第七章　线性变换

张彪

天津师范大学

zhang@tjnu.edu.cn

Outline

1. 线性变换的定义
2. 线性变换的运算
3. 线性变换的矩阵
4. 特征值与特征向量
5. 对角矩阵
6. 线性变换的值域与核
7. 不变子空间
8. 若尔当 (Jordan) 标准形介绍
9. 最小多项式

§1 线性变换的定义

下面如果不特别声明, 所考虑的都是某一固定的数域 P 上的线性空间.

定义

线性空间 V 的一个变换称为线性变换, 如果对于 V 中任意的元素 α, β 和数域 P 中任意数 k, 都有

$$\mathscr{A}(\alpha+\beta)=\mathscr{A}(\alpha)+\mathscr{A}(\beta)$$
$$\mathscr{A}(k\alpha)=k\mathscr{A}(\alpha)$$

注

- 一般用花体拉丁字母 $\mathscr{A}, \mathscr{B}, \cdots$ 或 $\sigma, \tau, \cdots$ 代表 V 的变换
- $\mathscr{A}(\alpha)$ 或 $\mathscr{A}\alpha$ 代表元素 α 在变换 $\mathscr{A}$ 下的像
- 线性变换保持向量的加法与数量乘法

下面我们来看几个简单的例子, 它们表明线性变换这个概念是有丰富的内容的.

例 1 (旋转变换)

平面上的向量构成实数域上的二维线性空间. 把平面围绕坐标原点按反时针方向旋转 θ 角, 就是一个线性变换, 我们用 $\mathscr{I}_\theta$ 表示. 如果平面上一个向量 α 在直角坐标系下的坐标是 (x_1, y_1), 那么像 $\mathscr{I}_\theta(\alpha)$ 的坐标, 即 α 旋转 θ 角之后的坐标 (x_2, y_2) 是按照公式

$$\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$$

来计算的. 同样地, 空间中绕轴的旋转也是一个线性变换.

例 2 (投影变换)

设 α 是几何空间中一固定的非零向量, 把每个向量 ζ 变到它在 α 上的内射影的变换也是一个线性变换, 以 Π_α 表示它. 用公式表示就是

$$\Pi_\alpha(\zeta) = \frac{(\alpha, \zeta)}{(\alpha, \alpha)}\alpha$$

这里 $(\alpha, \zeta), (\alpha, \alpha)$ 表示内积.

例 3

线性空间 V 中的恒等变换或称单位变换 $\mathscr{E}$, 即

$$\mathscr{E}(\alpha) = \alpha \quad (\alpha \in V)$$

以及零变换 $\mathscr{O}$, 即

$$\mathscr{O}(\alpha) = \mathbf{0} \quad (\alpha \in V)$$

都是线性变换.

例 4

设 V 是数域 P 上的线性空间, k 是 P 中某个数, 定义 V 的变换如下:

$$\alpha \to k\alpha, \quad \alpha \in V$$

可以证明, 这是一个线性变换, 称为由数 k 决定的数乘变换, 可用 $\mathscr{K}$ 表示. 当 $k = 1$ 时, 我们便得恒等变换, 当 $k = 0$ 时, 便得零变换.

例 5 (导数)

在线性空间 $P[x]$ 或者 $P[x]_n$ 中, 求导数是一个线性变换. 这个变换通常用 $\mathscr{D}$ 代表, 即

$$\mathscr{D}(f(x)) = f'(x)$$

例 6 (积分)

定义在闭区间 $[a, b]$ 上的全体连续函数组成实数域上一线性空间, 以 $C(a, b)$ 代表. 在这个空间中, 变换

$$\mathscr{S}(f(x)) = \int_a^x f(t)\mathrm{d}t \qquad (\text{这里}\mathscr{S}\text{是花体}S)$$

是一个线性变换.

从定义推出线性变换的以下简单性质:

1 设 $\mathscr{A}$ 是线桂空间 V 的一个线性变换, 则

$$\mathscr{A}(\mathbf{0})=\mathbf{0}, \quad \mathscr{A}(-\alpha)=-\mathscr{A}(\alpha).$$

证明

$$\mathscr{A}(\mathbf{0})=\mathscr{A}(0\,\alpha)=0\,\mathscr{A}(\alpha)=\mathbf{0},$$

$$\mathscr{A}(-\alpha)=A((-1)\alpha)=-A(\alpha)=-A(\alpha).$$

2 线性变换保持线性组合与线性关系式不变.

换句话说, 如果 β 是 $\alpha_1, \alpha_2, \cdots, \alpha,$ 的线性组合:

$$\beta = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_r\alpha_r$$

那么 $\mathscr{A}(\beta)$ 是 $\mathscr{A}(\alpha_1), \mathscr{A}(\alpha_2), \cdots, \mathscr{A}(\alpha)$ 的线性组合:

$$\mathscr{A}(\beta) = k_1\mathscr{A}(\alpha_1) + k_2\mathscr{A}(\alpha_2) + \cdots + k_r\mathscr{A}(\alpha_r)$$

又如果 $\alpha_1, \alpha_2, \cdots, \alpha,$ 之间有一线性关系式

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_r\alpha_r = \mathbf{0}$$

那么它们的像之间也有同样的关系

$$k_1\mathscr{A}(\alpha_1) + k_2\mathscr{A}(\alpha_2) + \cdots + k_r\mathscr{A}(\alpha_r) = \mathbf{0}$$

3 线性变换把线性相关的向量组变成线性相关的向量组.

注意：它的逆是不对的.

线性变换可能把线性无关的向量组也变成线性相关的向量组.

例如零变换就是这样.

§2 线性变换的运算

一、线性变换的乘积

首先, 线性空间的线性变换作为映射的特殊情形当然可以定义乘法. 设 $\mathscr{A},\mathscr{B}$ 是线性空间 V 的两个线性变换, 定义它们的**乘积** $\mathscr{A}\mathscr{B}$ 为

$$(\mathscr{A}\mathscr{B})(\alpha)=\mathscr{A}(\mathscr{B}(\alpha))\quad(\alpha\in V)$$

- 线性变换的乘积也是线性变换.

$$\begin{aligned}(\mathscr{A}\mathscr{B})(\alpha+\beta)&=\mathscr{A}(\mathscr{B}(\alpha+\beta))=\mathscr{A}(\mathscr{B}(\alpha)+\mathscr{B}(\beta))\\&=\mathscr{A}(\mathscr{B}(\alpha))+\mathscr{A}(\mathscr{P}(\beta))\\&=(\mathscr{A}\mathscr{B})(\alpha)+(\mathscr{A}\mathscr{B})(\beta)\\(\mathscr{A}\mathscr{B})(k\alpha)&=\mathscr{A}(\mathscr{B}(k\alpha))=\mathscr{A}(k\mathscr{B}(\alpha))\\&=k\mathscr{A}(\mathscr{B}(\alpha))=k(\mathscr{A}\mathscr{B})(\alpha)\end{aligned}$$

这说明 $\mathscr{A}\mathscr{B}$ 是线性的.

- 既然一般映射的乘法适合结合律, 线性变换的乘法当然也适合结合律, 即

$$(\mathscr{A}\mathscr{B})\mathscr{C} = \mathscr{A}(\mathscr{B}\mathscr{C})$$

- 但线性变换的乘法一般是不可交换的. 例如, 在实数域 R 上的线性空间 $R[x]$ 中, 线性变换

$$\mathscr{D}(f(x)) = f'(x)$$
$$\mathscr{S}(f(x)) = \int_0^x f(t)\mathrm{d}t$$

的乘积 $\mathscr{D}\mathscr{S} = \mathscr{E}$, 但一般 $\mathscr{S}\mathscr{D} \neq \mathscr{E}$.

- 对于乘法, 单位变换 $\mathscr{E}$ 有特殊的地位. 对于任意线性变换. 都有

$$\mathscr{E}\mathscr{A} = \mathscr{A}\mathscr{E} = \mathscr{A}$$

二、线性变换的加法

其次, 对于线性变换还可以定义加法. 设 $\mathscr{A}, \mathscr{B}$ 是线性空间 V 的两个线性变换, 定义它们的和为

$$(\mathscr{A}+\mathscr{B})(\alpha)=\mathscr{A}(\alpha)+\mathscr{B}(\alpha) \quad (\alpha \in V)$$

- 线性变换的和还是线性变换.

$$\begin{aligned}
(\alpha+\mathscr{B})(\alpha+\beta) &= \mathscr{A}(\alpha+\beta)+\mathscr{B}(\alpha+\beta) \\
&= (\mathscr{A}(\alpha)+\mathscr{A}(\beta))+(\mathscr{B}(\alpha)+\mathscr{B}(\beta)) \\
&= (\mathscr{A}(\alpha)+\mathscr{B}(\alpha))+(\mathscr{A}(\beta)+\mathscr{B}(\beta)) \\
&= (\mathscr{A}+\mathscr{B})(\alpha)+(\mathscr{A}+\mathscr{B})(\beta) \\
(\mathscr{A}+\mathscr{B})(k\alpha) &= \mathscr{A}(k\alpha)+\mathscr{B}(k\alpha) \\
&= k\mathscr{A}(\alpha)+k\mathscr{B}(\alpha)=k(\mathscr{A}(\alpha)+\mathscr{P}(\alpha)) \\
&= k(\mathscr{A}+\mathscr{B})(\alpha)
\end{aligned}$$

- 线性变换的加法适合结合律与交换律, 即

$$\mathscr{A}+(\mathscr{B}+\mathscr{C})=(\mathscr{A}+\mathscr{B})+\mathscr{C}$$

$$\mathscr{A}+\mathscr{B}=\mathscr{B}+\mathscr{A}$$

- 对于加法, 零变换 $\mathscr{O}$ 有着特殊的地位. 它与所有线性变换 $\mathscr{A}$ 的和仍等于 $\mathscr{A}$:

$$\mathscr{A}+\mathscr{O}=\mathscr{A}$$

对于每个线性变换 $-\mathscr{A}$, 我们可以定义它的负变换:

$$(-\mathscr{A})(\alpha)=-\mathscr{A}(\alpha)\quad(\alpha\in V)$$

- 负变换 $-\mathscr{A}$ 也是线性的, 且

$$\mathscr{A}+(-\mathscr{A})=\mathscr{O}$$

- 线性变换的乘法对加法有左右分配律, 即

$$\mathscr{A}(\mathscr{B}+\mathscr{C})=\mathscr{A}\mathscr{B}+\mathscr{A}\mathscr{C}$$
$$(\mathscr{B}+\mathscr{C})\mathscr{A}=\mathscr{B}\mathscr{A}+\mathscr{C}\mathscr{A}$$

三、线性变换的数量乘法

利用线性变换的乘法，可以定义数域 P 中的数与线性变换的数置乘法为

$$(k\mathscr{A})(\alpha) = (\mathscr{K}\mathscr{A})(\alpha) = \mathscr{K}(\mathscr{A}(\alpha))$$

- $k\mathscr{A}$ 还是线性变换.
- 线性变换的数量乘法适合以下的规律：
  - 数乘结合律 $(k\ell)\mathscr{A} = k(\ell\mathscr{A})$
  - 数乘分配律 $(k+\ell)\mathscr{A} = k\mathscr{A} + \ell\mathscr{A}$
  - 数乘分配律 $k(\mathscr{A}+\mathscr{B}) = k\mathscr{A} + k\mathscr{B}$
  - 1 的数乘　$1\mathscr{A} = \mathscr{A}$
- 由加法与数量乘法的性质可知，线性空间 V 上全体线性变换.
  对于如上定义的加法与数量乘法，也构成数域 P 上一个线性空间.

四、线性变换的逆变换

V 的变换 $\mathscr{A}$ 称为可逆的, 如果存在 V 的变换 $\mathscr{B}$, 使

$$\mathscr{A}\mathscr{B} = \mathscr{B}\mathscr{A} = \mathscr{E}$$

这时, 变换 $\mathscr{B}$ 称为 $\mathscr{A}$ 的逆变换, 记为 $\mathscr{A}^{-1}$ 可逆线性变换 $\mathscr{A}$ 的逆变换.

- 可逆线性变换 $\mathscr{A}$ 的逆变换 $\mathscr{A}^{-1}$ 也是线性变换

$$\begin{aligned}
\mathscr{A}^{-1}(k\alpha + \ell\beta) &= \mathscr{A}^{-1}\left(k\left(\mathscr{A}\mathscr{A}^{-1}\right)(\alpha) + \ell\left(\mathscr{A}\mathscr{A}^{-1}\right)(\beta)\right) \\
&= \mathscr{A}^{-1}\left(\mathscr{A}\left(k\mathscr{A}^{-1}(\alpha)\right) + \mathscr{A}\left(\ell A^{-1}(\beta)\right)\right) \\
&= \mathscr{A}^{-1}\left(\mathscr{A}\left(k\mathscr{A}^{-1}(\alpha) + \ell\mathscr{A}^{-1}(\beta)\right)\right) \\
&= \left(\mathscr{A}^{-1}\mathscr{A}\right)\left(k\mathscr{A}^{-1}(\alpha) + \ell\mathscr{A}^{-1}(\beta)\right) \\
&= k\mathscr{A}^{-1}(\alpha) + \ell\mathscr{A}^{-1}(\beta)
\end{aligned}$$

五、线性变换的多项式

当 n 个 (n 是正整数) 线性变换相乘时，我们就可以用

$$\overbrace{\mathscr{A}\mathscr{A}\cdots\mathscr{A}}^{n}$$

来表示，称为 $\mathscr{A}$ 的 n 次幂，简单地记作 $\mathscr{A}^n$. 此外，作为定义，令

$$\mathscr{A}^0=\mathscr{E}.$$

根据线性变换幂的定义，可以推出指数法则：

$$\mathscr{A}^{m+n}=\mathscr{A}^m\cdot\mathscr{A}^n,(\mathscr{A}^m)^n=\mathscr{A}^{mn}\quad(m,n\geqslant 0)$$

当线性变换 $\mathscr{A}$ 可逆时，定义 $\mathscr{A}$ 的负整数幂为

$$\mathscr{A}^{-n}=\left(\mathscr{A}^{-1}\right)^n\quad(n\text{ 是正整数})$$

值得注意的是，线性变换乘积的指数法则不成立，即一般说来

$$(\mathscr{A}\mathscr{B})^n\neq\mathscr{A}^n\mathscr{B}^n$$

设

$$f(x) = a_m x^m + a_{m-1} x^{m-1} + \cdots + a_0$$

是 $P[x]$ 中一多项式，$\mathscr{A}$ 是 V 的一线性变换, 我们定义

$$f(\mathscr{A}) = a_m \mathscr{A}^m + a_{m-1} \mathscr{A}^{m-1} + \cdots + a_0 \mathscr{E}$$

于是, $f(\mathscr{A})$ 是一线性变换, 它称为线性变换 $\mathscr{A}$ 的多项式.

