高等线性代数选讲

复数、复向量与复矩阵

复数

$$z = a + bi$$

共轭

$$\begin{aligned} z &= a + bi \\ \bar{z} &= a - bi \end{aligned}$$

运算性质：

$$\begin{aligned} \overline{z_1 + z_2} &= \overline{z_1} + \overline{z_2} \\ \overline{z_1 z_2} &= \overline{z_1}\ \overline{z_2} \\ |z| &= \sqrt{a^2 + b^2} = \sqrt{z\,\bar{z}} \end{aligned}$$

实矩阵有共轭特征根：

$$\begin{aligned} A \in M_n(\mathbb{R}), x \in \mathbb{R}^n, Ax = \lambda x \\ A\bar{x} = \overline{Ax} = \overline{\lambda x} = \bar{\lambda}\bar{x} \end{aligned}$$

极坐标表示

$$z = re^{i\theta} = r(\cos\theta + i\sin\theta)$$

> 欧拉公式 $e^{i\theta} = \cos\theta + i\sin\theta$

> 例：
>
> 由$\cos 3\theta + i\sin 3\theta = e^{i3\theta} = (e^i\theta)^3 = (\cos\theta + i\sin\theta)^3$导出三倍角公式。

单位根

方程$z^n = 1$有n个根，记作$1, \omega, \omega^2, \cdots, \omega^{n-1}$。

范德蒙行列式：

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ x_1^2 & x_2^2 & \cdots & x_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1} & x_2^{n-2} & \cdots & x_n^{n-1} \end{vmatrix}$$

> 傅里叶矩阵——特殊的范德蒙矩阵
>
> $$F_n = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \cdots & \omega^{n-1} \\ 1 & \omega^2 & \omega^4 & \cdots & \omega^{2(n-1)} \\ 1 & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{n-1} & \omega^{2(n-1)} & \cdots & \omega^{(n-1)^2} \end{bmatrix}$$

埃尔米特矩阵和酉矩阵

复向量空间

复向量与复矩阵：

$$z \in \mathbb{C}^n$$

$$z = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix}$$

$$z^{\mathrm{H}} := \bar{z}^{\mathrm{T}} = \begin{bmatrix} \overline{z_1} & \overline{z_2} & \cdots & \overline{z_n} \end{bmatrix}$$

$$||z||^2 = z^{\mathrm{H}} z$$

$$A \in \mathbb{C}_{n \times m}, B \in \mathbb{C}_{m \times s}$$
$$(AB)^{\mathrm{H}} = B^{\mathrm{H}} A^{\mathrm{H}}$$

向量内积：

$$(u, v) := u^{\mathrm{H}} v$$

> 注意：$u^{\mathrm{H}} v \neq v^{\mathrm{H}} u$，也即$(u, v) \neq (v, u)$
>
> 另一种可能的定义方式：$(u, v) := u^{\mathrm{T}} \bar{v}$

$$(Au, v) = (u, A^{\mathrm{H}} v)$$
$$(Au)^{\mathrm{H}} v = u^{\mathrm{H}} (A^{\mathrm{H}} v)$$

埃尔米特矩阵

$$S = S^{\mathrm{H}}$$

> 实对称矩阵$S = S^{\mathrm{T}}$的性质：
>
> 1. 所有特征值为实数
> 2. 属于不同特征值的特征向量两两相交
> 3. 可对角化。$S = Q \Lambda Q^{\mathrm{T}}$，$Q$为正交阵（$QQ^{\mathrm{T}} = I$）。

埃尔米特矩阵的性质：

1. 所有特征值为实数
2. $\forall z \in \mathbb{C}^n, z^{\mathrm{H}} S z \in \mathbb{R}$
3. 属于不同特征值的特征向量两两正交

> 反埃尔米特矩阵的非零特征值为纯虚数。

酉矩阵

$$Q^{\mathrm{H}} Q = I$$

> 正交阵$Q^{\mathrm{T}} Q = I$

酉矩阵的性质：

1. 保长度$||Qz|| = ||z||$
2. 特征值的绝对值为1
3. 行列式的模为1

> 例：
>
> $$F_3=\frac{1}{\sqrt{3}}\begin{bmatrix}1&1&1\\1&\omega&\omega^2\\1&\omega^2&\omega^4\end{bmatrix}$$
> $$\omega^3=1,1+\omega+\omega^2=0$$
>
> 列向量互相正交，（单位化的）傅里叶矩阵是酉矩阵，$F_3^{-1}=F_3^{\mathrm{H}}$。
>
> $$F_3=\frac{1}{\sqrt{3}}\begin{bmatrix}1&1&1\\1&e^{2\pi i/3}&e^{4\pi i/3}\\1&e^{4\pi i/3}&e^{2\pi i/3}\end{bmatrix}$$
> $$F_3^{-1}=\frac{1}{\sqrt{3}}\begin{bmatrix}1&1&1\\1&e^{-2\pi i/3}&e^{-4\pi i/3}\\1&e^{-4\pi i/3}&e^{-2\pi i/3}\end{bmatrix}$$

快速傅里叶变换（略）

酉空间与酉变换

欧几里得空间与酉空间

欧几里得空间

设V是**实数域**上的**线性空间**，若对于V中任意两个元素α,β，都定义了内积运算(α,β)，且满足：

1. $(\alpha,\beta)=(\beta,\alpha)$
2. $(k\alpha,\beta)=k(\alpha,\beta)$
3. $(\alpha+\beta,\gamma)=(\alpha,\gamma)+(\beta,\gamma)$
4. $(\alpha,\alpha)\geq 0$，且$(\alpha,\alpha)=0\Leftrightarrow\alpha=\theta$

则称V是欧几里得空间。

> 欧氏空间的例子：n维向量空间$\mathbb{R}^n$
>
> 可以定义长度$||\alpha||=\sqrt{(\alpha,\alpha)}$与夹角$<\alpha,\beta>=arccos\frac{(\alpha,\beta)}{||\alpha||\cdot||\beta||}$

酉空间

设V是**复数域**上的**线性空间**，若对于V中任意两个元素α,β，都定义了内积运算(α,β)，且满足：

1. $(\alpha,\beta)=\overline{(\beta,\alpha)}$
2. $(k\alpha,\beta)=k(\alpha,\beta)=(\alpha,\bar{k}\beta)$
3. $(\alpha+\beta,\gamma)=(\alpha,\gamma)+(\beta,\gamma)$
4. $(\alpha,\alpha)\geq 0$，且$(\alpha,\alpha)=0\Leftrightarrow\alpha=\theta$

则称V是酉空间。

> 酉空间的例子：
>
> 1. n维复向量空间$\mathbb{C}^n$，内积$(\alpha,\beta):=\alpha^{\mathrm{T}}\bar{\beta}$
> 2. n阶复矩阵空间$M_n(\mathbb{C})$，内积$(A,B):=tr(A\bar{B}^{\mathrm{T}})$

定义模：$|\alpha|=\sqrt{(\alpha,\alpha)}$。

柯西-施瓦兹不等式：

$$|(\alpha,\beta)|^2 \le (\alpha,\alpha)(\beta,\beta) = (|\alpha|\cdot|\beta|)^2$$

正交：$\alpha \perp \beta \Leftrightarrow (\alpha,\beta)=0$

两两正交的非零向量组线性无关。

施密特正交化

$\alpha_1,\alpha_2,\alpha_3,\cdots,\alpha_n$为酉空间$V$中的一组基。

$$\begin{aligned}
\beta_1 &= \alpha_1 \\
\beta_2 &= \alpha_2 - \frac{(\alpha_2,\beta_1)}{(\beta_1,\beta_1)}\beta_1 \\
\beta_3 &= \alpha_3 - \frac{(\alpha_3,\beta_2)}{(\beta_2,\beta_2)}\beta_2 - \frac{(\alpha_3,\beta_1)}{(\beta_1,\beta_1)}\beta_1 \\
&\vdots \\
\beta_n &= \alpha_n - \frac{(\alpha_n,\beta_{n-1})}{(\beta_{n-1},\beta_{n-1})}\beta_{n-1} - \frac{(\alpha_n,\beta_{n-2})}{(\beta_{n-2},\beta_{n-2})}\beta_{n-2} - \cdots - \frac{(\alpha_n,\beta_1)}{(\beta_1,\beta_1)}\beta_1
\end{aligned}$$

$$\epsilon_i = \frac{\beta_i}{||\beta_i||}, i=1,2,3,\cdots,n$$

$\epsilon_1,\epsilon_2,\epsilon_3,\cdots,\epsilon_n$为酉空间$V$中的一组**标准正交基**。

> 可逆复方阵的**UR分解**：
>
> 设$A\in M_n(\mathbb{C})$可逆，则$\exists U,R\in M_n(\mathbb{C})$，使
>
> $$A=UR$$
>
> 其中U为酉矩阵，R为上三角矩阵（对角线上为正实数）。

酉、正规、埃尔米特变换

酉变换

设σ是酉空间V的一个线性变换，若对$\forall\alpha,\beta\in V$，

$$(\sigma\alpha,\sigma\beta)=(\alpha,\beta)$$

则称σ为酉变换。

酉变换的等价命题：

1. σ是酉变换
2. $\forall\alpha\in V, |\sigma\alpha|=|\alpha|$
3. σ将一组标准正交基变为另一组标准正交基
4. σ在任意一组标准正交基下的矩阵是酉矩阵

