\sigma:\mathbb{R}^3\to V$ 如下:

$$\sigma:\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}\to\begin{pmatrix}0&x_1&x_2\\-x_1&0&x_3\\-x_2&-x_3&0\end{pmatrix}.$$

证明: σ 是欧几里得空间 $\mathbb{R}^3$ (指定标准内积) 到 V 的一个同构映射. 并求 V 的一组标准正交基.

解

(1) 显然内积的几个条件满足。

(2) σ 是线性映射是自然。我们需要验证 σ 保内积。首先

$$\left\|\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\right\|^2 = x_1^2 + x_2^2 + x_3^2.$$

其次

$$\left\|\begin{pmatrix} 0 & x_1 & x_2 \\ -x_1 & 0 & x_3 \\ -x_2 & -x_3 & 0 \end{pmatrix}\right\|^2 = \frac{1}{2}\mathrm{Tr}\begin{pmatrix} 0 & x_1 & x_2 \\ -x_1 & 0 & x_3 \\ -x_2 & -x_3 & 0 \end{pmatrix}\begin{pmatrix} 0 & -x_1 & -x_2 \\ x_1 & 0 & -x_3 \\ x_2 & x_3 & 0 \end{pmatrix} = x_1^2+x_2^2+x_3^2.$$

于是

$$\left\|\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\right\| = \left\|\begin{pmatrix} 0 & x_1 & x_2 \\ -x_1 & 0 & x_3 \\ -x_2 & -x_3 & 0 \end{pmatrix}\right\|.$$

再根据平行四边形法则，可得

$$\langle X, Y\rangle = \langle \sigma(X), \sigma(Y)\rangle, \quad \forall X, Y \in \mathbb{R}^3.$$

想求 V 的正交基，只需选 $\mathbb{R}$ 的正交基再映过去即可。

12. 设矩阵 $A = \begin{pmatrix} -2 & 0 & 0 \\ 2 & a & 2 \\ 3 & 1 & 1 \end{pmatrix}$，且 -1 是 A 的一个特征值

(1) 求 a 的值

(2) 证明 A 可对角化, 并求可逆矩阵 T 使得 $T^{-1}AT$ 为对角阵.

(3) 设 B 是一个 3 阶矩阵且满足 $AB = BA$, 证明: B 也可对角化.

解

(1) 由于 -1 是特征值，那么有

$$\det(-I_3 - A) = \det\begin{pmatrix} 1 & 0 & 0 \\ -2 & -1-a & -2 \\ -3 & -1 & -2 \end{pmatrix} = -2a = 0.$$

从而 $a = 0$。

(2) 计算 A 的特征多项式

$$f(\lambda) = \det(\lambda I_3 - A) = \det\begin{pmatrix} \lambda+2 & 0 & 0 \\ -2 & \lambda & -2 \\ -3 & -1 & \lambda-1 \end{pmatrix} = (\lambda+1)(\lambda-2)(\lambda+2).$$

于是有三个不同的特征值，从而 A 一定相似与对角阵。为了求出这个 T, 就是求特

征向量。对应于 $\lambda_1=-2,\lambda_2=2,\lambda_3=-1$ 的特征向量是

$$\alpha_1=\begin{pmatrix}-1\\0\\1\end{pmatrix},\quad \alpha_2=\begin{pmatrix}0\\1\\1\end{pmatrix},\quad \alpha_3=\begin{pmatrix}0\\-2\\1\end{pmatrix}.$$

于是所求的矩阵为

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

(3) 注意到如果 $A\alpha=\lambda\alpha$, 则

$$AB\alpha=BA\alpha=\lambda B\alpha.$$

从而 $B\alpha$ 也在 A 的属于特征值 λ 的特征子空间 V_λ 中。这样的话因为每一个 A 的特征子空间 V_λ 是一维的，而 B 在 V_λ 上的限制一定有特征向量。从而 B 一定有三个特征向量，从而 B 可以对角化。

13. 设 V 是实数域上所有 n 阶对称矩阵构成的线性空间, 对任意的 $A,B\in V$, 定义

$$(A,B)=\operatorname{tr}(AB).$$

其中 $\operatorname{tr}(AB)$ 表示 AB 的迹.

(1) 证明: V 构成一欧氏空间.

(2) 求使 $\operatorname{tr}(A)=0$ 的子空间 S 的维数.

(3) 求 S 的正交补空间 $S^\perp$ 的维数.

14. 已知二次型

$$f(x_1,x_2,x_3)=(x_1,x_2,x_3)\begin{pmatrix}1&4&-1\\0&2&-3\\1&-1&3\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}.$$

试将 f 化为标准形.

15. 设 A 是一个二阶实对称矩阵, A 的特征值为 $\lambda_1=2,\lambda_2=3$. 并设 $\alpha=\binom{1}{1}$ 是属于特征值 $\lambda_1=2$ 的特征向量, 求矩阵 A.

