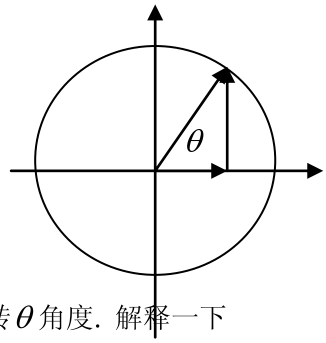
第七章　线性变换

§1 线性变换的定义

注：本章的资料中，数域是用F表示的，不是用的P.

1. 定义: 取线性空间V,从V到自身的变换$\sigma:V\to V$若满足对任意$\alpha,\beta\in V$及任意$k\in F$,都有

$\sigma(\alpha+\beta)=\sigma(\alpha)+\sigma(\beta)$,$\sigma(k\alpha)=k\sigma(\alpha)$,则称$\sigma$是$V$的一个线性变换.



2. 例子: (1) $\sigma:V\to V;\sigma(\alpha)=\alpha$:恒等变换.单位变换.

(2) $\sigma:V\to V;\sigma(\alpha)=0$:零变换.

(3) 固定$k\in F$,$\sigma:V\to V;\sigma(\alpha)=k\alpha$,数乘变换.

(4) 取$V=\mathbf{R}^2=\{(a,b)\mid a,b\in\mathbf{R}\}$,定义$\sigma:V\to V$为按照逆时针方向旋转$\theta$角度. 解释一下

看$\varepsilon_1=(1,0),\varepsilon_2=(0,1)$旋转$\theta$角度后变成什么.$\sigma(\varepsilon_1)=(\cos\theta,\sin\theta)$,$\sigma(\varepsilon_2)=(-\sin\theta,\cos\theta)$,

取$\alpha=(a,b)=a\varepsilon_1+b\varepsilon_2$,

$\sigma(\alpha)=a\sigma(\varepsilon_1)+b\sigma(\varepsilon_2)=a(\cos\theta,\sin\theta)+b(-\sin\theta,\cos\theta)=(a\cos\theta-b\sin\theta,a\sin\theta+b\cos\theta)$.

$$\sigma:V\to V,\alpha=\begin{pmatrix}a\\b\end{pmatrix}\mapsto\sigma(\alpha)=\begin{pmatrix}a\cos\theta-b\sin\theta\\a\sin\theta+b\cos\theta\end{pmatrix}=\begin{pmatrix}\cos\theta&-\sin\theta\\\sin\theta&\cos\theta\end{pmatrix}\alpha.$$

(5) 正交投影.固定一个向量α,$\sigma:F^3\to F^3,\xi\mapsto\sigma(\xi)=|\xi|\cos<\xi,\alpha>\dfrac{\alpha}{|\alpha|}=\dfrac{(\xi,\alpha)}{|\alpha|^2}\alpha=\dfrac{(\alpha,\xi)}{(\alpha,\alpha)}\alpha$.

(6) 求微商$\sigma:F[x]\to F[x],f(x)\mapsto f'(x)$.

3. 性质. (1) 设线性变换$\sigma:V\to V$,则$\sigma(0)=0,\sigma(-\alpha)=-\sigma(\alpha)$

(2) 线性变换保持线性组合和线性关系式不变.**解释**

(3) 线性变换把线性相关的向量组变为线性相关的向量组. 但是对线性无关的向量组不成立.**解释**

§2 线性变换的运算

取线性空间${}_FV$的两个线性变换$\sigma,\tau:V\to V$.这一节来定义线性变换的运算.

1 线性变换的和(加法运算)

给出V的任意两个线性变换$\sigma,\tau:V\to V$,任给$\alpha\in V$,定义$\sigma+\tau$为:$(\sigma+\tau)(\alpha)=\sigma(\alpha)+\tau(\alpha)$.

线性变换的和仍然是线性变换.

$(\sigma+\tau)(\alpha+\beta)=\sigma(\alpha+\beta)+\tau(\alpha+\beta)=\sigma(\alpha)+\sigma(\beta)+\tau(\alpha)+\tau(\beta)=\sigma(\alpha)+\tau(\alpha)+\sigma(\beta)+\tau(\beta)$

$=(\sigma+\tau)(\alpha)+(\sigma+\tau)(\beta)$

$(\sigma+\tau)(k\alpha)=\sigma(k\alpha)+\tau(k\alpha)=k\sigma(\alpha)+k\tau(\alpha)=k(\sigma(\alpha)+\tau(\alpha))=k(\sigma+\tau)(\alpha)$.

性质: 加法的交换律: $\sigma+\tau=\tau+\sigma$. 结合律: $(\sigma_1+\sigma_2)+\sigma_3=\sigma_1+(\sigma_2+\sigma_3)$. $\sigma+0=0+\sigma=\sigma$

2 线性变换的乘积(乘法运算)

给出线性空间V的任意两个线性变换$\sigma,\tau:V\to V$,任给$\alpha\in V$,定义$\sigma\tau$为: $(\sigma\tau)(\alpha)=\sigma(\tau(\alpha))$

即线性变换的合成, $V\xrightarrow{\tau}V\xrightarrow{\sigma}V$. 线性变换的乘积仍然是线性变换.

$(\sigma\tau)(k\alpha+l\beta)=\sigma(\tau(k\alpha+l\beta))=\sigma(k\tau(\alpha)+l\tau(\beta))=k\sigma(\tau(\alpha))+l\sigma(\tau(\beta))=k(\sigma\tau)(\alpha)+l(\sigma\tau)(\beta)$

性质: 结合律: $(\sigma_1\sigma_2)\sigma_3=\sigma_1(\sigma_2\sigma_3)$,一般不满足交换律.

如微分和求导的运算的合成.设$\sigma(f(x))=f'(x),\tau(f(x))=\int f(x)dx$.

则$(\sigma\tau)(f(x))=\left(\int f(x)dx\right)'=f(x)$,但是$(\tau\sigma)(f(x))=\int f'(x)dx=f(x)+c$.

对于单位变换, $id\cdot\sigma=\sigma\cdot id$. 左右分配律: $(\sigma_1+\sigma_2)\tau=\sigma_1\tau+\sigma_2\tau$, $\tau(\sigma_1+\sigma_2)=\tau\sigma_1+\tau\sigma_2$.

3 线性变换的负变换. 任给$\sigma:V\to V$,定义$(-\sigma)(\alpha)=-\sigma(\alpha)$.定义线性变换的差: $\sigma-\tau=\sigma+(-\tau)$.从而 $\sigma-\sigma=0$.

4 数量乘法: 任取数$k\in F$,任给线性变换$\sigma:V\to V$,则定义数k和σ的数乘变换$(k\sigma)(\alpha)=k\sigma(\alpha)$.

性质: $1\sigma=\sigma$, $(k+l)\sigma=k\sigma+l\sigma$, $k(\sigma+\tau)=k\sigma+k\tau$, $k(l\sigma)=(kl)\sigma$.

以上定义的加法和数量乘积是$L(V)=\{\sigma:V\to V\mid \sigma\text{是}V\text{的线性变换}\}$上的两个运算,集合$L(V)$关于这两个运算作成数域$F$上的一个线性空间.

5 线性变换的逆.

任给线性变换$\sigma:V\to V$,若存在$\tau:V\to V$,使得$\sigma\tau=\tau\sigma=id$,则称σ可逆.τ为σ的逆,记$\tau=\sigma^{-1}$.

线性变换的逆变换也是线性变换.

$\sigma^{-1}(\alpha+\beta)=\sigma^{-1}(\sigma\sigma^{-1}(\alpha)+\sigma\sigma^{-1}(\beta))=\sigma^{-1}(\sigma(\sigma^{-1}(\alpha)+\sigma^{-1}(\beta)))=\sigma^{-1}(\alpha)+\sigma^{-1}(\beta)$

$\sigma^{-1}(k\alpha)=\sigma^{-1}(k\sigma\sigma^{-1}(\alpha))=\sigma^{-1}(\sigma(k\sigma^{-1}(\alpha)))=k\sigma^{-1}(\alpha)$.

6 线性变换的方幂.

设线性变换$\sigma:V\to V$,任给$n\in\mathbf{Z}_{>0}$,定义$\sigma^n=\sigma^{n-1}\sigma$, 即$n$个$\sigma$的合成.特别的,定义$\sigma^0=id$.

若σ可逆,则定义$\sigma^{-n}=(\sigma^{-1})^n$,实际上$\sigma^{-n}=(\sigma^n)^{-1}$,则

$\sigma^{mn}=(\sigma^m)^n,\sigma^m\sigma^n=\sigma^{m+n}$.注: $(\sigma\tau)^{mn}\neq\sigma^{mn}\tau^{mn}$.

7 线性变换的多项式.

设多项式 $f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$,任给线性变换$\sigma:V\to V$,定义

$f(\sigma)=a_n\sigma^n+a_{n-1}\sigma^{n-1}+\cdots+a_1\sigma+a_0$,也是线性空间$V$的一个线性变换,实际上$f(\sigma)$是线性变换

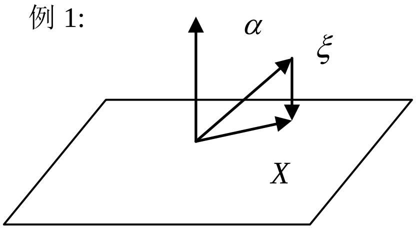
$\sigma^n,\sigma^{n-1},\cdots,\sigma,id$ 的一个线性组合.

性质: 若 $f(x)g(x)=h(x)$,则 $f(\sigma)g(\sigma)=h(\sigma)$,且 $f(\sigma)g(\sigma)=g(\sigma)f(\sigma)$.

若 $f(x)+g(x)=h(x)$,则 $f(\sigma)+g(\sigma)=h(\sigma)$.

线性组合的例子

例 1:



考察ξ关于α的内摄影$\pi_\alpha(\xi)=\dfrac{(\alpha,\xi)}{(\alpha,\alpha)}\alpha$,

ξ在平面X上的内摄影$\pi_X(\xi)$.

由图,得到$\pi_X(\xi)=\xi-\pi_\alpha(\xi)=(id-\pi_\alpha)(\xi)$,从而$\pi_X=id-\pi_\alpha$.

若考察ξ关于平面的投影$R_X(\xi)$,则$R_X(\xi)=\xi-2\pi_\alpha(\xi)=(id-2\pi_\alpha)(\xi)$,则$R_X=id-2\pi_\alpha$.

例 2: 对线性空间$F[x]_n$.取线性变换$D(f(x))=f'(x)$,则$D^n=0$.固定$a\in F$,做平移变换

$\tau_a(f(x))=f(x+a)$. $f(x+a)=f(x)+f'(x)a+\dfrac{f''(x)}{2!}a^2+\cdots+\dfrac{f^{(n-1)}(x)}{(n-1)!}a^{n-1}$.