共轭变换

设σ是酉空间V的一个线性变换，若V上线性变换σ^*满足

$$(\sigma\alpha,\beta)=(\alpha,\sigma^*\beta)$$

则称σ^*是σ的共轭变换。

设σ和σ^*在一组标准正交基下的矩阵分别为A,A^*，则$A^*=A^H$。

> 恒等变换的共轭变换是恒等变换

> 酉变换的共轭变换是其逆变换
>
> 埃尔米特变换的共轭变换是其自身

共轭变换的性质：

1. $(\sigma^*)^* = \sigma$
2. $(k\sigma)^* = \bar{k}\sigma^*$
3. $(\sigma+\tau)^* = \sigma^* + \tau^*$
4. $(\sigma\tau)^* = \tau^*\sigma^*$

正规变换

设σ是酉空间V的一个线性变换，若

$$\sigma\sigma^* = \sigma^*\sigma$$

则称σ是正规变换。

> 恒等变换、酉变换、埃尔米特变换都是正规变换。

若σ是酉空间V的正规变换，λ是σ的特征值，ζ是对应的一个特征向量。则ζ也是σ^*的属于$\bar{\lambda}$的特征向量。

> 证明：
>
> $(\sigma-\lambda\epsilon)^* = \sigma^* - \bar{\lambda}\epsilon$，$((\sigma-\lambda\epsilon)\zeta, (\sigma-\lambda\epsilon)\zeta) = 0$。
>
> 由$\sigma\sigma^* = \sigma^*\sigma$，有$(\sigma-\lambda\epsilon)(\sigma-\lambda\epsilon)^* = (\sigma-\lambda\epsilon)^*(\sigma-\lambda\epsilon)$，则
>
> $$\begin{aligned}&((\sigma^*-\bar{\lambda}\epsilon)\zeta, (\sigma^*-\bar{\lambda}\epsilon)\zeta)\\ =&(\zeta, (\sigma^*-\bar{\lambda}\epsilon)^*(\sigma^*-\bar{\lambda}\epsilon)\zeta)\\ =&(\zeta, (\sigma-\lambda\epsilon)(\sigma-\lambda\epsilon)^*\zeta)\\ =&(\zeta, (\sigma-\lambda\epsilon)^*(\sigma-\lambda\epsilon)\zeta)\\ =&((\sigma-\lambda\epsilon)\zeta, (\sigma-\lambda\epsilon)\zeta)\\ =&0\end{aligned}$$
>
> $$\sigma^*\zeta = \bar{\lambda}\zeta$$

设σ是正规变换，则σ属于不同特征值的特征向量相互正交。

设σ是n维酉空间中的正规变换，则V中存在一组标准正交基，使得σ在这组基下的矩阵是对角矩阵。

> 证明：
>
> 使用归纳法，设σ在V中有特征值λ_1，η_1是对应的特征向量。
>
> 设M为$L(\lambda_1)$的正交补$L(\lambda_1)^{\perp}$。
>
> 证明M是σ和σ^*的不变子空间。
>
> > $$\begin{gathered}\forall\alpha\in M\\ (\sigma\alpha, \eta_1) = (\alpha, \sigma^*\eta_1) = (\alpha, \bar{\lambda}_1\eta_1) = \lambda_1(\alpha, \eta_1) = 0\\ \therefore \sigma\alpha\perp\eta_1, \sigma\alpha\in M\end{gathered}$$
>
> 将σ限制到M上，$\sigma|_M$是正规变换。
>
> 由归纳假设，M存在一组标准正交基$\eta_2, \eta_3, \cdots, \eta_n$，使得
>
> $$\sigma(\eta_2, \eta_3, \cdots, \eta_n) = (\eta_2, \eta_3, \cdots, \eta_n)diag(\lambda_2, \lambda_3, \cdots, \lambda_n)$$
>
> 而$\eta_1, \eta_2, \eta_3, \cdots, \eta_n$也是$V$中的一组标准正交基。

> $$\sigma(\eta_1,\eta_2,\cdots,\eta_n)=(\eta_1,\eta_2,\cdots,\eta_n)diag(\lambda_1,\lambda_2,\cdots,\lambda_n)$$

正规变换在标准正交基下的矩阵称为正规矩阵。

正规矩阵A有性质$AA^{\mathrm{H}}=A^{\mathrm{H}}A$。

任意正规矩阵A，存在酉矩阵U，使得

$$U^{\mathrm{H}}AU=diag(\lambda_1,\lambda_2,\cdots,\lambda_n)$$

> 正规矩阵A的对角化方法
>
> 1. 求特征值λ_i
> 2. 对每个特征值求特征向量，即$(\lambda_i I-A)x=0$的基础解系$\alpha_{i1},\alpha_{i2},\cdots,\alpha_{im_i}$。
> 3. 对每组特征向量做施密特正交化，得到特征子空间V_{λ_i}的标准正交基。
> 4. 各特征子空间正交，合并标准正交基，令酉矩阵
> $$U=(\alpha_{11},\alpha_{12},\cdots,\alpha_{1m_1},\alpha_{21},\cdots,\alpha_{2m_2},\cdots,\alpha_{s1},\cdots,\alpha_{sm_s})$$
> $$U^{-1}AU=diag(\lambda_1,\lambda_1,\cdots,\lambda_1,\lambda_2,\cdots,\lambda_2,\cdots,\lambda_s,\cdots,\lambda_s)$$

埃尔米特变换

设σ是酉空间V的一个线性变换，若对$\forall\alpha,\beta\in V$，

$$(\sigma\alpha,\beta)=(\alpha,\sigma\beta)$$

也即

$$\sigma=\sigma^*$$

则称σ为埃尔米特变换。

埃尔米特变换的特征值都是实数。

设σ是酉空间V上的埃尔米特变换，W是σ的不变子空间，则$W^{\perp}$也是σ的不变子空间。

埃尔米特二次型

埃尔米特二次型

n个复变量$x_1,x_2,\cdots,x_n$的二次齐次函数

$$f=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}\overline{x_i}x_j,\text{其中}a_{ij}=\overline{a_{ji}}$$

称为埃尔米特二次型。

令$A=(a_{ij})$，A是埃尔米特矩阵。

$$f=\begin{bmatrix}\overline{x_1} & \overline{x_2} & \cdots & \overline{x_n}\end{bmatrix}A\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}=x^{\mathrm{H}}Ax$$

$$\overline{f}=\overline{x^{\mathrm{H}}Ax}=(x^{\mathrm{H}}A^{\mathrm{H}}x)=x^{\mathrm{H}}Ax=f$$

故f为实值函数。

任意埃尔米特二次型$f=x^{\mathrm{H}}Ax$，存在满秩线性替换$x=Cy$，化为

$$f=\overline{y_1}y_1+\overline{y_2}y_2+\cdots+\overline{y_p}y_p-\overline{y_{p+1}}y_{p+1}-\overline{y_{p+2}}y_{p+2}-\cdots-\overline{y_r}y_r$$

埃尔米特二次型的规范形。其中r为埃尔米特二次型矩阵A的秩，也称为埃尔米特二次型的秩p，p称为正惯性指数，$r-p$称为负惯性指数，$p-(r-p)=2p-r$称为符号差。

正定、半正定、负定

设$f=x^{\mathrm{H}}Ax$是埃尔米特二次型。

1. 若对于$\forall x\neq 0,x^{\mathrm{H}}Ax>0$，则称$f$是正定的。
2. 若对于$\forall x\neq 0,x^{\mathrm{H}}Ax<0$，则称$f$是负定的。
3. 若对于$\forall x\neq 0,x^{\mathrm{H}}Ax\geq 0$，则称$f$是半正定的。
4. 若对于$\forall x\neq 0,x^{\mathrm{H}}Ax\leq 0$，则称$f$是半负定的。
5. 否则，称f是不定的。

设A是n阶埃尔米特矩阵，以下命题等价：

1. A正定
2. 以A为矩阵的埃尔米特二次型正惯性指数等于矩阵的秩等于n
3. 存在可逆复方阵P，使P复合同于I，即$P^{\mathrm{H}}AP=I$
4. 存在可逆复方阵C，使$A=C^{\mathrm{H}}C$
5. A的特征值全是正实数
6. 存在主对角线上均是正实数的上三角复矩阵R，使$A=R^{\mathrm{H}}R$
7. A的所有主子式都大于0
8. A的所有顺序主子式都大于0

> 设A是n阶埃尔米特矩阵，以下命题等价：
>
> 1. A半正定
> 2. 以A为矩阵的埃尔米特二次型正惯性指数p等于矩阵的秩但小于n
> 3. 存在可逆复方阵P，使$P^{\mathrm{H}}AP=diag(I_r,\mathbf{0})$
> 4. 存在复方阵C，使$A=C^{\mathrm{H}}C$
> 5. A的特征值全是非负实数

若尔当标准型

极小多项式、商空间和三角化

极小多项式

设F为数域，$A\in M_n(F)$，$f(x)\in F[x]$为F上的多项式。

若$f(A)=0$，则称$f(x)$是A的化零多项式。

> Hamilton-Cayley定理：
>
> A的特征多项式$f_A(\lambda)=|\lambda I-A|$是$A$的化零多项式。

设$A\in M_n(F)$，A的化零多项式中次数最低的首项系数为1的多项式叫做A的极小多项式，记作$m_A(x)$。

设$A\in M_n(F)$，$\forall f(x)\in F[x]$，

$$f(A)=0\Leftrightarrow m_A(x)|f(x)$$

> $\Rightarrow$：

> 由带余除法，$f(x)=q(x)m_A(x)+r(x),\deg r(x)<\deg m_A(x)$
>
> $$f(A)=q(A)m_A(A)+r(A)=0$$
> $$r(A)=0$$
>
> 而$r(A)$不是化零多项式，$r(x)=0$