解 因为属于不同特征值的特征向量正交，那么特征值 $\lambda_2=3$ 有一个特征向量是

$$\beta=(1,-1).$$

化为标准正交基为

$$X_1=\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix},\quad X_1=\frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}.$$

令矩阵

$$P=\frac{1}{\sqrt{2}}\begin{pmatrix}1&1\\1&-1\end{pmatrix}$$

则

$$P^TAP=\begin{pmatrix}2&0\\0&3\end{pmatrix}.$$

从而

$$A=P\begin{pmatrix}2&0\\0&3\end{pmatrix}P^T=\frac{1}{2}\begin{pmatrix}5&-1\\-1&5\end{pmatrix}.$$

16. 设 $A=\begin{pmatrix}1&2\\2023&2022\end{pmatrix}$, 计算 A^{2024}.

解 令 $a=2022$。我们计算 A 的特征多项式为

$$f(\lambda)=(\lambda-1)(\lambda-a)-2\times(a+1)=(\lambda-(a+2))(\lambda+1).$$

于是特征值为

$$\lambda_1=a+2,\quad,\lambda_2=-1.$$

下面计算特征向量

$$(\lambda_1I_2-A)X=\begin{pmatrix}a+1&-2\\-(a+1)&2\end{pmatrix}X=0.$$

所以对应的特征向量是 $X_1=(2,a+1)^T$。再看

$$(\lambda_2I_2-A)X=\begin{pmatrix}-2&-2\\-(a+1)&-1-a\end{pmatrix}X=0.$$

所以对应的特征向量是 $X_2=(1,-1)^T$。于是我们选取矩阵

$$P=\begin{pmatrix}2&1\\a+1&-1\end{pmatrix}.$$

则

$$P^{-1}AP=\begin{pmatrix}a+2&0\\0&-1\end{pmatrix}.$$

于是

$$A^m=P\begin{pmatrix}(a+2)^m&0\\0&(-1)^m\end{pmatrix}P^{-1}.$$

17. 利用正交变换把下面的实二次型化为标准型.

$$7x_1^2+7x_2^2-5x_3^2+32x_1x_2-16x_1x_3-16x_2x_3.$$

解 对应的矩阵是

$$A=\begin{pmatrix}7&16&-8\\16&7&-8\\-8&-8&-5\end{pmatrix}.$$

特征多项式为

$$f(\lambda)=\det\begin{pmatrix}\lambda-7 & -16 & 8\\ -16 & \lambda-7 & 8\\ 8 & 8 & \lambda+5\end{pmatrix}=-(\lambda+9)^2(\lambda-27).$$

于是有二重特征值 $\lambda_1=-9$ 和一重根 $\lambda_2=-27$。可求得对应 $\lambda_1=9$ 的两个线性无关向量为

$$X_1=\begin{pmatrix}1\\0\\2\end{pmatrix},\quad X_2=\begin{pmatrix}-1\\1\\0\end{pmatrix}$$

正交化以后可得

$$\alpha_1=\frac{1}{\sqrt{5}}\begin{pmatrix}1\\0\\2\end{pmatrix},\quad \alpha_2=\frac{1}{\sqrt{2}}\begin{pmatrix}-1\\1\\0\end{pmatrix}.$$

可求得对应 $\lambda_2=27$ 的特征向量为

$$X_3=\begin{pmatrix}-2\\-2\\1\end{pmatrix}.$$

单位化以后可得

$$\alpha_3=\frac{1}{3}\begin{pmatrix}-2\\-2\\1\end{pmatrix}.$$

令

$$P=\begin{pmatrix}\frac{1}{\sqrt{5}} & -\frac{1}{\sqrt{2}} & -\frac{2}{3}\\ 0 & \frac{1}{\sqrt{2}} & -\frac{2}{3}\\ \frac{2}{\sqrt{2}} & 0 & \frac{1}{3}\end{pmatrix}.$$

则有

$$P^TAP=\begin{pmatrix}9 & 0 & 0\\ 0 & 9 & 0\\ 0 & 0 & 27\end{pmatrix}.$$

做变量替换

$$Y=P^{-1}X=P^TX=\begin{pmatrix}\frac{1}{\sqrt{5}}x_1+\frac{2}{\sqrt{5}}x_3\\ -\frac{1}{\sqrt{2}}x_1+\frac{1}{\sqrt{2}}x_2\\ -\frac{2}{3}x_1-\frac{2}{3}x_2+\frac{1}{3}x_3\end{pmatrix}.$$

就有

$$f(x_1,x_2,x_3)=-9y_1^2-9y_2^2+27y_3^2.$$

18. 设 A 是 n 阶实方阵, 证明 A 为实对称阵当且仅当 $AA^T = A^2$, 其中 A^T 表示矩阵 A 的转置.

解 如果你学过若当标准型的话，可以证。这里是一个巧的方法。如果 $AA^T = A^2$，则 $A^TA = (A^T)^2$, 于是我们有

$$\mathrm{Tr}((A^T - A)(A - A^T)) = \mathrm{Tr}(A^TA - (A^T)^2 - A^2 + AA^T) = 0.$$

因为 A 是实矩阵，从而 $A^T - A = 0$。

19. 已知 $A = \begin{pmatrix} -1 & -2 & 6 \\ -1 & 0 & 3 \\ -1 & -1 & 4 \end{pmatrix}$, 求 A^{2024}.