从而$\tau_a(f(x))=id(f(x))+D(f(x))+\dfrac{a^2}{2!}D^2(f(x))+\cdots+\dfrac{a^{n-1}}{(n-1)!}D^{n-1}(f(x))$.

则$\tau_a=id+D+\dfrac{a^2}{2!}D^2+\cdots+\dfrac{a^{n-1}}{(n-1)!}D^{n-1}$.

§3 线性变换的矩阵

由第二节,我们已知线性空间V的线性变换的全体作成数域F上的一个线性空间.我们来看这个线性空间的维数和一组基是什么?

一 $L(V)\cong F^{n\times n}$.

1 两个结论

设线性空间${}_FV$, $\dim V=n$,取一组基$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$,任给线性变换$\sigma\in L(V)$,任给$\xi\in V$,设

$$\xi=x_1\varepsilon_1+x_2\varepsilon_2+\cdots+x_n\varepsilon_n=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)X,$$

则$\sigma\xi\in V$,而$\sigma\xi=x_1\sigma\varepsilon_1+x_2\sigma\varepsilon_2+\cdots+x_n\sigma\varepsilon_n=(\sigma\varepsilon_1,\sigma\varepsilon_2,\cdots,\sigma\varepsilon_n)X$.

从而要想求任意一个向量在σ下的像,只要知道给定的一组基的像即可.

结论 1 设$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$是线性空间${}_FV$的一组基,若${}_FV$的两个线性变换$\sigma$与$\tau$满足$\sigma(\varepsilon_i)=\tau(\varepsilon_i)$,则$\sigma=\tau$.即线性变换完全由它在基上的作用所决定.

证明 任给$\alpha\in V$,证明$\sigma(\alpha)=\tau(\alpha)$.设$\alpha=x_1\varepsilon_1+x_2\varepsilon_2+\cdots+x_n\varepsilon_n$,则

$\sigma(\alpha)=x_1\sigma\varepsilon_1+x_2\sigma\varepsilon_2+\cdots+x_n\sigma\varepsilon_n=x_1\tau\varepsilon_1+x_2\tau\varepsilon_2+\cdots+x_n\tau\varepsilon_n=\tau(\alpha)$.

结论 2 设$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$是线性空间${}_FV$的一组基,对于$V$中任意一组向量$\alpha_1,\alpha_2,\cdots,\alpha_n$,存在一个线性变换$\sigma\in L(V)$,使得$\sigma(\varepsilon_i)=\alpha_i$, $\forall i=1,2,\cdots,n$.

证明 任给$\alpha\in V$,设$\alpha=x_1\varepsilon_1+x_2\varepsilon_2+\cdots+x_n\varepsilon_n$,定义$\sigma:V\to V;\sigma(\alpha)=x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n$. 则$\sigma\in L(V)$,且满足$\sigma(\varepsilon_i)=\alpha_i$, $\forall i=1,2,\cdots,n$.

定理 1: 设$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$是线性空间${}_FV$的一组基, $\alpha_1,\alpha_2,\cdots,\alpha_n$是$V$中任意一组向量,则存在唯一的一个线性变换$\sigma\in L(V)$,使得$\sigma(\varepsilon_i)=\alpha_i$, $\forall i=1,2,\cdots,n$.

2 线性变换在一组基下的矩阵

取线性空间${}_FV$,取定一组基$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$,任给$\sigma\in L(V)$,则$\sigma\varepsilon_1,\sigma\varepsilon_2,\cdots,\sigma\varepsilon_n$仍然是$V$中的向量.

从而$\sigma\varepsilon_1,\sigma\varepsilon_2,\cdots,\sigma\varepsilon_n$可由$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$线性表出.设

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

形式的写法$(\sigma\varepsilon_1,\sigma\varepsilon_2,\cdots,\sigma\varepsilon_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{pmatrix}$,

记$A=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{pmatrix}$,称$A$为线性变换$\sigma$在基$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$下的矩阵.简记$\sigma\xrightarrow{\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n}A$.

从而$\sigma\xrightarrow{\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n}A$等价于$(\sigma\varepsilon_1,\sigma\varepsilon_2,\cdots,\sigma\varepsilon_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)A$.即矩阵$A$的列向量组是$\sigma\varepsilon_1,\sigma\varepsilon_2,\cdots,\sigma\varepsilon_n$在基$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$下的坐标..

例 1 设微分变换$D:F[x]_n\to F[x]_n; f(x)\mapsto f'(x)$.

取一组基为$1,x,x^2,\cdots,x^{n-1}$,在D下的象为$0,1,2x,\cdots,(n-1)x^{n-2}$,从而

$$(0,1,2x,\cdots,(n-1)x^{n-2})=(1,x,x^2,\cdots,x^{n-1})\begin{pmatrix}0&1&0&\cdots&0&0\\0&0&2&\cdots&0&0\\0&0&0&\cdots&0&0\\\vdots&\vdots&\vdots&&\vdots&\vdots\\0&0&0&\cdots&0&n-1\\0&0&0&\cdots&0&0\end{pmatrix}$$

取一组基为$1,x,\frac{1}{2!}x^2,\cdots,\frac{1}{(n-1)!}x^{n-1}$,象为$0,1,x,\frac{1}{2!}x^2\cdots,\frac{1}{(n-2)!}x^{n-2}$,得

$$(0,1,x,\frac{1}{2!}x^2\cdots,\frac{1}{(n-2)!}x^{n-2})=(1,x,\frac{1}{2!}x^2,\cdots,\frac{1}{(n-1)!}x^{n-1})\begin{pmatrix}0&1&0&\cdots&0&0\\0&0&1&\cdots&0&0\\0&0&0&\cdots&0&0\\\vdots&\vdots&\vdots&&\vdots&\vdots\\0&0&0&\cdots&0&1\\0&0&0&\cdots&0&0\end{pmatrix}.$$

取一组基$1,x-a,(x-a)^2,\cdots,(x-a)^{n-1}$,象为$0,1,2(x-a),\cdots,(n-1)(x-a)^{n-2}$,得

$$(0,1,2(x-a),\cdots,(n-1)(x-a)^{n-2})=(1,x-a,(x-a)^2,\cdots,(x-a)^{n-1})\begin{pmatrix}0&1&0&\cdots&0&0\\0&0&2&\cdots&0&0\\0&0&0&\cdots&0&0\\\vdots&\vdots&\vdots&&\vdots&\vdots\\0&0&0&\cdots&0&n-1\\0&0&0&\cdots&0&0\end{pmatrix}$$

例 2 设n维线性空间${}_FV$,W是一个m维子空间,$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_m$是$W$的一组基,可以扩充为${}_FV$的一组基$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_m,\varepsilon_{m+1},\cdots,\varepsilon_n$,定义$\sigma$为$\sigma(\varepsilon_i)=\varepsilon_i,i=1,2,\cdots,m,\sigma(\varepsilon_j)=0,j=m+1,\cdots,n$,$\sigma$是${}_FV$的一个线性变换,且$\sigma$在$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_m,\varepsilon_{m+1},\cdots,\varepsilon_n$下的矩阵为$\begin{pmatrix}E_m&0\\0&0\end{pmatrix}$.

这个线性变换σ限制在子空间W上是W的一个线性变换,是W的单位变换.

例 3 设3维线性空间${}_FV$,$\varepsilon_1,\varepsilon_2,\varepsilon_3$是一组基,定义$\sigma:V\to V:\varepsilon_1\mapsto\varepsilon_2,\varepsilon_2\mapsto\varepsilon_3,\varepsilon_3\mapsto\varepsilon_1$,

则σ在$\varepsilon_1,\varepsilon_2,\varepsilon_3$下的矩阵为$(\sigma\varepsilon_1,\sigma\varepsilon_2,\sigma\varepsilon_3)=(\varepsilon_1,\varepsilon_2,\varepsilon_3)\begin{pmatrix}0&0&1\\1&0&0\\0&1&0\end{pmatrix}$.

3 $L(V)\cong F^{n\times n}$

设n维线性空间V,固定一组基$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$,任取$\sigma\in L(V)$,设$\sigma$在$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$下的矩阵为$A$,定义

$f:L(V)\to F^{n\times n}; f(\sigma)=A$.可证明这一映射是个同构.

定理 2 设$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$是$n$维线性空间$V$的一组基,线性变换在$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$下的矩阵定义了$L(V)$到$F^{n\times n}$的一个对应$f:L(V)\to F^{n\times n}; f(\sigma)=A$,这个对应满足:

(1) 线性变换的和对应于矩阵的和:$\sigma+\tau\mapsto A+B$,
(2) 线性变换的乘积对应于矩阵的乘积:$\sigma\tau\mapsto AB$,
(3) 可逆的线性变换对应的矩阵可逆,且线性变换的逆对应于矩阵的逆:$\sigma^{-1}\mapsto A^{-1}$,
(4) 线性变换的数乘对应于矩阵的数乘:$k\sigma\mapsto kA$.

证明 (1) 线性变换的和对应于矩阵的和. $\sigma+\tau\mapsto A+B$

$$(\sigma+\tau)(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)=\sigma(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)+\tau(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)A+(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)B$$

$$=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)(A+B)$$

(2) $(\sigma\tau)(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)=\sigma(\tau(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n))=\sigma((\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)B)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)AB$

从而同构$L(V)\cong F^{n\times n}$,故$\dim L(V)=n^2$,有n^2个线性变换组成的向量组线性无关.

一组基如何取?实际上$F^{n\times n}$的一组基的原象即可.

4 坐标变换公式

同一个向量在两组基之下的坐标有相应的坐标变换公式.这里类似的公式如下.

取n维线性空间V,取一组基$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$,任给$\sigma\in L(V)$,设$\sigma$在$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$下的矩阵为$A$.

即$(\sigma\varepsilon_1,\sigma\varepsilon_2,\cdots,\sigma\varepsilon_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)A$.

现在任给$\xi\in V$,设$\xi=x_1\varepsilon_1+x_2\varepsilon_2+\cdots+x_n\varepsilon_n=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)X$.

则$\sigma\xi\in V$,设$\sigma\xi=y_1\varepsilon_1+y_2\varepsilon_2+\cdots+y_n\varepsilon_n=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)Y$, $Y=?$.

$\sigma\xi=x_1\sigma\varepsilon_1+x_2\sigma\varepsilon_2+\cdots+x_n\sigma\varepsilon_n=(\sigma\varepsilon_1,\sigma\varepsilon_2,\cdots,\sigma\varepsilon_n)X=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)AX$,从而$Y=AX$.

定理:设σ在$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$下的矩阵为$A$,向量$\xi$在$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$下的坐标是$X$,则$\sigma\xi$在$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$下的坐标是$AX$.