相似矩阵有相同的极小多项式。

> $A\sim B, A=C^{-1}BC$。
>
> 证明$m_A(x)|m_B(x), m_B(x)|m_A(x)$

商空间

V是数域F上的线性空间，W是V的子空间。对于$\alpha,\beta\in V$，若$\alpha-\beta\in W$，则称α与β模W同余，记作$\alpha\equiv\beta\pmod W$。

> 模W同余是一种等价关系，因为它满足反身性、对称性、传递性。

定义模W的同余类$\alpha+W=\{\alpha+\beta|\beta\in W\}$，$\alpha$叫做这个同余类的一个代表。

同余类的性质：

1. 同余类与选取的代表无关
2. 两个模W的同余类或者完全相等，或者没有公共元素

按模W同余这一等价关系定义的等价类的集合$\overline{V}=\{\alpha+W|\alpha\in V\}$称为线性空间$V$的商集。

$\overline{V}$中元素定义加法与数乘运算：

1. $(\alpha+W)+(\beta+W)=(\alpha+\beta)+W$
2. $k(\alpha+W)=k\alpha+W$

$\overline{V}$中运算满足线性空间的七条公理，称为V的商空间，记作V/W。

> $\dim V/W=\dim V-\dim W$

诱导变换

设σ是线性空间V上的一个线性变换，W是σ的不变子空间。

在商空间V/W上定义变换$\tau:V/W\to V/W$，$\tau(\alpha+W)=\sigma(\alpha)+W$。

易验证τ是线性变换，称为σ诱导的变换。

> 诱导变换的矩阵
>
> 设$\dim W=r$，
>
> $$\begin{aligned}&\sigma(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_r,\varepsilon_{r+1},\cdots,\varepsilon_n)\\=&(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_r,\varepsilon_{r+1},\cdots,\varepsilon_n)\begin{bmatrix}a_{11}&\cdots&a_{1r}&a_{1r+1}&\cdots&a_{1n}\\\vdots&&\vdots&\vdots&&\vdots\\a_{r1}&\cdots&a_{rr}&a_{rr+1}&\cdots&a_{rn}\\0&\cdots&0&a_{r+1r+1}&\cdots&a_{r+1n}\\\vdots&&\vdots&\vdots&&\vdots\\0&\cdots&0&a_{nr+1}&\cdots&a_{nn}\end{bmatrix}\\=&(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_r,\varepsilon_{r+1},\cdots,\varepsilon_n)\begin{bmatrix}A_1&A_3\\\mathbf{0}&A_2\end{bmatrix}\end{aligned}$$

> $$
\begin{aligned}
\tau(\varepsilon_{r+1}+W) &= \sigma\varepsilon_{r+1}+W \\
&= a_{1r+1}\varepsilon_1+a_{2r+1}\varepsilon_2+\cdots+a_{nr+1}\varepsilon_n+W \\
&= a_{r+1r+1}\varepsilon_{r+1}+a_{r+2r+1}\varepsilon_{r+2}+\cdots+a_{nr+1}\varepsilon_n+W \\
&= a_{r+1r+1}(\varepsilon_{r+1}+W)+\cdots+a_{nr+1}(\varepsilon_n+W) \\
&\vdots \\
\tau(\varepsilon_n+W) &= a_{r+1n}(\varepsilon_{r+1}+W)+\cdots+a_{nn}(\varepsilon_n+W)
\end{aligned}
$$
>
> 所以
>
> $$
\begin{aligned}
&\tau(\varepsilon_{r+1}+W,\varepsilon_{r+2}+W,\cdots,\varepsilon_n+W) \\
=&(\varepsilon_{r+1}+W,\varepsilon_{r+2}+W,\cdots,\varepsilon_n+W)\begin{bmatrix} a_{r+1r+1} & a_{r+2r+1} & \cdots & a_{nr+1} \\ a_{r+2r+1} & a_{r+2r+2} & \cdots & a_{nr+2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{nr+1} & a_{nr+2} & \cdots & a_{nn} \end{bmatrix} \\
=&(\varepsilon_{r+1}+W,\varepsilon_{r+2}+W,\cdots,\varepsilon_n+W)A_2
\end{aligned}
$$

矩阵的三角化

设σ是复数域$\mathbb{C}$上的线性空间V的一个线性变换，则σ的任意真不变子空间必包含在一个维数增加1的不变子空间中。

> 证明:
>
> 取V的真不变子空间W，设$\tau : V/W \to V/W$为σ的诱导变换。
>
> 设λ是τ的特征值，$\tau(\alpha+W)=\lambda(\alpha+W)=(\lambda\alpha+W)$。
>
> 于是$\sigma\alpha=\lambda\alpha+w, w\in W$。
>
> 令$U=L(\alpha\cup W)$。$\forall\gamma\in U,\gamma=k\alpha+\beta,\beta\in W$。
>
> $\therefore \sigma\gamma=\sigma(k\alpha+\beta)=k(\lambda\alpha+w)+\sigma\beta=k\lambda\alpha+kw+\sigma\beta\in U$

从子空间$W_0=\{\theta\}$开始，依次构造维数递增1的不变子空间
$\{\theta\}=W_0\subset W_1\subset W_2\subset\cdots\subset W_n=V$。

可以在V中取一组基$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$，使$\sigma$在这组基下的矩阵为三角阵。

Schur定理：任意n阶复方阵总相似于上（下）三角矩阵。

循环矩阵和空间直和分解

子空间的直和

若W_1,W_2是V的子空间，$\forall\alpha\in W_1+W_2$，$\alpha=\alpha_1+\alpha_2$分解是唯一的，其中$\alpha_1\in W_1,\alpha_2\in W_2$。

则称W_1+W_2是W_1和W_2的直和，记作$W_1\oplus W_2$。

> 直和的例子：平面$\mathbb{R}^2$中的x轴与y轴。

W_1+W_2是直和的充要条件:

1. 零向量表示唯一（$\theta=\alpha_1+\alpha_2,\alpha_1\in W_1,\alpha_2\in W_2 \;\Rightarrow\; \alpha_1=\alpha_2=\theta$)
2. $W_1\cap W_2=\theta$
3. $\dim(W_1+W_2)=\dim(W_1)+\dim(W_2)$

幂零变换和循环变换

设$\sigma\in L(V)$，若$\exists m\in\mathbb{N},\sigma^m=0$，则称$\sigma$是幂零变换，满足性质的最小的$m$称为$\sigma$的幂零次数。

幂零变换的所有特征值都是0。非零变换的幂零变换不可对角化。

设$A\in M_n$，若$\exists m\in\mathbb{N},A^m=0$，则称$A$是幂零矩阵。

设$\sigma\in L(V),\alpha\in V$。若$\alpha,\sigma\alpha,\sigma^2\alpha,\cdots,\sigma^{n-1}\alpha$构成$V$的一组基，且$\sigma^n\alpha=\theta$。则称$\sigma$是一个循环变换，$\alpha,\sigma\alpha,\cdots,\sigma^{n-1}\alpha$，由循环基生成的子空间叫循环子空间。

循环变换在循环基下的矩阵称为循环矩阵，形如

$$N=\begin{bmatrix}0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0\end{bmatrix}$$

> 循环变换是特殊的幂零变换，幂零次数等于n。
>
> 循环子空间是不变子空间。

对于复数域上线性空间V的任意幂零变换σ，V必可以分解成若干循环子空间的直和：

$$V=T_1\oplus T_2\oplus\cdots\oplus T_t$$

使σ限制在每一个循环子空间上的变换$\sigma|_{T_i}$是循环变换。

> 记$W_i=Im\ \sigma^i,i=0,1,2,\cdots,m$。
>
> $\because\sigma^{i+1}\alpha=\sigma^i(\sigma\alpha)$
>
> $\therefore Im\ \sigma^{i+1}\subset Im\ \sigma^i$
>
> $\{\theta\}=W_m\subset W_{m-1}\subset\cdots\subset W_0=V$
>
> ---
>
> $W_m=\{\theta\},W_{m-1}\subset\ker\sigma$。
>
> 记$\dim W_{m-1}=p_{m-1}$，在W_{m-1}中取一组基$\varepsilon_1^{(m-1)},\varepsilon_2^{(m-1)},\cdots,\varepsilon_{p_{m-1}}^{(m-1)}$。
>
> 则$\sigma\varepsilon_i^{(m-1)}=\theta,i=1,2,\cdots,p_{m-1}$。
>
> 在W_{m-2}中有一组元素$\varepsilon_1^{(m-2)},\varepsilon_2^{(m-2)},\cdots,\varepsilon_{p_{m-1}}^{(m-2)}$满足$\sigma\varepsilon_i^{(m-2)}=\varepsilon_i^{(m-1)},i=1,2,\cdots,p_{m-1}$。
>
> 可以证明
>
> $$\begin{cases}\varepsilon_1^{(m-1)},\varepsilon_2^{(m-1)},\cdots,\varepsilon_{p_{m-1}}^{(m-1)}\\ \varepsilon_1^{(m-2)},\varepsilon_2^{(m-2)},\cdots,\varepsilon_{p_{m-1}}^{(m-2)}\end{cases}$$
>
> 线性无关。
>
> 将$\varepsilon_1^{(m-2)},\varepsilon_2^{(m-2)},\cdots,\varepsilon_{p_{m-1}}^{(m-2)}$扩充为$W_{m-2}$的一组基，使得被扩充的基元素在$\ker\sigma$里。
>
> 设$\dim W_{m-2}-\dim W_{m-1}=p_{m-2}$，则
>
> $$\begin{cases}\varepsilon_1^{(m-1)},\varepsilon_2^{(m-1)},\cdots,\varepsilon_{p_{m-1}}^{(m-1)}\\ \varepsilon_1^{(m-2)},\varepsilon_2^{(m-2)},\cdots,\varepsilon_{p_{m-1}}^{(m-2)},\varepsilon_{p_{m-1}+1}^{(m-2)},\cdots,\varepsilon_{p_{m-2}}^{(m-2)}\end{cases}$$
>
> 是W_{m-2}的基，其中$\sigma\varepsilon_{p_{m-1}+j}^{(m-2)}=\theta,j=1,2,\cdots,p_{m-2}-p_{m-1}$。
>
> 同理一步步构造出$W_{m-3},W_{m-4},\cdots,W_0=V$的基：