如果在 $P[x]$ 中

$$h(x) = f(x) + g(x), \quad p(x) = f(x)g(x)$$

那么

$$h(\mathscr{A}) = f(\mathscr{A}) + g(\mathscr{A}), \quad p(\mathscr{A}) = f(\mathscr{A})g(\mathscr{A})$$

特别地,

$$f(\mathscr{A})g(\mathscr{A}) = g(\mathscr{A})f(\mathscr{A})$$

即同一个线性变换的多项式的乘法是可交换的.

例 1

在线性空间 $P[\lambda]_n$ 中, 求微商是一个线性变换, 用 $\mathscr{D}$ 表示. 显然有

$$\mathscr{D}^n = \mathscr{O}$$

其次, 变数的平移

$$f(\lambda) \to f(\lambda + a) \quad (a \in P)$$

$$f(\lambda + a) = f(\lambda) + af'(\lambda) + \frac{a^2}{2!}f''(\lambda) + \cdots + \frac{a^{n-1}}{(n-1)!}f^{(n-1)}(\lambda)$$

因之 $\mathscr{S}_a$ 实质上是 $\mathscr{D}$ 的多项式:

$$\mathscr{S}_a = \mathscr{E} + a\mathscr{D} + \frac{a^2}{2!}\mathscr{D}^2 + \cdots + \frac{a^{n-1}}{(n-1)!}\mathscr{D}^{n-1}$$

§3 线性变换的矩阵

设 V 是数域 P 上 n 维线性空间 $,\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是 V 的一组基, 现在我们来建立线性变换与矩阵的关系.

- 空间 V 中任一向量 ξ 可以由基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 线性表出, 即

$$\xi = x_1\varepsilon_1 + x_2\varepsilon_2 + \cdots + x_n\varepsilon_n$$

  其中系数是唯一确定的, 它们就是 ξ 在这组基下的坐标.

- 由于线性变换保持线性关系不变, 因而在 ξ 的像 $\mathscr{A}\xi$ 与基的像 $\mathscr{A}\varepsilon_1,\mathscr{A}\varepsilon_2,\cdots,\mathscr{A}\varepsilon_n$ 之间也必然有相同的关系

$$\begin{aligned}\mathscr{A}\xi &= \mathscr{A}\left(x_1\varepsilon_1 + x_2\varepsilon_2 + \cdots + x_n\varepsilon_n\right)\\ &= x_1\mathscr{A}\left(\varepsilon_1\right) + x_2\mathscr{A}\left(\varepsilon_2\right) + \cdots + x_n\mathscr{A}\left(\varepsilon_n\right)\end{aligned}$$

- 上式表明, 如果我们知道了基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 的像, 那么线性空间中任意一个向量 ξ 的像也就知道了

1. 设 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是线性空间 V 的一组基. 如果线性变换 $\mathscr{A}$ 与 $\mathscr{B}$ 在这组基上的作用相同, 即 $\mathscr{A}\varepsilon_i = \mathscr{B}\varepsilon_i, \quad i = 1, 2, \cdots, n$ 那么 $\mathscr{A} = \mathscr{B}$.

证明 对 V 中任意向量 $\xi = x_1\varepsilon_1 + x_2\varepsilon_2 + \cdots + x_n\varepsilon_n$, 有

$$\begin{aligned}
\mathscr{A}\xi &= \mathscr{A}\left(x_1\varepsilon_1 + x_2\varepsilon_2 + \cdots + x_n\varepsilon_n\right) \\
&= x_1\mathscr{A}\varepsilon_1 + x_2\mathscr{A}\varepsilon_2 + \cdots + x_n\mathscr{A}\varepsilon_n \\
&= x_1\mathscr{B}\varepsilon_1 + x_2\mathscr{B}\varepsilon_2 + \cdots + x_n\mathscr{B}\varepsilon_n \\
&= \mathscr{B}\left(x_1\varepsilon_1 + x_2\varepsilon_2 + \cdots + x_n\varepsilon_n\right) \\
&= \mathscr{B}\xi.
\end{aligned}$$

因此, $\mathscr{A} = \mathscr{B}$. ∎

结论 1 的意义就是, 一个线性变换完全被它在一组基上的作用所决定.

下面我们进一步指出, 基向量的像却完全可以是任意的, 也就是说,

2. 设 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是线性空间 V 的一组基. 对任意一组向量 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 一定有一个线性变换 $\mathscr{A}$ 使

$$\mathscr{A}\varepsilon_i = \alpha_i, \quad i = 1, 2, \cdots, n$$

证明　我们来作出所要的线性变换. 设

$$\xi = \sum_{i=1}^{n} x_i \varepsilon_i$$

是线性空间 V 的任意一个向量, 我们定义 V 的变换 $\mathscr{A}$ 为

$$\mathscr{A}\xi = \sum_{i=1}^{n} x_i \alpha_i$$

下面来证明变换 $\mathscr{A}$ 是线性的. 在 V 中任取两个向量,

$$\beta=\sum_{i=1}^{n} b_i\,\varepsilon_i,\quad \gamma=\sum_{i=1}^{n} c_i\,\varepsilon_i$$

有

$$\beta+\gamma=\sum_{i=1}^{n}(b_i+c_i)\,\varepsilon_i,\qquad k\beta=\sum_{i=1}^{n} kb_i\,\varepsilon_i,\quad k\in P$$

于是,

$$\begin{aligned}
\mathscr{A}(\beta+\gamma)&=\sum_{i=1}^{n}(b_i+c_i)\,\alpha_i\\
&=\sum_{i=1}^{n} b_i\alpha_i+\sum_{i=1}^{n} c_i\alpha_i=\mathscr{A}\beta+\mathscr{A}\gamma\\
\mathscr{A}(k\beta)&=\sum_{i=1}^{n}(kb_i)\alpha_i=k\sum_{i=1}^{n} b_i\alpha_i=k\beta
\end{aligned}$$

因此, $\mathscr{A}$ 是线性变换.

再来证 $\mathscr{A}$ 满足 $\mathscr{A}\varepsilon_i = \alpha_i, \quad i = 1, 2, \cdots, n$. 因为

$$\varepsilon_i = 0\varepsilon_1 + \cdots + 0\varepsilon_{i-1} + 1\varepsilon_i + 0\varepsilon_{i+1} + \cdots + 0\varepsilon_n, \quad i = 1, 2, \cdots, n$$

所以

$$\mathscr{A}\varepsilon_i = 0\alpha_1 + \cdots + 0\alpha_{i-1} + 1\alpha_i + 0\alpha_{i+1} + \cdots + 0\alpha_n = \alpha_i \quad i = 1, 2, \cdots, n$$

证毕. ∎

综合以上两点, 得

定理 1

设 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是线性空间 V 的一组基, $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是 V 中任意 n 个向量, 存在唯一的线性变换 $\mathscr{A}$ 使

$$\mathcal{A}\varepsilon_i = \alpha_i, \quad i = 1, 2, \cdots, n.$$

二、线性变换在一组基下的矩阵

有了以上的讨论, 我们就可以来建立线性变换与矩阵的联系.

- 设 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是数域 P 上 n 维线性空间 V 的一组基, $\mathscr{A}$ 是 V 中的一个线性变换, 基向量的像可以被基线性表出:

$$\begin{cases} \mathscr{A}\varepsilon_1 & = a_{11}\varepsilon_1 + a_{21}\varepsilon_2 + \cdots + a_{n1}\varepsilon_n \\ \mathscr{A}\varepsilon_2 & = a_{12}\varepsilon_1 + a_{22}\varepsilon_2 + \cdots + a_{n2}\varepsilon_n \\ & \cdots\cdots \\ \mathscr{A}\varepsilon_n & = a_{1n}\varepsilon_1 + a_{2n}\varepsilon_2 + \cdots + a_{nn}\varepsilon_n \end{cases}$$

- 写成矩阵的形式

$$(\mathscr{A}\varepsilon_1, \mathscr{A}\varepsilon_2, \ldots, \mathscr{A}\varepsilon_n) = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n) \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

定义

矩阵

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

称为线性变换 $\mathscr{A}$ 在基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 下的矩阵.

记 $\mathscr{A}(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n) := (\mathscr{A}\varepsilon_1, \mathscr{A}\varepsilon_2, \ldots, \mathscr{A}\varepsilon_n)$. 于是,

$$\mathscr{A}(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n) = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n) A. \tag{1}$$

例 1

设 V 是数域 P 上 n 维线性空间, 则

- 恒等变换在任一组基下的矩阵是________
- 零变换在任一组基下的矩阵是________
- 由 k 决定的数乘变换在任一组基下的矩阵是________

例 1

设 V 是数域 P 上 n 维线性空间, 则

- 恒等变换在任一组基下的矩阵是________
- 零变换在任一组基下的矩阵是________
- 由 k 决定的数乘变换在任一组基下的矩阵是________

- n 阶单位矩阵 E
- n 阶零矩阵 O
- n 阶数量矩阵 kE

例 2

设 P^3 的线性变换 $\mathscr{A}$ 为 $\mathscr{A}(x_1, x_2, x_3) = (x_1, x_2, x_1 + x_2)$ 取一组基 $\varepsilon_1 = (1,0,0), \varepsilon_2 = (0,1,0), \varepsilon_3 = (0,0,1)$ 则

$$\begin{aligned}
\mathscr{A}\varepsilon_1 &= \varepsilon_1 \qquad\qquad + \varepsilon_3 \\
\mathscr{A}\varepsilon_2 &= \qquad\quad \varepsilon_2 \quad + \varepsilon_3 \\
\mathscr{A}\varepsilon_3 &= 0
\end{aligned}$$

所以, $\mathscr{A}$ 在基 $\varepsilon_1, \varepsilon_2, \varepsilon_3$ 的矩阵为

例 2

设 P^3 的线性变换 $\mathscr{A}$ 为 $\mathscr{A}(x_1, x_2, x_3) = (x_1, x_2, x_1 + x_2)$ 取一组基 $\varepsilon_1 = (1,0,0), \varepsilon_2 = (0,1,0), \varepsilon_3 = (0,0,1)$ 则

$$\begin{aligned}
\mathscr{A}\varepsilon_1 &= \varepsilon_1 \qquad\quad + \varepsilon_3 \\
\mathscr{A}\varepsilon_2 &= \qquad\ \varepsilon_2 \ + \varepsilon_3 \\
\mathscr{A}\varepsilon_3 &= 0
\end{aligned}$$

所以, $\mathscr{A}$ 在基 $\varepsilon_1, \varepsilon_2, \varepsilon_3$ 的矩阵为

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

例 3

在空间 $P[x]_n$ 中, 线性变换

$$\mathscr{D}f(x) = f'(x)$$

在基 $1, x, \frac{x^2}{2!}, \dots, \frac{x^{n-1}}{(n-1)!}$ 下的矩阵是

例 3

在空间 $P[x]_n$ 中, 线性变换

$$\mathscr{D}f(x) = f'(x)$$

在基 $1, x, \frac{x^2}{2!}, \ldots, \frac{x^{n-1}}{(n-1)!}$ 下的矩阵是

$$D = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

例 4

设 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_m$ 是 $n(n > m)$ 维线性空间 V 的子空间 W 的一组基, 把它扩充为 V 的一组基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$. 指定线性变换 $\mathscr{A}$ 如下

$$\begin{cases} \mathscr{A}\varepsilon_i = \varepsilon_i, \text{ 当} i = 1, 2, \cdots, m \\ \mathscr{A}\varepsilon_i = 0, \text{ 当} i = m+1, \cdots, n \end{cases}$$

如此确定的线性变换 $\mathscr{A}$ 称为对子空间 W 的一个投影. 于是

$$\mathscr{A}^2 = \mathscr{A}$$

投影 $\mathscr{A}$ 在基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 下的矩阵是

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

- 在取定一组基之后，我们就建立了由数域 P 上的 n 维线性空间 V 的线性变换到数域 P 上的 $n \times n$ 矩阵的一个映射.
- 前面的结论 1 说明这个映射是单射，结论 2 说明这个映射是满射. 换句话说，我们在这二者之间建立了一个双射.
- 这个对应的重要性表现在它保持运算，即有

定理 2

设 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是数域 P 上 n 维线性空间 V 的一组基，在这组基下，每个线性变换按公式(1)对应一个 $n \times n$ 矩阵. 这个对应具有以下的性质:

1. 线性变换的和对应于矩阵的和,
2. 线性变换的乘积对应于矩阵的乘积,
3. 线性变换的数量乘积对应于矩阵的数量乘积,
4. 可逆的线性变换与可逆矩阵对应，且逆变换对应于逆矩阵.

证明　设 $\mathscr{A}, \mathscr{B}$ 是两个线性变换, 它们在基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 下的矩阵分别是 A, B, 即

$$\mathscr{A}\left(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n\right) = \left(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n\right) A$$
$$\mathscr{B}\left(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n\right) = \left(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n\right) B$$

1) 由

$$\begin{aligned}
&(\mathscr{A} + \mathscr{B})\left(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n\right) \\
=&\mathscr{A}\left(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n\right) + \mathscr{P}\left(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n\right) \\
=&\left(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n\right) A + \left(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n\right) B \\
=&\left(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n\right)(A + B)
\end{aligned}$$

可知, 在基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 下, 线性变换 $\mathscr{A} + \mathscr{B}$ 的矩阵是 $A + B$

2) 相仿地,

$$\begin{aligned}
&(\mathscr{A}\mathscr{B})\left(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n\right) = \mathscr{A}\left(\mathscr{B}\left(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n\right)\right) \\
=&\mathscr{A}\left(\left(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n\right) B\right) = \left(\mathscr{A}\left(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n\right)\right) B \\
=&\left(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n\right) AB
\end{aligned}$$

3) 因为

$$(k\varepsilon_1, k\varepsilon_2, \cdots, k\varepsilon_n) = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)\, kE$$

所以数乘变换从在任何一组基下都对应于数量矩阵 kE. 由此可知, 数量乘积 $k\mathscr{A}$ 对应于矩阵的数量乘积 kA

4) 单位变换 $\mathscr{E}$ 对应于单位矩阵, 因之等式

$$\mathscr{A}\mathscr{B} = \mathscr{B}\mathscr{A} = \mathscr{E}$$

与等式

$$AB = BA = E$$

相对应, 从而可逆线性变换与可逆矩阵对应, 而且逆变换与逆矩阵对应. ∎

用 $L(V)$ 或者 $End(V)$ 表示数域 P 上 n 维线性空间 V 的全部线性变换组成的集合.