二 矩阵的相似

1 定义 设$A,B\in F^{n\times n}$,若存在可逆矩阵$C\in F^{n\times n}$,使得$B=C^{-1}AC$,则称A相似于B,记为$A\sim_X B$.

性质: 相似是一种等价关系.

$A=E^{-1}AE$;若$B=C^{-1}AC$,则$A=CBC^{-1}$;若$B=X_1^{-1}AX_1, C=X_2^{-1}BX_2$, 则$C=(X_1X_2)^{-1}AX_1X_2$

2 定理 线性变换在不同基下的矩阵是相似的.反之,若两个矩阵相似,则此矩阵可看作是同一个线性变换在两组基下的矩阵.

证明 设两组基$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$与$\eta_1,\eta_2,\cdots,\eta_n$,取$\sigma\in L(V)$,设

$\sigma(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)A,\sigma(\eta_1,\eta_2,\cdots,\eta_n)=(\eta_1,\eta_2,\cdots,\eta_n)B$.

并设从$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$到$\eta_1,\eta_2,\cdots,\eta_n$的过渡矩阵为$X:(\eta_1,\eta_2,\cdots,\eta_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)X$.代入上式有

$\sigma((\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)X)=((\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)X)B$,即$(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)AX=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)XB$,从而有

$AX=XB$,即$B=X^{-1}AX$.

反之,若有$B=X^{-1}AX$,对矩阵A,由于$L(V)\cong F^{n\times n}$,固定线性空间的一组基$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$,存在$\sigma\in L(V)$,使得$\sigma$在$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$下的矩阵为$A$.令$(\eta_1,\eta_2,\cdots,\eta_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)X$,则$\eta_1,\eta_2,\cdots,\eta_n$也是线性空间的一组基,且$\sigma$在$\eta_1,\eta_2,\cdots,\eta_n$下的矩阵为

$$\sigma(\eta_1,\eta_2,\cdots,\eta_n)=\sigma(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)X=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)AX=(\eta_1,\eta_2,\cdots,\eta_n)X^{-1}AX,$$

即σ在$\eta_1,\eta_2,\cdots,\eta_n$下的矩阵为$B$.

3. 相似的性质.

(1) 设$A_1\sim_X B_1,A_2\sim_X B_2$,则有$A_1+A_2\sim_X B_1+B_2,A_1A_2\sim_X B_1B_2$.

事实上,$B_1=X^{-1}A_1X,B_2=X^{-1}A_2X$,则$B_1+B_2=X^{-1}A_1X+X^{-1}A_2X=X^{-1}(A_1+A_2)X$.

$B_1B_2=X^{-1}A_1XX^{-1}A_2X=X^{-1}A_1A_2X$.

(2) 若$B=X^{-1}AX$,则$B^n=(X^{-1}AX)^n=X^{-1}A^nX$,$kB=X^{-1}(kA)X$.

设$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$,则$f(A)=a_nA^n+a_{n-1}A^{n-1}+\cdots+a_1A+a_0E$.从而有

$f(B)=a_nB^n+a_{n-1}B^{n-1}+\cdots+a_1B+a_0E=X^{-1}f(A)X$.

例 设V是F上2维线性空间,一组基$\varepsilon_1,\varepsilon_2$,$\sigma\in L(V)$在基$\varepsilon_1,\varepsilon_2$下的矩阵为$A=\begin{pmatrix}2&1\\-1&0\end{pmatrix}$,求$A^n$.

解 令$(\eta_1,\eta_2)=(\varepsilon_1,\varepsilon_2)\begin{pmatrix}1&-1\\-1&2\end{pmatrix}$,计算$\sigma\in L(V)$在基$\eta_1,\eta_2$下的矩阵$B$.记$X=\begin{pmatrix}1&-1\\-1&2\end{pmatrix}$.则

$\sigma(\eta_1,\eta_2)=\sigma(\varepsilon_1,\varepsilon_2)X=(\varepsilon_1,\varepsilon_2)AX=(\eta_1,\eta_2)X^{-1}AX$.

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

从而 $A=X\begin{pmatrix}1&1\\0&1\end{pmatrix}X^{-1}$.则 $A^n=X\begin{pmatrix}1&n\\0&1\end{pmatrix}X^{-1}=\begin{pmatrix}n+1&n\\-n&1-n\end{pmatrix}$.

补充相似的性质.

1. 若 $A\sim B$,则 $A^n\sim B^n$, $kA\sim kB$. $|A|=|B|$

2. 若 $A_i\sim_{X_i} B_i$, $i=1,2,\cdots,s$,则 $\begin{pmatrix}A_1&&&\\&A_2&&\\&&\ddots&\\&&&A_s\end{pmatrix}\sim_{\begin{pmatrix}X_1&&&\\&X_2&&\\&&\ddots&\\&&&X_s\end{pmatrix}}\begin{pmatrix}B_1&&&\\&B_2&&\\&&\ddots&\\&&&B_s\end{pmatrix}$

3. 若 $A\sim B$,且 A 可逆,则 B 也可逆,且 $A^{-1}\sim B^{-1}$.

§4 特征值与特征向量

这一节是关于线性变换的特征值与特征向量,但是由于线性变换与矩阵之间的一一对应,故也可以定义矩阵的特征值与特征向量.

这一节的目的:我们知道,线性空间取定一组基后,线性变换在这组基下有个矩阵,同一个线性变换在另外一组基下也有个矩阵,这两个矩阵是相似的.

现在就问:对于一个线性变换,能不能找到一组基,使得线性变换在此组基下的矩阵形式比较简单.

1. **特征值与特征向量的定义**:

设线性空间 V,取线性变换 $\sigma\in L(V)$,对数 $\lambda_0\in F$,若存在非零向量 $\xi\in V$,使得 $\sigma\xi=\lambda_0\xi$,则称 λ_0 是 σ 的一个特征值, ξ 是 σ 的属于特征值 λ_0 的一个特征向量.

几何意义: 线性变换对于特征向量的作用实际上是向量的增长或者缩短.

注:(1) σ 的特征值 λ_0: $\lambda_0\in F$,可正,可负,可为零,若 $\lambda_0=0$,则 $\sigma\xi=0$.

(2) 设 ξ 是 σ 的属于特征值 λ_0 的一个特征向量, $\sigma\xi=\lambda_0\xi$,任取 $k\in F$,则 $\sigma(k\xi)=k\sigma\xi=\lambda_0 k\xi$,从而 $k\xi$ 也是 σ 的属于特征值 λ_0 的一个特征向量.

设 $\sigma\xi=\lambda_0\xi,\sigma\eta=\lambda_0\eta$,则 $\sigma(\xi+\eta)=\sigma\xi+\sigma\eta=\lambda_0\xi+\lambda_0\eta=\lambda_0(\xi+\eta)$,

这说明,线性变换的属于某个相同特征值的特征向量的非零线性组合也是属于这个特征值的特征向量.从而属于同一个特征值的特征向量可以不唯一.即特征向量不是由特征值唯一决定.

(3) $\sigma\xi=\lambda_1\xi,\sigma\xi=\lambda_2\xi$,则 $\lambda_1\xi=\lambda_2\xi$,由于 $\xi\neq 0$,则 $\lambda_1=\lambda_2$.说明一个特征向量不能属于两个特征值.

属于同一个特征值的多个特征向量:存在.线性组合就是. 属于不同特征值的同一个特征向量:不存在.

2. 特征值与特征向量满足的条件.

取线性空间 V,固定一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$,取 $\sigma\in L(V)$,设 $\sigma\xi=\lambda_0\xi$.假设 σ 在 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵为 A, $\sigma(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)A$.

对ξ,设$\xi = x_{01}\varepsilon_1 + x_{02}\varepsilon_2 + \cdots + x_{0n}\varepsilon_n = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)X_0$,则

$$\sigma\xi = \sigma(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)X_0 = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)AX_0 = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)\lambda_0 X_0,$$

从而$AX_0 = \lambda_0 X_0$.改写一下,得到$0 = \lambda_0 X_0 - AX_0 = (\lambda_0 E - A)X_0$,即特征向量$\xi$在基$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$下的坐标$X_0$是齐次线性方程组$(\lambda_0 E - A)X = 0$的一个解.由于特征向量非零,坐标非零,则这个解是非零解.

既然齐次线性方程组$(\lambda_0 E - A)X = 0$有非零解,则线性方程组的系数矩阵的行列式应该等于零,即$|\lambda_0 E - A| = 0$,即说明特征值λ_0满足$|\lambda E - A| = 0$.

看$|\lambda E - A|$.

$$|\lambda E - A| = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix} = (\lambda - a_{11})(\lambda - a_{22})\cdots(\lambda - a_{nn}) + \text{展开式中的其余各项}$$

$$= \lambda^n - (\sum_{i=1}^{n} a_{ii})\lambda^{n-1} + \text{次数} \leq n-2 \text{的项}.$$

是关于λ的首1的n次多项式$= \lambda^n - (\sum_{i=1}^{n} a_{ii})\lambda^{n-1} + \cdots + (-1)^n |A|$.故特征值$\lambda_0$是$|\lambda E - A| = 0$的一个根.

3. 特征多项式的定义.设A是一个n阶方阵,

$$f_A(\lambda) = |\lambda E - A| = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix}$$称为矩阵A的特征多项式.这是一个关于λ的首1的n次多项式.

综上分析可知:假设λ_0是σ的一个特征值, ξ是σ的属于特征值λ_0的一个特征向量,即$\sigma\xi = \lambda_0\xi$.任意取定线性空间的一组基$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$,则首先$\sigma$在$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$下有矩阵$A$,从而特征值满足特征多项式$|\lambda E - A| = 0$.即$|\lambda E - A| = 0$的根就是$\sigma$的特征值.假设$\lambda_i$是其一个根,求属于$\lambda_i$的特征向量,这样的特征向量在$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$下的坐标是齐次线性方程组$(\lambda_i E - A)X = 0$的解,从而求得$(\lambda_i E - A)X = 0$的解,就可得到属于$\lambda_i$的特征向量.

4. 求特征值和特征向量的步骤.

(1). 对所给的线性变换,任给一组基$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$,先求线性变换在这一组基下的矩阵$A$.

(2). 求 A 的特征多项式 $|\lambda E-A|=0$ 的所有根.不妨设为 $\lambda_1,\lambda_2,\cdots,\lambda_s$,其重数分别为 $n_1,n_2,\cdots,n_s$.