> $$\begin{cases} \varepsilon_1^{(m-1)},\cdots,\varepsilon_{p_{m-1}}^{(m-1)} \\ \varepsilon_1^{(m-2)},\cdots,\varepsilon_{p_{m-1}}^{(m-2)},\varepsilon_{p_{m-1}+1}^{(m-2)},\cdots,\varepsilon_{p_{m-2}}^{(m-2)} \\ \varepsilon_1^{(m-3)},\cdots,\varepsilon_{p_{m-1}}^{(m-3)},\varepsilon_{p_{m-1}+1}^{(m-3)},\cdots,\varepsilon_{p_{m-2}}^{(m-3)},\varepsilon_{p_{m-2}+1}^{(m-3)},\cdots,\varepsilon_{p_{m-3}}^{(m-3)} \\ \vdots \\ \varepsilon_1^{(0)},\cdots,\varepsilon_{p_{m-1}}^{(0)},\varepsilon_{p_{m-1}+1}^{(0)},\cdots,\varepsilon_{p_{m-2}}^{(0)},\varepsilon_{p_{m-2}+1}^{(0)},\cdots,\varepsilon_{p_{m-3}}^{(0)},\cdots,\varepsilon_{p_1+1}^{(0)},\cdots,\varepsilon_{p_0}^{(0)} \end{cases}$$
>
> 每一列的基元素构成一个σ的循环子空间$T_i, i=1,2,\cdots,p_0$，$\sigma|_{T_i}$是循环变换。

将基$\{\varepsilon_i^{(j)}\}$从上到下，从左到右顺序排列，σ在这组基下的矩阵就是$diag(N_1,N_2,\cdots,N_{p_0})$，其中$N_i$为循环矩阵，阶数为循环子空间$T_i$的维数。

根子空间与空间分解定理

根向量与根子空间

设$\sigma\in L(V)$，λ是σ的特征值，$\alpha\in V$。若存在$m\in\mathbb{N}$，使得

$$(\sigma-\lambda\varepsilon)^m\alpha=\theta$$

则称α是属于特征值λ的根向量。

> 设$\sigma\in L(V)$，任意λ，令
>
> $$U_\lambda=\{\alpha\in V|\exists m\in\mathbb{N},(\sigma-\lambda\varepsilon)^m\alpha=\theta\}$$
>
> 则：
>
> 1. U_λ是V的子空间
> 2. U_λ是σ的不变子空间
> 3. $U_\lambda\neq\{\theta\}\Leftrightarrow\lambda$是$\sigma$的特征值

设$\sigma\in L(V)$，λ是σ的特征值，U_λ由所有属于λ的根向量构成，称为特征值λ的根子空间。

$(\sigma-\lambda\varepsilon)|_{U_\lambda}$是幂零变换。

> 当$\mu\neq\lambda$时，$(\sigma-\mu\varepsilon)|_{U_\lambda}$是可逆变换。
>
> 可逆变换⇔双射⇔$\dim\ker\sigma=0$（有限维线性空间V上线性变换σ满足$\dim\ker\sigma+\dim\mathrm{Im}\sigma=\dim V$）
>
> 设$(\sigma-\mu\varepsilon)\alpha=\theta$，$\sigma\alpha=\mu\alpha$。
>
> 则$(\sigma-\lambda\varepsilon)\alpha=(\mu-\lambda)\alpha$。
>
> 又$(\sigma-\lambda\varepsilon)|_{U_\lambda}$是幂零变换，特征值为0。
>
> $\therefore\alpha=\theta,\dim\ker\sigma=0$

设σ有s个不同的特征值$\lambda_1,\lambda_2,\cdots,\lambda_s$，对应的根子空间为$U_1,U_2,\cdots,U_s$，则根子空间$U_1,U_2,\cdots,U_s$的和$U_1+U_2+\cdots+U_s$是直和。

> 证明：
>
> 证零元素表示唯一，即
>
> $$\alpha_i\in U_i,i=1,2,\cdots,s$$
> $$\alpha_1+\alpha_2+\cdots+\alpha_s=\theta\Leftrightarrow\alpha_i=\theta,i=1,2,\cdots,s$$
>
> 数学归纳法，$\exists m\in\mathbb{N}$，使得
>
> $$(\sigma-\lambda_s\varepsilon)^m\alpha_s=\theta$$

> 令
>
> $$\begin{aligned}\beta_i = (\sigma - \lambda_s\varepsilon)^m\alpha_i, i = 1, 2, \cdots, s-1 \\ \beta_i \in U_i \\ \beta_1 + \beta_2 + \cdots + \beta_{s-1} &= (\sigma - \lambda_s\varepsilon)^m(\alpha_1 + \alpha_2 + \cdots + \alpha_{s-1}) \\ &= (\sigma - \lambda_s\varepsilon)^m(\alpha_1 + \alpha_2 + \cdots + \alpha_{s-1} + \alpha_s) \\ &= \theta\end{aligned}$$
>
> 由归纳假设，
>
> $$\beta_i = (\sigma - \lambda_s\varepsilon)^m\alpha_i = \theta, i = 1, 2, \cdots, s-1$$
>
> 又由$(\sigma - \lambda_s\varepsilon)|_{U_i}$是可逆变换，
>
> $$\alpha_i = \theta, i = 1, 2, \cdots, s-1$$
>
> 故
>
> $$\alpha_s = (\alpha_1 + \alpha_2 + \cdots + \alpha_s) - (\alpha_1 + \alpha_2 + \cdots + \alpha_{s-1}) = \theta$$

对于σ的特征值λ，根子空间U_λ的维数等于λ的代数重数。

> 证明：
>
> 设特征值λ_0的代数重数为n_0，$\dim U_\lambda = s$。
>
> 在U_λ中选一组基，扩充为V的基。
>
> 在这组基下σ的矩阵
>
> $$A = \begin{bmatrix} B & C \\ \mathbf{0} & D \end{bmatrix}$$
>
> σ的特征多项式
>
> $$\det(\lambda I - A) = \det(\lambda I - B)\det(\lambda I - D) = (\lambda - \lambda_0)^s \det(\lambda I - D)$$
>
> 自然地，$s \le n_0$。
>
> 若$s < n_0$，$(\lambda - \lambda_0)|\det(\lambda I - D)$。记$\tau : V/U_{\lambda_0} \to V/U_{\lambda_0}$是$\sigma$的诱导变换，$\lambda_0$是$\tau$的特征值，设$\alpha + U_{\lambda_0}$是对应的特征向量。
>
> $$\begin{aligned}\tau(\alpha + U_{\lambda_0}) &= \lambda_0(\alpha + U_{\lambda_0}) \\ &= \sigma\alpha + U_{\lambda_0} \\ (\sigma - \lambda_0\varepsilon)\alpha = \gamma, \gamma \in U_{\lambda_0}\end{aligned}$$
>
> 因此
>
> $$\begin{gathered}\exists m \in \mathbb{N}, (\sigma - \lambda_0\varepsilon)^m\gamma = \theta \\ (\sigma - \lambda_0\varepsilon)^{m+1}\alpha = (\sigma - \lambda_0\varepsilon)^m\gamma = \theta \\ \alpha \in U_{\lambda_0}\end{gathered}$$
>
> 矛盾。

空间分解定理和若尔当标准形定理

设V是复数域上n维线性空间，$\sigma \in L(V)$，$\lambda_1, \lambda_2, \cdots, \lambda_s$是$\sigma$的全部相异特征值，代数重数分别为$n_0, n_1, \cdots, n_s$。特征多项式$f(\lambda) = \det(\lambda\varepsilon - \sigma) = (\lambda - \lambda_1)^{n_1}(\lambda - \lambda_2)^{n_2}\cdots(\lambda - \lambda_s)^{n_s}$，则

$$V = U_{\lambda_1} \oplus U_{\lambda_2} \oplus \cdots \oplus U_{\lambda_s}$$

形如

$$\begin{bmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix}$$

的矩阵称为若尔当块，由若尔当块构成的准对角矩阵称为若尔当标准形。

设V是复数域$\mathbb{C}$上的线性空间，$\sigma \in L(V)$，则存在一组基，使σ在这组基下的矩阵是若尔当标准形。

> 证明:
>
> V可以分解为根子空间$U_{\lambda_1}, U_{\lambda_2}, \cdots, U_{\lambda_s}$的直和。
>
> 设$\rho_i = (\sigma - \lambda_i \varepsilon)|_{U_{\lambda_i}}$是幂零变换，$U_{\lambda_i}$可以分解为循环子空间$T_{i1}, T_{i2}, \cdots, T_{it_i}$的直和，使$\rho_i|_{T_{ij}}$是循环变换。
>
> 在T_{ij}上选取一组循环基，ρ_i对应的矩阵为循环矩阵N。
>
> $\sigma = \rho_i + \lambda_i \varepsilon$，在循环基下对应的矩阵为
>
> $$N + \lambda I = \begin{bmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix}$$
>
> 即若尔当块。
>
> 把所有循环基按顺序排列，即得若尔当标准形。