解 特征多项式为

$$f(\lambda) = \det \begin{pmatrix} \lambda+1 & 2 & -6 \\ 1 & \lambda & -3 \\ 1 & 1 & \lambda-4 \end{pmatrix} = (\lambda-1)^3.$$

所以只有一个特征值 $\lambda = 1$, 解方程

$$(\lambda I_3 - A)X = \begin{pmatrix} 2 & 2 & -6 \\ 1 & 1 & -3 \\ 1 & 1 & -3 \end{pmatrix} X = 0$$

可得

$$X_1 = \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}, \quad X_2 = \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix}.$$

计算

$$(\lambda I_3 - A)^2 = \begin{pmatrix} 2 & 2 & -6 \\ 1 & 1 & -3 \\ 1 & 1 & -3 \end{pmatrix}^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

取

$$X_3 = \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}.$$

则

$$AX_3 = \begin{pmatrix} -1 & -2 & 6 \\ -1 & 0 & 3 \\ -1 & -1 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix} = \begin{pmatrix} -7 \\ -1 \\ -4 \end{pmatrix} = -\frac{8}{3}X_1 - \frac{4}{3}X_2 + X_3.$$

于是令

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

从而有

$$P^{-1}AP=\Lambda=\begin{pmatrix}1&0&-\frac{8}{3}\\0&1&-\frac{4}{3}\\0&0&1\end{pmatrix}=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}+\begin{pmatrix}0&0&-\frac{8}{3}\\0&0&-\frac{4}{3}\\0&0&0\end{pmatrix}=I_3+B.$$

因为 $B^2=0$, 所以

$$(I_3+B)^k=I_3+kB,\quad k=1,2,3,\cdots.$$

于是

$$P^{-1}A^kP=(I_3+B)^k=I_3+kB=\begin{pmatrix}1&0&-\frac{8}{3}k\\0&1&-\frac{4}{3}k\\0&0&1\end{pmatrix}$$

于是

$$\begin{aligned}A^k&=P\begin{pmatrix}1&0&-\frac{8}{3}k\\0&1&-\frac{4}{3}k\\0&0&1\end{pmatrix}P^{-1}\\&=\begin{pmatrix}3&0&1\\0&3&3\\1&1&0\end{pmatrix}\begin{pmatrix}1&0&-\frac{8}{3}k\\0&1&-\frac{4}{3}k\\0&0&1\end{pmatrix}\begin{pmatrix}\frac{1}{4}&-\frac{1}{12}&\frac{1}{4}\\-\frac{1}{4}&\frac{1}{12}&\frac{3}{4}\\\frac{1}{4}&\frac{1}{4}&-\frac{3}{4}\end{pmatrix}\\&=\begin{pmatrix}\frac{1}{4}(1-8k)+\frac{3}{4}&\frac{1}{4}(1-8k)-\frac{1}{4}&\frac{3}{4}-\frac{3}{4}(1-8k)\\\frac{1}{4}(3-4k)-\frac{3}{4}&\frac{1}{4}(3-4k)+\frac{1}{4}&\frac{9}{4}-\frac{3}{4}(3-4k)\\-k&-k&3k+1\end{pmatrix}\\&=\begin{pmatrix}1-2k&-2k&6k\\-k&1-k&3k\\-k&-k&3k+1\end{pmatrix}.\end{aligned}$$

20. 若 $\mathbb{C}$ 上的两个 n 阶矩阵 A,B 无公共特征根, $f(\lambda)$ 为 A 的特征多项式. 证明:

(1) $f(B)$ 可逆.

(2) $AX=XB$ 只有零解.

解

(1) 假定 B 的特征值是 $\{\lambda_1,\lambda_2,\cdots,\lambda_n\}$, 则根据习题 2.41, $f(B)$ 的特征值是 $\{f(\lambda_k):1\le k\le n\}$。因为 A 和 B 没有公共的特征值, 从而 $f(\lambda_k)\neq 0$ 对任意的 $1\le k\le n$

都成立。于是

$$\det f(B)=\prod_{k=1}^{n} f(\lambda_k)\neq 0.$$

所以 $f(B)$ 可逆。

(2) 如果 $AX=XB$，则对任意的 $k\geq 1$ 有

$$A^kX=A^{k-1}AX=A^{k-1}XB=A^{k-2}AXB=A^{k-2}XB^2=\cdots=XB^k.$$

于是

$$0=f(A)X=Xf(B).$$

根据第一问知道 $f(B)$ 可逆，所以 $X=0$.

21. 已知二次型 $f(x_1,x_2,x_3)=5x_1^2+5x_2^2+\beta x_3^2-2x_1x_2+6x_1x_3-6x_2x_3$ 的秩为 2 .

(1) 求 β 的值.

(2) 求一实正交变换, 将上述二次型化为标准形, 并求出标准形.

解

(1) 我们先写出这个二次型的矩阵为

$$A=\begin{pmatrix}5 & -1 & 3\\ -1 & 5 & -3\\ 3 & -3 & \beta\end{pmatrix}.$$

因为 A 的秩为 2, 所以算一下 A 的行列式

$$\det(A)=5(5\beta-9)+(-\beta+9)-36=0.$$

于是可得

$$\beta=3.$$

(2) 又到了求特征值的环节。略。

22. 已知矩阵 $A=\begin{pmatrix}2 & 2 & -2\\ 2 & 5 & -4\\ -2 & -4 & 5\end{pmatrix}$.

(1) 求 A 的特征多项式, 并确定其是否有重根.

(2) 求一个正交矩阵 P 使得 PAP^{-1} 为对角阵.