(3). 对每一个特征值 λ_i,求齐次线性方程组 $(\lambda_i E-A)X=0$ 的一组基础解系,设为 $\eta_{i1},\eta_{i2},\cdots,\eta_{im_i}$,其中 $m_i=r(\lambda_i E-A)$.以基础解系为坐标的向量 $\xi_{i1},\xi_{i2},\cdots,\xi_{im_i}$ 就是属于 λ_i 的线性无关的一组特征向量.即 $\xi_{ij}=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)\eta_{ij}$, $j=1,2,\cdots,m_i$.而 $\xi_{i1},\xi_{i2},\cdots,\xi_{im_i}$ 的非零线性组合 $k_{i1}\xi_{i1}+k_{i2}\xi_{i2}+\cdots+k_{im_i}\xi_{im_i}$ (其中 $k_{i1},k_{i2},\cdots,k_{im_i}$ 不全为零)就是对应于 λ_i 的所有特征向量.

5. 例子

1) 取线性变换为数乘变换. 任给 $k\in F$,定义 $\sigma:V\to V;\sigma\alpha=k\alpha$.由定义,特征值为 k,任给非零向量都是对应的特征向量.事实上,取一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$, $\sigma\varepsilon_i=k\varepsilon_i, i=1,2,\cdots,n$,从而

$$\sigma(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)=(k\varepsilon_1,k\varepsilon_2,\cdots,k\varepsilon_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)\begin{pmatrix}k&&&\\&k&&\\&&\ddots&\\&&&k\end{pmatrix},$$ 矩阵为对角阵.

$$|\lambda E-A|=\begin{vmatrix}\lambda-k&&&\\&\lambda-k&&\\&&\ddots&\\&&&\lambda-k\end{vmatrix}=(\lambda-k)^n,$$ 从而特征值为 $\lambda=k$ (n 重).

齐次线性方程组 $(kE-A)X=0$ 的解集为 F^n,从而任一非零向量都是属于 $\lambda=k$ 的特征向量.

2) 取3线性空间 V, $\sigma\in L(V)$,设 σ 在基 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 下的矩阵为 $A=\begin{pmatrix}1&2&2\\2&1&2\\2&2&1\end{pmatrix}$,求 σ 的特征值与特征向量.

解 (1) 特征多项式为 $|\lambda E-A|=\begin{vmatrix}\lambda-1&-2&-2\\-2&\lambda-1&-2\\-2&-2&\lambda-1\end{vmatrix}=(\lambda+1)^2(\lambda-5)$, $\lambda_1=\lambda_2=-1,\lambda_3=5$.

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

$\begin{cases}-2x_1-2x_2-2x_3=0\\-2x_1-2x_2-2x_3=0\\-2x_1-2x_2-2x_3=0\end{cases}$, 得基础解系为 $X_1=\begin{pmatrix}1\\-1\\0\end{pmatrix}, X_2=\begin{pmatrix}1\\0\\-1\end{pmatrix}$,则属于 $\lambda_1=-1$ 的两个线性无关的特征向量为 $\xi_1=\varepsilon_1-\varepsilon_2,\xi_2=\varepsilon_1-\varepsilon_3$.属于特征值 $\lambda_1=-1$ 的全部特征向量为 $k_1\xi_1+k_2\xi_2$,其中 k_1,k_2 不全为零.

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 $X_3=\begin{pmatrix}1\\1\\1\end{pmatrix}$, 则属于 $\lambda_3=5$ 的线性无关的特征向量为 $\xi_3=\varepsilon_1+\varepsilon_2+\varepsilon_3$.

属于特征值 $\lambda_3=5$ 的全部特征向量为 $k_3\xi_3$,其中 $k_3\neq 0$.

3) 取微分变换 $\sigma: F[x]_n\to F[x]_n$, σ 在基 $1,x,\frac{1}{2!}x^2,\cdots,\frac{1}{(n-1)!}x^{n-1}$ 下的矩阵为 $A=\begin{pmatrix}0&1&\cdots&0\\0&0&\ddots&0\\ \vdots&\vdots&\cdots&1\\0&0&\cdots&0\end{pmatrix}$,则

$|\lambda E-A|=\lambda^n$,特征值为 $\lambda=0$ (n 重).

则方程组 $(\lambda E-A)X=0$,即 $AX=0$.基础解系为 $\eta=(1,0,0,\cdots,0)$,从而属于 $\lambda=0$ 的线性无关的特征向量为 $\xi=1$,从而全部特征向量为 k,其中 k 为非零常数.

4) 平面上的旋转变换. $\sigma:\mathbf{R}^2\to\mathbf{R}^2, \alpha=\begin{pmatrix}a\\b\end{pmatrix}\mapsto\sigma(\alpha)=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\alpha$.

取基 $\varepsilon_1=(1,0)^T,\varepsilon_2=(0,1)^T$,则 σ 在基 $\varepsilon_1,\varepsilon_2$ 下的矩阵为 $A=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$.

从而特征多项式 $|\lambda E-A|=\begin{vmatrix}\lambda-\cos\theta & \sin\theta\\ -\sin\theta & \lambda-\cos\theta\end{vmatrix}=(\lambda-\cos\theta)^2+\sin^2\theta\geq 0$.

则 $|\lambda E-A|=0\Leftrightarrow \lambda-\cos\theta=0,\sin\theta=0$,即 $\theta=k\pi$,从而 $\lambda=\pm 1$.

只有旋转 $\theta=k\pi$ 角度,才会出现旋转后的向量是原来向量的一个常数倍,且此常数为 $\lambda=\pm 1$.

6. 矩阵的特征值与特征向量:

对于线性空间 F^n,对 n 阶方阵 A,对于数 $\lambda_0\in F$,若存在非零向量 $X\in F^n$,使得 $AX=\lambda_0 X$,则称 λ_0 是 A 的一个特征值, X 是 A 的属于特征值 λ_0 的一个特征向量.

例子 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) 求特征值. $|\lambda E-A|=\begin{vmatrix}\lambda-1&-2&-2\\-2&\lambda-1&-2\\-2&-2&\lambda-1\end{vmatrix}=(\lambda+1)^2(\lambda-5)$, $\lambda_1=\lambda_2=-1,\lambda_3=5$.

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

$\begin{cases}-2x_1-2x_2-2x_3=0\\-2x_1-2x_2-2x_3=0\\-2x_1-2x_2-2x_3=0\end{cases}$, 即 $x_1+x_2+x_3=0$,得基础解系为 $X_1=\begin{pmatrix}1\\-1\\0\end{pmatrix}, X_2=\begin{pmatrix}1\\0\\-1\end{pmatrix}$, 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.

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 $X_3=\begin{pmatrix}1\\1\\1\end{pmatrix}$, 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 $k_3\neq 0$.

7 特征多项式的性质

任给 n 阶方阵 A,其特征多项式为 $f_A(\lambda)=|\lambda E-A|=\begin{vmatrix}\lambda-a_{11} & -a_{12} & \cdots & -a_{1n}\\ -a_{21} & \lambda-a_{22} & \cdots & -a_{2n}\\ \vdots & \vdots & & \vdots\\ -a_{n1} & -a_{n2} & \cdots & \lambda-a_{nn}\end{vmatrix}$,

$|\lambda E-A|=\lambda^n+a_1\lambda^{n-1}+\cdots+a_n$. 是一个关于 λ 的首1的 n 次多项式.

$|\lambda E-A|=\lambda^n-(\sum_{i=1}^n a_{ii})\lambda^{n-1}+\cdots+(-1)^n|A|$. λ^{n-1} 项的系数为 $\sum_{i=1}^n a_{ii}$,常数项为 $(-1)^n|A|$.

$|\lambda E-A|=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)=\lambda^n-(\sum_{i=1}^n\lambda_i)\lambda^{n-1}+\cdots+(-1)^n\prod_{i=1}^n\lambda_i$.

性质: (1) 首 1, n 次.

(2) λ^{n-1} 项的系数: $-(\sum_{i=1}^n a_{ii})=-(\sum_{i=1}^n\lambda_i)$,从而 $\sum_{i=1}^n a_{ii}=\sum_{i=1}^n\lambda_i$.

常数项: $(-1)^n|A|=(-1)^n\prod_{i=1}^n\lambda_i$,从而 $|A|=\prod_{i=1}^n\lambda_i$

定义: 设 n 阶方阵 $A=(a_{ij})$,主对角线元素的和 $\sum_{i=1}^n a_{ii}$ 称为矩阵的迹,记为 $\mathrm{tr}A=\sum_{i=1}^n a_{ii}$.

(3) 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 $|\lambda E-A|=|\lambda E-A_1||\lambda E-A_3|$.

(4) 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 $\sum_{i=1}^{n}a_{ii}=\sum_{i=1}^{n}\lambda_i$.

(5) $f_A(\lambda)$ 在所考虑的数域范围内不一定有解,从而可以没有特征值.

8. 特征子空间.

考察集合: $V_{\lambda_0}=\{\xi\in V\mid\sigma\xi=\lambda_0\xi\}$,集合非空,$0\in V_{\lambda_0}$.实际上 V_{λ_0} 是 σ 的属于特征值 λ_0 的特征向量的全体,再添加零向量所得的集合.

V_{λ_0} 关于 V 的加法和数乘运算封闭. 从而 V_{λ_0} 是 V 的一个子空间,称为 σ 的一个特征子空间.此子空间的位数 $\dim V_{\lambda_0}=n-r(\lambda_0E-A)$,其中 A 是线性变换 σ 在一组基下的矩阵.

定理: (1) 相似的矩阵有相同的特征多项式. (2) 相似的矩阵有相同的迹.

证明 (1) 设 $A\sim B$, $B=X^{-1}AX$,则

$$|\lambda E-B|=|\lambda E-X^{-1}AX|=|\lambda X^{-1}EX-X^{-1}AX|=|X^{-1}(\lambda E-A)X|=|\lambda E-A|.$$

(2) $|\lambda E-B|=|\lambda E-A|$,展开比较相应的系数即可.

但是定理 6 的逆不成立,即特征多项式虽然相同,但是不一定相似.

例子 $E=\begin{pmatrix}1&0\\0&1\end{pmatrix}, A=\begin{pmatrix}1&1\\0&1\end{pmatrix}$,则 $|\lambda E-E|=|\lambda E-A|=(\lambda-1)^2$,但 E 与 A 不相似,与 E 相似的只能是 E .