任意复矩阵与若尔当标准形相似。

若尔当标准形的极小多项式与计算

若尔当标准形的极小多项式

设A是准对角矩阵$A = diag(A_1, A_2, \cdots, A_s)$，则$A$的极小多项式是$A_i$的极小多项式的最小公倍式。

$$m_A(x) = [m_{A_1}(x), m_{A_2}(x), \cdots, m_{A_s}(x)]$$

> 证明:
>
> $$f(A) = diag(f(A_1), f(A_2), \cdots, f(A_s))$$
> $$m_A(A) = 0 \Leftrightarrow m_{A_i}(A) = 0$$

k阶若尔当块

$$J_k = \begin{bmatrix} \lambda_0 & 1 & & \\ & \lambda_0 & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_0 \end{bmatrix}$$

的极小多项式为

$$m_{J_k}(x) = (x - \lambda_0)^k$$

若尔当标准形的极小多项式为各若尔当块的极小多项式的最小公倍式。

> 矩阵A可对角化$\Leftrightarrow A$的极小多项式无重根

若尔当标准形的计算

特征值λ_i对应的特征子空间U_{λ_i}中j维循环子空间的数目

$$\begin{aligned} d_j &= 2\dim\ker(\sigma-\lambda_i\varepsilon)^j - \dim\ker(\sigma-\lambda_i\varepsilon)^{j-1} - \dim\ker(\sigma-\lambda_i\varepsilon)^{j+1} \\ &= rank(A-\lambda_i I)^{j-1} + rank(A-\lambda_i I)^{j+1} - 2rank(A-\lambda_i I)^j \end{aligned}$$

若尔当块的数量

$$t_i = \dim\ker(\sigma-\lambda_i\varepsilon) = n - rank(A-\lambda_i I)$$

> 例：
>
> 设$A=\begin{bmatrix} 1 & 2 & 2 & & \\ 2 & 1 & 2 & & \\ 2 & 2 & 1 & & \\ & & & 5 & 0 \\ & & & 2 & 5 \end{bmatrix}$，求$A$的若尔当标准形。
>
> $$|\lambda I - A| = (\lambda+1)^2(\lambda-5)^3$$
>
> $\lambda_1=-1, n_1=2$：
>
> $$A+I=\begin{bmatrix} 2 & 2 & 2 & & \\ 2 & 2 & 2 & & \\ 2 & 2 & 2 & & \\ & & & 6 & 0 \\ & & & 2 & 6 \end{bmatrix}, rank(A+I)=3$$
>
> $t_1=2$，基图：
>
> $$\varepsilon_1, \varepsilon_2$$
>
> 两个一阶若尔当块。
>
> $\lambda_2=5, n_2=3$：
>
> $$A-5I=\begin{bmatrix} -4 & 2 & 2 & & \\ 2 & -4 & 2 & & \\ 2 & 2 & -4 & & \\ & & & 0 & 0 \\ & & & 2 & 0 \end{bmatrix}, rank(A+I)=3$$
>
> $t_2=2$，基图：
>
> $$\begin{matrix} \varepsilon_1, \varepsilon_2 \\ \varepsilon_3 \end{matrix}$$
>
> 一个一阶若尔当块，一个二阶若尔当块。
>
> $$J=\begin{bmatrix} -1 & & & & \\ & -1 & & & \\ & & 5 & 1 & \\ & & & 5 & \\ & & & & 5 \end{bmatrix}$$

P矩阵的计算

由线性方程组$(A-\lambda_i I)x=0$求得第一行向量$x_1^{(j)}$，由$(A-\lambda_i I)x=x_1^{(j)}$求得对应的2级广义特征向量$x_2^{(j)}$，如此求得所有的广义特征向量构成的若干若尔当链，排列即得矩阵P，使

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

矩阵分析初步

矩阵函数的微积分

函数矩阵

以实变量x的函数$a_{ij}(x)$为元素的矩阵

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

称为函数矩阵。其中$A_{ij}(x)$是定义在$[a,b]$上的实函数。

函数矩阵的运算：

1. 函数矩阵加法$A(x)+B(x)$
2. 函数与函数矩阵的数乘$k(x)A(x)$
3. 函数矩阵乘法$A(x)B(x)$
4. 函数矩阵的转置$A^{\mathrm{T}}(x)$

对于n阶函数矩阵$A(x)$，若存在n阶函数矩阵$B(x)$，使得对于任意$x\in[a,b]$，都有

$$A(x)B(x)=B(x)A(x)=I$$

则称$A(x)$在$[a,b]$上可逆，$B(x)=A^{-1}(x)$是$A(x)$的逆矩阵。

n阶方阵$A(x)$在区间$[a,b]$上可逆$\Leftrightarrow \det A(x)$在区间$[a,b]$**处处不为**$0$

且若$A(x)$可逆，则

$$A^{-1}(x)=\frac{1}{\det A(x)}\mathrm{adj}\ A(x)$$

> 伴随矩阵
>
> $$\mathrm{adj}\ A(x)=\begin{bmatrix} C_{11}(x) & C_{21}(x) & \cdots & C_{n1}(x) \\ C_{12}(x) & C_{22}(x) & \cdots & C_{n2}(x) \\ \vdots & \vdots & \ddots & \vdots \\ C_{1n}(x) & C_{2n}(x) & \cdots & C_{nn}(x) \end{bmatrix}$$
>
> 其中$C_{ij}(x)$是$a_{ij}(x)$的代数余子式。

函数矩阵$A(x)$在$[a,b]$上**不恒等于**0的子式的最高阶数称为$A(x)$的秩。

> 函数矩阵：可逆 $\Rightarrow$ 满秩，满秩 $\nRightarrow$ 可逆

函数矩阵的微积分

若$A(x)=(a_{ij}(x))_{m\times n}$的所有元素$a_{ij}(x)$在$x=x_0$处有极限，且

$$\lim_{x\to x_0}a_{ij}(x)=a_{ij}$$

则称$A(x)$在$x=x_0$处有极限，记作

$$\lim_{x\to x_0}A(x)=A$$

其中$A=(a_{ij})_{m\times n}$。

设$\lim_{x\to x_0}A(x)=A,\lim_{x\to x_0}B(x)=B$，有性质：

1. $\lim_{x\to x_0}(A(x)\pm B(x))=A\pm B$
2. $\lim_{x\to x_0}(kA(x))=kA,k\in\mathbb{R}$
3. $\lim_{x\to x_0}(A(x)B(x))=AB$

若$A(x)=(a_{ij}(x))_{m\times n}$的所有元素$a_{ij}(x)$在$x=x_0$处连续，即

$$\lim_{x\to x_0}a_{ij}(x)=a_{ij}(x_0)$$

则称$A(x)$在$x=x_0$处连续，记作

$$\lim_{x\to x_0}A(x)=A(x_0)$$

其中$A(x_0)=(a_{ij}(x_0))_{m\times n}$。

若$A(x)=(a_{ij}(x))_{m\times n}$的所有元素$a_{ij}(x)$在$x=x_0$处可导，则称$A(x)$在$x=x_0$处可导，记作

$$A'(x_0)=\left.\frac{\mathrm{d}A(x)}{\mathrm{d}x}\right|_{x=x_0}=\lim_{\Delta x\to 0}\frac{A(x_0+\Delta x)-A(x_0)}{\Delta x}=(a'_{ij}(x_0))$$

导数性质：

1. $A(x)$是常数矩阵$\Leftrightarrow\frac{\mathrm{d}A(x)}{\mathrm{d}x}=\mathbf{0}$
2. 若$A(x),B(x)$可导，则$A(x)\pm B(x)$也可导，且
$$\frac{\mathrm{d}}{\mathrm{d}x}(A(x)\pm B(x))=\frac{\mathrm{d}A(x)}{\mathrm{d}x}\pm\frac{\mathrm{d}B(x)}{\mathrm{d}x}$$
3. 若$k(x),A(x)$可导，则$k(x)A(x)$也可导，且
$$\frac{\mathrm{d}}{\mathrm{d}x}(k(x)A(x))=\frac{\mathrm{d}k(x)}{\mathrm{d}x}A(x)+k(x)\frac{\mathrm{d}A(x)}{\mathrm{d}x}$$
4. 若$A(x),B(x)$可导，$A(x)B(x)$可乘，则$A(x)B(x)$也可导，且
$$\frac{\mathrm{d}}{\mathrm{d}x}(A(x)B(x))=\frac{\mathrm{d}A(x)}{\mathrm{d}x}B(x)+A(x)\frac{\mathrm{d}B(x)}{\mathrm{d}x}$$

> 矩阵乘法不满足交换律：
>
> $$\frac{\mathrm{d}}{\mathrm{d}x}(A^2(x))=\frac{\mathrm{d}A(x)}{\mathrm{d}x}A(x)+A(x)\frac{\mathrm{d}A(x)}{\mathrm{d}x}\neq 2A(x)\frac{\mathrm{d}A(x)}{\mathrm{d}x}$$

> 逆矩阵的导数：
>
> $$A(x)A^{-1}(x)=I$$
>
> 两边求导，
>
> $$\begin{aligned}\frac{\mathrm{d}A(x)}{\mathrm{d}x}A^{-1}(x)+A(x)\frac{\mathrm{d}A^{-1}(x)}{\mathrm{d}x}&=\mathbf{0}\\\frac{\mathrm{d}A^{-1}(x)}{\mathrm{d}x}&=-A^{-1}(x)\frac{\mathrm{d}A(x)}{\mathrm{d}x}A^{-1}(x)\end{aligned}$$