性质

$L(V)$ 对于线性变换的加法与数量乘法构成 P 上一个线性空间, 与数域 P 上 n 级方阵构成的线性空间 $P^{n\times n}$ 同构.

定理 3

设线性变换 $\mathscr{A}$ 在基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$, 下的矩阵是 A, ξ 在基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 下的坐标 $(x_1, x_2, \cdots, x_n)$, 则向量 $\mathscr{A}\xi$ 在基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 下的坐标 $(y_1, y_2, \cdots, y_n)$ 可以按公式

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

计算.

证明 由假设

$$\xi = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

于是

$$\begin{aligned} \mathscr{A}\xi &= (\mathscr{A}\varepsilon_1, \mathscr{A}\varepsilon_2, \cdots, \mathscr{A}\varepsilon_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \\ &= (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)\, A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \end{aligned}$$

另一方面, 由假设

$$\mathscr{A}\xi = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n) \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

由于 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 线性无关, 所以

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

∎

四、线性变换在不同基下的矩阵

- 线性变换的矩阵是与空间中一组基联系在一起的.
- 一般说来, 随着基的改变, 同一个线性变换就有不同的矩阵.
- 为了利用矩阵来研究线性变换, 我们有必要弄清楚线性变换的矩阵是如何随着基的改变而改变的.

定理 4

设线性空间 V 中线性变换 $\mathscr{A}$ 在两组基

$$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n \tag{I}$$

$$\eta_1, \eta_2, \cdots, \eta_n \tag{II}$$

下的矩阵分别为 A 和 B, 从基(I)到(II)的过渡矩阵是 X, 于是 $B = X^{-1}AX$

证明 已知

$$(\mathscr{A}\varepsilon_1, \mathscr{A}\varepsilon_2, \cdots, \mathscr{A}\varepsilon_n) = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n) A$$
$$(\mathscr{A}\eta_1, \mathscr{A}\eta_2, \cdots, \mathscr{A}\eta_n) = (\eta_1, \eta_2, \cdots, \eta_n) B$$
$$(\eta_1, \eta_2, \cdots, \eta_n) = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n) X$$

于是

$$\begin{aligned}
&(\mathscr{A}\eta_1, \mathscr{A}\eta_2, \ldots, \mathscr{A}\eta_n) = \mathscr{A}(\eta_1, \eta_2, \cdots, \eta_n) \\
= &\mathscr{A}((\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n) X) = (\mathscr{A}(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)) X \\
= &(\mathscr{A}\varepsilon_1, \mathscr{A}\varepsilon_2, \ldots, \mathscr{A}\varepsilon_n) X = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n) AX \\
= &(\eta_1, \eta_2, \ldots, \eta_n) X^{-1} AX
\end{aligned}$$

由此即得

$$B = X^{-1}AX.$$

∎

定理 4 告诉我们, 同一个线性变换 $\mathscr{A}$ 在不同基下的矩阵之间的关系. 这个基本关系在以后的讨论中是重要的.

现在, 我们对于矩阵引进相应的定义.

定义

设 A,B 为数域 P 上两个 n 级矩阵, 如果可以找到数域 P 上的 n 级可逆矩阵 X , 使得 $B = X^{-1}AX$, 就说 A 相似手 B, 记作 $A \sim B$

相似是矩阵之间的一种关系, 这种关系具有下面三个性质:

1. 反身性: $A \sim A$ 这是因为 $A = E^{-1}AE$
2. 对称性: 如果 $A \sim B$, 那么 $B \sim A$.
   如果 $A \sim B$, 那么有 X 使 $B = X^{-1}AX$.
   令 $Y = X^{-1}$, 就有 $A = XBX^{-1} = Y^{-1}BY$, 所以 $B \sim A$
3. 传递性: 如果 $A \sim B, B \sim C$, 那么 $A \sim C$.
   已知有 X, Y 使 $B = X^{-1}AX, C = Y^{-1}BY$.
   令 $Z = XY$, 就有 $C = Y^{-1}X^{-1}AXY = Z^{-1}AZ$.

因之有了矩阵相似的概念之后，定理 4 可以补充成：

定理 5

- 线性变换在不同基下所对应的矩阵是相似的；反过来，
- 如果两个矩阵相似，那么它们可以看作同一个线性变换在两组基下所对应的矩阵.

证明 前一部分已经为定理 4 证明.

现在证明后一部分. 设级矩阵 A 和 B 相似. A 可以看做是 n 维线性空间 V 中一个线性变换 $\mathscr{A}$ 在基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 下的矩阵. 因为 $B = X^{-1}AX$, 令

$$(\eta_1, \eta_2, \cdots, \eta_n) = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n) X$$

因为 X 可逆，所以 $\eta_1, \eta_2, \cdots, \eta_n$ 也是一组基，$\mathscr{A}$ 在这组基下的矩阵就是 B. ∎

矩阵的相似对于运算有下面的性质.

- 如果 $B_1 = X^{-1}A_1X, B_2 = X^{-1}A_2X$, 那么

$$B_1 + B_2 = X^{-1}(A_1 + A_2)X,$$
$$B_1B_2 = X^{-1}(A_1A_2)X.$$

- 如果 $B = X^{-1}AX$, 且 $f(x)$ 是数域 P 上一多项式, 那么

$$f(B) = X^{-1}f(A)X.$$

利用矩阵相似的这个性质可以简化矩阵的计算.

例 5

设 V 是数域 P 上一个二维线性空间, $\varepsilon_1, \varepsilon_2$ 是一组基, 线性变换 $\mathscr{A}$ 在 $\varepsilon_1, \varepsilon_2$ 下的矩阵是

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

现在来计算 $\mathscr{A}$ 在 V 的另一组基 η_1, η_2 下的矩阵, 这里

$$(\eta_1, \eta_2) = (\varepsilon_1, \varepsilon_2) \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$$

由定理 4, $\mathscr{A}$ 在 η_1, η_2 下的矩阵为

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

归纳可知,

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^k = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$$

再利用上面得到的关系

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

即

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

我们可以得到

$$\begin{aligned}
\begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix}^k &= \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^k \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}^{-1} \\
&= \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \\
&= \begin{pmatrix} k+1 & k \\ -k & -k+1 \end{pmatrix}.
\end{aligned}$$

例 6

取 P^3 的线性变换 $\sigma(a,b,c)=(2a-b,b+c,a)$

1. 求 σ 在基 $\varepsilon_1=(1,0,0),\varepsilon_2=(0,1,0),\varepsilon_3=(0,0,1)$ 下的矩阵;
2. 求 σ 在基 $\eta_1=(1,0,0),\eta_2=(1,1,0),\eta_3=(1,1,1)$ 下的矩阵;
3. 求向量 $\alpha=(1,2,3)$ 的像 $\sigma\alpha$ 分别在基 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 和 η_1,η_2,η_3 下的坐标.

解

(1) $\sigma\varepsilon_1=\sigma(1,0,0)=(2,0,1)=(\varepsilon_1,\varepsilon_2,\varepsilon_3)\begin{pmatrix}2\\0\\1\end{pmatrix}$,

$$\sigma\varepsilon_2=(\varepsilon_1,\varepsilon_2,\varepsilon_3)\begin{pmatrix}-1\\1\\0\end{pmatrix},\quad \sigma\varepsilon_3=(\varepsilon_1,\varepsilon_2,\varepsilon_3)\begin{pmatrix}0\\1\\0\end{pmatrix},$$

则

$$\sigma\left(\varepsilon_1,\varepsilon_2,\varepsilon_3\right)=\left(\varepsilon_1,\varepsilon_2,\varepsilon_3\right)A,\quad \text{其中}A=\begin{pmatrix}2&-1&0\\0&1&1\\1&0&0\end{pmatrix}$$

(2) $(\eta_1,\eta_2,\eta_3)=(\varepsilon_1,\varepsilon_2,\varepsilon_3)\,C,\quad$ 其中 $C=\begin{pmatrix}1&1&1\\0&1&1\\0&0&1\end{pmatrix}$, 则 σ 在基 η_1,η_2,η_3

下的矩阵为 $C^{-1}AC=\begin{pmatrix}2&0&-1\\-1&0&1\\1&1&1\end{pmatrix}$

(3) $\alpha = (\varepsilon_1, \varepsilon_2, \varepsilon_3) \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$,

$$\sigma\alpha = \sigma(\varepsilon_1, \varepsilon_2, \varepsilon_3) \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = (\varepsilon_1, \varepsilon_2, \varepsilon_3) A \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = (\varepsilon_1, \varepsilon_2, \varepsilon_3) \begin{pmatrix} 0 \\ 5 \\ 1 \end{pmatrix}$$

$$\sigma\alpha = (\eta_1, \eta_2, \eta_3) C^{-1} \begin{pmatrix} 0 \\ 5 \\ 1 \end{pmatrix} = (\eta_1, \eta_2, \eta_3) \begin{pmatrix} -5 \\ 4 \\ 1 \end{pmatrix}$$

§4 特征值与特征向量

- 我们知道, 在有限维线性空间中, 取了一组基之后, 线性变换就可以用矩阵来表示.
- 为了利用矩阵来研究线性变换, 对于每个给定的线性变换, 我们希望能找到一组基使得它的矩阵具有最简单的形式.
- 从现在开始, 我们主要地就来讨论, 在适当的选择基之后, 一个线性变换的矩阵可以化成什么样的简单形式.

设 A 是 n 阶方阵, P 为 n 阶可逆阵,

$$P^{-1}AP = \Lambda$$

其中 $\Lambda = \mathrm{diag}\left\{\lambda_1, \lambda_2, \cdots, \lambda_n\right\}$
令 $P = (p_1, p_2, \cdots, p_n)$

$$\begin{aligned}
&P^{-1}AP = \Lambda \\
\Leftrightarrow &A(p_1, p_2, \cdots, p_n) = (p_1, p_2, \cdots, p_n)\,\mathrm{diag}\left\{\lambda_1, \lambda_2, \cdots, \lambda_n\right\} \\
\Leftrightarrow &(Ap_1, Ap_2, \cdots, Ap_n) = (\lambda_1 p_1, \lambda_2 p_2, \cdots, \lambda_n p_n) \\
\Leftrightarrow &Ap_i = \lambda_i p_i, i = 1, 2, \cdots, n
\end{aligned}$$

此过程的逆推在最后一步要求矩阵 P 是可逆的。

为了这个目的, 先介绍特征值和特征向量的概念, 它们对于线性变换的研究具有基本的重要性.

定义

设 $\mathscr{A}$ 是数域 P 上线性空间 V 的一个线性变换, 如果对于数域 P 中一数 λ_0, 存在一个非零向量 ξ, 使得

$$\mathscr{A}\xi = \lambda_0\xi$$

个特征向量.

从几何上来看, 特征向量的方向经过线性变换后, 保持在同一条直线上, 这时或者方向不变 ($\lambda_0 > 0$) , 或者方向相反 $(\lambda_0 < 0)$, 至于 $\lambda_0 = 0$ 时, 特征向量就被线性变换变成 $\mathbf{0}$

如果 ξ 是线性变换 $\mathscr{A}$ 的属于特征值 λ_0 的特征向量, 那么的任何一个非零倍数 $k\xi$ 也是 $\mathscr{A}$ 的属于 λ_0 的特征向量. 因为 $\mathscr{A}\xi = \lambda_0\xi$, 所以

$$\mathscr{A}(k\xi) = \lambda_0(k\xi)$$

这说明特征向量不是被特征值所唯一确定定的. 相反, 特征值却是被特征向量所唯一确定的, 因为, 一个特征向量只能属于一个特征值.

现在来给出寻找特征值和特征向量的方法, 设 V 是数域 P 上 n 维线性空间, $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是它的一组基, 线性变换 $\mathscr{A}$ 在这组基下的矩阵是 A. 设 λ 是特征值, 它的一个特征向量 ξ 在 $\varepsilon_1\ \varepsilon_2, \cdots, \varepsilon_n$ 下的坐标是 $x_{01}, x_{02}, \cdots, x_{0n}$, 则 $\mathscr{A}\xi$ 的坐标是

$$A\begin{pmatrix} x_{01} \\ x_{02} \\ \vdots \\ x_{0n} \end{pmatrix}$$

$\lambda_0\xi$ 的坐标是

$$\lambda_0\begin{pmatrix} x_{01} \\ x_{02} \\ \vdots \\ x_{0n} \end{pmatrix}$$

因此, $\mathscr{A}\xi = \lambda_0\xi$, 相当于坐标之间的等式

$$A \begin{pmatrix} x_{01} \\ x_{02} \\ \vdots \\ x_{0n} \end{pmatrix} = \lambda_0 \begin{pmatrix} x_{01} \\ x_{02} \\ \vdots \\ x_{0n} \end{pmatrix}$$

或

$$(\lambda_0 E - A) \begin{pmatrix} x_{01} \\ x_{02} \\ \vdots \\ x_{0n} \end{pmatrix} = \mathbf{0} \qquad (*)$$

这说明特征向量 ξ 的坐标 $(x_{01}, x_{02}, \cdots, x_{0n})$ 满足齐次方程组即

$$\begin{cases} (\lambda_0 - a_{11})x_1 & -a_{12}x_2 & - & \cdots & -a_{1n}x_n & = 0 \\ -a_{21}x_1 & +(\lambda_0 - a_{22})x_2 & - & \cdots & -a_{2n}x_n & = 0 \\ \cdots & \cdots & & \cdots & \cdots & \\ -a_{n1}x_1 & -a_{n2}x_2 & - & \cdots & +(\lambda_0 - a_{nn})x_n & = 0 \end{cases}$$

由于 $\xi \neq \mathbf{0}$, 所以它的坐标 $x_{01}, x_{02}, \cdots, x_{0n}$ 不全为零, 即齐次方程组有非零解.

我们知道, 齐次线性方程组 $(\lambda E - A)X = \mathbf{0}$ 有非零解的充分必要条件是它的系数行列式为零, 即

$$|\lambda_0 E - A| = \begin{vmatrix} \lambda_0 - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda_0 - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda_0 - a_{nn} \end{vmatrix} = 0$$

我们引入以下的定义.

定义

$\lambda E - A$ 的行列式

$$|\lambda E - A| = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix}$$

称为 A 的特征多项式, 这是数域 P 上的一个 n 次多项式.