9. 哈密顿-凯莱定理:

设 A 是数域 F 上 n 阶方阵, $f_A(\lambda)=|\lambda E-A|$ 是 A 的特征多项式,设

$$f_A(\lambda)=\lambda^n+a_1\lambda^{n-1}+\cdots+a_{n-1}\lambda+a_n,\text{则 } f_A(A)=A^n+a_1A^{n-1}+\cdots+a_{n-1}A+a_nE=0.$$

思路: (1) $AA^*=|A|E$ (2) $\begin{pmatrix}f_1(\lambda)&f_2(\lambda)\\f_3(\lambda)&f_4(\lambda)\end{pmatrix}=\begin{pmatrix}*&*\\*&*\end{pmatrix}\lambda^n+\begin{pmatrix}*&*\\*&*\end{pmatrix}\lambda^{n-1}+\cdots.$

例子: $\begin{pmatrix}\lambda^3+\lambda^2+2\lambda+2&\lambda^4+1\\\lambda^2+2&\lambda^2+\lambda+3\end{pmatrix}=\begin{pmatrix}0&1\\0&0\end{pmatrix}\lambda^4+\begin{pmatrix}1&0\\0&0\end{pmatrix}\lambda^3+\begin{pmatrix}1&0\\1&1\end{pmatrix}\lambda^2+\begin{pmatrix}2&0\\0&1\end{pmatrix}\lambda+\begin{pmatrix}2&1\\1&3\end{pmatrix}.$

元素是多项式的矩阵可以写成系数是矩阵的多项式.

证明: 对矩阵 $\lambda E-A$,设其伴随矩阵为 $(\lambda E-A)^*=B(\lambda)$.

先看一下 $B(\lambda)$ 的具体元素的形式. $B(\lambda)$ 的各个元素是 $\lambda E-A$ 的相应元素的代数余子式.

主对角线元素的代数余子式是一个 $n-1$ 次的多项式,之外的元素的代数余子式则是最多 $n-2$ 次的多项式.

从而 $B(\lambda)$ 可以写成系数是矩阵的 $n-1$ 次多项式.

$B(\lambda)=B_0\lambda^{n-1}+B_1\lambda^{n-2}+\cdots+B_{n-2}\lambda+B_{n-1}$.则

$$B(\lambda)(\lambda E-A)=(B_0\lambda^{n-1}+B_1\lambda^{n-2}+\cdots+B_{n-2}\lambda+B_{n-1})(\lambda E-A)$$

$$=B_0\lambda^n+(B_1-B_0A)\lambda^{n-1}+(B_2-B_1A)\lambda^{n-2}+\cdots+(B_{n-2}-B_{n-3}A)\lambda^2+(B_{n-1}-B_{n-2}A)\lambda-B_{n-1}A$$

$$=|\lambda E-A|E=E\lambda^n+a_1E\lambda^{n-1}+\cdots+a_{n-1}E\lambda+a_nE.$$

从而$\begin{cases}B_0=E\\B_1-B_0A=a_1E\\B_2-B_1A=a_2E\\\cdots\\B_{n-2}-B_{n-3}A=a_{n-2}E\\B_{n-1}-B_{n-2}A=a_{n-1}E\\-B_{n-1}A=a_nE\end{cases}$,两边同乘$\begin{cases}A^n\\A^{n-1}\\A^{n-2}\\\cdots\\A^2\\A\\E\end{cases}$,则有$\begin{cases}B_0A^n=A^n\\B_1A^{n-1}-B_0A^n=a_1A^{n-1}\\B_2A^{n-2}-B_1A^{n-1}=a_2A^{n-2}\\\cdots\\B_{n-2}A^2-B_{n-3}A^3=a_{n-2}A^2\\B_{n-1}A-B_{n-2}A^2=a_{n-1}A\\-B_{n-1}A=a_nE\end{cases}$ 左右相加:

$0=A^n+a_1A^{n-1}+\cdots+a_{n-1}A+a_nE$,即 $f_A(A)=0$.

10. 几点说明:

假设$\sigma\xi=\lambda_0\xi$(或者 $AX=\lambda_0X$),则

A	kA	A^m	$f(A)$	A^{-1}	A^*	A^T	$P^{-1}AP$
X	X	X	X	X	X		$P^{-1}X$
λ_0	$k\lambda_0$	λ_0^m	$f(\lambda_0)$	λ_0^{-1}	$\frac{\lvert A\rvert}{\lambda_0}$	λ_0	λ_0
σ	$k\sigma$	σ^m	$f(\sigma)$	σ^{-1}			$\tau^{-1}\sigma\tau$
ξ	ξ	ξ	ξ	ξ			$\tau^{-1}\xi$
λ_0	$k\lambda_0$	λ_0^m	$f(\lambda_0)$	λ_0^{-1}			λ_0

$AX=\lambda_0X$,$kAX=k\lambda_0X$, $A^mX=\lambda_0^mX$, $f(A)X=f(\lambda_0)X$, $A^{-1}X=\lambda_0^{-1}X$

$A^*AX=A^*\lambda_0X=\lambda_0A^*X$, 则 $A^*X=\dfrac{1}{\lambda_0}|A|X=\dfrac{|A|}{\lambda_0}X$.

§5 对角矩阵(矩阵的相似对角化)

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.

例: 设3维线性空间的一个线性变换σ 在一组基$\varepsilon_1,\varepsilon_2,\varepsilon_3$下的矩阵为 $A=\begin{pmatrix}1&2&2\\2&1&2\\2&2&1\end{pmatrix}$.

已知$|\lambda E-A|=\begin{vmatrix}\lambda-1 & -2 & -2\\ -2 & \lambda-1 & -2\\ -2 & -2 & \lambda-1\end{vmatrix}=(\lambda+1)^2(\lambda-5)$,求得特征值为$\lambda_1=\lambda_2=-1,\lambda_3=5$.

对$\lambda_1=\lambda_2=-1$,方程组$(-E-A)X=0$的基础解系为$X_1=\begin{pmatrix}1\\-1\\0\end{pmatrix},X_2=\begin{pmatrix}1\\0\\-1\end{pmatrix}$,

对$\lambda_3=5$,方程组$(5E-A)X=0$的基础解系为$X_3=\begin{pmatrix}1\\1\\1\end{pmatrix}$,

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$\sigma\xi_1=-\xi_1,\sigma\xi_2=-\xi_2,\sigma\xi_3=5\xi_3$.

ξ_1,ξ_2,ξ_3是线性空间的一组基, σ在这组基下的矩阵为$B=\begin{pmatrix}-1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & 5\end{pmatrix}$.是个对角阵.

1. 定义:

(1) 设n阶方阵A,若存在一个可逆阵X,使得$X^{-1}AX=D$是一个对角阵,则称矩阵A可以相似对角化.

(2) 设线性变换$\sigma\in L(V)$,若存在V的一组基$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$,使得$\sigma$在这组基下的矩阵是一个对角阵,则称$\sigma$可以相似对角化.

2. 矩阵.线性变换可以相似对角化的条件.

(1) 假设A可以相似对角化,即存在一个可逆阵X,使得$X^{-1}AX=diag\left(\lambda_1,\lambda_2,\cdots,\lambda_n\right)$.

则$AX=Xdiag\left(\lambda_1,\lambda_2,\cdots,\lambda_n\right)$,取矩阵$X$的列分块$X=(X_1,X_2,\cdots,X_n)$,则

$$A(X_1,X_2,\cdots,X_n)=(X_1,X_2,\cdots,X_n)\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix}=(\lambda_1X_1,\lambda_2X_2,\cdots,\lambda_nX_n).$$

从而$AX_i=\lambda_iX_i,i=1,2,\cdots,n$,即可逆阵$X$的列向量组$X_1,X_2,\cdots,X_n$就是矩阵$A$的$n$个线性无关的特征向量.对应的特征值恰为对角阵的主对角线元素.

反之,假设矩阵A有n个线性无关的特征向量$X_1,X_2,\cdots,X_n$,设$AX_i=\lambda_iX_i,i=1,2,\cdots,n$.做

$X=(X_1,X_2,\cdots,X_n)$,则X是一个可逆阵,则

$$AX=A(X_1,X_2,\cdots,X_n)=(AX_1,AX_2,\cdots,AX_n)=(\lambda_1X_1,\lambda_2X_2,\cdots,\lambda_nX_n)=Xdiag\left(\lambda_1,\lambda_2,\cdots,\lambda_n\right)$$

即 $X^{-1}AX=diag\left(\lambda_1,\lambda_2,\cdots,\lambda_n\right)$, A 与对角阵 $diag\left(\lambda_1,\lambda_2,\cdots,\lambda_n\right)$ 相似,则 A 可以相似对角化.

从而我们有

定理 1: n 阶方阵 A 可以相似对角化的充要条件是 A 有 n 个线性无关的特征向量.

对应于线性变换,σ 可以相似对角化的充要条件是 σ 有 n 个线性无关的特征向量.

证明: 假设 σ 可以相似对角化,则存在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$,使得 σ 在此基下的矩阵是 $diag\left(\lambda_1,\lambda_2,\cdots,\lambda_n\right)$,

则 $\sigma(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)diag\left(\lambda_1,\lambda_2,\cdots,\lambda_n\right)=(\lambda_1\varepsilon_1,\lambda_2\varepsilon_2,\cdots,\lambda_n\varepsilon_n)$,

从而 $\sigma\varepsilon_i=\lambda_i\varepsilon_i, i=1,2,\cdots,n$.同理返回可得充分性.

定理: 属于不同特征值的特征向量线性无关.

证明: 假设 $\xi_1,\xi_2,\cdots,\xi_s$ 是分别属于特征值 $\lambda_1,\lambda_2,\cdots,\lambda_s$ 的特征向量,其中 $\lambda_i\neq\lambda_j(i\neq j)$.

假设 $k_1\xi_1+k_2\xi_2+\cdots+k_s\xi_s=0$,则

$k_1\sigma\xi_1+k_2\sigma\xi_2+\cdots+k_s\sigma\xi_s=0$,即 $k_1\lambda_1\xi_1+k_2\lambda_2\xi_2+\cdots+k_s\lambda_s\xi_s=0$. (1)

由 $k_1\xi_1+k_2\xi_2+\cdots+k_s\xi_s=0$ 得到 $k_1\lambda_s\xi_1+k_2\lambda_s\xi_2+\cdots+k_s\lambda_s\xi_s=0$. (2)

(1) (2)式子相减得到 $k_1(\lambda_1-\lambda_s)\xi_1+k_2(\lambda_2-\lambda_s)\xi_2+\cdots+k_{s-1}(\lambda_{s-1}-\lambda_s)\xi_{s-1}=0$

这样对 s 应用数学归纳法,应用归纳假设 $\xi_1,\xi_2,\cdots,\xi_{s-1}$ 线性无关,从而 $k_1=k_2=\cdots=k_{s-1}=0$.最后求得

$k_1=k_2=\cdots=k_s=0$.

推论: 若线性变换 σ 有 n 个互异的特征值,则 σ 可以相似对角化.但是反之不成立.

定理: 设 $\lambda_1,\lambda_2,\cdots,\lambda_s$ 是 σ 的互异特征值,则 $\xi_{i1},\xi_{i2},\cdots,\xi_{ir_i}$ 是属于 λ_i 的线性无关的特征向量,则

$\xi_{11},\cdots,\xi_{1r_1},\xi_{21},\cdots,\xi_{2r_2},\cdots,\xi_{s1},\cdots,\xi_{sr_s}$ 仍线性无关.

