1.2 线性变换及其矩阵

一. 线性变换及其运算

定义1.11 如果线性空间V的一个变换T具有性质

$$T(k\boldsymbol{x}+l\boldsymbol{y})=k(T\boldsymbol{x})+l(T\boldsymbol{y})$$

其中 $\boldsymbol{x}, \boldsymbol{y}\in \mathbf{V}$, $k, l\in \mathbf{K}$，则称T为V的一个线性变换或线性算子.

例1.10 T是 $\mathbf{R}^2$ 上的旋转变换，即 $T(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}$ 其中

$$\boldsymbol{x} = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \in \mathrm{R}^2, \boldsymbol{A} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.$$

例 1.11 在 $\boldsymbol{P}_n$中，求微分的变换. 即

$$Df(t) = f'(t), \ \forall f(t) \in P_n$$

例1.12 定义在闭区间[a,b]上的所有实连续函数的集合C(a,b)构成R上的一个线性空间. 在C(a,b)上定义变换J，即

$$J(f(t)) = \int_a^b f(t)dt, \quad \forall f(t) \in C(a,b)$$

线性变换的性质：

(1) $T(0)=0$; (2) $T(-x)=-T(x)$

(3) 如果$x_1,\cdots,x_m$线性相关，则$Tx_1,\cdots,Tx_m$也相关;

(4) 如果$x_1,\cdots x_m$线性无关，且T是单射，则

$Tx_1,\cdots,Tx_m$也无关.

证明 （3）因为$\mathbf{x}_1,\mathbf{x}_2,\ldots\mathbf{x}_m$线性相关，所以∃不全为零的$k_1,k_2,\ldots,k_m$使

$$k_1\mathbf{x}_1+k_2\mathbf{x}_2+\ldots+k_m\mathbf{x}_m=0$$

作变换T，有 $T(k_1\mathbf{x}_1+k_2\mathbf{x}_2+\ldots+k_m\mathbf{x}_m)=0$

即 $k_1T\mathbf{x}_1+k_2T\mathbf{x}_2+\ldots+k_mT\mathbf{x}_m=0$

故 $T\mathbf{x}_1,T\mathbf{x}_2,\ldots,T\mathbf{x}_m$线性相关。

（4）设一组数$k_1,k_2,\ldots,k_m$使得

$$k_1T\mathbf{x}_1+k_2T\mathbf{x}_2+\ldots+k_mT\mathbf{x}_m=0\text{，即}$$

$$T(k_1\mathbf{x}_1+k_2\mathbf{x}_2+\ldots+k_m\mathbf{x}_m)=0,$$

由T是单射可得 $k_1\mathbf{x}_1+k_2\mathbf{x}_2+\ldots+k_m\mathbf{x}_m=0$，因为$\mathbf{x}_1,\mathbf{x}_2,\ldots\mathbf{x}_m$线性无关，所以有$k_1=k_2=\ldots=k_m=0$，故$T\mathbf{x}_1,T\mathbf{x}_2,\ldots,T\mathbf{x}_m$线性无关。 证毕

定义1.12 设T是V的线性变换，集合 $\{Tx \mid x \in V\}$ 称为T的值域，用**R(*T*)** 表示，即 $R(T) = \{Tx \mid x \in V\}$

集合$\{x \mid Tx = 0, x \in V\}$，称为$T$的核，用 **N(*T*)** 表示，即

$$N(T) = \{x \mid Tx = 0, x \in V\}$$

定理1.8 线性空间V的线性变换T的值域和核都是V的线性子空间.

定义1.13　象子空间的维数 dim$\mathbf{R}(T)$ 称为T的秩；核子空间的维数 dim$\mathbf{N}(T)$ 称为T 的亏（零度）.

例1.13　设 V^n 的一个基为 $x_1, x_2, \cdots, x_n$，T是 V^n 中的线性变换，则 $R(T) = L(Tx_1, Tx_2, \cdots, Tx_n)$.

证　$\forall \beta \in R(T)$, 有$T(\alpha)=\beta$, $\alpha \in V^n$, 由

$$\alpha = k_1 \mathbf{x}_1 + k_2 \mathbf{x}_2 + \ldots + k_n \mathbf{x}_n$$

所以　$\beta = T\alpha = k_1 T\mathbf{x}_1 + k_2 T\mathbf{x}_2 + \ldots + k_n T\mathbf{x}_n$,

故$R(T) \subset L(T\mathbf{x}_1, T\mathbf{x}_2, \ldots, T\mathbf{x}_n)$, 另一边显然，故成立.

例1.14 设 $x=(\xi_1,\xi_2,\xi_3,\xi_4)\in R^4$，试证明

$$T(\xi_1,\xi_2,\xi_3,\xi_4)=(\xi_1+\xi_2-3\xi_3-\xi_4,\ 3\xi_1-\xi_2-3\xi_3+4\xi_4,0,0)$$

是 V^n 的线性变换，并求T的秩与亏.

证明 省。由于$\boldsymbol{T}(\boldsymbol{e}_1)=(1,3,0,0),\boldsymbol{T}(\boldsymbol{e}_2)=(1,-1,0,0)$,
$\boldsymbol{T}(\boldsymbol{e}_3)=(-3,-3,0,0),\boldsymbol{T}(\boldsymbol{e}_4)=(-1,4,0,0)$, 所以$\text{rank}\boldsymbol{T}=2$.