(3) 令 V 是所有与 A 可交换的实矩阵全体, 证明 V 是一个实数域上的线性空间, 并确定 V 的维数.

解

(1) 这是一个对称矩阵，我们走一遍它的对角化过程。特征多项式为

$$
\begin{aligned}
f(\lambda) =&\det(\lambda I_3 - A) = \begin{vmatrix} \lambda-2 & -2 & 2 \\ -2 & \lambda-5 & 4 \\ 2 & 4 & \lambda-5 \end{vmatrix} \\
=&(\lambda-2)(\lambda-5+4)(\lambda-5-4)+2(-2(\lambda-5)-8)+2(-8-2(\lambda-5)) \\
=&(\lambda-2)(\lambda-1)(\lambda-9)-8(\lambda-1)=(\lambda-1)(\lambda^2-11\lambda+10) \\
=&(\lambda-1)^2(\lambda-10).
\end{aligned}
$$

所以有特征值 $\lambda_1 = 1$ 是二重根，特征值 $\lambda_2 = 10$ 是一重根。

(2) 下面计算特征向量。当 $\lambda = 1$ 时，有

$$
\begin{pmatrix} \lambda-2 & -2 & 2 \\ -2 & \lambda-5 & 4 \\ 2 & 4 & \lambda-5 \end{pmatrix} X = \begin{pmatrix} -1 & -2 & 2 \\ -2 & -4 & 4 \\ 2 & 4 & -4 \end{pmatrix} X = 0.
$$

有两个线性无关的解

$$
X_1 = (0,1,1)^T, \quad X_2 = (2,0,1).
$$

做正交化为

$$
\alpha_1 = \frac{X_1}{\|X_1\|} = \left(0, \frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)
$$

和

$$
\alpha_2 = \frac{X_2 - \langle X_2, \alpha_1\rangle \alpha_1}{\|X_2 - \langle X_2, \alpha_1\rangle \alpha_1\|} = \left(\frac{2\sqrt{2}}{3}, -\frac{\sqrt{2}}{6}, \frac{\sqrt{2}}{6}\right).
$$

当 $\lambda = 10$ 时，有

$$
\begin{pmatrix} \lambda-2 & -2 & 2 \\ -2 & \lambda-5 & 4 \\ 2 & 4 & \lambda-5 \end{pmatrix} X = \begin{pmatrix} 8 & -2 & 2 \\ -2 & 5 & 4 \\ 2 & 4 & 5 \end{pmatrix} X = 0.
$$

有一个线性无关的解

$$
X_3 = (1,2,-2)^T.
$$

归一化后得到

$$
\alpha_3 = \frac{X_3}{\|X_3\|} = \left(\frac{1}{3}, \frac{2}{3}, -\frac{2}{3}\right).
$$

于是取

$$
P^{-1} = \begin{pmatrix} 0 & \frac{2\sqrt{2}}{3} & \frac{1}{3} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{6} & \frac{1}{3} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{6} & -\frac{2}{3} \end{pmatrix}
$$

则有

$$
PAP^{-1} = \operatorname{diag}(1,1,10).
$$

(3) 验证 V 是一个线性空间是容易的。只需要计算可以跟

$$\Lambda = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 10 \end{pmatrix}$$

可交换的矩阵形成的空间的维数。此时

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 10 \end{pmatrix} \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = \begin{pmatrix} a & b & c \\ d & e & f \\ 10g & 10h & 10i \end{pmatrix}$$

而

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 10 \end{pmatrix} = \begin{pmatrix} a & b & 10c \\ d & e & 10f \\ g & h & 10i \end{pmatrix}.$$

所以

$$\begin{pmatrix} a & b & c \\ d & e & f \\ 10g & 10h & 10i \end{pmatrix} = \begin{pmatrix} a & b & 10c \\ d & e & 10f \\ g & h & 10i \end{pmatrix}.$$

于是 $g = h = c = f = 0$. 我们有

$$\begin{pmatrix} a & b & 0 \\ d & e & 0 \\ 0 & 0 & i \end{pmatrix}.$$

显然这种类型的矩阵组成的向量空间的维数是 5.

23. 设 A, B 是两个 n 阶复方阵, $n > 1$.

(1) 如果 $AB = BA$, 证明 A, B 有公共的特征向量.

(2) 如果 $AB - BA = \mu B$, 其中 μ 是一个非零复数, 那么 A, B 是否会有公共的特征向量? 回答“是”请给出证明; 回答“否”请给出反例.

解

(1) 参见习题 2.51。

(2) 假定 A 的互不相同的特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_k$, 对应的特征子空间为

$$V_{\lambda_1}, V_{\lambda_2}, \cdots, V_{\lambda_k}.$$

如果对任意的特征子空间 $V_{\lambda_i}, 1 \leq i \leq k$, 都存在 $\alpha_i \in V_{\lambda_i}$ 使得 $B(\alpha_i) \neq 0$. 那么根据

$$AB(\alpha_i) - BA(\alpha_i) = \mu B(\alpha_i)$$

可得

$$AB(\alpha_i) = (\lambda_i + \mu) B(\alpha_i).$$

所以 $B(\alpha_i)$ 是特征值 $\lambda_i + \mu$ 的特征向量。于是 A 有特征值

$$\{\lambda_1 + \mu, \lambda_2 + \mu, \cdots, \lambda_k + \mu\}.$$

因为 A 的特征值是

$$\{\lambda_1, \lambda_2, \cdots, \lambda_k\}.$$

所以

$$\{\lambda_1+\mu, \lambda_2+\mu, \cdots, \lambda_k+\mu\} = \{\lambda_1, \lambda_2, \cdots, \lambda_k\}.$$