3. 代数重数与几何重数.

对方阵 A,特征多项式 $f_A(\lambda)=|\lambda E-A|=(\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\cdots(\lambda-\lambda_s)^{n_s}$, $\lambda_1,\lambda_2,\cdots,\lambda_s$ 两两互异.

其中 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. $m_i=n-r(\lambda_iE-A)$.

结论: 矩阵的任一特征值的几何重数小于等于其代数重数.

证明: 对 A,任给线性空间 V 及一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$,存在线性变换 σ ,使得 $\sigma\xrightarrow{\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n}A$.

假设 λ_0 是 A 的一个特征值,代数重数为 n_0 .设 $\dim V_{\lambda_0}=m_0$,不妨设 $\xi_1,\xi_2,\cdots,\xi_{m_0}$ 是 V_{λ_0} 的一组基,扩充这一

组基为V的一组基，$\xi_1,\cdots,\xi_{m_0},\xi_{m_0+1},\cdots,\xi_n$,则由于$\sigma\xi_i=\lambda_0\xi_i,i=1,2,\cdots,m_0$.则

$\sigma\xrightarrow{\xi_1,\cdots,\xi_{m_0},\xi_{m_0+1},\cdots,\xi_n}\begin{pmatrix}\lambda_0E & A_2\\ 0 & A_3\end{pmatrix}$.$\begin{pmatrix}\lambda_0E & A_2\\ 0 & A_3\end{pmatrix}$与$A$是同一个线性变换$\sigma$在不同基下的矩阵,则

$\begin{pmatrix}\lambda_0E & A_2\\ 0 & A_3\end{pmatrix}$与$A$相似,从而特征多项式相同. $\left|\lambda E-\begin{pmatrix}\lambda_0E & A_2\\ 0 & A_3\end{pmatrix}\right|=|\lambda E-A|$.

右边因式$\lambda-\lambda_0$是n_0重.左边中$\lambda-\lambda_0$也应该是n_0重.但是左边展开后

$\left|\lambda E-\begin{pmatrix}\lambda_0E & A_2\\ 0 & A_3\end{pmatrix}\right|=\begin{vmatrix}(\lambda-\lambda_0)E & A_2\\ 0 & \lambda E-A_3\end{vmatrix}=(\lambda-\lambda_0)^{m_0}|\lambda E-A_3|$.从而$m_0\le n_0$.

定理: 方阵A可相似对角化当且仅当任一特征值的代数重数与几何重数相等.

定理: 设A是一n阶方阵,A的特征多项式为$f_A(\lambda)=(\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\cdots(\lambda-\lambda_s)^{n_s}$.

则A可相似对角化$\Leftrightarrow n_i=n-r(\lambda_iE-A),i=1,2,\cdots,s\Leftrightarrow r(\lambda_iE-A)=n-n_i,i=1,2,\cdots,s$.

4. 判断相似对角化的方法,步骤.

判断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.

(1) 对A,求所有的特征值. $f_A(\lambda)=|\lambda E-A|=(\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\cdots(\lambda-\lambda_s)^{n_s}$.

(2) 是否可以相似对角化$\Leftrightarrow r(\lambda_iE-A)=n-n_i,i=1,2,\cdots,s$.若都成立,则可以;若有一个不成立,则不能.

对单根λ_i不用判断$r(\lambda_iE-A)=n-1$,这自然成立.主要是对重根λ_i,来判断是否$r(\lambda_iE-A)=n-n_i$.

(3) 若可以,求方程组$(\lambda_iE-A)X=0$的基础解系$X_{i1},X_{i2},\cdots,X_{in_i}$,

令$X=(X_{11},\cdots,X_{1n_1},X_{21},\cdots,X_{2n_2},\cdots,X_{s1},\cdots,X_{sn_s})$,则$X^{-1}AX=\begin{pmatrix}\lambda_1E_{n_1} & & & \\ & \lambda_2E_{n_2} & & \\ & & \ddots & \\ & & & \lambda_sE_{n_s}\end{pmatrix}$.

例子:

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

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$1\ne 2-2$,故不能相似对角化.

例 2. 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: $|\lambda E-A|=\begin{vmatrix}\lambda-1&-1&-1&-1\\-1&\lambda-1&-1&-1\\-2&-2&\lambda-2&-2\\-3&-3&-3&\lambda-3\end{vmatrix}=\begin{vmatrix}\lambda&-\lambda&0&0\\-1&\lambda-1&-1&-1\\0&-2\lambda&\lambda&0\\0&-3\lambda&0&\lambda\end{vmatrix}=\begin{vmatrix}\lambda&0&0&0\\-1&\lambda-7&-1&-1\\0&0&\lambda&0\\0&0&0&\lambda\end{vmatrix}$

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

$r(-A)=1=4-3$,有 3 个,可以相似对角化.

例 3. $A=\begin{pmatrix}1&2&2\\0&2&2\\0&0&3\end{pmatrix}$, 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 $\lambda_1=1$, $(E-A)X=0$ 的解: $X_1=(1,0,0)^T$.　对 $\lambda_2=2$, $(2E-A)X=0$ 的解: $X_2=(2,1,0)^T$.

对 $\lambda_3=3$, $(3E-A)X=0$ 的解: $X_3=(3,2,1)^T$.令 $X=\begin{pmatrix}1&2&3\\0&1&2\\0&0&1\end{pmatrix}$,则 $X^{-1}AX=\begin{pmatrix}1&&\\&2&\\&&3\end{pmatrix}$.

例 4. $A=\begin{pmatrix}1&2&2\\0&1&2\\0&0&2\end{pmatrix}$, 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. 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 1 特征值, $(E-A)X=0$ 的基础解系含有 $n-(n-r)=r$ 个向量 $\xi_1,\xi_2,\cdots,\xi_r$.

对 -1 特征值, $(-E-A)X=0$ 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 6. 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 1 特征值,$(E-A)X=0$ 的基础解系含有 $n-(n-r)=r$ 个向量 $\xi_1,\xi_2,\cdots,\xi_r$.

对 0 特征值,$(-A)X=0$ 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 $X^{-1}AX=\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$.

§6 线性变换的值域与核

1. 定义: 设线性空间 V 的一个线性变换 σ ,V 中向量在 σ 作用下的象的全体称为 σ 的值域,记为 $\sigma(V)$,所有被 σ 变成零向量的向量的全体称为 σ 的核,记为 $\ker\sigma=\sigma^{-1}(0)$.

即: $\sigma(V)=\sigma V=\{\sigma\alpha \mid \alpha\in V\}$, $\ker\sigma=\{\alpha\in V \mid \sigma\alpha=0\}$.

例子: 单位变换 $id:V\to V$,则 $\sigma(V)=V$, $\ker\sigma=\{0\}$.零变换: $0:V\to V$,则 $\sigma(V)=\{0\}$, $\ker\sigma=V$.

性质: (1) $\sigma(V)$ 与 $\ker\sigma=\sigma^{-1}(0)$ 都是 V 的子空间.

取 $\alpha,\beta\in\ker\sigma$,则 $\sigma\alpha=0,\sigma\beta=0$,则 $\sigma(k\alpha+l\beta)=k\sigma\alpha+l\sigma\beta=0$,从而 $k\alpha+l\beta\in\ker\sigma$.

$\sigma\alpha,\sigma\beta\in\sigma V$, $k\sigma\alpha+l\sigma\beta=\sigma(k\alpha+l\beta)\in\sigma V$.

(2) $\dim\sigma(V)$ 称为 σ 的秩. $\dim\sigma^{-1}(0)$ 称为 σ 的零度.

例子: 设 $F[x]_n$ 的微分变换 σ . $\sigma F[x]_n=F[x]_{n-1}$, $\sigma^{-1}(0)=F$,从而 $\dim\sigma(V)=n-1$, $\dim\sigma^{-1}(0)=1$.

2. 求线性变换值域与核.

定理: 设 $\sigma\in L(V)$, $\dim V=n$, $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是一组基,并设 σ 在 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵为 A ,则

(1) $\sigma(V)=L(\sigma\varepsilon_1,\sigma\varepsilon_2,\cdots,\sigma\varepsilon_n)$,即 $\sigma(V)$ 是由基的象所生成的子空间.

(2) σ 的秩等于矩阵 A 的秩,即 $\dim\sigma(V)=r(A)$.

证明: (1) 集合相等. $\sigma(V)\supseteq L(\sigma\varepsilon_1,\sigma\varepsilon_2,\cdots,\sigma\varepsilon_n)$ 显然.

任给 $\sigma\alpha\in\sigma(V)$,则 $\alpha\in V$,从而 $\alpha=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)X$,则

$\sigma\alpha=(\sigma\varepsilon_1,\sigma\varepsilon_2,\cdots,\sigma\varepsilon_n)X\in L(\sigma\varepsilon_1,\sigma\varepsilon_2,\cdots,\sigma\varepsilon_n)$,从而 $\sigma(V)\subseteq L(\sigma\varepsilon_1,\sigma\varepsilon_2,\cdots,\sigma\varepsilon_n)$,则

$\sigma(V)=L(\sigma\varepsilon_1,\sigma\varepsilon_2,\cdots,\sigma\varepsilon_n)$

(2) $(\sigma\varepsilon_1,\sigma\varepsilon_2,\cdots,\sigma\varepsilon_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)A$,则 $\dim(\sigma\varepsilon_1,\sigma\varepsilon_2,\cdots,\sigma\varepsilon_n)=r(A)$.

求值域与核.

(1) $\sigma(V)=L(\sigma\varepsilon_1,\sigma\varepsilon_2,\cdots,\sigma\varepsilon_n)$,求 $\sigma\varepsilon_1,\sigma\varepsilon_2,\cdots,\sigma\varepsilon_n$ 的极大无关组即可.

(2) 任给 $\alpha\in\ker\sigma$,则 $\sigma\alpha=0$,假设 $\sigma\xrightarrow{\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n}A$,并设 $\alpha=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)X$,则

$0=\sigma\alpha=\sigma(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)X=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)AX$,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 $\ker\sigma$ 的一组基,即 $\ker\sigma=L(\xi_1,\xi_2,\cdots,\xi_r)$.

例子: 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 $X=\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}$,则 $\sigma(X)=\begin{pmatrix}2&1&0\\0&1&1\\2&0&-1\end{pmatrix}X$,取单位向量 $\varepsilon_1,\varepsilon_2,\varepsilon_3$,则 $\sigma\xrightarrow{\varepsilon_1,\varepsilon_2,\varepsilon_3}A=\begin{pmatrix}2&1&0\\0&1&1\\2&0&-1\end{pmatrix}$.