由$\boldsymbol{Tx}=\mathbf{0}$,有

$$\begin{aligned}\xi_1+\xi_2-3\xi_3-\xi_4&=0\\ 3\xi_1-\xi_2-3\xi_3+4\xi_4&=0\end{aligned}$$

基础解系有4−2=2个向量，故T 的亏为**2.** 且解得 $(3,3,2,0),(-3,7,0,4)$ 是**N(*T*)**的基。

特殊的线性变换：

（1）单位变换或恒等变换 T_e，即 $T_e x = x,\ \forall x \in V$

（2）零变换 T_0，即 $T_0 x = 0,\ \forall x \in V$

线性变换的运算：

1. 加法

设 T_1, T_2 是V 的两个线性变换，定义它的和 T_1+T_2 为 $(T_1+T_2)x = T_1 x + T_2 x,\ \forall x \in V$

线性变换T的负变换定义为 $(-T)\boldsymbol{x} = -(T\boldsymbol{x})$

运算规律：

(1) $T_1 + T_2 = T_2 + T_1$ (2) $(T_1+T_2)+T_3 = T_1+(T_2+T_3)$

(3) $T + T_0 = T$ (4) $T + (-T) = T_0$

2. 线性变换与数的乘法

设 $k\in$**K**，$\boldsymbol{T}$为**V**中的线性变换，定义数 k 与$\boldsymbol{T}$的乘积 $k\boldsymbol{T}$ 为

$$(kT)\boldsymbol{x} = k(T\boldsymbol{x}),\ \forall \boldsymbol{x} \in V$$

运算规律：

$$(1)\ \ k(T_1+T_2)=kT_1+kT_2$$

$$(2)\ \ (k+l)T=kT+lT$$

$$(3)\ \ (kl)T=k(lT) \qquad (4)\ \ 1T=T$$

性质：

(1) $\boldsymbol{T}_1+\boldsymbol{T}_2$ **是V上的线性变换.**

(2) $k\boldsymbol{T}$ **是V上的线性变换.**

(3) V上的所有线性变换的集合，形成一个线性空间，记为 $Hom(V,V)$

3. 线性变换的乘法

设 T_1, T_2 是V的两个线性变换，定义T_1与T_2 的乘积为

$$(T_1T_2)x = T_1(T_2(x)), \quad \forall x \in V$$

运算规律： (1) $(T_1T_2)T_3 = T_1(T_2T_3)$

(2) $T_1(T_2 + T_3) = T_1T_2 + T_1T_3$

(3) $(T_1 + T_2)T_3 = T_1T_3 + T_2T_3$

注：$T_1T_2 \neq T_2T_1$

4. 逆变换

若T是V的线性变换，且存在线性变换S，使

$$(ST)x = (TS)x = x, \ \forall x \in V$$

成立，则称S是T的逆变换，记为 $S=T^{-1}$.且有

$$T^{-1}T = TT^{-1} = T_e$$

5. 线性变换的多项式

设n是正整数，T是V的线性变换. 定义T的n次幂为

$$T^n = TT\cdots T$$

定义T的零次幂为 $T^0 = T_e$

运算规律：

当T是可逆时，$T^{m+n} = T^m T^n$, $(T^m)^n = T^{mn}$, $m,n \in N^+$.

$$T^{-n} = (T^{-1})^n,\ n \in N^+$$

设 $f(t) = a_m t^m + a_{m-1} t^{m-1} + \cdots + a_0$ 是t的m次多项式，

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

称为线性变换T 的多项式.

性质： 如果 $h(t) = f(t)g(t), p(t) = f(t) + g(t)$, 则

$$h(T) = f(T)g(T), p(T) = f(T) + g(T).$$

二. 线性变换的矩阵表示

设T是 V^n的线性变换，$x_1, x_2, \cdots, x_n$是V^n的一个基，则有

$$\begin{cases} Tx_1 = a_{11}x_1 + a_{21}x_2 + \cdots + a_{n1}x_n \\ Tx_2 = a_{12}x_1 + a_{22}x_2 + \cdots + a_{n2}x_n \\ \cdots\cdots\cdots\cdots \\ Tx_n = a_{1n}x_1 + a_{2n}x_2 + \cdots + a_{nn}x_n \end{cases}$$

采用矩阵形式表示 $T(x_1, x_2, \cdots, x_n) = (Tx_1, Tx_2, \cdots, Tx_n)$

$$= (x_1, x_2, \cdots, x_n)\boldsymbol{A}$$

其中

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

定义1.14 矩阵A称为T在 V^n 的基 $x_1, x_2, \cdots, x_n$ 下的矩阵，简称A为T的矩阵.

例如 T_0 的矩阵为 O ， T_e的矩阵为 I， T_m的矩阵为 mI (数量矩阵).

例1.15 在n+1维线性空间 P_n中，其两个基分别为

$$\text{(I)}\ f_0(t)=1, f_1(t)=t, f_2(t)=\frac{t^2}{2!}, \cdots, f_n(t)=\frac{t^n}{n!}$$

$$\text{(II)}\ g_0(t)=1, g_1(t)=t, g_2(t)=t^2, \cdots, g_n(t)=t^n$$

求微分变换D在两个基下的矩阵.

解　因为　$T(f_0)=0, T(f_1)=1, \cdots, T(f_n)=t^{n-1}/(n-1)!$

$T(g_0)=0, T(g_1)=1, \cdots, T(g_n)=nt^{n-1}$

所以T在基（Ⅰ）与基（Ⅱ）下的矩阵分别为

$$\boldsymbol{A}=\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}, \quad \boldsymbol{B}=\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & n \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

例1.16 设A是 V^n 的线性变换T 的矩阵，则

$$\dim R(T)=\dim R(\boldsymbol{A}),\ \dim N(T)=\dim N(\boldsymbol{A})$$

证　设 V^n 的基为 $x_1,x_2,\cdots,x_n$，　$\dim R(\boldsymbol{A})=r$

由于　$R(T)=L(Tx_1,Tx_2,\cdots,Tx_n)$

所以　$\dim R(T)=\operatorname{rank}(Tx_1,Tx_2,\cdots,Tx_n)=\operatorname{rank}(\boldsymbol{A})$

因而　$\dim R(T)=\dim R(\boldsymbol{A})$

设$\dim N(\boldsymbol{A})=s$，当$s=0$时，显然有$\dim N(T)\geq s$;

而当 $s\neq 0$时，设 $y_1,\cdots,y_s$ (列向量)为 $N(\boldsymbol{A})$ 的一个基

记 $\tilde{x}_i=(x_1,x_2,\cdots,x_n)y_i\ (i=1,\cdots,s)$．由

$$T\tilde{x}_i=(x_1,x_2,\cdots,x_n)\boldsymbol{A}y_i=0$$

知 $\tilde{x}_i\in N(T)$　因此　$\dim N(T)\geq s=\dim N(\boldsymbol{A})$

若设 $\dim N(T)=r$,则当$r=0$时，有$\dim N(\boldsymbol{A})\geq r$

而当 $\boldsymbol{r\neq 0}$ 时，设 $N(T)$的一个基$\tilde{x}_1,\tilde{x}_2,\cdots,\tilde{x}_r$,它 们在$\mathbf{V^n}$