对所有元素求和得到

$$\sum_{i=1}^{k} \lambda_i = \sum_{i=1}^{k} \lambda_i + k\mu.$$

于是 $\mu=0$, 与题意矛盾! 因此一定存在 $i \in \{1,2,\cdots,k\}$ 使得对任意的 $\beta \in V_{\lambda_i}$ 都有 $B(\beta)=0$. 于是只需取 $\beta \in V_{\lambda_i} \setminus \{0\}$, 就满足

$$A(\beta)=\lambda_i \beta, \quad B(\beta)=0.$$

即二者有相同的特征向量。

24. 证明: 任何复数方阵 A 都与它的转置矩阵 A^T 相似.

25. 在二阶实数矩阵构成的线性空间 $\mathbb{R}^{2\times 2}$ 中定义:

$$(A, B)=\operatorname{tr}\left(A^T B\right), \forall A, B \in \mathbb{R}^{2\times 2},$$

其中, A^T 表示矩阵 A 的转置, $\operatorname{tr}(X)$ 表示矩阵 X 的迹.

(1) 证明 (A, B) 是线性空间 $\mathbb{R}^{2\times 2}$ 的内积.

(2) 设 W 是由 $A_1=\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, A_2=\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ 生成的子空间. 求 $W^{\perp}$ 的一组标准正交基.

26. 设二次型

$$f=a\sum_{i=1}^{n} x_i^2+b\sum_{i=1}^{n} x_i x_{n-i+1},$$

其中 a, b 为实数, 问 a, b 满足什么条件时, 二次型 f 是正定的.

27. 设 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 是有限维欧式空间的标准正交向量组, 对 $\forall \alpha \in V$, 均有

$$\sum_{i=1}^{s}\left(\alpha, \alpha_i\right)^2=|\alpha|^2,$$

那么 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 是 V 的一组基.

28. 已知 n 维欧式空间 V 的一个标准正交基为 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 且 $\alpha_0=\alpha_1+2\alpha_2+\cdots+n\alpha_n$, 定义:

$$\mathscr{A}(\alpha)=\alpha+k\left(\alpha, \alpha_0\right)\alpha_0, \alpha \in V, 0 \neq k \in \mathbb{R}.$$

(1) 验证: $\mathscr{A}$ 是线性变换.

(2) 求 $\mathscr{A}$ 在基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的矩阵.

(3) 证明: $\mathscr{A}$ 为正交变换的充要条件为 $k=-\frac{2}{1+2^2+3^2+\cdots+n^2}$.

解

(1) 直接验证。

(2) 直接计算对 $1 \le j \le n$.

$$\mathscr{A}(\alpha_j) = \alpha_j + k(\alpha_j, \alpha_0)\alpha_0 = \alpha_j + kj\alpha_0 = \alpha_j + kj\sum_{l=1}^{n} l\alpha_l.$$

所以矩阵 A 的第 j 列是

$$\underbrace{(kj, 2kj, \cdots, (j-1)kj, 1+kj^2, (j+1)kj, \cdots, nkj)^T}_{\text{第 } j \text{ 个位置是 } 1+kj^2}.$$

于是

$$\begin{aligned} A =& k\begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ 2 & 2\times 2 & 2\times 3 & \cdots & 2\times n \\ 3 & 3\times 2 & 3\times 3 & \cdots & 3\times n \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ n & n\times 2 & n\times 3 & \cdots & n\times n \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \\ =& k\begin{pmatrix} 1 \\ 2 \\ 3 \\ \vdots \\ n \end{pmatrix}(1, 2, 3, \cdots, n) + I_n. \end{aligned}$$

(3) 令

$$\alpha = \sqrt{k}(1, 2, 3, \cdots, n)^T.$$

则 $A = \alpha\alpha^T + I_n$. 当 $\mathscr{A}$ 是正交变换时，有 $A^T A = I_n$。于是

$$(\alpha\alpha^T + I_n)^2 = I_n.$$

也即

$$\alpha\alpha^T\alpha\alpha^T + 2\alpha\alpha^T + I_n = I_n.$$

因为

$$\alpha^T\alpha = p = k(1^2 + 2^2 + \cdots + n^2)$$

于是

$$(p+2)\alpha\alpha^T = 0.$$

所以 $p = -2$, 也即

$$k = -\frac{2}{1 + 2^2 + \cdots + n^2}.$$

29. 设 V 是有理数域 $\mathbb{Q}$ 上的线性空间, $\mathscr{A}$ 是 V 上的线性变换, 设

$$g(x) = x\left(x^2 + x + 1\right).$$

证明: 如果 $\mathscr{A}$ 的多项式 $g(\mathscr{A}) = 0$, 则 V 是 $\mathscr{A}$ 的核与值域的直和.