- 上面的分析说明, 如果 λ_0 是线性变换 $\mathscr{A}$ 的特征值, 那么 λ_0 一定是矩阵 A 的特征多项式的一个根;
- 反过来, 如果 λ_0 是知阵 A 的特征多项式在数域 P 中的一个根, 即 $|\lambda_0 E - A| = 0$, 那么齐线性方程组 $(\lambda E - A)X = \mathbf{0}$ 就有非零解.
- 这时, 如果 $(x_{01}; x_{02}, \cdots, x_{0n})$ 是方程组 $(\lambda E - A)X = \mathbf{0}$ 的一个非零解, 那么非零向量

$$\xi = x_{01}\varepsilon_1 + x_{02}\varepsilon_2 + \cdots + x_{0n}\varepsilon_n$$

λ_0 的一个特征向量.

求一个线性变换 $\mathscr{A}$ 的特征值与特征向量的方法可以分成以下几步:

1. 在线性空间 V 中取一组基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$, 写出 $\mathscr{A}$ 在这组基下的矩阵 A;
2. 求出 $\mathscr{A}$ 的特征多项式 $|\lambda E - A|$ 在数域 P 中全部的根, 它们也就是线性变换 $\mathscr{A}$ 的全部特征值.
3. 把所求得的特征值逐个地代入方程组 $(\lambda E - A)X = \mathbf{0}$, 对于每一个特征值, 解方程组 $(\lambda E - A)X = \mathbf{0}$, 求出一组基础解系, 它们就是属于这个特征值的几个线性无关的特征向量在基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 下的坐标, 这样, 我们也就求出了属于每个特征值的全部线性无关的特征向量.

矩阵的特征值与特征向量

定义

- 矩阵 A 的特征多项式的根称为 A 的特征值,
- 相应的线性方程组 $(\lambda E - A)X = \mathbf{0}$ 的解称为 A 的属于这个特征值的特征向量.

例 1

在 n 维线性空间中, 数乘变换 $\mathscr{K}$ 在任意一组基下的矩阵都且 kE, 它的特征多项式是

$$|\lambda E - kE| = (\lambda - k)^n$$

因此, 数乘变换从的特征值只有 k. 由定义可知, 每个非零向量都是属于数乘变换 $\mathscr{K}$ 的特征向量.

例 2

设 $A=\begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix}$, 求 A 的特征值与对应的特征向量.

解 1 由矩阵 A 的特征方程，求出特征值。

$$|\lambda E-A|=\begin{vmatrix} \lambda-1 & -2 & -2 \\ -2 & \lambda-1 & -2 \\ -2 & -2 & \lambda-1 \end{vmatrix}=(\lambda+1)^2(\lambda-5),$$

得

$$\lambda_1=\lambda_2=-1, \lambda_3=5.$$

2、把每个特征值 λ 代入线性方程组 $(A-\lambda E)X=0$ 求出基础解系。

- 再求属于 $\lambda_1=\lambda_2=-1$ 的特征向量, 即求方程组 $(-E-A)X=0$ 的解.

$$\begin{cases} -2x_1-2x_2-2x_3=0 \\ -2x_1-2x_2-2x_3=0, \end{cases}$$

即

$$x_1+x_2+x_3=0,$$

得基础解系为 $X_1=(1,-1,0)'$, $X_2=(1,0,-1)'$,
X_1,X_2 就是属于 $\lambda_1=-1$ 的两个线性无关的特征向量,
属于 $\lambda_1=-1$ 的全部特征向量为 $k_1X_1+k_2X_2$, 其中 k_1,k_2 不全为零.

- 最后求属于 $\lambda_3 = 5$ 的特征向量, 即求方程组 $(5E - A)X = 0$ 的解.

$$\begin{cases} 4x_1 - 2x_2 - 2x_3 = 0 \\ -2x_1 + 4x_2 - 2x_3 = 0 \\ -2x_1 - 2x_2 + 4x_3 = 0 \end{cases},$$

  基础解系为 $X_3 = (1, 1, 1)'$

  X_3 就是属于 $\lambda_3 = 5$ 的线性无关的特征向量,

  属于 $\lambda_3 = 5$ 的全部特征向量为 k_3X_3, 其中 $k_3 \neq 0$.

注

- 审题时注意：题目要的求是矩阵的特征值与特征向量，还是线性变换的特征值与特征向量.
- 两者的特征值都是一样的，但线性变换的的特征向量是 V 中的向量，不是坐标；而矩阵的特征向量是 $\mathbb{R}^n$ 中的向量.

例 3

求矩阵 $A = \begin{pmatrix} -1 & 1 & 0 \\ -4 & 3 & 0 \\ 1 & 0 & 2 \end{pmatrix}$ 的特征值和特征向量.

解 1 由矩阵 A 的特征多项式，求出特征值。

$$|A - \lambda E| = \begin{vmatrix} -1-\lambda & 1 & 0 \\ -4 & 3-\lambda & 0 \\ 1 & 0 & 2-\lambda \end{vmatrix} = (2-\lambda)\begin{vmatrix} -1-\lambda & 1 \\ -4 & 3-\lambda \end{vmatrix}$$

$$= (2-\lambda)(\lambda-1)^2 = 0$$

特征值为 $\lambda = 2, 1.$

2 把每个特征值 λ 代入线性方程组 $(A-\lambda E)X=\mathbf{0}$ 求出基础解系。

- 当 $\lambda=2$ 时，解线性方程组 $(A-2E)X=\mathbf{0}$

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

解得

$$\begin{cases} x_1 = 0 \\ x_2 = 0 \end{cases}$$

得基础解系

$$p_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

- 当 $\lambda = 1$ 时，解线性方程组 $(A-E)X = \mathbf{0}$

$$(A-E) = \begin{pmatrix} -2 & 1 & 0 \\ -4 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

解得

$$\begin{cases} x_1 + x_3 = 0 \\ x_2 + 2x_3 = 0 \end{cases}$$

得基础解系

$$p_2 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}.$$

例 4

在空间 $P[x]_n$ 中, 线性变换

$$\mathscr{D}f(x) = f'(x)$$

在基 $1, x, \frac{x^2}{2!}, \ldots, \frac{x^{n-1}}{(n-1)!}$ 下的矩阵是

$$D = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

$\mathscr{D}$ 的特征多项式是 $|\lambda E - D| = \begin{vmatrix} \lambda & -1 & 0 & \cdots & 0 \\ 0 & \lambda & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & -1 \\ 0 & 0 & 0 & \cdots & \lambda \end{vmatrix} = \lambda^n$

$\mathscr{D}$ 的特征值只有 0.

通过解相应的齐次线性方程组知道, 属于特征值 0 的线性无关的特征向量只能是任一非零常数.

例 5

平面上全体向量构成实数域上一个二维线性空间, 第一节例 3 中旋转变换 $\mathscr{I}_\theta$ 在直角坐标系下的矩阵为

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

它的特征多项式为 $\begin{vmatrix} \lambda-\cos\theta & \sin\theta \\ -\sin\theta & \lambda-\cos\theta \end{vmatrix} = \lambda^2 - 2\lambda\cos\theta + 1$

当 $\theta \neq k\pi$ 时, 这个多项式没有实根.

因之, 当 $\theta \neq k\pi$ 时, $\mathscr{I}_\theta$ 没有特征值.

- 属于同一特征值 λ_0 的特征向量全体连同零向量构成 V 的一个子空间 V_{λ_0}, 称其为特征子空间. 用集合记号可号为

$$V_{\lambda_0} = \{\alpha | \mathscr{A}\alpha = \lambda_0\alpha, \alpha \in V\}.$$

它的维数等于属于同一特征值 λ_0 的线性无关特征向量的最大个数, 称为特征值 λ_0 的几何重数.

特征多项式的系数

前两项

$$
|\lambda E - A| = \begin{vmatrix} \lambda - a_{11} & -a_{12} & & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{21} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix}
$$

- 展开式中有一项是主对角线上元素的连乘积

$$
(\lambda - a_{11})(\lambda - a_{22}) \cdots (\lambda - a_{nn}).
$$

- 展开式中其余各项, 至多包含 $n-2$ 个主对角线上的元素, 它对 λ 的次数最多是 $n-2$.

因此, 特征多项式中含 λ 的 n 次与 $n-1$ 次的项只能出现在主对角线上元素的连乘积中, 它们是

$$\lambda^n - (a_{11} + a_2 + \cdots + a_{nn})\,\lambda^{n-1}$$

在 $|\lambda E - A|$ 中, 令 $\lambda = 0$, 得到常数项系数为

$$|-A| = (-1)^n |A|$$

性质

- 如果只写出特征多项式的前两项和常数项, 就有

$$|\lambda E - A| = \lambda^n - (a_{11} + a_{22} + \cdots + a_{nn})\lambda^{n-1} + \cdots + (-1)^n |A|.$$

- 由多项式的根与系数的关系可知

$$\sum_{i=1}^{n} \lambda_i = tr(A), \quad \prod_{i=1}^{n} \lambda_i = |A|,$$

其中 $tr(A) := \sum_{i=1}^{n} a_{ii}$ 称为矩阵 A 的迹.

对于相似矩阵我们有

定理 6

相似的矩阵有相同的特征多项式.

证明 设 $A \sim B$, 即有可逆矩阵 X, 使 $B = X^{-1}AX$. 于是

$$\begin{aligned}|\lambda E - B| &= \left|\lambda E - X^{-1}AX\right| = \left|X^{-1}(\lambda E - A)X\right| \\ &= \left|X^{-1}\right| |\lambda E - A||X| = |\lambda E - A|.\end{aligned}$$

∎

- 线性变换的矩阵的特征多项式与基的选择无关, 它是直接被线性变换决定的.
- 定理 6 的逆是不对的, 特征多项式相同的矩阵不一定是相似的. 例如

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

它们的特征多项式都是 $(\lambda - 1)^2$, 但 A 和 B 不相似.

这是因为与单位矩阵 A 相似的矩阵只能是 A 本身.

- 设 $A=\begin{pmatrix} A_1 & A_2 \\ 0 & A_3 \end{pmatrix}$, 则 $f_A(\lambda)=f_{A_1}(\lambda)f_{A_3}(\lambda)$, 即 $|\lambda E-A|=|\lambda E-A_1||\lambda E-A_3|$
- 根据 $|A|=\prod_{i=1}^{n}\lambda_i$, 则
  $|A|=0 \Leftrightarrow A$ 一定有零特征值.
  $|A|\neq 0 \Leftrightarrow A$ 无零特征值, 即特征值非零.
  例子: 设 $A=\begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$, 则 $|A|=0$, 从而有一个零特征值, 另一个为 3. 依据公式 $\sum_{i=1}^{n}a_{ii}=\sum_{i=1}^{n}\lambda_i$.
- $f_A(\lambda)$ 在所考虑的数域范围内不一定有解, 从而可以没有特征值.

- 线性变换的矩阵的特征多项式与基的选择无关, σ 在的矩阵 A 的特征多项式称为 σ 的特征多项式, 并记作 $f_\sigma(\lambda)$.
- 相似矩阵有相同的行列式, 故可定义线性变换 σ 的行列式为 σ 在任意一组基下的矩阵的行列式.
- 矩阵的特征值、迹、行列式都是相似不变量.

例 6

设 $A = \begin{pmatrix} 3 & 2 & -1 \\ 2 & 3 & -1 \\ -4 & b & a \end{pmatrix}$, 已知 A 有特征值 $\lambda_1 = \lambda_2 = 1, \lambda_3 = 5$. 求 a, b 的值.

解 因为 A 有特征值 $\lambda_1 = \lambda_2 = 1, \lambda_3 = 5$,

所以有 $\mathrm{Tr}(A) = 6 + a = 7$, 即 $a = 1$.

由 $|A| = b + 1 = \lambda_1\lambda_2\lambda_3$ 得 $b = 4$.

特征多项式的性质

首先介绍一个概念. 形如

$$B(\lambda) = \lambda^m B_0 + \lambda^{m-1} B_1 + \cdots + B_m$$

多项式, 其中 $B_0, B_1, \cdots, B_m$ 都是 $n \times n$ 数字矩阵, 叫做一个矩阵多项式, n 叫做的级数, 当 $B_0 \neq 0$ 时, m 叫做它的次数. 如

$$\begin{aligned}\begin{pmatrix} \lambda^3 + 3\lambda + 1 & \lambda + 2 \\ \lambda^3 + 2\lambda & 3 \end{pmatrix} &= \begin{pmatrix} \lambda^3 & 0 \\ \lambda^3 & 0 \end{pmatrix} + \begin{pmatrix} 3\lambda & \lambda \\ 2\lambda & 0 \end{pmatrix} + \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix} \\ &= \lambda^3 \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} + \lambda \begin{pmatrix} 3 & 1 \\ 2 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}\end{aligned}$$

哈密顿 - 凯莱 (Hamilton-Caylay) 定理

- 设 A 是数域 P 一个 $n \times n$ 矩阵, $f(\lambda) = |\lambda E - A|$ 是 A 的特征多项式, 则

$$f(A) = A^n - (a_{11} + a_{22} + \cdots + a_{nn}) A^{n-1} + \cdots + (-1)^n |A| E = 0$$

- 设 $\mathscr{A}$ 是有限维线性空间 V 的线性变换, $f(\lambda)$ 是 $\mathscr{A}$ 的特征多项式, 那么

$$f(\mathscr{A}) = \mathscr{O}.$$

性质

若 λ 是 A 的特征值，即 $AX=\lambda X(X\neq 0)$, 则

- $k\lambda$ 是 kA 的特征值 (k 是常数), 且 $kAX=k\lambda X$.
- λ^m 是 A^m 的特征值 (m 是正整数)，且$A^mX=\lambda^mX$.
- 若 A 可逆，则
  - λ^{-1} 是 A^{-1} 的特征值, 且 $A^{-1}X=\lambda^{-1}X$,
  - $\lambda^{-1}|A|$ 是 A^* 的特征值，且 $A^*X=\lambda^{-1}|A|X$.
- 设 $\varphi(t)$ 为 t 的多项式，则 $\varphi(\lambda)$ 是 $\varphi(A)$ 的特征值, 且 $\varphi(A)X=\varphi(\lambda)X$.
- 矩阵 A 和 A^{T} 的特征值相同，特征多项式相同.

本节的目的: 研究什么样的线性变换, 在一组基下的矩阵形式是对角阵.

问题

设 3 维线性空间的一个线性变换 $\mathscr{A}$ 在一组基 $\varepsilon_1, \varepsilon_2, \varepsilon_3$ 下的矩阵为

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

是否存在在一组基，使得 $\mathscr{A}$ 在这组基下的矩阵形式是对角阵.

解 令 $X = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{pmatrix}$, 可知 X 可逆.