的基下的坐标依次为 $y_1,y_2,\cdots,y_r$，由$T\tilde{x}_i=0$

可得 $T\tilde{x}_i=(x_1,x_2,\cdots,x_n)\boldsymbol{A}y_i=0$ 即$\boldsymbol{A}y_i=0$

所以 $y_i\in N(\boldsymbol{A})$. 因此$\dim N(\boldsymbol{A})\geq r=\dim N(T)$

综上所述，便得 $dim\,N(\boldsymbol{A})=dim\,N(T)$

证毕

定理1.9 设 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$ 是V^{n}的一个基，线性变换 $\boldsymbol{T_1},\boldsymbol{T_2}$ 在该基下依次有$\boldsymbol{n}$阶矩阵$\boldsymbol{A}$，$\boldsymbol{B}$. 则有

$$(1)\quad (T_1+T_2)(x_1,x_2,\cdots,x_n)=(x_1,x_2,\cdots,x_n)(\boldsymbol{A}+\boldsymbol{B})$$

$$(2)\quad (kT_1)(x_1,x_2,\cdots,x_n)=(x_1,x_2,\cdots,x_n)(k\boldsymbol{A})$$

$$(3)\quad (T_1T_2)(x_1,x_2,\cdots,x_n)=(x_1,x_2,\cdots,x_n)(\boldsymbol{AB})$$

$$(4)\quad T_1^{-1}(x_1,x_2,\cdots,x_n)=(x_1,x_2,\cdots,x_n)\boldsymbol{A}^{-1}$$

推论 设 $f(t)=a_mt^m+a_{m-1}t^{m-1}+\cdots+a_0$是$t$的多项式，$\boldsymbol{T}$为 V^n 的线性变换，且对 V^n 的基 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$ 有$T(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n)=(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n)\boldsymbol{A}$，则有 V^n 的线性变换 $\boldsymbol{f(T)}$ 在该基下的矩阵是

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

定理1.10 设线性变换T在 V^n 的基 $x_1, x_2, \cdots, x_n$ 下的矩阵 $A=(a_{ij})$， 向量 x 在该基下的坐标是 $(\xi_1,\xi_2,\cdots,\xi_n)^{\mathrm{T}}$, 则 Tx 在该基下的坐标 $(\eta_1,\cdots,\eta_n)^{\mathrm{T}}$ 满足

$$(\eta_1,\eta_2,\cdots,\eta_n)^{\mathrm{T}} = \boldsymbol{A}(\xi_1,\xi_2,\cdots,\xi_n)^{\mathrm{T}}$$

证　因为

$$\boldsymbol{x} = (\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n)\begin{pmatrix}\xi_1\\ \vdots\\ \xi_n\end{pmatrix}$$

$$T\boldsymbol{x} = T(\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n)\begin{pmatrix}\xi_1\\ \vdots\\ \xi_n\end{pmatrix} = (\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n)\boldsymbol{A}\begin{pmatrix}\xi_1\\ \vdots\\ \xi_n\end{pmatrix}$$

所以 $$(\eta_1,\eta_2,\cdots,\eta_n)^{\mathrm{T}} = \boldsymbol{A}(\xi_1,\xi_2,\cdots,\xi_n)^{\mathrm{T}}$$ 证毕

定理1.11 设 V^n的线性变换T，对于 V^n的两个基 $x_1,x_2,\cdots,x_n$和$y_1,y_2,\cdots,y_n$ 的矩阵依次是A和B，并且

$$(y_1,y_2,\cdots,y_n)=(x_1,x_2,\cdots,x_n)C$$

那么 $$B=C^{-1}AC$$

证 因为 $T(x_1,\cdots,x_n)=(x_1,\cdots,x_n)A$, $T(y_1,\cdots,y_n)=(y_1,\cdots,y_n)B$

$$T(y_1,\cdots,y_n)=T(x_1,\cdots,x_n)C=(x_1,\cdots x_n)AC$$

$$=(y_1,\cdots,y_n)C^{-1}AC$$

所以 $B=C^{-1}AC$ 证毕

相似的性质: (1) $A\sim A$; (2)如果$A\sim B$， 那么$B\sim A$;

(3)如果$A\sim B$，$B\sim C$，那么$A\sim C$;

(4)如果$A\sim B$，那么 $f(A)\sim f(B)$.

三. 特征值与特征向量

定义1.16 设T是 V^n 的线性变换，且对K中的某一数λ_0，存在非零向量 $\boldsymbol{x}\in V^n$，使得 $Tx=\lambda_0 x$ 成立，则称λ_0为T的特征值，$\boldsymbol{x}$ 为T 的属于λ_0 的特征向量.

设 V^n的一个基为 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$ ，则有

$$\boldsymbol{x}=\xi_1\boldsymbol{x}_1+\xi_1\boldsymbol{x}_2+,\cdots,+\xi_n\boldsymbol{x}_n$$

由于 $Tx=\lambda_0 x$ 所以 $A(\xi_1,\xi_2,\cdots,\xi_n)^{\mathrm{T}}=\lambda_0(\xi_1,\xi_2,\cdots,\xi_n)^{\mathrm{T}}$

即 $(A-\lambda_0 I)(\xi_1,\xi_2,\cdots,\xi_n)^{\mathrm{T}}=0$

结论：A的特征值与T的特征值相同，且A的特征向量是T的特征向量的坐标.

计算T的特征值和特征向量的步骤：

（1）取定 V^n 的一个基，写出T 在该基下的矩阵A；

（2）求出A 的特征值，它们就是T的特征值.

（3）求出A的全部线性无关的特征向量；

（4）以A的特征向量为V^n中取定基下的坐标,得到T的特征向量.

例1.18 设线性变换T在 V^3的基 x_1, x_2, x_3 下的矩阵是

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

求T 的特征值和特征向量.