取 $\alpha_1=(2,0,2),\alpha_2=(1,1,0)$,则 $\sigma(V)=L(\alpha_1,\alpha_2)$.

解 $AX=0$ 的基础解系: $\begin{pmatrix}2&1&0\\0&1&1\\2&0&-1\end{pmatrix}\to\begin{pmatrix}2&1&0\\0&1&1\\0&0&0\end{pmatrix}$,基础解系为 $\eta=(1,-2,2)^T$,则 $\ker\sigma=L(\eta)$.

3.定理: 设 $\sigma\in L(V)$, $\dim V=n$,则 $\sigma(V)$ 的一组基的原象及 $\ker\sigma$ 的一组基合起来是 V 的一组基,从而有

$\dim\sigma(V)+\dim\sigma^{-1}(0)=\dim V=n$,即 σ 的秩+σ 的零度$=\dim V$.

证明: 设 $\sigma(V)$ 的一组基为 $\xi_1,\xi_2,\cdots,\xi_r$,其原象设为 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_r$,即 $\sigma\varepsilon_i=\xi_i,i=1,2,\cdots,r$.

设 $\ker\sigma$ 的一组基为 $\varepsilon_{r+1},\varepsilon_{r+2},\cdots,\varepsilon_s$,证明: $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_r,\varepsilon_{r+1},\varepsilon_{r+2},\cdots,\varepsilon_s$ 是 V 的一组基.

假设线性组合为零. $k_1\varepsilon_1+\cdots+k_r\varepsilon_r+k_{r+1}\varepsilon_{r+1}+\cdots+k_s\varepsilon_s=0$.

用 σ 作用一下, $k_1\sigma\varepsilon_1+\cdots+k_r\sigma\varepsilon_r=0$,即 $k_1\xi_1+\cdots+k_r\xi_1=0$,从而 $k_1=\cdots=k_r=0$.

则 $k_{r+1}\varepsilon_{r+1}+\cdots+k_s\varepsilon_s=0$，$\varepsilon_{r+1},\varepsilon_{r+2},\cdots,\varepsilon_s$ 为一组基,从而 $k_{r+1}=\cdots=k_s=0$.

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

任给 $\alpha\in V$,则 $\sigma\alpha\in\sigma(V)=L(\xi_1,\xi_2,\cdots,\xi_r)$,从而 $\sigma\alpha=k_1\xi_1+k_2\xi_2+\cdots+k_r\xi_r$,

$\sigma\alpha=k_1\xi_1+k_2\xi_2+\cdots+k_r\xi_r=k_1\sigma\varepsilon_1+k_2\sigma\varepsilon_2+\cdots+k_r\sigma\varepsilon_r=\sigma(k_1\varepsilon_1+k_2\varepsilon_2+\cdots+k_r\varepsilon_r)$,

即 $\sigma(\alpha-k_1\varepsilon_1-k_2\varepsilon_2-\cdots-k_r\varepsilon_r)=0$,则 $\alpha-k_1\varepsilon_1-k_2\varepsilon_2-\cdots-k_r\varepsilon_r\in\ker\sigma$,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，$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_r,\varepsilon_{r+1},\varepsilon_{r+2},\cdots,\varepsilon_s$ 是 V 的一组基.

另外的证明：σ 的秩+σ 的零度$=\dim V$.
由于 σ 的秩等于 σ 在一组基下矩阵 A 的秩,而 σ 的零度等于核的维数,即等于 $AX=0$ 的基础解系所含向量的个数 $n-r(A)$,从而 σ 的秩+σ 的零度$=r(A)+n-r(A)=n$.

推论：设 V 是一个有限维线性空间，$\sigma\in L(V)$,则 σ 是单射当且仅当 σ 是满射.

注:虽然 σV 与 $\ker\sigma$ 的维数之和为 n，但是 $\sigma V+\ker\sigma$ 不一定是直和.

例子：设 $F[x]_n$ 的微分变换 σ.

$\sigma F[x]_n=F[x]_{n-1}$，$\sigma^{-1}(0)=F$，从而 $\dim\sigma(V)=n-1,\dim\sigma^{-1}(0)=1$.$\sigma V\cap\ker\sigma=F$.

例子：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 σ,使得 σ 在 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵为

A,则 $\sigma^2=\sigma$,对 σV,取一组基 $\xi_1,\xi_2,\cdots,\xi_r$,其原象设为 $\alpha_1,\alpha_2,\cdots,\alpha_r$,即 $\sigma\alpha_i=\xi_i,i=1,2,\cdots,r$.

则 $\sigma\xi_i=\sigma^2\alpha_i=\sigma\alpha_i=\xi_i,i=1,2,\cdots,r$,即 $\xi_1,\xi_2,\cdots,\xi_r$ 得原象就是 $\xi_1,\xi_2,\cdots,\xi_r$.取 $\ker\sigma$ 的一组基

$\xi_{r+1},\xi_{r+2},\cdots,\xi_n$,则 $\xi_1,\xi_2,\cdots,\xi_r,\xi_{r+1},\xi_{r+2},\cdots,\xi_n$ 是 V 的一组基，σ 在 $\xi_1,\xi_2,\cdots,\xi_n$ 下的矩阵为

$$\sigma(\xi_1,\xi_2,\cdots,\xi_n)=(\xi_1,\cdots,\xi_r,0,\cdots,0)=(\xi_1,\xi_2,\cdots,\xi_n)\begin{pmatrix}E_r & 0\\ 0 & 0\end{pmatrix}.$$

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

1. 不变子空间.

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

2) 例子:

(1) (非不变子空间) 取3维平面$V=\mathbf{R}^3$,取线性变换σ为绕X轴线从Y轴线到Z轴线旋转90°角的变换.

即$\sigma(x,y,z)=(x,-z,y)$,取子空间$W=XOY$面,则W不是不变子空间,应为$\sigma(0,1,0)=(0,0,1)$不在W中,

(2) 平凡子空间$\{0\}$与V是任一线性变换的不变子空间.

(3) 任一线性变换σ的值域σV与核$\ker\sigma$是σ的不变子空间.

任给$\sigma\alpha\in\sigma V$则$\sigma(\sigma\alpha)\in\sigma V$.

任给$\alpha\in\ker\sigma$,则$\sigma\alpha=0$,则$\sigma(\sigma\alpha)=\sigma(0)=0$,从而$\sigma\alpha\in\ker\sigma$.

(4) 取$\sigma,\tau\in L(V)$,满足$\sigma\tau=\tau\sigma$,则τV与$\ker\tau$都是$\sigma-$子空间.

τV:任给$\tau\alpha\in\tau V$,证$\sigma(\tau\alpha)\in\tau V$? $\sigma(\tau\alpha)=(\sigma\tau)\alpha=(\tau\sigma)\alpha=\tau(\sigma\alpha)\in\tau V$.

$\ker\tau$: 任给$\alpha\in\ker\tau$,证$\sigma\alpha\in\ker\tau$,即证$\tau(\sigma\alpha)=0$, $\tau(\sigma\alpha)=\sigma(\tau\alpha)=0$.

(5) 任一子空间都是数乘变换的不变子空间.

3) 考察特征向量与一维不变子空间的关系.

取$\sigma\xi=\lambda_0\xi$,设$W=L(\xi)$,则$\dim W=1$, $\sigma(k\xi)=\lambda_0 k\xi\in W$,从而$W=L(\xi)$是$\sigma-$子空间.

反之,取一维不变子空间$W=L(\xi)$, $\sigma W\subseteq W$,则$\sigma\xi\in W=L(\xi)$,从而存在一个数λ,使得$\sigma\xi=\lambda\xi$,即生成元ξ是σ的属于λ的一个特征向量.

4) 取$\sigma-$子空间W_1,W_2,则W_1+W_2与$W_1\bigcap W_2$都是$\sigma-$子空间.

5) 设$\sigma-$子空间W,由于$\sigma W\subseteq W$,则σ就可以看做是W上的一个线性变换,从而定义$\sigma|_W$为σ限制在不变子空间W上的线性变换.它的作用为

任给$\alpha\in W$, $\sigma|_W(\alpha)=\sigma(\alpha)$,但是$\sigma$对$W$之外的向量没有作用.

6) W是$\sigma-$子空间当且仅当σ对W的基作用的象仍在W中.

设$W=L(\alpha_1,\alpha_2,\cdots,\alpha_m)$,若$\sigma W\subseteq W$,则$\sigma\alpha_1,\sigma\alpha_2,\cdots,\sigma\alpha_m\in W$.反之,若$\sigma\alpha_1,\sigma\alpha_2,\cdots,\sigma\alpha_m\in W$,则任给$\alpha\in W$,设$\alpha=(\alpha_1,\alpha_2,\cdots,\alpha_m)X$,则$\sigma\alpha=(\sigma\alpha_1,\sigma\alpha_2,\cdots,\sigma\alpha_m)X\in W$,从而$\sigma W\subseteq W$. W是$\sigma-$子空间.

2. 不变子空间与线性变换的矩阵.

(1) 取W是一个$\sigma-$子空间, $\dim W=m$,设$W=L(\alpha_1,\alpha_2,\cdots,\alpha_m)$,则$\sigma W\subseteq W$,从而$\sigma\alpha_1,\sigma\alpha_2,\cdots,\sigma\alpha_m$

可由$\alpha_1,\alpha_2,\cdots,\alpha_m$线性表出,将$\alpha_1,\alpha_2,\cdots,\alpha_m$扩充为$V$的一组基$\alpha_1,\alpha_2,\cdots,\alpha_m,\alpha_{m+2},\cdots,\alpha_n$.看

$\sigma \xrightarrow{\alpha_1,\alpha_2,\cdots,\alpha_n} B$.

$\sigma\alpha_1 = b_{11}\alpha_1 + b_{21}\alpha_2 + \cdots + b_{m1}\alpha_m + 0\alpha_{m+1} + \cdots + 0\alpha_n$, $\cdots$

$\sigma\alpha_m = b_{1m}\alpha_1 + b_{2m}\alpha_2 + \cdots + b_{mm}\alpha_m + 0\alpha_{m+1} + \cdots + 0\alpha_n$,

$\sigma\alpha_{m+1} = b_{1m+1}\alpha_1 + b_{2m+1}\alpha_2 + \cdots + b_{mm+1}\alpha_m + b_{mm+1}\alpha_{m+1} + \cdots + b_{nm+1}\alpha_n$, $\cdots$

$\sigma\alpha_n = b_{1n}\alpha_1 + b_{2n}\alpha_2 + \cdots + b_{mn}\alpha_m + b_{mn}\alpha_{m+1} + \cdots + b_{nn}\alpha_n$.