令

$$(\xi_1, \xi_2, \xi_3) = (\varepsilon_1, \varepsilon_2, \varepsilon_3)\, X,$$

即令

$$\xi_1 = \varepsilon_1 - \varepsilon_2, \xi_2 = \varepsilon_1 - \varepsilon_3, \xi_3 = \varepsilon_1 + \varepsilon_2 + \varepsilon_3,$$

则

$$\mathscr{A}\xi_1 = -\xi_1, \mathscr{A}\xi_2 = -\xi_2, \mathscr{A}\xi_3 = 5\xi_3,$$

且 ξ_1, ξ_2, ξ_3 是线性空间的一组基.

因此, $\mathscr{A}$ 在这组基下的矩阵

$$X^{-1}AX = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 5 \end{pmatrix}$$

是对角阵.

定理 7

设 $\mathscr{A}$ 是 n 维线性空间 V 的一个线性变换 $\mathscr{A}$ 的矩阵.

$\mathscr{A}$ 在某一组基下的矩阵为对角矩阵 $\Longleftrightarrow \mathscr{A}$ 有 n 个线性无关的特征向量.

证明

- 设 $\mathscr{A}$ 在基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 下具有对角矩阵

  $$\operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n),$$

  这就是说,

  $$\mathscr{A}\varepsilon_i = \lambda_i \varepsilon_i, i = 1, 2, \cdots, n$$

  因此, $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 就是 $\mathscr{A}$ 的 n 个线性无关的特征向量.

- 反过来, 如果 $\mathscr{A}$ 有 n 个线性无关的特征向量 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$, 那么就取 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 为基, 在这组基下 $\mathscr{A}$ 的矩阵是对角矩阵.

定理 8

属于不同特征值的特征向量是线性无关的.

证明 对特征值的个数作数学归纳法.

- 由于特征向量是不为零的, 所以单个的特征向量必然线性无关.
- 现在设属于 k 个不同特征值伺特征向量线性无关.
- 我们证明属于 $k+1$ 个不同特征值 $\lambda_1, \lambda_2, \cdots, \lambda_{k+1}$ 的特征向量 $\xi_1, \xi_2, \cdots, \xi_{k+1}$ 也线性无关.

假设有关系式

$$a_1\xi_1 + a_2\xi_2 + \cdots + a_k\xi_k + a_{k+1}\xi_{k+1} = 0 \tag{1}$$

成立. 等式两端乘以 λ_{k+1}, 得

$$a_1\lambda_{k+1}\xi_1 + a_2\lambda_{k+1}\xi_2 + \cdots + a_k\lambda_{k+1}\xi_k + a_{k+1}\lambda_{k+1}\xi_{k+1} = 0 \tag{2}$$

(1) 式两端同时施行变换 $\mathscr{A}$, 即有

$$a_1\lambda_1\xi_1 + a_2\lambda_2\xi_2 + \cdots + a_k\lambda_k\xi_k + a_{k+1}\lambda_{k+1}\xi_{k+1} = 0 \tag{3}$$

(3) 减去 (2) 得到

$$a_1(\lambda_1-\lambda_{k+1})\xi_1+\cdots+a_k(\lambda_k-\lambda_{k+1})\xi_k=\mathbf{0}$$

根据归纳法假设, $\xi_1,\xi_2,\cdots,\xi_k$ 线性无关, 于是

$$a_i(\lambda_i-\lambda_{k+1})=0, i=1,2,\cdots,k$$

但 $\lambda_i-\lambda_{k+1}\neq 0(i\leqslant k)$, 所以 $a_i=0, i=1,2,\cdots,k$.
这时 (1) 式变成

$$a_{k+1}\xi_{k+1}=\mathbf{0}.$$

又因为 $\xi_{k+1}\neq 0$, 所以只有 $a_{k+1}=0$.
这就证明了 $\xi_1\ \xi_2,\cdots,\xi_{k+1}$ 线性无关.
根据归纳法原理, 定理得证. ∎

从上面这两个定理就得到

推论

在 n 维线性空间 V 中, 线性变换 $\mathscr{A}$ 的特征多项式在数域 P 中有 n 个不同的根, 即 $\mathscr{A}$ 有 n 个不同的特征值, 那么 $\mathscr{A}$ 在某组基下的矩阵是对角形的.

复数域中任一个 n 次多项式都有 n 个根. 因此, 上面的论断可以改写为

推论

在复数域上的线性空间中, 如果线性变换 $\mathscr{A}$ 的特征多项式没有重根, 那么 $\mathscr{A}$ 在某组基下的矩阵是对角形的.

在一个线性变换没有 n 个不同的特征值的情形, 要判别这个线性变换的矩阵能不能成为对角形, 问题就要复杂些.

为了利用定理 7, 我们把定理 8 推广为

定理 9

如果 $\lambda_1, \cdots, \lambda_k$ 是线性变换 $\mathscr{A}$ 的不同的特征值, 而 $\alpha_{i1}, \cdots, \alpha_{ir_i}$ 是属于特征值 λ_i 的线性无关的特征向量 $i = 1, \cdots, k$, 那么向量组

$$\alpha_{11}, \cdots, a_{1r_1}, \cdots, a_{k1}, \cdots, a_{kr_n}$$

也线性无关.

例 1

在 F^3 中, 设线性变换 σ 在基 $\varepsilon_1, \varepsilon_2, \varepsilon_3$ 下的矩阵为

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

其中 $\varepsilon_1, \varepsilon_2, \varepsilon_3$ 为 3 维单位向量组.

问: 线性变换 σ 在此基下的矩阵是否为对角形?

例 1

在 F^3 中, 设线性变换 σ 在基 $\varepsilon_1, \varepsilon_2, \varepsilon_3$ 下的矩阵为

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

其中 $\varepsilon_1, \varepsilon_2, \varepsilon_3$ 为 3 维单位向量组.

问: 线性变换 σ 在此基下的矩阵是否为对角形?

解　特征多项式 $f_A(\lambda) = |\lambda E - A| = (\lambda - 1)(\lambda - 5)(\lambda + 5)$,

所以 $\lambda_1 = 1, \lambda_2 = 5, \lambda_3 = -5.$

由推论可知，线性变换 σ 在此基下的矩阵为对角形.

根据这个定理, 对于一个线性变换, 求出属于每个特征值的线性无关的特征向量, 把它们合在一起还是线性无关的.

- 如果它们的个数等于空间的维数, 那么这个线性变换在一组合适的基下的矩阵是对角矩阵;
- 如果它们的个数小于空间的维数, 那么这个线性变换在任何一组基下的矩阵都不能是对角形的.

设 $\mathscr{A}$ 的全部不同的特征值是 $\lambda_1, \cdots, \lambda_s$.
对每个 $1 \leq i \leq s$, 设 $\dim V_{\lambda_i} = m_i, \{\alpha_{i1}, \cdots, \alpha_{i_m}\}$ 是 V_i 的一组基.
则各特征子空间 V_{λ_i} 的基 M_i 所含向量共同组成的集合

$$\{\alpha_{ij} \,|\, 1 \leq i \leq s, 1 \leq j \leq m_i\}$$

线性无关，它包含 $m_1 + m_2 + \cdots + m_s$ 个线性无关的特征向量, 是 $\mathscr{A}$ 的特征向量集合的一个极大线性无关量组.

性质

$\mathscr{A}$ 在某一组基下的矩阵成对角形 $\Leftrightarrow$ $\mathscr{A}$ 的各特征子空间 $V_{\lambda_1}, \ldots, V_{\lambda_s}$ 的维数之和等于空间的维数 $\dim V$.

性质

当线性变换 $\mathscr{A}$ 在一组基下的矩阵 A 是对角形

$$diag\left(\lambda_1, \lambda_2, \ldots, \lambda_n\right),$$

时, $\mathscr{A}$ 的特征多项式为

$$|\lambda E - A| = \left(\lambda - \lambda_1\right)\left(\lambda - \lambda_2\right)\cdots\left(\lambda - \lambda_n\right)$$

因此, 如果线性变换线上的元素除排列次序外是确定的, 它们正是 $\mathscr{A}$ 的特征多项式全部的根 (重根按重数计算).

根据 § 3 定理 5, 一个线性变换的矩阵能不能在某一组基下是对角形的问题就相当于一个矩阵是不是相似于一个对角矩阵的问题.

因此, 这一节的讨论也就解决了后一个问题.

总结：矩阵的可对角化问题

- 如果数域 P 上的 n 级矩阵 A 可相似于对角矩阵, 则称 A 可对角化.
- 数域 P 上 n 级矩阵 A 可对角化的条件如下：
  1. (充分必要条件) A 有 n 个线性无关的特征向量.
  2. (充分条件) A 有 n 个互异的特征值.
  3. (充分必要条件) A 的所有重特征值对应的线性无关特征向量的个数等于其重数.

总结：线性变换的可对角化问题

设 $\mathscr{A}$ 是数域 P 上 n 维线性空间 V 的一个线性变换，$\mathscr{A}$ 的矩阵可以在某一组基下为对角矩阵有下列条件：

1. (充分必要条件）$\mathscr{A}$ 有 n 个线性无关的特征向量.
2. (充分条件）$\mathscr{A}$ 在数域 P 中有 n 个不同的特征值.
3. (充分必要条件）设 $\mathscr{A}$ 的全部互异的特征值为 $\lambda_1\ \lambda_2\ \ldots, \lambda_s$, 则 $\mathscr{A}$ 的特征子空间 $V_1, V_2, \cdots, V_s$ 的维数之和等于 n.

代数重数与几何重数

对方阵 A, 特征多项式

$$f_A(\lambda) = |\lambda E - A| = (\lambda - \lambda_1)^{n_1} (\lambda - \lambda_2)^{n_2} \cdots (\lambda - \lambda_s)^{n_s},$$

其中 $\lambda_1, \lambda_2, \cdots, \lambda_s$ 两两互异.

- n_i 称为特征值 λ_i 的代数重数.
- 再看特征值 λ_i 的特征子空间 V_{λ_i}, 其维数 $\dim V_{\lambda_i} = m_i$ 称为 λ_i 的几何重数. $\quad m_i = n - r(\lambda_i E - A)$

性质

矩阵的任一特征值的几何重数小于等于其代数重数.

证明需要不变子空间的性质，会在本章 §7 给出。

性质

矩阵的任一特征值的几何重数小于等于其代数重数.

性质

方阵 A 可相似对角化 $\Leftrightarrow$ 任一特征值的代数重数与几何重数相等.

性质

设 A 是一 n 阶方阵, A 的特征多项式为

$$f_A(\lambda) = (\lambda - \lambda_1)^{n_1} (\lambda - \lambda_2)^{n_2} \cdots (\lambda - \lambda_s)^{n_i}$$

则

$$\begin{aligned} A \text{ 可相似对角化} &\Leftrightarrow n_i = n - r(\lambda_i E - A), \quad i = 1, 2, \cdots, s. \\ &\Leftrightarrow r(\lambda_i E - A) = n - n_i, \quad i = 1, 2, \cdots, s. \end{aligned}$$

判断相似对角化的方法

(1) 对 A, 求所有特征值 $f_A(\lambda) = |\lambda E - A| = (\lambda - \lambda_1)^{n_1} \cdots (\lambda - \lambda_s)^{n_s}$

(2) 是否可以相似对角化 $\Leftrightarrow r(\lambda_i E - A) = n - n_i, i = 1, 2, \cdots, s$.

- 若都成立, 则可以;
- 若有一个不成立, 则不能.
  - 对单根 λ_i 自然成立, 不用判断 $r(\lambda_i E - A) = n - 1$.
  - 主要是对重根 λ_i, 来判断是否 $r(\lambda_i E - A) = n - n_i$.

(3) 若可以, 求方程组 $(\lambda_i E - A) X = 0$ 的基础解系 $X_{i1}, \cdots, X_{in_i}$

令 $X = (X_{11}, \cdots, X_{1n_1}, X_{21}, \cdots, X_{sn_s})$, 则

$$X^{-1}AX = \begin{pmatrix} \lambda_1 E_{n_1} & & \\ & \ddots & \\ & & \lambda_s E_{n_s} \end{pmatrix}$$

例 2

判断 $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ 是否可以相似对角化.

若可以, 求可逆阵 X , 使得 $X^{-1}AX$ 为对角阵.

解 $|\lambda E - A| = \begin{vmatrix} \lambda & -1 \\ 0 & \lambda \end{vmatrix} = \lambda^2 = 0$, 则 $\lambda_1 = \lambda_2 = 0$.

$0E - A = -A$, 秩为 $1 \neq 2 - 2$, 故不能相似对角化.

例 3

判断 $A=\begin{pmatrix} 1 & 2 & 2 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{pmatrix}$ 是否可以相似对角化.

若可以, 求可逆阵 X , 使得 $X^{-1}AX$ 为对角阵.

解 A 的特征值为 $\lambda_1=1, \lambda_2=2, \lambda_3=3$, 有 3 个互异的特征值, 可以相似对角化.

对 $\lambda_1=1$, 线性方程组 $(E-A)X=0$ 的解: $X_1=(1,0,0)'$,

对 $\lambda_2=2$, 线性方程组 $(2E-A)X=0$ 的解: $X_2=(2,1,0)'$,

对 $\lambda_3=3$, 线性方程组 $(3E-A)X=0$ 的解: $X_3=(3,2,1)'$.

令 $X=\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$,

则 $X^{-1}AX=\mathrm{diag}(1,2,3)$.

例 4

判断 $A=\begin{pmatrix}1&2&2\\0&1&2\\0&0&2\end{pmatrix}$ 是否可以相似对角化.

若可以，求可逆阵 X，使得 $X^{-1}AX$ 为对角阵.

解 A 的特征值为 $\lambda_1=\lambda_2=1,\lambda_3=2$

对 $\lambda_1=1, E-A=\begin{pmatrix}0&-2&-2\\0&0&-2\\0&0&-1\end{pmatrix}$，秩为 2, $2\neq 3-2$,

不能相似对角化

例 5

判断 $A=\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 \end{pmatrix}$ 是否可以相似对角化

解 $\begin{vmatrix} \lambda-1 & -1 & -1 & -1 \\ -1 & \lambda-1 & -1 & -1 \\ -2 & -2 & \lambda-2 & -2 \\ -3 & -3 & -3 & \lambda-3 \end{vmatrix}=\lambda^3(\lambda-7)$,

得到特征值 $\lambda_1=0$(3 重), $\lambda_2=7$.

$r(-A)=1=4-3$, 有 3 个.

可以相似对角化.

例 6 (*)

设 n 阶方阵 A 满足 $A^2=E$, 判断 A 是否可以相似对角化

解 $A^2-E=(A+E)(A-E)=0.$

首先 $|A+E||A-E|=0$, 则 A 有特征值 1 或者 -1.