解　因为A 的特征多项式是

$$\varphi(\lambda)=det(\lambda \boldsymbol{I}-\boldsymbol{A})=(\lambda+1)^2(\lambda-5)$$

因此T 的特征值是　$\lambda_{1,2}=-1, \lambda_3=5.$

特征方程$(\lambda_{1,2}\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=0$ 的一个基础解系为

$$(1,0,-1)^T,\ (0,1,-1)^T$$

T 的属于 $\lambda_{1,2}$ 的两个线性无关的特征向量为

$$\boldsymbol{y}_1=\boldsymbol{x}_1-\boldsymbol{x}_3,\quad \boldsymbol{y}_2=\boldsymbol{x}_2-\boldsymbol{x}_3$$

T 的属于$\lambda_{1,2}$ 的全体特征向量为

$$k_1\boldsymbol{y}_1+k_2\boldsymbol{y}_2\quad (k_1,k_2\in K\text{不同时为零})$$

特征方程$(\lambda_3\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=0$ 的一个基础解系为 $(1,1,1)^{T}$，则T 的属于 λ_3 的线性无关的特征向量为

$$\boldsymbol{y}_3=\boldsymbol{x}_1+\boldsymbol{x}_2+\boldsymbol{x}_3$$

所以T的属于 λ_3 的全部特征向量为$k_3\boldsymbol{y}_3$ （$k_3\in \mathbf{K}$不等于零）

特征子空间 V_{λ_0}: $V_{\lambda_0} = \{\boldsymbol{x} \mid T\boldsymbol{x} = \lambda_0 \boldsymbol{x}, \boldsymbol{x} \in V^n\}$

定理1.12 设 $A=(a_{ij})_{m\times n}, B=(b_{ij})_{n\times m}$，则 $tr(\boldsymbol{AB})=tr(\boldsymbol{BA})$

定理1.13 相似矩阵有相同的迹.

定理1.14 相似矩阵有相同的特征多项式，因此也有相同的特征值.

定理1.15 设 $\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_m$均为方阵 $\boldsymbol{A} = diag(\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_n)$,
则

$$det\,(\lambda \boldsymbol{I} - \boldsymbol{A}) = \prod_{i=1}^{m} det(\lambda \boldsymbol{I}_i - \boldsymbol{A}_i)$$

定理1.16 设 $A \in \mathrm{R}^{m\times n}, B \in \mathrm{R}^{n\times m}$ ，又设 AB 的特征多项式为 $\varphi_{AB}(\lambda)$， BA 的特征多项式为 $\phi_{BA}(\lambda)$，则

$$\lambda^n \varphi_{AB}(\lambda) = \lambda^m \varphi_{BA}(\lambda)$$

证明　构造行列式 $\begin{vmatrix} \lambda I_m & A \\ \lambda B & \lambda I_n \end{vmatrix}$

因为 $$\begin{pmatrix} I_m & \\ B & I_n \end{pmatrix} \begin{pmatrix} \lambda I_m & A \\ \lambda B & \lambda I_n \end{pmatrix} = \begin{pmatrix} \lambda I_m & A \\ O & \lambda I_n - BA \end{pmatrix}$$

$$\begin{pmatrix} I_m & A \\ B & \lambda I_n \end{pmatrix} \begin{pmatrix} I_m & O \\ B & I_n \end{pmatrix} = \begin{pmatrix} \lambda I_m - AB & A \\ O & \lambda I_n \end{pmatrix}$$

故 $$\begin{vmatrix} \lambda I_m & A \\ \lambda B & \lambda I_n \end{vmatrix} = \lambda^m \left| \lambda I_n - BA \right| = \lambda^n \left| \lambda I_m - AB \right|$$ 证毕

例 已知列向量$\alpha=(1,2,\ldots,n)^{\mathrm{T}}$, $\beta=(1,1,\ldots 1)^{\mathrm{T}}$,求行列式$\det(I_n+\alpha\beta^{\mathrm{T}})$.

解 因为 $\det(I_n+\alpha\beta^{\mathrm{T}})=\det(1+\beta^{\mathrm{T}}\alpha)$,

所以 $\det(I_n+\alpha\beta^{\mathrm{T}}) = 1+\dfrac{n(n+1)}{2}.$

定理1.17 任意n阶矩阵与三角矩阵相似.

定理1.18 (Hamilton-Cayley)n 阶矩阵A 是其特征多项式的根（零点），即令

$$\varphi(\lambda)=det(\lambda \boldsymbol{I}-\boldsymbol{A})=\lambda^n+a_{n-1}\lambda^{n-1}+\cdots+a_0$$

则
$$\varphi(\boldsymbol{A})=\boldsymbol{A}^n+a_{n-1}\boldsymbol{A}^{n-1}+\cdots+a_0\boldsymbol{I}=\mathrm{O}$$

证　改写 $\varphi(\lambda)$ 为 $\varphi(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$

因为

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{pmatrix}\lambda_1 & * & \cdots & * \\ & \lambda_2 & \ddots & \vdots \\ & & \ddots & * \\ & & & \lambda_n\end{pmatrix}$$

于是
$$\varphi(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P})=(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}-\lambda_1\boldsymbol{I})\cdots(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}-\lambda_n\boldsymbol{I})$$

$$=\begin{pmatrix}0 & * & \cdots & * \\ & \lambda_2-\lambda_1 & \ddots & \vdots \\ & & \ddots & * \\ & & & \lambda_n-\lambda_1\end{pmatrix}\times$$

$$\begin{pmatrix} \lambda_1-\lambda_2 & * & \cdots & * \\ & 0 & \ddots & \vdots \\ & & \ddots & * \\ & & & \lambda_n-\lambda_2 \end{pmatrix} \times\cdots\times(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}-\lambda_n\boldsymbol{I})=$$

$$\begin{pmatrix} 0 & 0 & * & \cdots & * \\ 0 & 0 & * & \cdots & * \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & & & * \end{pmatrix} \times \begin{pmatrix} \lambda_1-\lambda_3 & * & * & \cdots & * \\ & \lambda_2-\lambda_3 & * & \cdots & * \\ & & 0 & \ddots & \vdots \\ & & & \ddots & * \\ & & & & \lambda_n-\lambda_3 \end{pmatrix} \times$$

$$\cdots\times(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}-\lambda_n\boldsymbol{I})=O$$

即 $$\varphi(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P})=\boldsymbol{P}^{-1}\varphi(\boldsymbol{A})\boldsymbol{P}=O$$