若$A(x)$为函数矩阵，$x=f(t)$是t的函数，$A(x)$与$f(t)$可导，则

$$\frac{\mathrm{d}A(x)}{\mathrm{d}t}=\frac{\mathrm{d}A(x)}{\mathrm{d}x}\frac{\mathrm{d}f(t)}{\mathrm{d}t}=\frac{\mathrm{d}f(t)}{\mathrm{d}t}\frac{\mathrm{d}A(x)}{\mathrm{d}x}$$

$A(x)$的k阶导数：

$$A^{(k)}(x)=\frac{\mathrm{d}^kA(x)}{\mathrm{d}x^k}=\frac{\mathrm{d}}{\mathrm{d}x}(\frac{\mathrm{d}^{k-1}A(x)}{\mathrm{d}x^{k-1}}),k\geq 2$$

若$A(x)=(a_{ij}(x))_{m\times n}$的所有元素$a_{ij}(x)$在区间$[a,b]$上可积，则称$A(x)$在区间$[a,b]$上可积，且

$$\int_a^b A(x)\mathrm{d}x=\left(\int_a^b a_{ij}(x)\mathrm{d}x\right)_{m\times n}$$

定积分性质：

1. $\int_a^b(A(x)\pm B(x))\mathrm{d}x=\int_a^b A(x)\mathrm{d}x\pm\int_a^b B(x)\mathrm{d}x$
2. $\int_a^b(kA(x))\mathrm{d}x=k\int_a^b A(x)\mathrm{d}x,k\in\mathbb{R}$

函数向量的线性相关性

设$\alpha_i(x)=(a_{i1}(x),a_{i2}(x),\cdots,a_{in}(x))^{\mathrm{T}},i=1,2,\cdots,m$是在区间$[a,b]$上连续的函数向量。若存在不全为零的实数$k_1,k_2,\cdots,k_m$，使得对于$\forall x\in[a,b]$，

$$k_1\alpha_1(x)+k_2\alpha_2(x)+\cdots+k_m\alpha_m(x)=\mathbf{0}$$

成立，则称在区间$[a,b]$上，$\alpha_1(x),\alpha_2(x),\cdots,\alpha_m(x)$线性相关。否则称为线性无关。

设$\alpha_1(x),\alpha_2(x),\cdots,\alpha_m(x)$是在区间$[a,b]$上连续的函数向量，记

$$g_{ij}=\int_a^b\alpha_i^{\mathrm{T}}(x)\alpha_j(x)\mathrm{d}x$$

则矩阵$G=(g_{ij})_{m\times m}$称为$\alpha_1(x),\alpha_2(x),\cdots,\alpha_m(x)$的格拉姆矩阵，$\det G$称为格拉姆行列式。

在区间$[a,b]$上连续的函数向量$\alpha_1(x),\alpha_2(x),\cdots,\alpha_m(x)$线性无关的充要条件是格拉姆矩阵满秩。

> 证明：
>
> 设
>
> $$k_1\alpha_1(x)+k_2\alpha_2(x)+\cdots+k_m\alpha_m(x)=\mathbf{0}$$
>
> 两边左乘$\alpha_i^{\mathrm{T}}(x)$，对x在$[a,b]$上积分
>
> $$\begin{aligned}\int_a^b\alpha_i^{\mathrm{T}}(x)(k_1\alpha_1(x)+k_2\alpha_2(x)+\cdots+k_m\alpha_m(x))\mathrm{d}x=0\\=\sum_{j=1}^m k_j\int_a^b\alpha_i^{\mathrm{T}}(x)\alpha_j(x)\mathrm{d}x\\=\sum_{j=1}^m k_jg_{ij}\end{aligned}$$
>
> 令$k=(k_1,k_2,\cdots,k_m)^{\mathrm{T}}$，
>
> $$Gk=\mathbf{0}$$
>
> G满秩 $\Leftrightarrow$ k只有零解。
>
> ---
>
> 若G不满秩，k存在非零解。令

> $$\alpha(x)=\sum_{i=1}^{m}k_i\alpha_i(x)$$
> $$\int_a^b\alpha^{\mathrm{T}}(x)\alpha(x)=0$$
> $$\alpha(x)\equiv 0$$

设$\alpha_1(x),\alpha_2(x),\cdots,\alpha_m(x)$是在区间$[a,b]$上有$m-1$阶导数的函数向量，记

$$A(x)=\begin{bmatrix}\alpha_1^{\mathrm{T}}(x)\\ \alpha_2^{\mathrm{T}}(x)\\ \vdots\\ \alpha_m^{\mathrm{T}}(x)\end{bmatrix}=(a_{ij})_{m\times n}$$

令

$$W(x)=(A(x),A'(x),\cdots,A^{(m-1)}(x))_{m\times nm}$$

称为$\alpha_1(x),\alpha_2(x),\cdots,\alpha_m(x)$的朗斯基矩阵。

设$W(x)$是$\alpha_1(x),\alpha_2(x),\cdots,\alpha_m(x)$的朗斯基矩阵。若存在某点$x_0\in[a,b]$，使$W(x_0)$的秩等于$m$，则向量$\alpha_1(x),\alpha_2(x),\cdots,\alpha_m(x)$在$[a,b]$上线性无关。

> 反证法：
>
> 若$\alpha_1(x),\alpha_2(x),\cdots,\alpha_m(x)$线性相关，则存在不全为零的实数$k_i$，使得
>
> $$k_1\alpha_1(x)+k_2\alpha_2(x)+\cdots+k_m\alpha_m(x)=\mathbf{0}$$
>
> 逐次求导，
>
> $$\begin{cases}k_1\alpha_1(x)+k_2\alpha_2(x)+\cdots+k_m\alpha_m(x)=\mathbf{0}\\ k_1\alpha_1'(x)+k_2\alpha_2'(x)+\cdots+k_m\alpha_m'(x)=\mathbf{0}\\ k_1\alpha_1''(x)+k_2\alpha_2''(x)+\cdots+k_m\alpha_m''(x)=\mathbf{0}\\ \quad\vdots\\ k_1\alpha_1^{(m-1)}(x)+k_2\alpha_2^{(m-1)}(x)+\cdots+k_m\alpha_m^{(m-1)}(x)=\mathbf{0}\end{cases}$$
>
> 代入$x=x_0$，写成矩阵形式，
>
> $$\begin{bmatrix}\alpha_1(x)&\alpha_2(x)&\cdots&\alpha_m(x)\\ \alpha_1'(x)&\alpha_2'(x)&\cdots&\alpha_m'(x)\\ \vdots&\vdots&\ddots&\vdots\\ \alpha_1^{(m-1)}(x)&\alpha_2^{(m-1)}(x)&\cdots&\alpha_m^{(m-1)}(x)\end{bmatrix}\begin{bmatrix}k_1\\ k_2\\ \vdots\\ k_m\end{bmatrix}=0$$
>
> 记$k=(k_1,k_2,\cdots.k_m)^{\mathrm{T}}$，
>
> $$W^{\mathrm{T}}(x_0)k=\mathbf{0}$$
>
> 由$k\neq\mathbf{0}$，$W^{\mathrm{T}}(x_0)$秩小于m，矛盾。

矩阵序列和矩阵级数

矩阵序列

设矩阵序列$\{A_k\}\subseteq M_n(\mathbb{C})$，其中$A_k=(a_{ij}^{(k)})_{n\times n}$。如果$n^2$个数列$\{a_{ij}^{(k)}\}$都收敛，则称矩阵序列$\{A_k\}$是收敛的。以数列$\{a_{ij}^{(k)}\}$的极限$a_{ij}$为元素的矩阵$A=(a_{ij})\in M_n(\mathbb{C})$称为矩阵序列$\{A_k\}$的极限，记作

$$\lim_{k\to\infty} A_k = A$$

否则称$\{A_k\}$是发散的。

若$A_k \to A, B_k \to B$，有性质：

1. $aA_k + bB_k \to aA + bB, a, b \in C$
2. $A_k B_k \to AB$
3. $PA_kQ \to PAQ, P, Q \in M_n(\mathbb{C})$是可逆阵
4. 若A与A_k均可逆，$A_k^{-1} \to A^{-1}$

设$A_k \in M_n(\mathbb{C})$，则

$$\lim_{k\to\infty} A^k = \mathbf{0}$$

的充要条件是A的所有特征值的模都小于1。

> 证明：
>
> 设A的若尔当标准形$J = diag(J_1, J_2, \cdots, J_s)$。
>
> 存在可逆阵P，使
>
> $$\begin{gathered} P^{-1}AP = J \\ A = PJP^{-1} \\ A^k = PJ^kP^{-1} \end{gathered}$$
>
> 则
>
> $$\lim_{k\to\infty} A^k = \mathbf{0} \Leftrightarrow \lim_{k\to\infty} J^k = \mathbf{0} \Leftrightarrow \lim_{k\to\infty} J_i^k = \mathbf{0}, i = 1, 2, \cdots, s$$
>
> 而
>
> $$J_i^k = (\lambda_i I + N)^k = \begin{bmatrix} \lambda_i^k & C_k^1\lambda_i^{k-1} & C_k^1\lambda_i^{k-2} & \cdots & C_k^{d-1}\lambda^{k-d+1} \\ & \lambda_i^k & C_k^1\lambda_i^{k-1} & \cdots & C_k^{d-1}\lambda^{k-d+2} \\ & & \lambda_i^k & \cdots & C_k^{d-1}\lambda^{k-d+3} \\ & & & \ddots & \vdots \\ & & & & \lambda_i^k \end{bmatrix}_{d\times d}$$
>
> 于是
>
> $$\lim_{k\to\infty} J_i^k = \mathbf{0} \Leftrightarrow |\lambda_i| < 1, i = 1, 2, \cdots, s$$