解 任取 $\alpha \in \mathrm{Ker}(\mathscr{A}) \cap \mathrm{Im}(\mathscr{A})$, 则存在 $\beta \in V$ 使得 $\alpha = \mathscr{A}(\beta)$ 并且 $\mathscr{A}(\alpha) = 0$。因此

$$\mathscr{A}^2(\beta) = 0.$$

因为 $g(\mathscr{A}) = \mathscr{A}^3 + \mathscr{A}^2 + \mathscr{A} = 0$。作用在 β 可得

$$\mathscr{A}(\beta) = 0.$$

因此 $\alpha = \mathscr{A}(\beta) = 0$。于是 $\mathrm{Ker}(\mathscr{A}) \cap \mathrm{Im}(\mathscr{A}) = \{0\}$。
任取 $\alpha \in V$, 因为

$$\mathscr{A}(\mathscr{A}^2(\alpha) + \mathscr{A}(\alpha) + \alpha) = 0,$$

所以

$$\mathscr{A}^2(\alpha) + \mathscr{A}(\alpha) + \alpha \in \mathrm{Ker}(\mathscr{A}).$$

而

$$-(\mathscr{A}^2(\alpha) + \mathscr{A}(\alpha)) \in \mathrm{Im}(\mathscr{A}).$$

于是

$$\alpha = \alpha + \mathscr{A}^2(\alpha) + \mathscr{A}(\alpha) - (\mathscr{A}^2(\alpha) + \mathscr{A}(\alpha)) \in \mathrm{Ker}(\mathscr{A}) + \mathrm{Im}(\mathscr{A}).$$

30. 设 n 阶复矩阵 $A = (a_{ij})$ 的特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$. 证明: 下面等式成立

$$\sum_{i=1}^{n} \lambda_i^2 = \sum_{i=1}^{n} \sum_{k=1}^{n} a_{ik} a_{ki}.$$

解 取 trace 即可。

31. 已知实矩阵

$$A = \begin{pmatrix} 2 & 2 \\ 2 & a \end{pmatrix}, B = \begin{pmatrix} 4 & b \\ 3 & 1 \end{pmatrix}.$$

证明:

(1) A 相似于 B 的充要条件是 $a = 3, b = \frac{2}{3}$.

(2) 矩阵方程 $AX = B$ 有解但 $BY = A$ 无解的充分必要条件是 $a \neq 2, b = \frac{4}{3}$.

解

(1) 由于 A 和 B 相似，那么他们的迹和行列式相等，于是

$$(2 + a = 4 + 12a - 4 = 4 - 3b$$

解得 $a \neq 2, b = \frac{4}{3}$.

(2) 意思是 B 的列能被 A 的列线性表示，而 A 的列不能被 B 的列线性表示。因为 B 的第一列 $\begin{pmatrix} 4 \\ 3 \end{pmatrix}$ 与 A 的第一列线性无关。因此只能是 A 的两个列线性无关，于是只能 $a \neq 2$. 又因为 A 的列不能被 B 的列线性表示，只能是 B 的两个列线性相关，也即 $b = \frac{4}{3}$。

32. 设

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

求 A^{101}.

解 还是老办法。计算 A 的特征多项式

$$f(\lambda)=\det\begin{pmatrix}\lambda+\frac{1}{3}&0&-\frac{2}{3}\\0&\lambda-\frac{1}{3}&-\frac{2}{3}\\-\frac{2}{3}&-\frac{2}{3}&\lambda\end{pmatrix}=\lambda(\lambda+1)(\lambda-1).$$

对应的特征值为

$$\lambda_1=0,\quad\lambda_2=1,\quad\lambda_3=-1.$$

与之对应的特征向量是

$$\alpha_1=\begin{pmatrix}2\\-2\\1\end{pmatrix},\quad\alpha_2=\begin{pmatrix}\frac{1}{2}\\1\\1\end{pmatrix},\quad\alpha_3=\begin{pmatrix}-1\\-\frac{1}{2}\\1\end{pmatrix}.$$

归一化后得

$$\beta_1=\begin{pmatrix}\frac{2}{3}\\-\frac{2}{3}\\\frac{1}{3}\end{pmatrix},\quad\beta_2=\begin{pmatrix}\frac{1}{3}\\\frac{2}{3}\\\frac{2}{3}\end{pmatrix},\quad\beta_3=\begin{pmatrix}-\frac{2}{3}\\-\frac{1}{3}\\\frac{2}{3}\end{pmatrix}.$$

令

$$P=(\beta_1,\beta_2,\beta_3)=\begin{pmatrix}\frac{2}{3}&\frac{1}{3}&-\frac{2}{3}\\-\frac{2}{3}&\frac{2}{3}&-\frac{1}{3}\\\frac{1}{3}&\frac{2}{3}&\frac{2}{3}\end{pmatrix}.$$

则有

$$P^TAP=\begin{pmatrix}0&0&0\\0&1&0\\0&0&-1\end{pmatrix}.$$

于是

$$A^k=P\begin{pmatrix}0&0&0\\0&1&0\\0&0&(-1)^k\end{pmatrix}P^T.$$

33. 设 $\alpha=\begin{pmatrix}1\\2\\3\\4\end{pmatrix}$, $A=\alpha\alpha^T$.

(1) 求 A 的特征值和特征向量;

(2) 求正交变换 $X=PY$ 将二次型 $f(X)=X'AX$ 化为标准形;

(3) 写出 $f(X)$ 的标准形.