故 $$\varphi(\boldsymbol{A})=O$$

证毕

例1.20 已知

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

求 $A^{100}+2A^{50}$

解 令 $f(\lambda)=\lambda^{100}+2\lambda^{50}$，因为

所以 $\varphi(\lambda)=det(\lambda I-A)=(\lambda-1)^2(\lambda-2)$

$$f(\lambda)=\varphi(\lambda)q(\lambda)+b_0+b_1\lambda+b_2\lambda^2$$

因为

$$\begin{cases} f(1)=b_0+b_1+b_2=3 \\ f(2)=b_0+2b_1+4b_2=2^{100}+2^{51} \\ f'(1)=b_1+2b_2=200 \end{cases}$$

所以
$$\begin{cases} b_0 = 2^{100} + 2^{51} - 400 \\ b_1 = 606 - 2^{101} - 2^{52} \\ b_2 = -203 + 2^{100} + 2^{51} \end{cases}$$

于是 $\boldsymbol{A}^{100} + 2\boldsymbol{A}^{50} = f(\boldsymbol{A}) = b_0\boldsymbol{I} + b_1\boldsymbol{A} + b_2\boldsymbol{A}^2$

定义 1.19 **首项系数为1，次数最小，且以矩阵*A* 为根的多项式，称为*A* 的最小多项式，常用 $\boldsymbol{m(\lambda)}$ 表示.**

定理1.19 矩阵$\boldsymbol{A}$ 的最小多项式 $\boldsymbol{m(\lambda)}$ 可整除以$\boldsymbol{A}$ 为根的任意多项式 $\boldsymbol{f(\lambda)}$，且$\boldsymbol{m(\lambda)}$ 唯一.

证 假若 $\boldsymbol{m(\lambda)}$ 不能整除 $\boldsymbol{f(\lambda)}$，则有

$$f(\lambda)=m(\lambda)q(\lambda)+r(\lambda)$$

其中$\deg(r(\lambda))<\deg(m(\lambda))$，于是由

$$f(\boldsymbol{A})=m(\boldsymbol{A})q(\boldsymbol{A})+r(\boldsymbol{A}) \quad \text{知} \quad r(\boldsymbol{A})=O$$

这与 $m(\lambda)$ 是$\boldsymbol{A}$ 的最小多项式相矛盾.

唯一性. 设$\boldsymbol{A}$ 有两个不同的最小多项式 $m_1(\lambda)$ $m_2(\lambda)$，取 $g(\lambda)=m_1(\lambda)-m_2(\lambda)$，由于 $\deg(g(\lambda))<\deg(m(\lambda))$ 且 $g(\boldsymbol{A})=O$，这与 $m_1(\lambda)$ 是最小多项式矛盾. 所以唯一. 证毕

定理1.20 矩阵A 的最小多项式 $m(\lambda)$ 与其特征多项式 $\varphi(\lambda)$ 的零点相同（不计重数）.

证　设 λ_0 是 $\varphi(\lambda)$ 的零点，即 $\varphi(\lambda_0)=0$

因为 $\boldsymbol{Ax}=\lambda_0\boldsymbol{x}(\boldsymbol{x}\neq 0)$，所以$m(\boldsymbol{A})\boldsymbol{x}=m(\lambda_0)\boldsymbol{x}=0$

因而 $m(\lambda_0)=0$　　证毕

***定理1.21** 设n 阶矩阵A 的特征多项式为$\varphi(\lambda)$，特征矩阵 $\lambda I-A$ 的全体 $n-1$ 阶子式的最大公因式为 $d(\lambda)$，则A 的最小多项式为

$$m(\lambda)=\frac{\varphi(\lambda)}{d(\lambda)}$$

例 求矩阵 $A=\begin{pmatrix}1 & 1 & & \\ & 1 & & \\ & & 2 & \\ & & 1 & 1\end{pmatrix}$ 的最小多项式.

解 因为 $\varphi(\lambda)=(\lambda-1)^3(\lambda-2)$，所以$A$ 的最小多项式可能是$m(\lambda)=(\lambda-1)(\lambda-2)$，由于$m(A)\neq O$，所以$A$ 的最小多项式可能是

$$m(\lambda)=(\lambda-1)^2(\lambda-2),$$

由于$m(A)=O$，所以A 的最小多项式是

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

例1.22 证明相似矩阵有相同的最小多项式.

证　设 $\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$，$m_{\boldsymbol{A}}(\lambda)$与$m_{\boldsymbol{B}}(\lambda)$ 分别表示$\boldsymbol{A}$与$\boldsymbol{B}$的最小多项式，

由$m_{\boldsymbol{B}}(\boldsymbol{B})=\boldsymbol{O}$可得：

$$m_B(\boldsymbol{A})=m_B(\boldsymbol{P}\boldsymbol{B}\boldsymbol{P}^{-1})=\boldsymbol{P}m_{\boldsymbol{B}}(\boldsymbol{B})\boldsymbol{P}^{-1}=\boldsymbol{O}$$

所以$m_A(\lambda)|m_B(\lambda)$，又由$m_{\boldsymbol{A}}(\boldsymbol{A})=\boldsymbol{O}$可得：

$$m_A(\boldsymbol{B})=m_A(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P})=\boldsymbol{P}^{-1}m_A(\boldsymbol{A})\boldsymbol{P}=\boldsymbol{O}$$

所以$m_B(\lambda)|m_A(\lambda)$，因而$m_A(\lambda)=m_B(\lambda)$

注：矩阵相似是其最小多项式相等的充分条件，不是必要条件。
例：矩阵diag(1, 1, 2)与diag(2, 2, 1)不相似，但最小多项式相同。

例 若矩阵

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

则A的最小多项式为$(\lambda-1)(\lambda-2)$;B的最小多项式为$(\lambda-1)^2(\lambda-2)$.

定理1.22 如果 $\lambda_1, \ldots, \lambda_m$是矩阵$A$ 的互不相同的特征值，$\boldsymbol{x}_1,\ldots,\boldsymbol{x}_m$是分别属于它们的特征向量，那么 $\boldsymbol{x}_1,\ldots,\boldsymbol{x}_m$ 线性无关.