则$\sigma \xrightarrow{\alpha_1,\alpha_2,\cdots,\alpha_n} \begin{pmatrix} b_{11} & \cdots & b_{1m} & b_{1m+1} & \cdots & b_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ b_{m1} & \cdots & b_{mm} & b_{mm+1} & \cdots & b_{mn} \\ 0 & \cdots & 0 & b_{m+1m+1} & \cdots & b_{m+1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & b_{nm+1} & \cdots & b_{nn} \end{pmatrix} = \begin{pmatrix} B_1 & B_2 \\ 0 & B_3 \end{pmatrix}$

其中B_1实际上是$\sigma|_W$在$\alpha_1,\alpha_2,\cdots,\alpha_m$下的矩阵.

反之,若一线性变换σ在一组基$\alpha_1,\alpha_2,\cdots,\alpha_n$下的矩阵为$\begin{pmatrix} B_1 & B_2 \\ 0 & B_3 \end{pmatrix}$,其中$B_1$是一个$m$阶方阵,则$\alpha_1,\alpha_2,\cdots,\alpha_n$前$m$个向量所生成的子空间$W = L(\alpha_1,\alpha_2,\cdots,\alpha_m)$就是一个$\sigma-$子空间.

(2) 若一线性变换σ在一组基$\alpha_1,\alpha_2,\cdots,\alpha_n$下的矩阵为$\begin{pmatrix} B_1 & 0 \\ 0 & B_3 \end{pmatrix}$,其中$B_1$是一个$m$阶方阵,$B_2$是一个$n-m$阶方阵,则$\alpha_1,\alpha_2,\cdots,\alpha_n$中前$m$个向量所生成的子空间$W_1 = L(\alpha_1,\alpha_2,\cdots,\alpha_m)$与后$n-m$个向量所生成的子空间$W_2 = L(\alpha_{m+1},\alpha_{m+2},\cdots,\alpha_n)$都是$\sigma-$子空间.且$W_1 \oplus W_2 = V$

(3) 一般的, 设$A = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_s \end{pmatrix}$是一个准对角阵,$A_i$分别是$n_i$阶方阵,并设$\sigma \xrightarrow{\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n} A$

则$W_1 = L(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_{n_1})$, $W_2 = L(\varepsilon_{n_1+1},\varepsilon_{n_1+2},\cdots,\varepsilon_{n_1+n_2})$,$\cdots$, $W_s = L(\varepsilon_{n_1+\cdots+n_{s-1}+1},\varepsilon_{n_1+\cdots+n_{s-1}+2},\cdots,\varepsilon_n)$都是不变子空间.且$W_1 \oplus W_2 \oplus \cdots \oplus W_s = V$.反之,亦成立.

定理: 线性空间V ,$\dim V = n$,取线性变换σ ,特征多项式为 $f(\lambda) = (\lambda-\lambda_1)^{r_1}(\lambda-\lambda_2)^{r_2}\cdots(\lambda-\lambda_s)^{r_s}$.

则V可分解为不变子空间的直和$V_1 \oplus V_2 \oplus \cdots \oplus V_s = V$,其中$V_i = \{\xi \mid (\sigma-\lambda_i id)^{r_i}\xi = 0, \xi \in V\}$.

证明:令 $f_i(\lambda)=\dfrac{f(\lambda)}{(\lambda-\lambda_i)^{r_i}}$,令 $V_i = f_i(\sigma)V$.

(1) 首先:$V_i = f_i(\sigma)V$ 是 $f_i(\sigma)$ 的值域,从而是 σ 的不变子空间,且满足 $(\sigma-\lambda_i id)^{r_i}V_i = f(\sigma)V=0$.

(2) 证明: $V_1\oplus V_2\oplus\cdots\oplus V_s=V$,

$(f_1(\lambda),f_2(\lambda),\cdots,f_s(\lambda))=1$,则存在多项式 $u_1(\lambda),u_2(\lambda),\cdots,u_s(\lambda)$,使得

$u_1(\lambda)f_1(\lambda)+u_2(\lambda)f_2(\lambda)+\cdots+u_s(\lambda)f_s(\lambda)=1$,从而 $u_1(\sigma)f_1(\sigma)+u_2(\sigma)f_2(\sigma)+\cdots+u_s(\sigma)f_s(\sigma)=id$,

则任给 $\alpha\in V$,有 $\alpha=u_1(\sigma)f_1(\sigma)\alpha+\cdots+u_s(\sigma)f_s(\sigma)\alpha$,其中 $u_i(\sigma)f_i(\sigma)\alpha\in f_i(\sigma)V=V_i$.

假设 $\beta_1+\beta_2+\cdots+\beta_s=0$,其中 β_i 满足 $(\sigma-\lambda_i id)^{r_i}\beta_i=0$,证明 $\beta_i=0$. 由于 $(\lambda-\lambda_j)^{r_j}\mid f_i\ (j\neq i)$,故

$f_i(\sigma)\beta_j=0(j\neq i)$,用 $f_i(\sigma)$ 作用于 $\beta_1+\beta_2+\cdots+\beta_s=0$.得到 $f_i(\sigma)\beta_i=0$,又 $(f_i(\lambda),(\lambda-\lambda_i)^{r_i})=1$,故

有多项式 $u(\lambda),v(\lambda)$,使得 $u(\lambda)f_i(\lambda)+v(\lambda)(\lambda-\lambda_i)^{r_i}=1$,则 $\beta_i=u(\sigma)f_i(\sigma)\beta_i+v(\sigma)(\sigma-\lambda_i id)^{r_i}\beta_i=0$.

现在若 $\beta_i\in f_i(\sigma)V=V_i$,自然有 $(\sigma-\lambda_i id)^{r_i}\beta_i=0$,故 $V_1\oplus V_2\oplus\cdots\oplus V_s$ 是直和.

(3) 现在设 $\alpha\in W_i=\{\xi\mid(\sigma-\lambda_i id)^{r_i}\xi=0,\xi\in V\}$,由直和设 $\alpha=\alpha_1+\alpha_2+\cdots+\alpha_s$,其中 $\alpha_i\in V_i$.从而

$\alpha_1+\alpha_2+\cdots+\alpha_i-\alpha+\cdots+\alpha_s=0$,而 $(\sigma-\lambda_i id)^{r_i}(\alpha_i-\alpha)=0$,由(2) , $\alpha=\alpha_i\in V_i$.即 $V_i\supseteq W_i$,反之,

$(\sigma-\lambda_i id)^{r_i}V_i=f(\sigma)V=0$,则 $V_i\subseteq W_i$.故 $V_i=\{\xi\mid(\sigma-\lambda_i id)^{r_i}\xi=0,\xi\in V\}$.

§8 若当标准形

对方阵 A, A 可相似对角化的条件 $\Leftrightarrow$ A 有 n 个线性无关的特征向量 $\Leftrightarrow$ A 的任一特征值的代数重数和几何重数相等. $\Leftarrow$ A 有 n 个互异的特征值.

并不是任一方阵都可以相似对角化,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.

1. 若当标准形.

1) 若当快:形式为 $\begin{pmatrix}\lambda&0&\cdots&0&0&0\\1&\lambda&\cdots&0&0&0\\\vdots&\vdots&&\vdots&\vdots&\vdots\\0&0&\cdots&1&\lambda&0\\0&0&\cdots&0&1&\lambda\end{pmatrix}$ 的矩阵称为若当块,其中 λ 是一个复数.

2) 由若当块组成的准对角阵称为若当形矩阵.

定理: 设复数域上线性空间 V 的线性变换 σ ,则存在 V 的一组基,使得 σ 在此组基下的矩阵为若当形矩阵.

§9 最小多项式

任给方阵 A,存在多项式 $f(x)$,使得 $f(A)=0$.对 $f(x)$ 而言,使得 $f(A)=0$ 成立的 A 称为多项式的根,

$f(x)$ 以 A 为根.则称首1的次数最低的以 A 为根的多项式称为 A 的最小多项式.

性质: 引理 1: A 的最小多项式是唯一的.

证明: 设 $f(x)$ 与 $g(x)$ 都是 A 的最小多项式,设 $f(x)=g(x)q(x)+r(x)$,其中 $r(x)=0$,或者

$\partial(r(x))<\partial(g(x))$.代入 A,得到 $r(A)=0$,从而必有 $r(x)=0$,即 $g(x)\mid f(x)$.同理 $f(x)\mid g(x)$.

从而 $f(x)=cg(x)$,又首1,从而 $f(x)=g(x)$.

引理 2: 设 $g(x)$ 是 A 的最小多项式,对多项式 $f(x)$, $f(A)=0\Leftrightarrow g(x)\mid f(x)$.

由哈密顿-凯莱定理,最小多项式是特征多项式的一个因式.

例子: 对 $A=kE$,最小多项式为 $f(x)=x-k$.单位阵的最小多项式为 $x-1$.零矩阵的最小多项式为 x.

例子: 设 $A=\begin{pmatrix}1&1&0\\0&1&0\\0&0&1\end{pmatrix}$,求 A 的最小多项式.

解: 特征多项式为 $|\lambda E-A|=(\lambda-1)^3$.最小多项式为 $g(x)=(\lambda-1)^2$.

相似矩阵有相同的最小多项式.但是反之不成立.

例子: 设 $A=\begin{pmatrix}1&1&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&2\end{pmatrix}$, $B=\begin{pmatrix}1&1&0&0\\0&1&0&0\\0&0&2&0\\0&0&0&2\end{pmatrix}$.最小多项式为 $(\lambda-1)^2(\lambda-2)$.但是 A,B 不相似.

引理: 设 $A=\begin{pmatrix}A_1&0\\0&A_2\end{pmatrix}$.设 A_i 的最小多项式为 $g_i(x)$,则 A 的最小多项式为 $[g_1(x),g_2(x)]$.

证明: 设 $g(x)=[g_1(x),g_2(x)]$,首先 $g(A)=\begin{pmatrix}g(A_1)&0\\0&g(A_2)\end{pmatrix}=0$.

其次,若 $h(x)=0$,则 $h(A)=\begin{pmatrix}h(A_1)&0\\0&h(A_2)\end{pmatrix}=0$,则 $g_1(x)\mid h(x), g_2(x)\mid h(x)$,从而 $g(x)\mid h(x)$.

引理 4: k 级若当块 $J=\begin{pmatrix}a&&&\\1&\ddots&&\\&\ddots&\ddots&\\&&1&a\end{pmatrix}$ 的最小多项式为 $(x-a)^k$.

定理: 数域 F 上 n 阶阵 A 与对角阵相似的充分必要条件是 A 的最小多项式是 F 上互素的一次因式的乘积.

推论: 复矩阵 A 与对角阵相似的充分必要条件是 A 的最小多项式没有重根.