解

(1)

(2)

(3)

34. 设 4 元二次型

$$f\left(x_1,x_2,x_3,x_4\right)=2x_1x_2+2x_3x_4.$$

(1) 写出二次型 $f\left(x_1,x_2,x_3,x_4\right)$ 的矩阵表达式 $f\left(x_1,x_2,x_3,x_4\right)=X'AX$; 8

(2) 求 A 的特征值和特征向量;

(3) 求正交阵 P, 使得 $P^{-1}AP=\Lambda$, 其中 Λ 是对角阵;

(4) 写出二次型 $f\left(x_1,x_2,x_3,x_4\right)$ 的标准形.

35. 设实矩阵 $A=\begin{pmatrix}1 & a & 1\\ a & 1 & 2b\\ 1 & 2b & 1\end{pmatrix}$. 如果有正交矩阵 T, 使

$$T^{-1}AT=\begin{pmatrix}0 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 2\end{pmatrix}$$

求 a,b 及 T.

36. 设 $A\in M_n(\mathbb{R})$ 且 $\det(A)\neq 0$, 证明存在唯一的 Q 和 T, 使得 $A=QT$. 其中 Q 为正交矩阵, T 为上三角矩阵, 且对角元素 $ent_{ii}T>0$.

37. 设有二次型

$$f\left(x_1,x_2,x_3\right)=x_1^2+x_2^2+3x_3^2-2x_1x_3+2ax_2x_3.$$

(1) a 满足什么条件时, f 正定.

(2) 写出 f 所对应的矩阵 A, 当 A 正定时, 求矩阵 C, 使 $A=C'C$.

38. 设有二次型

$$f\left(x_1,x_2,x_3\right)=x_1^2+x_2^2+3x_3^2-2x_1x_3+2x_2x_3.$$

(1) 证明 f 是正定;

(2) 写出 f 所对应的矩阵 A, 求矩阵 C, 使 $A=C^TC$.

39. 欧式空间 V 中的线性变换 $\mathscr{A}$ 称为反对称的, 如果对任意 $\alpha,\beta\in V$, 有

$$(\mathscr{A}\alpha,\beta)=-(\alpha,\mathscr{A}\beta).$$

(1) $\mathscr{A}$ 为反对称的充分必要是 $\mathscr{A}$ 在 V 的一组标准正交基下的矩阵是反对称的;

(2) 如果 W 是反对称变换 $\mathscr{A}$ 的不变子空间, 则 W 的正交补 $W^{\perp}$ 也是反对称变换的不变子空间.

40. 设 $\boldsymbol{\alpha}_1=\begin{pmatrix}1\\1\\2\end{pmatrix}, \boldsymbol{\alpha}_2=\begin{pmatrix}-1\\-2\\0\end{pmatrix}, \boldsymbol{\alpha}_3=\begin{pmatrix}1\\5\\-1\end{pmatrix}, \boldsymbol{\alpha}_4=\begin{pmatrix}-1\\-7\\5\end{pmatrix}$.

(1) 求向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 的秩;

(2) $\boldsymbol{\alpha}_4$ 是否能用 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表示? 若能, 写出用 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表示 $\boldsymbol{\alpha}_4$ 的表达式.

41. 设 $\boldsymbol{A}=\begin{pmatrix}1&3&3\\-3&x&-3\\3&3&1\end{pmatrix}$ 与 $\boldsymbol{B}=\begin{pmatrix}y&0&0\\0&-2&0\\0&0&-2\end{pmatrix}$ 相似, 求 x 与 y 的值.

42. 假定 V 是不超过 6 次的多项式构成的空间，求线性映射 $\mathcal{A}:V\to V$,

$$[\mathcal{A}(p)](x)=p(x)-p(-x),p(x)\in V$$

的特征值.

43. 设 A 是 n 阶实方阵, $\mathcal{L}$ 是矩阵 A^TA 的列向量维生成的线性空间. 论证: 对任: 意实 n 维列向量 β, 都有向量 $\mathrm{A}^T\beta\in\mathcal{L}$.

44. 求 $A=\begin{pmatrix}1&-3&-3\\-3&1&3\\3&-3&-5\end{pmatrix}$ 的特征值及重数, 说明 A 是否可以对角化

45. $\mathbb{R}^4$ 中的子空间 W_1 的一组基为 $\alpha_1=(1,-1,-1,1)^T,\alpha_2=1,-1,1,-3)^T, W_2$ 的一组基为 $\beta_1=(2,6,1,6)^T,\beta_2=(1,3,1,2)^T$, 求 W_1+W_2 的维数

46. 计算 $\begin{pmatrix}0&1&0\\0&0&0\\1&0&0\end{pmatrix}^n$ 和 $\begin{pmatrix}2&1&0\\0&2&0\\1&0&2\end{pmatrix}^n, (n\geq 3)$

47. 方阵 $A\in\mathbb{R}^{2\times 2}$ 满足 $A^2+aA+bE_2=O$, 其中 $a,b\in\mathbb{R},a^2-4b<0$. 试证明

(1) 对于任意的非零向量 $\beta\in\mathbb{R}^2$, 向量组 $(\beta,A\beta)$ 线性无关（6 分）

(2) 矩阵 A 相似于 $\begin{pmatrix}0&-b\\1&-a\end{pmatrix}$ (6 分)

48. 用正交变换将实二次型 $f(x_1,x_2,x_3)=x_1x_2+x_2x_3$ 化成一个标准型并求 $f(x_1,x_2,x_3)$ 的正惯性指数和负惯性指数.