$r(A+E)+r(A-E)\le n$. 同时 $n=r(A+E-A+E)\le r(A+E)+r(A-E)$

则 $r(A+E)+r(A-E)=n$. 设 $r(A+E)=r$, 则 $r(A-E)=n-r$

对特征值 1, 线性方程组 $(E-A)X=0$ 的基础解系含有 $n-(n-r)=r$ 个向量 $\xi_1,\xi_2,\cdots,\xi_r$

对特征值-1, 线性方程组 $(-E-A)X=0$ 的基础解系含有 $n-r$ 个向量 $\xi_{r+1},\xi_{r+2},\cdots,\xi_n$

又属于不同特征值的特征向量线性无关, 从而 A 有 n 个线性无关的特征向量 $\xi_1,\xi_2,\cdots,\xi_n$

令 $X=(\xi_1,\xi_2,\cdots,\xi_n)$ 则 $X^{-1}AX=\begin{pmatrix} E_r & 0 \\ 0 & -E_{n-r} \end{pmatrix}$.

例 7 (*)

设 n 阶方阵 A 满足 $A^2 = A$, 判断 A 是否可以相似对角化.

解 $A^2 - A = A(A-E) = 0$, 首先 $|A||A-E| = 0$, 则 A 有特征值 0 或者 1 且 $r(A) + r(A-E) = n$. 设 $r(A) = r$,

则 $r(A-E) = n - r$

对特征值 1，线性方程组 $(E-A)X = 0$ 的基础解系含有 $n-(n-r) = r$ 个向量 $\xi_1, \xi_2, \cdots, \xi_r$

对特征值 0, 线性方程组 $(-A)X = 0$ 的基础解系含有 $n-r$ 个向量 $\xi_{r+1}, \xi_{r+2}, \cdots, \xi_n$

又属于不同特征值的特征向量线性无关, 从而 A 有 n 个线性无关的特征向量 $\xi_1, \xi_2, \cdots, \xi_n$

令 $X = (\xi_1, \xi_2, \cdots, \xi_n)$. 则 $X^{-1}AX = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$.

定义

- 设 $\mathscr{A}$ 是线性空间 V 的一个线性变换, 则的全体像组成的集合称为 A 的**值域**, 用 $\mathscr{A}V$ 或者 $\mathrm{Im}\,\mathscr{A}$ 表示.
- 所有被 $\mathscr{A}$ 变成零向量的向量组成的集合称为 $\mathscr{A}$ 的**核**, 用 $\mathscr{A}^{-1}(\mathbf{0})$ 或 $\mathscr{A}^{-1}(\mathbf{0})$ 表示.

若用集合的集合, 则

$$\mathscr{A}V = \{\mathscr{A}\xi \mid \xi \in V\},$$

$$\mathscr{A}^{-1}(\mathbf{0}) = \{\xi \mid \xi \in V, \mathscr{A}\xi = \mathbf{0}\}.$$

性质

线性变换的值域与核都是 V 的子空间.

证明

- 首先 $\mathscr{A}V$ 非空, 并且对于 V 中任何向量 α, β 和 $k \in P$, 都有

$$\mathscr{A}\alpha + \mathscr{A}\beta = \mathscr{A}(\alpha + \beta), \quad k\mathscr{A}\alpha = \mathscr{A}(k\alpha)$$

即 $\mathscr{A}V$ 对 V 的加法和数乘封闭, 故 $\mathscr{A}V$ 是 V 的子空间.

- 同样, 由于 $\mathscr{A}(\mathbf{0}) = \mathbf{0}$, 故 $\mathbf{0} \in \mathscr{A}^{-1}(\mathbf{0})$ $\mathscr{A}^{-1}(\mathbf{0})$ 非空.
设 α, β 是 $\mathscr{A}^{-1}(\mathbf{0})$ 中任何向量和 $k \in P$ 由

$$\mathscr{A}\alpha = \mathbf{0}, \quad \mathscr{A}\beta = \mathbf{0}$$

可知

$$\mathscr{A}(\alpha + \beta) = \mathbf{0}, \quad \mathscr{A}(k\alpha) = \mathbf{0}$$

所以, $\mathscr{A}^{-1}(\mathbf{0})$ 对加法和数乘封闭, 故 $\mathscr{A}^{-1}(\mathbf{0})$ 是 V 的子空间.

定义

- $\mathscr{A}V$ 的维数称为 $\mathscr{A}$ 的秩, 记为 $rank(\mathscr{A})$ 或 $r(\mathscr{A})$.
- $\mathscr{A}^{-1}(\mathbf{0})$ 的维数称为 $\mathscr{A}$ 的零度, 记为 $null(\mathscr{A})$.

例 1

在线性空间 $P[x]_n$ 中, 令

$$\mathscr{D}(f(x)) = f'(x)$$

则 $\mathscr{D}$ 的值域就是 $P[x]_{n-1}$, $\mathscr{D}$ 的核就是子空间 P.
从而 $rank(\mathscr{A}) = n - 1$, $null(\mathscr{A}) = 1$.

定理 10

设 $\mathscr{A}$ 是 n 维向量空间 V 的线性变换, $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ 是 V 的一组基, $\mathscr{A}$ 在这组基下的矩阵是 A, 则

1. $\mathscr{A}$ 的值域 $\mathscr{A}V$ 是由基像组生成的子空间, 即
$$\mathscr{A}V = L\left(\varepsilon_1, \varepsilon_2, \cdots, \mathscr{A}\varepsilon_n\right).$$
2. $\mathscr{A}$ 的秩 $=$ A 的秩.

证明 (1) 集合相等. 首先, $\mathscr{A}V \supseteq L\left(\mathscr{A}\varepsilon_1, \mathscr{A}\varepsilon_2, \cdots, \mathscr{A}\varepsilon_n\right)$.

其次, 任给 $\mathscr{A}\alpha \in \mathscr{A}V$, 则 $\alpha \in V$, 从而 $\alpha = \left(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n\right) X$,

则 $\mathscr{A}\alpha = \left(\mathscr{A}\varepsilon_1, \mathscr{A}\varepsilon_2, \cdots, \mathscr{A}\varepsilon_n\right) X \in L\left(\mathscr{A}\varepsilon_1, \mathscr{A}\varepsilon_2, \cdots, \mathscr{A}\varepsilon_n\right)$,

从而 $\mathscr{A}V \subseteq L\left(\mathscr{A}\varepsilon_1, \mathscr{A}\varepsilon_2, \cdots, \mathscr{A}\varepsilon_n\right)$.

因此, $\mathscr{A}V = L\left(\mathscr{A}\varepsilon_1, \mathscr{A}\varepsilon_2, \cdots, \mathscr{A}\varepsilon_n\right)$.

(2) 因为 $\left(\mathscr{A}\varepsilon_1, \mathscr{A}\varepsilon_2, \cdots, \mathscr{A}\varepsilon_n\right) = \left(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n\right) A$, 所以

$$\dim L\left(\mathscr{A}\varepsilon_1, \mathscr{A}\varepsilon_2, \cdots, \mathscr{A}\varepsilon_n\right) = r(A).$$

求值域与核

- $\mathscr{A}V = L(\mathscr{A}\varepsilon_1, \mathscr{A}\varepsilon_2, \cdots, \mathscr{A}\varepsilon_n)$, 求 $\mathscr{A}\varepsilon_1, \mathscr{A}\varepsilon_2, \cdots, \mathscr{A}\varepsilon_n$ 的极大无关组即可.
- 任给 $\alpha \in \mathscr{A}^{-1}(\mathbf{0})$, 则 $\mathscr{A}\alpha = 0$, 假设 $\mathscr{A}$ 在基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 下的矩阵为 A, 并设 $\alpha = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)X$, 则

$$0 = \mathscr{A}\alpha = \mathscr{A}(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)X = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)AX,$$

从而 $AX = 0$, 则 α 在基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 下的坐标是齐次线性方程组 $AX = 0$ 的解. 从而求得 $AX = 0$ 的一组基础解系

$$\eta_1, \eta_2, \cdots, \eta_r,$$

以基础解系为坐标的向量

$$\xi_1, \xi_2, \cdots, \xi_r$$

就是 $\mathscr{A}^{-1}(\mathbf{0})$ 的一组基, 即 $\mathscr{A}^{-1}(\mathbf{0}) = L(\xi_1, \xi_2, \cdots, \xi_r)$.

例 2

设 $\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4$ 是线性空间 V 的一组基，已知线性变换 σ 在此基下的矩阵为 $A = \begin{pmatrix} 1 & 0 & 2 & 1 \\ -1 & 2 & 1 & 3 \\ 1 & 2 & 5 & 5 \\ 2 & -2 & 1 & -2 \end{pmatrix}$. 求 σ 的值域与核的维数与基.

解 先求 $\sigma(V)$. 对矩阵 A 做初等行变换，得

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

于是，σ 的秩为 2, 即 $\sigma(V)$ 为 2 维的, 且

$$\begin{aligned} \sigma\left(\varepsilon_1\right) = & \quad \varepsilon_1 - \varepsilon_2 + \varepsilon_3 + 2\varepsilon_4, \\ \sigma\left(\varepsilon_2\right) = & \qquad 2\varepsilon_2 + 2\varepsilon_3 - 2\varepsilon_4. \end{aligned}$$

是 $\sigma(V)$ 的一组基.

求 $\sigma^{-1}(\mathbf{0})$. 设 $\xi \in \sigma^{-1}(\mathbf{0})$, 它在 $\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4$ 下的坐标为 (x_1, x_2, x_3, x_4) 由于 $\sigma(\xi) = \mathbf{0}$, 有 $\sigma(\xi)$ 在 $\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4$ 下的坐标为 $(0, 0, 0, 0)$. 解齐次线性方程组

$$\begin{pmatrix} 1 & 0 & 2 & 1 \\ -1 & 2 & 1 & 3 \\ 1 & 2 & 5 & 5 \\ 2 & -2 & 1 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

得它的一个基础解系 $(-2, -2/3, 1, 0)$, $(-1, -2, 0, 1)$. 于是，$\sigma^{-1}(\mathbf{0})$ 的维数为 2，且

$$\begin{array}{rrrr} -2\varepsilon_1 & -\frac{2}{3}\varepsilon_2 & +\varepsilon_3, & \\ -\varepsilon_1 & -2\varepsilon_2 & & +\varepsilon_4 \end{array}$$

是 $\sigma^{-1}(\mathbf{0})$ 的一组基. ∎

例 3

设 $V = F^3$, 变换: $\sigma(x_1, x_2, x_3) = (2x_1 + x_2, x_2 + x_3, 2x_1 - x_3)$. 求 σ 的值域与核.

解

- 设 $X = (x_1, x_2, x_3)'$, 取单位向量 $\varepsilon_1, \varepsilon_2, \varepsilon_3$, 则 $\mathscr{A}$ 在基 $\varepsilon_1, \varepsilon_2, \varepsilon_3$ 下的矩阵为 $A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & 1 \\ 2 & 0 & -1 \end{pmatrix}$.

  将 A 做初等行变换化成阶梯形矩阵: $A \to \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$

  则 σ 的值域的维数为 2, 一组基为 $\alpha_1 = (2, 0, 2)', \alpha_2 = (1, 1, 0)',$.
- 解方程组 $AX = 0$, 得基础解系 $\eta = (1, -2, 2)'$.

  则 σ 的核的维数为 1, 一组基为 η.

定理 11

设 $\mathscr{A} \in L(V), \dim V = n$, 则

- $\mathscr{A}(V)$ 的一组基的原象及 $\mathscr{A}^{-1}(\mathbf{0})$ 的一组基合起来是 V 的一组基,
- $\mathscr{A}$的秩 $+\ \mathscr{A}$ 的零度 $= n$.

例 4

设 $F[x]_n$ 的微分变换 $\mathscr{D}$, $\mathscr{D}F[x]_n = F[x]_{n-1}, \mathscr{D}^{-1}(\mathbf{0}) = F$, 从而 $\dim \mathscr{A}(V) = n-1, \dim \mathscr{A}^{-1}(\mathbf{0}) = 1$.

注

$\mathscr{A}V$ 与 $\mathscr{A}^{-1}(\mathbf{0})$ 的维数之和为 n, 但是 $\mathscr{A}V + \mathscr{A}^{-1}(\mathbf{0})$ 不一定是直和.

证明 设 $\mathscr{A}(V)$ 的一组基为 $\eta_1, \eta_2, \cdots, \eta_r$, 其原象设为 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_r$, 即 $\mathscr{A}\varepsilon_i = \eta_i$, $i = 1, 2, \cdots, r$. 设 $\mathscr{A}^{-1}(\mathbf{0})$ 的一组基为 $\varepsilon_{r+1}, \varepsilon_{r+2}, \cdots, \varepsilon_s$, 下证

$$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_r, \varepsilon_{r+1}, \varepsilon_{r+2}, \cdots, \varepsilon_s$$

是 V 的一组基.

1) 先证 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_r, \varepsilon_{r+1}, \varepsilon_{r+2}, \cdots, \varepsilon_s$ 线性无关. 设

$$k_1\varepsilon_1 + \cdots + k_r\varepsilon_r + k_{r+1}\varepsilon_{r+1} + \cdots + k_s\varepsilon_s = \mathbf{0}$$

用 $\mathscr{A}$ 作用一下,

$$k_1\mathscr{A}\varepsilon_1 + \cdots + k_r\mathscr{A}\varepsilon_r = \mathbf{0},$$

即

$$k_1\eta_1 + \cdots + k_r\eta_1 = \mathbf{0}.$$

因为 $\eta_1, \eta_2, \cdots, \eta_r$ 是 $\mathscr{A}(V)$ 的一组基, 所以

$$k_1 = \cdots = k_r = 0.$$

于是,

$$k_{r+1}\varepsilon_{r+1}+\cdots+k_s\varepsilon_s=\mathbf{0}.$$

因为 $\varepsilon_{r+1},\varepsilon_{r+2},\cdots,\varepsilon_s$ 为 $\mathscr{A}^{-1}(\mathbf{0})$ 的一组基, 所以

$$k_{r+1}=\cdots=k_s=0.$$

因此,$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_r,\varepsilon_{r+1},\varepsilon_{r+2},\cdots,\varepsilon_s$ 线性无关.