定理1.23 如果$\lambda_1,\lambda_2,\ldots,\lambda_m$ 是矩阵A 的不同特征值，而 $\boldsymbol{x}_{i1},\cdots,\boldsymbol{x}_{ir_i}\,(i=1,\cdots,m)$是属于$\lambda_i$ 的线性无关的特征向量，那么向量组 $\boldsymbol{x}_{11},\cdots,\boldsymbol{x}_{1r_1},\cdots,\boldsymbol{x}_{m1},\cdots,\boldsymbol{x}_{mr_m}$ 也线性无关.

四．对角矩阵

定理1.24 设T是V^n的线性变换，T在某一基下的矩阵A可以为对角矩阵$\Leftrightarrow T$有n个线性无关的特征向量.

定理1.25 n阶矩阵A与对角矩阵相似$\Leftrightarrow A$有n个线性无关的特征向量，或A有完备的特征向量系.

例1.23 已知R^3上的线性变换

$$T(\xi_1,\xi_2,\xi_3)=(\xi_1+\xi_2+\xi_3,\xi_1+\xi_2+\xi_3,\xi_1+\xi_2+\xi_3)$$

求R^3上的一个基使T在该基下的矩阵为对角矩阵.

解　取$\mathbf{R}^3$的一个基 $\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3$，$\boldsymbol{T}$ 在该基下的矩阵为

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

因为$\det(\lambda \boldsymbol{I}-\boldsymbol{A})=\lambda^2(\lambda-3)=0$，所以$\lambda_{1,2}=0,\lambda_3=3$

由$\boldsymbol{A}\boldsymbol{x}=0$得基础解系

$$p_1=(1,-1,0)^T,\ p_2=(1,0,-1)^T$$

由　$(3\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=0$　得基础解系　$\boldsymbol{p}_3=(1,1,1)^{\mathrm{T}}$

所以存在基　$\boldsymbol{e}_1-\boldsymbol{e}_2,\boldsymbol{e}_1-\boldsymbol{e}_3,\boldsymbol{e}_1+\boldsymbol{e}_2+\boldsymbol{e}_3$

$\boldsymbol{T}$ 在该基下的矩阵是对角阵　$diag(0,0,3)$

五. 不变子空间

定义1.20 如果T 是V 的线性变换， V_1是V的子空间，并且对于任意一个$x\in V_1$，都有$Tx\in V_1$，则称 V_1是T的不变子空间.

例如：

$$T_m\boldsymbol{x}=m\boldsymbol{x}, D(f(t))=f'(t), \mathrm{W}=P_n; \mathrm{V}_{\lambda_0}=\{\boldsymbol{x}\mid T\boldsymbol{x}=\lambda_0\boldsymbol{x}\}.$$

结论：如果V_1,V_2都是T 的不变子空间，那么 $V_1\cap V_2$，V_1+V_2 也都是T 的不变子空间.

定理1.27 设T是V^n的线性变换，且V^n可分解为s个不变子空间的直和

$$V^n = V_1 \oplus V_2 \oplus \cdots \oplus V_s$$

又在每个不变子空间V_i中取基

$$x_{i1}, \cdots, x_{in_i} \quad (i = 1, \cdots, s)$$

将它们合并作为V^n的基，则T在该基下的矩阵为

$$\boldsymbol{A} = diag(\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_s)$$

其中A_i(i=1,2,…,s) 就是T在V_i的基下的矩阵.

证 因为 $V_1, V_2, \ldots, V_s$ 都是T的不变子空间，所以当 $x_{ij} \in V_i$时有 $T\boldsymbol{x}_{ij} \in V_i \quad (i = 1, \cdots, s;\ j = 1, \cdots, n_i)$

于是
$$\begin{cases} T\boldsymbol{x}_{i1} = a_{11}^{(i)}\boldsymbol{x}_{i1} + a_{21}^{(i)}\boldsymbol{x}_{i2} + \cdots + a_{n_i1}^{(i)}\boldsymbol{x}_{in_i} \\ T\boldsymbol{x}_{i2} = a_{12}^{(i)}\boldsymbol{x}_{i1} + a_{22}^{(i)}\boldsymbol{x}_{i2} + \cdots + a_{n_i2}^{(i)}\boldsymbol{x}_{in_i} \\ \vdots \\ T\boldsymbol{x}_{in_i} = a_{1n_i}^{(i)}\boldsymbol{x}_{i1} + a_{2n_i}^{(i)}\boldsymbol{x}_{i2} + \cdots + a_{n_in_i}^{(i)}\boldsymbol{x}_{in_i} \end{cases}$$

因此，在该基下，T 的矩阵是A，其中

$$\boldsymbol{A}_i = \begin{pmatrix} a_{11}^{(i)} & a_{12}^{(i)} & \cdots & a_{1n_i}^{(i)} \\ a_{21}^{(i)} & a_{22}^{(i)} & \cdots & a_{2n_i}^{(i)} \\ \vdots & \vdots & & \vdots \\ a_{n_i1}^{(i)} & a_{n_i2}^{(i)} & \cdots & a_{n_in_i}^{(i)} \end{pmatrix} \quad (i = 1,2,\cdots,s)$$

推论 V^n的线性变换T 在V^n的某个基下的矩阵A为对角矩阵 $\Leftrightarrow$ V^n可分解为n 个T 的一维特征子空间的直和.

六. Jordan 标准形介绍

定义1.21 方阵J 称为Jordan 标准形，即

$$J = diag\left(J_1(\lambda_1), J_2(\lambda_2), \cdots, J_s(\lambda_s)\right)$$

其中

$$J_i(\lambda_i) = \begin{pmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \lambda_i & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_i \end{pmatrix}_{m_i \times m_i} \quad (i = 1,2,\cdots,s)$$

定理1.28 设T是复数域C上的线性空间V的线性变换，任取V的一个基，T在该基下的矩阵是A，T(或A)的特征多项式可分解因式为

$$\varphi(\lambda)=(\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\cdots(\lambda-\lambda_l)^{m_s},(m_1+m_2+\cdots+m_s=n)$$

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$V_i=\left\{x \,\middle|\, (T-\lambda_i T_e)^{m_i}x=0, x\in V\right\}$.

定理1.29 设 $A \in \mathrm{C}^{n\times n}$，则存在 $P \in \mathrm{C}_n^{n\times n}$，使

$$P^{-1}AP = J$$

例1.24 设 $A \in \mathrm{C}^{6\times 6}$ 的特征多项式是

$$\varphi(\lambda) = det(\lambda I - A) = \lambda^2(\lambda+2)(\lambda-1)^3$$

则 A 的Jordan 标准形为下列6个矩阵之一.