49. 矩阵 $A,B\in\mathbb{R}^{n\times n}$ 满足 $AB-BA=aB,(a\neq 0)$. 求证

(1) 对任意的正整数 m, 我们有 $AB^m-B^mA=maB^m$ (8 分)

(2) 令 α 为 A 的一个特征向量, 则存在正整数使得 $B^k\alpha=0$ (4 分)

解

(1) 第一问直接演算即可，或者归纳一下。

(2)

$$AB^m\alpha-\lambda B^m\alpha=maB^m\alpha.$$

也即

$$A(B^m(\alpha))=(\lambda+ma)B^m(\alpha).$$

如果对任意的 $k>0, B^k\neq 0$, 则 A 有特征值

$$\{\lambda+ma: m=1,2,3,4,\cdots\}$$

这是不可能的。因此存在正整数使得 $B^k\alpha=0$。

50. 设 $\mathbb{R}^{2\times 2}$ 上的线性变换 $\mathscr{A}(X)=AX-XA$, 其中 $A=\begin{pmatrix}1&1\\0&1\end{pmatrix}$.

(1) 求 $\mathscr{A}$ 的所有特征值及其特征向量.

(2) 求 $\mathscr{A}$ 的最小多项式.

解

$$\mathscr{A}:\begin{pmatrix}x_1&x_2\\x_3&x_4\end{pmatrix}\mapsto\begin{pmatrix}x_3&x_4-x_1\\0&-x_3\end{pmatrix}$$

在 $\mathbb{R}^{2\times 2}$ 的基 $E_{11},E_{12},E_{21},E_{22}$ 下矩阵

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

(1) $\varphi_A(x)=x^4$, 特征值 0 的特征向量

$$X=\begin{pmatrix}x_1&x_2\\0&x_1\end{pmatrix}, x_1,x_2\in\mathbb{R} \text{ 不全为 } 0\,.$$

(2) $A^2\neq O, A^3=O\Rightarrow d_A(x)=x^3$.

51. 设 A,B 都是 n 阶实方阵.

(1) 证明: 若 A,B 无公共特征值, 则存在唯一的实方阵 X 使得 $AX-XB=I$.

(2) 证明：A 与 B 正交相抵当且仅当 A^TA 与 B^TB 正交相抵.

(3) 设 P 是正交方阵, 求 $\|PA-B\|_F$ 的最小值, 其中 $\|X\|_F=\sqrt{\sum_{i,j}x_{ij}^2}$.

解

(1) 首先根据习题 20. 知道 $AX-XB=0$ 只有零解。考虑线性变换

$$\mathscr{A}: M_n(\mathbb{R})\mapsto M_n(\mathbb{R}), X\mapsto AX-XB.$$

则有

$$\mathrm{Ker}(\mathscr{A})=\{0\}.$$

因为 $M_n(\mathbb{R})$ 是有限维的，从而 $\mathscr{A}$ 是同构，因此存在唯一的 X 使得 $\mathscr{A}(X)=I_n$。

(2)

52. 设 $\alpha=2^{\frac{1}{3}}, V=\{f(\alpha)\mid f\in\mathbb{Q}[x]\}$.

(1) 证明：V 在实数运算下构成 $\mathbb{Q}$ 上的线性空间，并且 $\{1,\alpha,\alpha^2\}$ 是 V 的基.

(2) 求 $\beta=3\alpha^2+2\alpha+1$ 在 V 的基 $\{1,\alpha^2+1,\alpha^2+3\alpha+3\}$ 下的坐标.

(3) 求 V 上线性变换 $\mathscr{A}(v)=(\alpha+1)v$ 的特征多项式.

解

(1) 易验证 $V = \{a + b\alpha + c\alpha^2 \mid a, b, c \in \mathbf{Q}\}$ 构成 $\mathbb{Q}$ 上的线性空间. 由 $x^3 = 2$ 在 $\mathbb{Q}$ 中无解, 得 $p = x^3 - 2$ 在 $\mathbb{Q}[x]$ 中不可约. 若 $\{1, \alpha, \alpha^2\}$ 线性相关, 则存在非零多项式 $f \in \mathbb{Q}_3[x]$ 使得 $f(\alpha) = 0$ ，与 p 不可约矛盾. 因此, $\{1, \alpha, \alpha^2\}$ 线性无关，构成 V 的基.

(2) $\beta = -\frac{10}{3} + \frac{7}{3}(\alpha^2 + 1) + \frac{2}{3}(\alpha^2 + 3\alpha + 3)$. 故 β 的坐标是 $\left(-\frac{10}{3}, \frac{7}{3}, \frac{2}{3}\right)$.

(3)

$$\mathscr{A}\begin{pmatrix} 1 & \alpha & \alpha^2 \end{pmatrix} = \begin{pmatrix} 1 & \alpha & \alpha^2 \end{pmatrix}\begin{pmatrix} 1 & 0 & 2 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

于是特征多项式为

$$\varphi_{\mathscr{A}}(x) = (x - 1)^3 - 2.$$