2)　再证 V 中任意向量都可由 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_r,\varepsilon_{r+1},\varepsilon_{r+2},\cdots,\varepsilon_s$ 线性表示.
任给 $\alpha\in V$, 则 $\mathscr{A}\alpha\in\mathscr{A}(V)=L\left(\eta_1,\eta_2,\cdots,\eta_r\right)$, 从而

$$\begin{aligned}\mathscr{A}\alpha&=k_1\eta_1+k_2\eta_2+\cdots+k_r\eta_r\\&=k_1\mathscr{A}\varepsilon_1+k_2\mathscr{A}\varepsilon_2+\cdots+k_r\mathscr{A}\varepsilon_r\\&=\mathscr{A}\left(k_1\varepsilon_1+k_2\varepsilon_2+\cdots+k_r\varepsilon_r\right)\end{aligned}$$

即

$$\mathscr{A}\left(\alpha-k_1\varepsilon_1-k_2\varepsilon_2-\cdots-k_r\varepsilon_r\right)=\mathbf{0},$$

则

$$\alpha - k_1\varepsilon_1 - k_2\varepsilon_2 - \cdots - k_r\varepsilon_r \in \mathscr{A}^{-1}(\mathbf{0}),$$

即有

$$\alpha - k_1\varepsilon_1 - k_2\varepsilon_2 - \cdots - k_r\varepsilon_r = k_{r+1}\varepsilon_{r+1} + k_{r+2}\varepsilon_{r+2} + \cdots + k_s\varepsilon_s$$

从而

$$\alpha = k_1\varepsilon_1 + k_2\varepsilon_2 + \cdots + k_r\varepsilon_r + k_{r+1}\varepsilon_{r+1} + k_{r+2}\varepsilon_{r+2} + \cdots + k_s\varepsilon_s.$$

综上,

$$\underbrace{\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_r,}_{\mathscr{A}(V)\text{的一组基的原象}} \quad \underbrace{\varepsilon_{r+1}, \varepsilon_{r+2}, \cdots, \varepsilon_s}_{\mathscr{A}^{-1}(\mathbf{0})\text{的一组基}}$$

是 V 的一组基, 从而 $s = n$, 且 $\dim \mathscr{A}(V) + \dim \mathscr{A}^{-1}(\mathbf{0}) = \dim V = n$. ∎

推论

设 V 是一个有限维线性空间, $\mathscr{A} \in L(V)$, 则 $\mathscr{A}$ 是单射 $\Leftrightarrow$ $\mathscr{A}$ 是满射.

证明

$$\mathscr{A}\text{ 是单射} \Leftrightarrow \mathscr{A}^{-1}(\mathbf{0}) = \mathbf{0} \Leftrightarrow \dim \mathscr{A}^{-1}(\mathbf{0}) = 0$$

$$\mathscr{A}\text{ 是满射} \Leftrightarrow \mathscr{A}(V) = V \quad \Leftrightarrow \dim \mathscr{A}(V) = n$$

例 5 (*)

设 n 阶方阵 A, 满足 $A^2 = A$, 证明 A 可以相似于对角阵 $\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$.

证明 取线性空间 V 及一组基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$, 则存在线性变换 $\mathscr{A}$, 使得 $\mathscr{A}$ 在 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 下的矩阵为 A, 则 $\mathscr{A}^2 = \mathscr{A}$.

对 $\mathscr{A}V$, 取一组基 $\eta_1, \eta_2, \cdots, \eta_r$, 其原象设为 $\alpha_1, \alpha_2, \cdots, \alpha_r$,

即 $\mathscr{A}\alpha_i = \eta_i, i = 1, 2, \cdots, r$ 则 $\mathscr{A}\eta_i = \mathscr{A}^2\alpha_i = \mathscr{A}\alpha_i = \eta_i, i = 1, 2, \cdots, r$,

即 $\eta_1, \eta_2, \cdots, \eta_r$ 得原象就是 $\eta_1, \eta_2, \cdots, \eta_r$.

取 $\mathscr{A}^{-1}(\mathbf{0})$ 的一组基 $\eta_{r+1}, \eta_{r+2}, \cdots, \eta_n$,

则 $\eta_1, \eta_2, \cdots, \eta_r, \eta_{r+1}, \eta_{r+2}, \cdots, \eta_n$ 是 V 的一组基,

$\mathscr{A}$ 在 $\eta_1, \eta_2, \cdots, \eta_n$ 下的矩阵为

$$\mathscr{A}(\eta_1, \eta_2, \cdots, \eta_n) = (\eta_1, \cdots, \eta_r, 0, \cdots, 0) = (\eta_1, \eta_2, \cdots, \eta_n)\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

从而由于一个线性变换在不同基下的矩阵是相似的.

故 A 与 $\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$ 相似.

§7 不变子空间

定义

取线性空间 $V, \mathscr{A} \in L(V)$, 设 W 是 V 的一个子空间, 若 W 中的向量在 $\mathscr{A}$ 作用下的象仍在 W 中, 即任给 $\alpha \in W$, 则有 $\mathscr{A}\alpha \in W$, 则称 W 是 $\mathscr{A}$ 的一个不变子空间, 简称为 $\mathscr{A}-$ 子空间.

例 1

非不变子空间 取 3 维平面 $V = \mathbb{R}^3$, 取线性变换 $\mathscr{A}$ 为绕 X 轴线从 Y 轴线到 Z 轴线旋转 $90°$ 角的变换即 $\mathscr{A}(x, y, z) = (x, -z, y)$, 取子空间 $W = XOY$ 面, 则 W 不是不变子空间, 因为 $\mathscr{A}(0, 1, 0) = (0, 0, 1) \notin W$.

例 2

平凡子空间 $\mathbf{0}$ 与 V 是任一线性变换的不变子空间.

例 3

任一线性变换 $\mathscr{A}$ 的值域 $\mathscr{A}V$ 与核 $\mathscr{A}^{-1}(\mathbf{0})$ 是 $\mathscr{A}$ 的不变子空间.

证明

- 任给 $\mathscr{A}\alpha \in \mathscr{A}V$, 则 $\mathscr{A}(\mathscr{A}\alpha) \in \mathscr{A}V$.
- 任给 $\alpha \in \mathscr{A}^{-1}(\mathbf{0})$, 则 $\mathscr{A}\alpha = \mathbf{0}$, 从而 $\mathscr{A}(\mathscr{A}\alpha) = \mathscr{A}(\mathbf{0}) = \mathbf{0}$, 因此 $\mathscr{A}\alpha \in \mathscr{A}^{-1}(\mathbf{0})$.

例 4

取 $\mathscr{A}, \mathscr{B} \in L(V)$, 满足 $\mathscr{A}\mathscr{B} = \mathscr{B}\mathscr{A}$, 则 $\mathscr{B}V$ 与 $\mathscr{B}^{-1}(\mathbf{0})$ 都是 $\mathscr{A}$-子空间 $\mathscr{B}V$.

证明

- 任给 $\mathscr{B}\alpha \in \mathscr{B}V$, 证 $\mathscr{A}(\mathscr{B}\alpha) \in \mathscr{B}V$.
  $\mathscr{A}(\mathscr{B}\alpha) = (\mathscr{A}\mathscr{B})\alpha = (\mathscr{B}\mathscr{A})\alpha = \mathscr{B}(\mathscr{A}\alpha) \in \mathscr{B}V$.
- 任给 $\alpha \in \mathscr{B}^{-1}(\mathbf{0})$, 证 $\mathscr{A}\alpha \in \mathscr{B}^{-1}(\mathbf{0})$, 即证 $\mathscr{B}(\mathscr{A}\alpha) = \mathscr{A}(\mathscr{B}\alpha) = \mathbf{0}$.

例 5

任一子空间都是数乘变换的不变子空间, 这是因为子空间对数乘封闭.

注

- 特征向量与一维不变子空间的关系.
  - 取 $\mathscr{A}\xi=\lambda_0\xi$, 设 $W=L(\xi)$, 则 $\dim W=1, \mathscr{A}(k\xi)=\lambda_0 k\xi\in W$, 从而 $W=L(\xi)$ 是 $\mathscr{A}-$ 子空间.
  - 反之, 取一维不变子空间 W = L(ξ), $\mathscr{A}W\subseteq W$, 则 $\mathscr{A}\xi\in W=L(\xi)$, 从而存在一个数 λ, 使得 $\mathscr{A}\xi=\lambda\xi$, 即生成元 ξ 是 $\mathscr{A}$ 的属于 λ 的一个特征向量.
- $\mathscr{A}$ 的属于特征值 λ_0 的特征子空间 V_{λ_0} 也是 $\mathscr{A}$ 的不变子空间.
- 取 $\mathscr{A}$-子空间 W_1, W_2, 则 W_1+W_2 与 $W_1\cap W_2$ 都是 $\mathscr{A}$-子空间.

定义

设 $\mathscr{A}$ -子空间 W, 由于 $\mathscr{A}W \subseteq W$, 则 $\mathscr{A}$ 就可以看做是 W 上的一个线性变换, 从而定义 $\mathscr{A}|_W$ 为 $\mathscr{A}$ **限制**在不变子空间 W 上的线性变换.

注意 $\mathscr{A}$ 和 $\mathscr{A}|_W$ 的异同:

- $\mathscr{A}$ 是 V 的线性变换, V 中每个向量在 $\mathscr{A}$ 下都有确定的像;
- $\mathscr{A}|_W$ 是不变子空间 W 上的线性变换. 它的作用为任给 $\alpha \in W$, $\mathscr{A}|_m(\alpha) = \mathscr{A}(\alpha)$, 但是 $\mathscr{A}$ 对 W 之外的向量没有作用.

例如, 对于任一线性变换 $\mathscr{A}$,

- $\mathscr{A}$ 在它的核空间上引起的变换是零变换, 即 $\mathscr{A}|_{\mathscr{A}^{-1}(\mathbf{0})} = \mathscr{O}$,
- $\mathscr{A}$ 在特征子空间 V_{λ_0} 上引起的变换是数乘变换 $\mathscr{A}|_{V_{\lambda_0}}(\xi) = \lambda_0\xi$.

性质

若 $W = L(\alpha_1, \alpha_2, \cdots, \alpha_m)$, 则

$$W \text{ 是 } \mathscr{A} - \text{子空间} \Leftrightarrow \mathscr{A}\alpha_1, \mathscr{A}\alpha_2, \cdots, \mathscr{A}\alpha_m \in W.$$

证明 $\Rightarrow$ 若 $\mathscr{A}W \subseteq W$, 则 $\mathscr{A}\alpha_1, \mathscr{A}\alpha_2, \cdots, \mathscr{A}\alpha_s \in W$.

$\Leftarrow$ 反之, 若 $\mathscr{A}\alpha_1, \mathscr{A}\alpha_2, \cdots, \mathscr{A}\alpha_s \in W$, 则任给 $\alpha \in W$, 设 $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_s)X$, 则 $\mathscr{A}\alpha = (\mathscr{A}\alpha_1, \mathscr{A}\alpha_2, \cdots, \mathscr{A}\alpha_s)X \in W$, 从而 $\mathscr{A}W \subseteq W$. W 是 $\mathscr{A}$-子空间.

不变子空间与线性变换的矩阵

1) 设 $\mathscr{A}$ 是 n 维线性空间 V 的线性变换， W 是 V 的 $\mathscr{A}$-子空间.
在 W 中取一组基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_k$, 并且把它扩充成 V 的一组基

$$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_k, \varepsilon_{k+1}, \cdots, \varepsilon_n, \tag{1}$$

那么 $\mathscr{A}$ 在这组基下的矩阵就具有下列形状

$$\begin{pmatrix} a_{11} & \cdots & a_{1k} & a_{1,k+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} & a_{k,k+1} & & a_{kn} \\ 0 & \cdots & 0 & a_{k+1,k+1} & \cdots & a_{k+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & a_{n,k+1} & \cdots & a_{nn} \end{pmatrix} = \begin{pmatrix} A_1 & A_3 \\ 0 & A_2 \end{pmatrix} \tag{2}$$

且左上角的 k 级矩阵 A_1 就是 $\mathscr{A}|_W$ 在 W 的基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_k$ 下的矩阵.

这是因为 W 是 $\mathscr{A}$-子空间, 所以像 $\mathscr{A}\varepsilon_1, \mathscr{A}\varepsilon_2, \cdots, \mathscr{A}\varepsilon_k$ 仍在 W 中. 它们可以通过 W 的基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_k$ 线性表示

$$
\begin{aligned}
&\mathscr{A}\varepsilon_1 = a_{11}\varepsilon_1 + a_{21}\varepsilon_2 + \cdots + a_{k1}\varepsilon_k \\
&\mathscr{A}\varepsilon_2 = a_{12}\varepsilon_1 + a_{22}\varepsilon_2 + \cdots + a_{k2}\varepsilon_k \\
&\cdots\cdots\cdots\cdots \\
&\mathscr{A}\varepsilon_k = a_{1k}\varepsilon_1 + a_{2k}\varepsilon_2 + \cdots + a_{kk}\varepsilon_k
\end{aligned}
$$

从而

- $\mathscr{A}$ 在基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_k, \varepsilon_{k+1}, \cdots, \varepsilon_n$ 下的矩阵具有形状 (2),
- $\mathscr{A}|_W$ 在 W 的基 $\varepsilon_1, \varepsilon_2, \cdots\ \varepsilon_k$ 下的矩阵是 A_1.

反之, 如果 $\mathscr{A}$ 在基 (1) 下的矩阵是 (2), 那么可以证明, 由 $\varepsilon_1\ \varepsilon_2, \cdots, \varepsilon_k$ 生成的子空间 W 是 $\mathscr{A}$ 的不变子空间.

2) 设 V 分解成若干个 $\mathscr{A}$-子空间的直和：

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_s$$

在每一个 $\mathscr{A}$-子空间 W 中取基

$$\varepsilon_{i\,1}, \varepsilon_{i\,2}, \cdots, \varepsilon_{i\,n_i} \quad (i = 1, 2, \cdots, s) \tag{3}$$

并把它们合并起来成为 V 的一组基, 则在这组基下的矩阵为准对角形

$$\operatorname{diag}(A_1, A_2, \ldots, A_s) \tag{4}$$

其中 $A_i (i = 1, 2, \cdots, s)$ 就是 $\mathscr{A}|W_i$ 在基(3)下的矩阵.
反之, 如果线性变换 A 在某组基下的矩阵是准对角形(4), 则由 (3) 生成的子空间 W_i 是 $\mathscr{A}$-子空间.
由此可知,矩阵分解为准对角形与空间分解为不变子空间的直和是相当的.

定义

设 λ_i 是线性变换 $\mathscr{A}$ 的一个特征值，则

- 特征子空间 V_i 维数称为 λ_i 的几何重数.
- 每个特征根 λ_i 在特征多项式 $f_{\mathscr{A}}(\lambda_i)$ 中的重数称为 代数重数.

同一个特征值的代数重数和几重数之间有如下关系:

定理

设 λ_i 是线性变换 $\mathscr{A}$ 的特征值，它的代数重数为 n_i, 几何重数为 m_i, 则

1. 每个特征值的几何重数小于或等于它的代数重数, 即 $1 \leq m_i \leq n_i$.
2. $\mathscr{A}$ 的矩阵可以在某一组基下为对角矩阵 $\Leftrightarrow$ 每个特征值的几何重数都等于代数重数.