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

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

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

1. λ-矩阵（多项式矩阵）的Smith 标准形

定义 矩阵

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

其中 $a_{ij}(\lambda)\ (i,j=1,2,\cdots,n)$ 为K 上的 λ 的多项式,

则称矩阵 $\mathbf{A}(\lambda)$为 λ-矩阵.

λ-矩阵的**初等变换**：

$$(1)\ r_i(c_i)\leftrightarrow r_j(c_j)$$

$$(2)\ \ kr_i(c_i),k\neq 0$$

$$(3)\ \ r_i(c_i)+f(\lambda)r_j(c_j)$$ （$f(\lambda)$是λ的多项式）

λ-矩阵Smith 标准形：$A(\lambda)$经过初等变换化为如下形式的 λ-矩阵

$$\boldsymbol{A}(\lambda) \to diag(d_1(\lambda), d_2(\lambda), \cdots, d_s(\lambda), 0, \cdots, 0)$$

其中 $d_1(\lambda) | d_2(\lambda), d_2(\lambda) | d_3(\lambda), \cdots, d_{s-1}(\lambda) | d_s(\lambda), s \le n,$

且$d_i(\lambda)(i=1,\cdots,s)$是首1多项式。

例1.25 试用初等变换化 λ-矩阵为$A(\lambda)$标准形.

$$\boldsymbol{A}(\lambda) = \begin{pmatrix} -\lambda+1 & 2\lambda-1 & \lambda \\ \lambda & \lambda^2 & -\lambda \\ \lambda^2+1 & \lambda^2+\lambda-1 & -\lambda^2 \end{pmatrix}$$

解

$$A(\lambda) \xrightarrow{c_3+c_1} \begin{pmatrix} -\lambda+1 & 2\lambda-1 & 1 \\ \lambda & \lambda^2 & 0 \\ \lambda^2+1 & \lambda^2+\lambda-1 & 1 \end{pmatrix} \xrightarrow{c_1 \leftrightarrow c_3} \begin{pmatrix} 1 & 2\lambda-1 & -\lambda+1 \\ 0 & \lambda^2 & \lambda \\ 1 & \lambda^2+\lambda-1 & \lambda^2+1 \end{pmatrix}$$

$$\xrightarrow{r_3-r_1} \begin{pmatrix} 1 & 2\lambda-1 & 1 \\ 0 & \lambda^2 & \lambda \\ 0 & \lambda^2-\lambda & \lambda^2+\lambda \end{pmatrix} \xrightarrow{c_2 \leftrightarrow c_3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda & \lambda^2 \\ 0 & \lambda^2+\lambda & \lambda^2-\lambda \end{pmatrix}$$

$$\xrightarrow{c_3-\lambda c_2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & \lambda^2+\lambda & -\lambda^3-\lambda \end{pmatrix} \xrightarrow[r_3-(\lambda+1)r_2]{(-1)c_3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda^3+\lambda \end{pmatrix}$$

由此 $d_1(\lambda)=1, d_2(\lambda)=\lambda, d_3(\lambda)=\lambda^3+\lambda$

结论： $\boldsymbol{A}(\lambda)$ 的标准形唯一，且称 $d_i(\lambda)(i=1,\cdots,s)$ 为 $\boldsymbol{A}(\lambda)$ 的不变因子.

$\boldsymbol{A}(\lambda)$ 的行列式因子：$\boldsymbol{D}_i(\lambda)$ $(i=1,\ldots,n)$ 表示$\boldsymbol{A}(\lambda)$ 的一切 i 阶子式的最大公因式.

结论：不变因子 $d_i(\lambda)(i=1,\ldots,s)$ 唯一，且

$$d_i(\lambda)=\frac{D_i(\lambda)}{D_{i-1}(\lambda)},\ \ D_0(\lambda)=1\ \ (i=1,\cdots,s)$$

$A(\lambda)$的初等因子：将 $A(\lambda)$的不变因子分解为不可约因式幂的积，所有的不可约因式的幂就是 $A(\lambda)$ 的初等因子组.

例如：若 $A(\lambda)$ 的不变因子是

$$d_1(\lambda)=1, d_2(\lambda)=\lambda^2+1, d_3(\lambda)=(\lambda^2+1)(\lambda+1)^2$$

则在R上初等因子是：$\lambda^2+1, \lambda^2+1, (\lambda+1)^2$

在C上初等因子是：$\lambda+i, \lambda-i, \lambda+i, \lambda-i, (\lambda+1)^2$

在C上，求 $A\in C^{n\times n}$的Jordan标准形的步骤：

（1）求特征矩阵 $\lambda I-A$ 的初等因子组，设为

$$(\lambda-\lambda_1)^{m_1}, (\lambda-\lambda_2)^{m_2}, \cdots, (\lambda-\lambda_s)^{m_s} \quad (m_1+\cdots+m_s=n)$$

（2）写出每个初等因子 $(\lambda-\lambda_i)^{m_i}\ (i=1,\cdots,s)$ 对应的 **Jordan 块**

$$J_i(\lambda_i)=\begin{pmatrix}\lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \lambda_i & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_i\end{pmatrix}\quad(i=1,\cdots,s)$$

（3）写出以这些Jordan块构成的Jordan标准形

$$\boldsymbol{J}=\begin{pmatrix}\boldsymbol{J}_1 & & & \\ & \boldsymbol{J}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{J}_s\end{pmatrix}$$

例1.28 求矩阵

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

的Jordan标准形.