矩阵级数

设有矩阵序列$A_1, A_2, \cdots, A_k, \cdots$，称$A_1 + A_2 + \cdots + A_k + \cdots$为矩阵级数，前$k$项的和$S_k = A_1 + A_2 + \cdots + A_k$称为级数的部分和。

若

$$\lim_{k\to\infty} S_k = S$$

则称级数$A_1 + A_2 + \cdots + A_k + \cdots$收敛，$S$是级数的和。否则称为发散。

显然，级数收敛的充要条件是n^2个数项级数$\sum_{k=1}^{\infty} a_{ij}^{(k)}$都收敛。

矩阵级数的性质：

1. 若$\sum_{k=1}^{\infty} A_k$收敛，则$\lim_{k\to\infty} A_k = \mathbf{0}$

2. 若$\sum_{k=1}^{\infty} A_k = S, \sum_{k=1}^{\infty} B_k = T$，则$\sum_{k=1}^{\infty}(A_k \pm B_k) = S \pm T$
3. 若$\sum_{k=1}^{\infty} A_k = S$，则$\sum_{k=1}^{\infty}(cA_k) = cS, c \in \mathbb{C}$
4. 若$P, Q \in M_n(\mathbb{C})$可逆，$\sum_{k=1}^{\infty} A_k$收敛，则$\sum_{k=1}^{\infty} PA_kQ$也收敛，且
$$\sum_{k=1}^{\infty} PA_kQ = P\left(\sum_{k=1}^{\infty} A_k\right)Q$$

> 若$\sum_{k=1}^{\infty} A_k$收敛，则A的特征值的模都小于1。

设矩阵级数$A_1 + A_2 + \cdots + A_k + \cdots$，其中$A_k = (a_{ij}^{(k)}) \in M_n(\mathbb{C})$。如果$n^2$个数项级数$\sum_{k=1}^{\infty} a_{ij}^{(k)}$都绝对收敛，则称矩阵级数绝对收敛。

矩阵级数$\sum_{k=1}^{\infty} A_k$绝对收敛的充要条件是级数$\sum_{k=1}^{\infty} ||A_k||$收敛，其中$||A_k|| = \max_{i,j} |(A_k)_{ij}|$。

若$P, Q \in M_n(\mathbb{C})$可逆，$\sum_{k=1}^{\infty} A_k$绝对收敛，则$\sum_{k=1}^{\infty} PA_kQ$也绝对收敛，且

$$\sum_{k=1}^{\infty} PA_kQ = P\left(\sum_{k=1}^{\infty} A_k\right)Q$$

设$\sum_{k=1}^{\infty} A_k, \sum_{k=1}^{\infty} B_k$是两个矩阵级数，且满足

1. B_k的所有元素$b_{ij}^{(k)} \geq 0$
2. A_k和B_k的对应元素满足$|a_{ij}^{(k)}| \leq b_{ij}^{(k)}$
3. $\sum_{k=1}^{\infty} B_k$收敛

则$\sum_{k=1}^{\infty} A_k$绝对收敛。

矩阵函数

矩阵多项式

设$f(x) \in \mathbb{C}[x], A \in M_n(\mathbb{C})$，若

$$f(x) = a_m x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0$$

则

$$f(A) = a_m A^m + a_{m-1}A^{m-1} + \cdots + a_1 A + a_0 I \quad (a_m \neq 0)$$

称为m次矩阵多项式。

计算矩阵多项式：

$$\begin{gathered} P^{-1}AP = J \\ f(A) = Pf(J)P^{-1} \end{gathered}$$

设$A \in M_n(\mathbb{C})$，A的所有相异特征值组成的集合叫做矩阵A的谱。

设$A \in M_n(\mathbb{C})$，$\lambda_1, \lambda_2, \cdots, \lambda_s$是$A$的互异特征值，$A$的极小多项式为

$$m_A(x) = (x-\lambda_1)^{d_1}(x-\lambda_2)^{d_2}\cdots(x-\lambda_s)^{d_s}$$

假设给定的函数$f(x)$有足够多阶导数，则

$$\begin{gathered} f(\lambda_1), f'(\lambda_1), \cdots, f^{(d_1-1)}(\lambda_1), \\ f(\lambda_2), f'(\lambda_2), \cdots, f^{(d_2-1)}(\lambda_2), \\ \vdots \\ f(\lambda_s), f'(\lambda_s), \cdots, f^{(d_s-1)}(\lambda_s) \end{gathered}$$

称为$f(x)$关于矩阵A的谱上的值。如果$f(x)$关于矩阵A的谱上的值都存在，则称函数$f(x)$在A的谱上有定义。

> A的极小多项式$m_A(x)$在A的谱上的值均为零。

若$g(x)$在A的谱上的值均为零，则$m_A(x)|g(x)$。

> 证明:
>
> $$\because g(\lambda_i)=g'(\lambda_i)=\cdots=g^{(d_i-1)}(\lambda_i)=0$$
> $$\therefore (x-\lambda_i)^{d_i}|g(x)$$

设$A\in M_n(\mathbb{C})$，$f(x),g(x)$是多项式，则

$$f(A)=g(A)\ \Leftrightarrow\ f(x)\text{与}g(x)\text{在}A\text{的谱上的值相同}$$

> $\Rightarrow$:
>
> 记
>
> $$h(x)=f(x)-g(x),h(A)=0$$
> $$\therefore m_A(x)|h(x)$$
> $$f(x)-g(x)=q(x)m_A(x)$$
> $$f^{(k)}(\lambda_i)-g^{(k)}(\lambda_i)=(q(x)m_A(x))^{(k)}\Big|_{x=\lambda_i}=0$$
>
> 即
>
> $$f^{(k)}(\lambda_i)=g^{(k)}(\lambda_i),i=1,2,\cdots,s,k=0,1,\cdots,d_i-1$$
>
> $\Leftarrow$:
>
> $$h(x)=f(x)-g(x)$$
>
> 在A的谱上的值为零。
>
> $$m_A(x)|h(x)$$
> $$f(x)-g(x)=q(x)m_A(x)$$
> $$f(A)=g(A)$$

推论：A的化零多项式在A的谱上的值为零。

矩阵函数

设函数$f(x)$在矩阵A的谱上有定义，$g(x)$是任意多项式。若$f(x)$与$g(x)$在A的谱上有相同的值

$$f^{(k)}(\lambda_i)=g^{(k)}(\lambda_i),\quad i=1,2,\cdots,s,k=0,1,\cdots,d_i-1$$

则矩阵函数$f(A)$定义为

$$f(A)=g(A)$$

$g(A)$称为$f(A)$的定义多项式。

> 例:
>
> 设$A=\begin{bmatrix}6&-1\\3&2\end{bmatrix}$，$f(x)=e^{2x}$，计算$f(A)$。
>
> $$m_A(x)=(x-3)(x-5)$$
>
> $\lambda_1=3,\lambda_2=5$，拉格朗日插值:

> $$
\begin{aligned}
p(x) &= f(\lambda_1)\frac{x-\lambda_2}{\lambda_1-\lambda_2}+f(\lambda_2)\frac{x-\lambda_1}{\lambda_2-\lambda_1} \\
&= -\frac{1}{2}e^6(x-5)+\frac{1}{2}e^{10}(x-3)
\end{aligned}
$$
>
> $$
\therefore f(A)=e^{2A}=-\frac{1}{2}e^6(A-5I)+\frac{1}{2}e^{10}(A-3I)=\frac{1}{2}\begin{bmatrix} 3e^{10}-e^6 & -e^{10}+e^6 \\ 3e^{10}-3e^6 & -e^{10}+3e^6 \end{bmatrix}
$$

矩阵函数性质：

1. $Af(A)=f(A)A$
2. 若$f(x)=f_1(x)+f_2(x)$，则$f(A)=f_1(A)+f_2(A)$
3. 若$f(x)=f_1(x)f_2(x)$，则$f(A)=f_1(A)f_2(A)$
4. 设$A=diag(A_1,A_2,\cdots,A_s)$，$f(A)=diag(f(A_1),f(A_2),\cdots,f(A_s))$
5. 若$B=P^{-1}AP$，则$f(B)=P^{-1}f(A)P$

> 矩阵函数$f(A)$的若尔当表示：
>
> $$
f(A)=Pdiag(f(J_1),f(J_2),\cdots,f(J_s))P^{-1}
$$

设若尔当块$J_i=\lambda_i I+N\in M_t(\mathbb{C})$，若$f(x)$在$\lambda_i$邻域内有$t-1$阶导数，则

$$
f(J_i)=\begin{bmatrix}
f(\lambda_i) & f'(\lambda_i) & \frac{1}{2!}f''(\lambda_i) & \cdots & \frac{1}{(t-1)!}f^{(t-1)}(\lambda_i) \\
 & f(\lambda_i) & f'(\lambda_i) & \cdots & \frac{1}{(t-2)!}f^{(t-2)}(\lambda_i) \\
 & & f(\lambda_i) & \cdots & \frac{1}{(t-3)!}f^{(t-3)}(\lambda_i) \\
 & & & \ddots & \vdots \\
 & & & & f(\lambda_i)
\end{bmatrix}
$$