证明 (1) 特征值 λ_i 的特征子空间 V_{λ_i} 的维数等于 m_i, 取 V_{λ_i} 的一组基扩充充为 V 的一组基，则 $\mathscr{A}$ 在这组基下的矩阵为上三角矩阵

$$B = \begin{pmatrix} \lambda_i E_{m_i} & A_3 \\ 0 & A_2 \end{pmatrix}$$

它的特征多项式 $(\lambda - \lambda_i)^{m_i} \det(\lambda E - A_2)$, 含有因子 $(\lambda - \lambda_i)^{m_i}$. 于是 λ_i 在其中的代数重数 $n_i \geq m_i$.

(2) 设 $\lambda_1, \quad \cdots, \lambda_t$ 是 $\mathscr{A}$ 的全部不同的特征值. 由 §5 中的推论知道 $\mathscr{A}$ 可对角化的充分必要条件是

$$m_1 + \cdots + m_s = n = n_1 + \cdots + n_s$$

成立. 但 $m_i \leqslant n_i$ 对 $1 \leqslant i \leqslant s$ 成立，故 $\mathscr{A}$ 可对角化当且仅当 $m_i = n_i$ 对所有的 $1 \leqslant i \leqslant s$ 成立。

下面我们应用哈密尔顿 - 凯莱定理将空间 V 按特征值分解成不变子空间的直和.

定理 12 (准素分解)

设线性变换 $\mathscr{A}$ 的特征多项式为 $f(\lambda)$, 它可分解成一次因式的乘积

$$f(\lambda) = (\lambda - \lambda_1)^{r_1} (\lambda - \lambda_2)^{r_2} \cdots (\lambda - \lambda_s)^{r_s}$$

则 V 可分解成不变子空间的直和

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_s,$$

其中

$$V_i = \left\{\xi \mid (\mathscr{A} - \lambda_i \mathscr{E})^{r_i} \xi = 0, \ \xi \in V\right\}.$$

通常将 V_i 称为根子空间.

一般来说, 并不是对于每一个线性变换都有一组基, 使它在这组基下的矩阵成为对角形.

这一节的目的就是考察一般的一个方阵相似最简形式是什么. 我们的讨论限制在复数域中.

定义

形式为

$$J(\lambda, t) = \begin{pmatrix} \lambda & 0 & \cdots & 0 & 0 & 0 \\ 1 & \lambda & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & \lambda & 0 \\ 0 & 0 & \cdots & 0 & 1 & \lambda \end{pmatrix}_{t\times t}$$

的矩阵称为 *Jordan* 块, 其中 λ 是复数.

定义

由 *Jordan* 块组成的准对角矩阵称为 *Jordan* 形矩阵, 其一般形状如

$$\begin{pmatrix} J(\lambda_1, k_1) & & & \\ & J(\lambda_2, k_2) & & \\ & & \ddots & \\ & & & J(\lambda_s, k_s) \end{pmatrix}$$

并且 $\lambda_1, \lambda_2, \cdots, \lambda_s$ 中有一些可以相等.

例如

$$\begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} i & 0 \\ 1 & i \end{pmatrix}$$

都是若尔当块, 而

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{pmatrix}$$

是一个若尔当形矩阵.

注

$$\begin{pmatrix} \lambda & 0 & \cdots & 0 & 0 & 0 \\ 1 & \lambda & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & \lambda & 0 \\ 0 & 0 & \cdots & 0 & 1 & \lambda \end{pmatrix}_t \text{与} \begin{pmatrix} \lambda & 1 & \cdots & 0 & 0 & 0 \\ 0 & \lambda & \cdots & 1 & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \lambda & 1 \\ 0 & 0 & \cdots & 0 & 0 & \lambda \end{pmatrix}_t \text{相似.}$$

证明

令 $P = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 & 1 \\ 0 & 0 & \cdots & 0 & 1 & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 1 & \cdots & 0 & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 & 0 \end{pmatrix}_t$, 则 $P^{-1}J(\lambda,t)P$ 等于右边的矩阵.

注

- 上面的 $\lambda_1, \lambda_2, \cdots, \lambda_s$ 中有一些可以相等.
- 一阶 *Jordan* 块就是一阶矩阵, 因此 *Jordan* 形矩阵中包括对角矩阵.
- 因为 *Jordan* 形矩阵是下三角形矩阵, 所以, *Jordan* 标准形中, 主对角线上的元素正是特征多项式的全部的根 *(*重根按重数计算*)*.

定理 13

设复数域上线性空间 V 的线性变换 $\mathscr{A}$, 则存在 V 的一组基, 使得 $\mathscr{A}$ 在此组基下的矩阵为 *Jordan* 形矩阵.

定理 14

设 A 是复数域上的一个方阵, 则 A 相似于一个 *Jordan* 形矩阵.

定义

设 A 是数域 P 上 n 阶方阵，$f(x) \in P[x]$.

- 如果 $f(A) = 0$, 那么称 $f(x)$ 是 A 的零化多项式.
- 如果 $f(x)$ 首 1，且 $f(x)$ 是 A 的零化多项式中次数最低者, 那么称 A 的零化多项式 $f(x)$ 是 A 的最小多项式.

引理 1

数域 P 上 n 阶方阵 A 的最小多项式唯一

证明 设 $g_1(x)$ 和 $g_2(x)$ 都是 A 的最小多项式, 根据带余除法，$g_1(x)$ 可表成

$$g_1(x) = q(x)g_2(x) + r(x)$$

其中 $r(x) = 0$ 或 $\deg(r(x)) < \deg(g_2(x))$, 于是

$$g_1(A) = q(A)g_2(A) + r(A) = O$$

因此 $r(A) = O$. 由最小多项式的定义 , $r(x) = 0$, 即 $g_2(x) | g_1(x)$.
同样可证 $g_1(x)|g_2(x)$. 因此 $g_1(x)$ 与 $g_2(x)$ 只能相差一个非零常数因子.
又因 $g_1(x)$ 与 $g_2(x)$ 的首项系数都等于 1, 所以 $g_1(x) = g_2(x)$ ∎

引理 2

设 $m(x)$ 是方阵 A 的最小多项式，$f(x) \in P[x]$, 则

$$f(x) \text{ 是 } A \text{ 的零化多项式} \Leftrightarrow m(x)|f(x).$$

证明 $\Rightarrow$ 根据带余除法，$f(x)$ 可表成

$$f(x) = q(x)m(x) + r(x)$$

其中 $r(x) = 0$ 或 $\deg(r(x)) < \deg(m(x))$. 若 $r(x) \neq 0$, 则由

$$r(A) = f(A) - q(A)m(A) = O$$

可知 $r(x)$ 也是 A 的零化多项式. 但 $\deg(r(x)) < \deg(m(x))$ 与设 $m(x)$ 是方阵 A 的最小多项式矛盾. 因此,$r(x) = 0$, 进而 $m(x)|f(x)$.

$\Leftarrow$ 因为 $m(x)|f(x)$, 所以存在 $q(x) \in P[x]$ 使得 $f(x) = q(x)m(x)$. 又因为 $m(x)$ 是 A 的零化多项式, 所以 $f(A) = q(A)m(A) = O$. 即 $f(x)$ 是 A 的零化多项式. ∎

推论

最小多项式整除特征多项式.

例 1

- 数量矩阵 kE 的最小多项式是 $x-k$. 反过来,
- 最小多项式是一次的矩阵必是数量矩阵.

例 2

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

求 A 的最小多项式.

解 因为 A 的特征多项式为 $|xE-A|=(x-1)^3$, 所以 A 的最小多项式为 $(x-1)^3$ 的因式. 因为 $A-E\neq O$ 而 $(A-E)^2=O$, 所以 A 的最小多项式为 $(x-1)^2$.

如果矩阵 A 与 B 相似, 即存在可逆矩阵 T, 使得 $B = T^{-1}AT$, 那么对任一多项式 $f(x)$, 有 $f(B) = T^{-1}f(A)T$.

性质

因此 $f(B) = O$ 当且仅当 $f(A) = O$ 这说明相似矩阵有相同的最小多项式.

但是需要注意, 这个条件并不是充分的, 即最小多项式相同的矩阵不一定是相似的. 下面的例子说明这个结论例 3 设

$$A = \begin{pmatrix} 1 & 1 & & \\ & 1 & & \\ & & 1 & \\ & & & 2 \end{pmatrix}, B = \begin{pmatrix} 1 & 1 & & \\ & 1 & & \\ & & 2 & \\ & & & 2 \end{pmatrix}$$

A 与 B 的最小多项式都等于 $(x-1)^2(x-2)$, 但是它们的特征手项式不同, 因此 A 和 B 不是相似的.

引理 3

设 A 是一个准对角矩阵

$$A = \begin{bmatrix} A_1 & \\ & A_2 \end{bmatrix}$$

并设 A_1 的最小多项式为 $g_1(x), A_2$ 的最小多项式为 $g_2(x)$, 那么 A 的最小多项式为 $g_1(x), g_2(x)$ 的最小公倍式 $[g_1(x), g_2(x)]$

证明 记 $g(x) = [g_1(x), g_2(x)]$, 首先

$$g(A) = \begin{bmatrix} g(A_1) & \\ & g(A_2) \end{bmatrix} = O$$

因此 $g(x)$ 能被 A 的最小各项式整除. 其次, 如果 $h(A) = O$, 那么

$$h(A) = \begin{bmatrix} h(A_1) & \\ & h(A_2) \end{bmatrix} = O$$

所以 $h(A_1) = O, h(A_2) = O$. 因而 $g_1(x) | h(x), g_2(x) | h(x)$ 并由此得 $g(x) | h(x)$. 这样就证明了 $g(x)$ 是 $\mathbf{A}$ 的最小多项式. ∎

这个结论可以推广到 A 为若干个矩阵组成的准对角矩阵的情形.

推论

如果

$$A = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A \end{pmatrix}$$

A_i 的最小多项式为 $g_i(x), i = 1, 2, \cdots, s$, 那么 A 的最小多项式为 $[g_1(x), g_2(x), \cdots, g_s(x)]$

引理 4

k 级 *Jordan* 块

$$J = \begin{pmatrix} a & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & a \end{pmatrix}$$

的最小多项式为 $(x-a)^k$.

证明　J 的特征多项式为 $(x-a)^k$, 而所以 J 的最小多项式为 $(x-a)^k$. ∎

定理 15

数域 P 上 n 级矩阵 A 与对角矩阵相似 $\Leftrightarrow$ A 的最小多项式是 P 上互素的一次因式的乘积.

推论

复数矩阵 A 与对角矩阵相似 $\Leftrightarrow$ A 的最小多项式没有重根.

例 3

设数域 P 上的 n 级矩阵 A 满足 $A^2 = E$. 证明 A 相似于对角形.

证明 由 $A^2 = E$ 可知 $x^2 - 1$ 是 A 的一个零化多项式. 因此, A 的最小多项式是 $x^2 - 1$ 的因式，从而可以分解为数域 P 上互素的一次因式的乘积，所以 A 相似于对角形. ∎

例 4

设数域 P 上的 n 级矩阵 A 满足 $A^2 = A$. 证明 A 相似于对角形.

证明 由 $A^2 = A$ 可知 $x^2 - x$ 是 A 的一个零化多项式. 因此, A 的最小多项式是 $x^2 - x$ 的因式，从而可以分解为数域 P 上互素的一次因式的乘积，所以 A 相似于对角形. ∎

性质

如果 $f(\lambda)$ 是 A 的零化多项式，则 A 的所有特征值 λ_i 都是 $f(\lambda)$ 的根。特别地，A 的所有特征值都是 A 的最小多项式 $d_A(\lambda)$ 的根.

证明 设 X 是属于特征值 λ_i 的任一特征向量, 则 $AX = \lambda_i X = \mathbf{0}$.
由此推出 $A^2X = A(AX) = A\lambda_i X = \lambda_i AX = \lambda_i(\lambda_i X) = \lambda_i^2 X$.
一般地，用数学归纳法可以证明 $A^m X = \lambda_i^m X$ 对所有的正整数 m 成立.
设 $f(\lambda) = a_0 + a_1\lambda + \cdots + a_m\lambda^m$，则

$$\begin{aligned} f(A)X &= a_0X + a_1AX + \cdots + a_mA^mX \\ &= a_0X + a_1\lambda_iX + \cdots + a_m\lambda_i^mX = f(\lambda_i)X. \end{aligned}$$

由 $f(A) = O$ 知 $f(A)X = \mathbf{0}$, 从而 $f(\lambda_i)X = \mathbf{0}$.
又因为 $X \neq \mathbf{0}$，所以 $f(\lambda_i) = 0$. ∎

例 1 (与 §3 的例6类似)

取 P^3 的线性变换 $\sigma(a,b,c)=(a-b,b-c,a+b)$

1. 求 σ 在基 $\varepsilon_1=(1,0,0),\varepsilon_2=(0,1,0),\varepsilon_3=(0,0,1)$ 下的矩阵
2. 求 σ 在基 $\eta_1=(1,0,0),\eta_2=(1,1,0),\eta_3=(1,1,1)$ 下的矩阵
3. 求向量 $\alpha=(1,2,3)$ 的像 $\sigma\alpha$ 分别在基 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 和 η_1,η_2,η_3 下的坐标.

例 2 (与 §6 的例3类似)

在线性空间 P^3 中，定义线性变换

$$\sigma(a,b,c) = (a+2b-c,\, b+c,\, a+b-2c),$$

求 σ 的值域与核.

例 3 (课本习题 10)

设 σ 是 n 维线性空间 V 上的一个线性变换，若有向量 ξ 满足 $\sigma^{n-1}\xi \neq 0, \sigma^n\xi = 0$, 证明 $\xi, \sigma\xi, \cdots, \sigma^{n-1}\xi$ 是 V 的一组基.

例 4

设 $A \in P^{n\times n}$, 且 $A^2 = E$. 记

$V_1 = \{X | X \in P^n, (A - E)X = 0\}, V_2 = \{X | X \in P^n, (A + E)X = 0\}$.

1. 证明 $V_1 \oplus V_2 = P^n$.
2. 证明存在可逆矩阵 C 使得 $C^{-1}AC$ 是对角阵.

例 5

设 $A \in P^{n\times n}$, 且 $A^2 = A$. 记

$V_1 = \{X | X \in P^n, (A - E)X = 0\}, V_2 = \{X | X \in P^n, AX = 0\}$.

1. 证明 $V_1 \oplus V_2 = P^n$.
2. 证明存在可逆矩阵 C, 使得 $C^{-1}AC$ 是对角阵.