解 求 $\lambda I-A$ 的初等因子组. 由于

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

$$\begin{pmatrix}-1 & 0 & 0\\ \lambda-3 & (\lambda+1)(\lambda-3)+4 & 0\\ 0 & -1 & \lambda-2\end{pmatrix}\to$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & (\lambda-1)^2 & 0 \\ 0 & -1 & \lambda-2 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & (\lambda-2)^2 & (\lambda-2)(\lambda-1)^2 \\ 0 & -1 & 0 \end{pmatrix} \to$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & (\lambda-2)(\lambda-1)^2 \end{pmatrix}$$

所以，所求的初等因子组为 $\lambda-2,(\lambda-1)^2$，于是有

$$\boldsymbol{A} \sim \boldsymbol{J} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

例1.27 求矩阵$\boldsymbol{A}$ 的Jordan 标准形，其中

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

解 求出 $\lambda\boldsymbol{I}-\boldsymbol{A}$ 的初等因子组.

$$\lambda\boldsymbol{I}-\boldsymbol{A}=\begin{pmatrix}\lambda-1&-2&-3&-4\\&\lambda-1&-2&-3\\&&\lambda-1&-2\\&&&\lambda-1\end{pmatrix}$$

$$D_4(\lambda)=(\lambda-1)^4$$

而
$$\begin{vmatrix} -2 & -3 & -4 \\ \lambda-1 & -2 & -3 \\ 0 & \lambda-1 & -2 \end{vmatrix} = -4\lambda(\lambda+1)$$

又 $D_3(\lambda)\big|D_4(\lambda)$，所以 $D_3(\lambda)=1$，从而 $D_2(\lambda)=D_1(\lambda)=1$，于是得 $\lambda\boldsymbol{I}-\boldsymbol{A}$ 不变因子

$$d_1(\lambda)=d_2(\lambda)=d_3(\lambda)=1, d_4(\lambda)=(\lambda-1)^4$$

$\lambda\boldsymbol{I}-\boldsymbol{A}$ 只有初等因子 $(\lambda-1)^4$，故

$$\boldsymbol{A} \sim \boldsymbol{J} = \begin{pmatrix} 1 & 1 & & \\ & 1 & 1 & \\ & & 1 & 1 \\ & & & 1 \end{pmatrix}$$

定理1.30 每个矩阵 $A \in C^{n\times n}$ 都与一个Jordan标准形相似，这个Jordan标准形除去其中Jordan块的排列次序外，是被A唯一确定的.

P 的求法：

假如

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & 1 \\ & & \lambda_2 \end{pmatrix}$$

其中$P = (x_1, x_2, x_3)$，于是有

$$A(x_1, x_2, x_3) = (x_1, x_2, x_3)\begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & 1 \\ & & \lambda_2 \end{pmatrix}$$

由此可得
$$\begin{cases}(\lambda_1 \boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}_1=0\\(\lambda_2 \boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}_2=0\\(\lambda_2 \boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}_3=-\boldsymbol{x}_2\end{cases}$$

x_3 称为A 的属于 λ_2 的广义特征向量.

求矩阵的Jordan标准形的方法：

方法1：(1) 求矩阵A-λI的初等因子；

(2) 根据初等因子写出A的Jordan标准形

方法2：(1) 求出矩阵A的全部特征值

$\lambda_1(r_1\text{重}),\lambda_2(r_2\text{重}),\ldots,\lambda_s(r_s\text{重})$

(2) 求出矩阵A的特征值λ_i对应的特征向量的线性无关的个数p_i;

(3) 基于λ_i对应p_i个Jordan块写出A的Jordan标准形。

例1.28 试计算使矩阵$\boldsymbol{A}$ 相似Jordan 标准形的相似变换矩阵$\boldsymbol{P}$，其中

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

解 （1）因为

$$\boldsymbol{A} \sim \begin{pmatrix} 2 & & \\ & 1 & 1 \\ & & 1 \end{pmatrix}$$

设 $\boldsymbol{P} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3)$

所以 $(2\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x}_1 = 0, (\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x}_2 = 0, (\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x}_3 = -\boldsymbol{x}_2$

由此得 $\boldsymbol{x}_1 = (0,0,1)^{\mathrm{T}}, \boldsymbol{x}_2 = (1,2,-1)^{\mathrm{T}}, \boldsymbol{x}_3 = (0,-1,1)^{\mathrm{T}}$

故得

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

（2）设 $\boldsymbol{P}=\left(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3,\boldsymbol{x}_4\right)$，因为

$$\boldsymbol{A} \sim \begin{pmatrix} 1 & 1 & & \\ & 1 & 1 & \\ & & 1 & 1 \\ & & & 1 \end{pmatrix}$$

所以

$$(\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}_1=0,(\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}_2=-\boldsymbol{x}_1,$$

$$(\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}_3=-\boldsymbol{x}_2,(\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}_4=-\boldsymbol{x}_3$$

由此得

$$\boldsymbol{x}_1=(8,0,0,0)^{\mathrm{T}},\boldsymbol{x}_2=(4,4,0,0)^{\mathrm{T}},$$

$$\boldsymbol{x}_3=(0,-1,2,0)^{\mathrm{T}},\boldsymbol{x}_4=(0,1,-2,1)^{\mathrm{T}}$$

于是得到

$$\boldsymbol{P}=\begin{pmatrix} 8 & 4 & 0 & 0 \\ 0 & 4 & -1 & 1 \\ 0 & 0 & 2 & -2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

应用：

例1.29 如果 $\lambda_1, \lambda_2, \cdots \lambda_s$ 是$\boldsymbol{A}$的特征值，证明 $\boldsymbol{A}^k$ 的特征值只能是 $\lambda_1^k, \lambda_2^k, \cdots \lambda_s^k$.

证　因为 $\boldsymbol{J} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ ，所以　$\boldsymbol{J}^k = \boldsymbol{P}^{-1}\boldsymbol{A}^k\boldsymbol{P}$

而　$\boldsymbol{J}^k = diag\left(\boldsymbol{J}_1^k, \boldsymbol{J}_2^k, \cdots, \boldsymbol{J}_s^k\right)$

$$\boldsymbol{J}_i^k = \begin{pmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix}^k = \begin{pmatrix} \lambda_i^k & * & \cdots & * \\ & \lambda_i^k & \ddots & \vdots \\ & & \ddots & * \\ & & & \lambda_i^k \end{pmatrix}$$

所以 $\lambda_1^k, \lambda_2^k, \cdots \lambda_s^k$ 是 $\boldsymbol{A}^k$的特征值.

例1.30 解微分方程组

$$\begin{cases} \dfrac{d\xi_1}{dt} = -\xi_1 + \xi_2 \\ \dfrac{d\xi_2}{dt} = -4\xi_1 + 3\xi_2 \\ \dfrac{d\xi_3}{dt} = \xi_1 + 2\xi_3 \end{cases}$$