> 证明：J_i的极小多项式$m_{J_i}(x)=(x-\lambda_i)^t$，$f(x)$在$J_i$的谱上的值为$f(\lambda_i),f'(\lambda_i),\cdots,f^{(t-1)}(\lambda_i)$。
>
> 令
>
> $$
g(x)=f(\lambda_i)+f'(\lambda_i)(x-\lambda_i)+\frac{1}{2!}f''(\lambda_i)(x-\lambda_i)^2+\cdots+\frac{1}{(t-1)!}f^{(t-1)}(\lambda_i)(x-\lambda_i)^{t-1}
$$
>
> 是$f(J_i)$的定义多项式，故
>
> $$
\begin{aligned}
f(J_i) &= f(\lambda_i)I+f'(\lambda_i)(J_i-\lambda_i I)+\frac{1}{2!}f''(\lambda_i)(J_i-\lambda_i I)^2+\cdots+\frac{1}{(t-1)!}f^{(t-1)}(\lambda_i)(J_i-\lambda_i I)^{t-1} \\
&= f(\lambda_i)I+f'(\lambda_i)N_t+\frac{1}{2!}f''(\lambda_i)N_t^2+\cdots+\frac{1}{(t-1)!}f^{(t-1)}(\lambda_i)N_t^{t-1}
\end{aligned}
$$

矩阵函数的幂级数表示

设复变量函数$f(x)$是在开圆$|x-\lambda_0|<r$内解析的函数，即在$|x-\lambda_0|<r$内可以展开成幂级数

$$
f(x)=\sum_{k=0}^{\infty}a_k(x-\lambda_0)^k
$$

则只要方阵A的所有特征值都在开圆$|x-\lambda_0|<r$内，就有

$$
f(A)=\sum_{k=0}^{\infty}a_k(A-\lambda_0 I)^k
$$

> 对若尔当块$J=\lambda I+N$证明即可。
>
> $f(x)$可以在λ点展开为

$$f(x)=\sum_{k=0}^{\infty}\frac{f^{(k)}(\lambda)}{k!}(x-\lambda)^k$$

记$b_k=\frac{f^{(k)}(\lambda)}{k!}$,

又由$f(x)$也可展开为

$$f(x)=\sum_{k=0}^{\infty}a_k(x-\lambda_0)^k$$

于是

$$\sum_{k=0}^{\infty}a_k(x-\lambda_0)^k=\sum_{k=0}^{\infty}b_k(x-\lambda)^k$$
$$\sum_{k=0}^{\infty}a_k(J-\lambda_0 I)^k=\sum_{k=0}^{\infty}b_k(J-\lambda I)^k$$

故

$$f(J)=\sum_{k=0}^{t-1}\frac{1}{k!}f^{(k)}(\lambda)N^k=\sum_{k=0}^{t-1}b_kN^k=\sum_{k=0}^{\infty}b_k(J-\lambda_0 I)^k=\sum_{k=0}^{\infty}a_k(J-\lambda_0 I)^k$$

结合初等函数的幂级数表示，可以求得一些常见初等函数的矩阵函数：

$$e^A=\sum_{k=0}^{\infty}\frac{1}{k!}A^k$$
$$\ln(I+A)=\sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k}A^k,\quad \rho(A)<1$$
$$(I+A)^{\alpha}=\sum_{k=0}^{\infty}\frac{\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-k+1)}{k!}A^k,\quad \rho(A)<1$$
$$sinA=\sum_{k=0}^{\infty}\frac{(-1)^k}{(2k+1)!}A^{2k+1}$$
$$cosA=\sum_{k=0}^{\infty}\frac{(-1)^k}{(2k)!}A^{2k}$$

其中谱半径$\rho(A)=\max|\lambda_i|$。

微分方程组的矩阵分析解法

一阶常系数线性微分方程组

$$\frac{\mathrm{d}x}{\mathrm{d}t}=Ax$$

其中$A\in M_n(\mathbb{C})$，x是t的n维函数向量$(x_1(t),x_2(t),\cdots,x_n(t))^{\mathrm{T}}$。

$$\frac{\mathrm{d}e^{At}}{\mathrm{d}t}=Ae^{At}$$

> 证明：
>
> $$\frac{\mathrm{d}e^{At}}{\mathrm{d}t}=\frac{\mathrm{d}}{\mathrm{d}t}(\sum_{k=0}^{\infty}\frac{1}{k!}A^kt^k)=\sum_{k=1}^{\infty}\frac{1}{(k-1)!}A^kt^{k-1}=A\sum_{k=1}^{\infty}\frac{1}{(k-1)!}A^{k-1}t^{k-1}=Ae^{At}$$

方程$\frac{\mathrm{d}x}{\mathrm{d}t}=Ax$的满足初始条件$x|_{t=0}=x_0$的解存在且唯一，且

$$x(t)=e^{At}x_0$$

> 唯一性证明：
>
> 设$x(t)=(x_1(t),x_2(t),\cdots,x_n(t))^{\mathrm{T}}$是方程的解，且$x(0)=x_0$。
>
> 将$x(t)$在$t=0$处展开，
>
> $$x(t)=\sum_{k=0}^{\infty}\frac{x^{(k)}(0)}{k!}t^k$$
>
> 由$x(t)$是方程的解，
>
> $$x^{(k)}(t)=A^kx(t),x^{(k)}(0)=A^kx(0)$$
>
> 故
>
> $$\begin{aligned}x(t)&=\sum_{k=0}^{\infty}\frac{A^kx(0)}{k!}t^k\\&=\sum_{k=0}^{\infty}\frac{A^kt^k}{k!}x(0)\\&=e^{At}x(0)\end{aligned}$$

用特征值与特征向量表示微分方程组的解

希望找到形如

$$x(t)=\alpha(t)e^{\lambda t}$$

的解。设

$$\alpha(t)=\frac{t^{k-1}}{(k-1)!}\alpha_1+\frac{t^{k-2}}{(k-2)!}\alpha_2+\cdots+\frac{t}{1!}\alpha_{k-1}+\alpha_k$$

其中$\alpha_i\in\mathbb{C}^n,\alpha_1\neq\mathbf{0}$。

> $$\begin{gathered}\frac{\mathrm{d}}{\mathrm{d}t}x(t)=\alpha'(t)e^{\lambda t}+\lambda\alpha(t)e^{\lambda t}=Ax(t)=A\alpha(t)e^{\lambda t}\\\therefore\alpha'(t)=(A-\lambda I)\alpha(t)\end{gathered}$$
>
> 关于t逐次求导，
>
> $$\begin{gathered}\alpha^{(i)}(t)=(A-\lambda I)^i\alpha(t),\quad i=1,2,\cdots,k\\\alpha^{(k-1)}(t)=\alpha_1=(A-\lambda I)^{k-1}\alpha(t)\end{gathered}$$
>
> 而
>
> $$(A-\lambda I)^k\alpha(t)=\alpha^{(k)}(t)=0$$
>
> 故
>
> $$(A-\lambda I)\alpha_1=(A-\lambda I)^k\alpha(t)=0$$
>
> λ是A的特征值，α_1是A属于λ的特征向量。
>
> 而
>
> $$\alpha^{(k-2)}(t)=t\alpha_1+\alpha_2=(A-\lambda I)^{k-2}\alpha(t)$$
>
> 两边左乘$(A-\lambda I)$，

> $$(A-\lambda I)(t\alpha_1+\alpha_2)=(A-\lambda I)\alpha_2=(A-\lambda I)^{k-1}\alpha(t)=\alpha_1$$
>
> 故α_2是由α_1生成的广义特征向量。
>
> 同理，$\alpha_1,\alpha_2,\cdots,\alpha_k$形成一条$A$的属于特征值$\lambda$的若尔当链。

$$x(t)=\alpha(t)e^{\lambda t}$$
$$\alpha(t)=\frac{t^{k-1}}{(k-1)!}\alpha_1+\frac{t^{k-2}}{(k-2)!}\alpha_2+\cdots+\frac{t}{1!}\alpha_{k-1}+\alpha_k$$

上式定义的$x(t)$是方程组$\frac{\mathrm{d}x}{\mathrm{d}t}=Ax$的解，当且仅当$\alpha_1,\alpha_2,\cdots,\alpha_k$构成$A$的一条属于特征值$\lambda$的若尔当链。

由若尔当链定义的解叫做方程组的基本解，它们的线性组合叫一般解。

> 例：
>
> 求$\frac{\mathrm{d}x(t)}{\mathrm{d}t}=\begin{bmatrix}4&2&2\\-2&0&-4\\0&0&4\end{bmatrix}x(t)$的一般解。
>
> $$A=\begin{bmatrix}4&2&2\\-2&0&-4\\0&0&4\end{bmatrix}$$
> $$f_A(\lambda)=(\lambda-4)(\lambda-2)^2$$
>
> 属于$\lambda=4$的特征向量$\alpha_1^{(1)}=(0,-1,1)^{\mathrm{T}}$，属于$\lambda=2$的特征向量$\alpha_2^{(1)}=(2,-2,0)^{\mathrm{T}}$以及广义特征向量$\alpha_2^{(2)}=(1,0,0)^{\mathrm{T}}$。
>
> 所求的一般解为
>
> $$x(t)=k_1\begin{bmatrix}0\\-1\\1\end{bmatrix}e^{4t}+k_2\begin{bmatrix}2\\-2\\0\end{bmatrix}e^{2t}+k_3\left(\begin{bmatrix}2\\-2\\0\end{bmatrix}t+\begin{bmatrix}1\\0\\0\end{bmatrix}\right)e^{2t}$$