矩阵论复习笔记

修改时间：2018.12.26　　E-mail: zhushuai0403@163.com

1. 线性空间与线性变换

（1）**线性空间**的定义：

> 以$\alpha, \beta, \gamma, \ldots$为元素的非空集合$V$，数域$F$，定义**两种运算**：加法$\forall \alpha, \beta \in V$，$\alpha + \beta \in V$；数乘$\forall k \in F, \alpha \in V, k\alpha \in V$。**满足8条：加法交换律、加法结合律、数乘结合律、两个分配律，0元存在，1元存在，负元存在。称V为数域F上的线性空间。**

（2）证明一组向量是线性空间的基，两步走：

- 证明这组向量线性无关；
- 证明线性空间任意向量可由这组向量表示。

（3）如果$\{E_{ij}, i = 1, 2, \ldots, m; j = 1, 2, \ldots, n\}$是矩阵空间$R^{m \times n}$的一组基，则$\dim R^{m \times n} = m \times n$。

> **注：这里有前提条件，实际上$\dim R^{m \times n}$并不是总等于$m \times n$，如2015年填空题第2题。**

（4）$\alpha_1, \alpha_2, \ldots, \alpha_n$是线性空间$V_n(F)$的一组基，对于$\forall \beta \in V$，

$$\beta = (\alpha_1\ \alpha_2\ \ldots\ \alpha_n) \begin{bmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{bmatrix} = (\alpha_1\ \alpha_2\ \ldots\ \alpha_n) X$$

其中X称为向量β在基$(\alpha_1\ \alpha_2\ \ldots\ \alpha_n)$下对应的坐标。

$V_n(F)$中向量组$\{\beta_1\ \beta_2\ \ldots\ \beta_m\}$线性相关的充要条件是坐标向量组$\{X_1, X_2, \ldots, X_m\}$是线性相关组。

（5）设$(\alpha_1\ \alpha_2\ \ldots\ \alpha_n)$和$(\beta_1\ \beta_2\ \ldots\ \beta_n)$是$n$维线性空间$V_n(F)$中的两组基，则有$C \in F^{m \times n}$

$$(\beta_1\ \beta_2\ \ldots\ \beta_n) = (\alpha_1\ \alpha_2\ \ldots\ \alpha_n) C$$

其中C称为从基设$(\alpha_1\ \alpha_2\ \ldots\ \alpha_n)$到$(\beta_1\ \beta_2\ \ldots\ \beta_n)$的**过渡矩阵。**

> **重要推论：如果向量$\alpha \in V_n(F)$，α在两组基下对应坐标分别是X和Y，则有：**

$$\alpha = (\alpha_1\ \alpha_2\ \ldots\ \alpha_n) X$$

$$\alpha = (\beta_1\ \beta_2\ \ldots\ \beta_n) Y$$

> **显然有：$X = CY$。**

---

(6) 设W是线性空间$V_n(F)$的非空子集合，则W是$V_n(F)$的子空间的充要条件是：

- **若$\alpha,\beta\in W$，则$\alpha+\beta\in W$**
- **若$\alpha\in W$，$k\in F$，则$k\alpha\in W$**

也就是说只需要验证**对加法和数乘封闭**即可。

---

(7) 设W_1，W_2是线性空间V的子空间，则：

- W_1与W_2的**交空间**为：$W_1\cap W_2=\{\alpha|\alpha\in W_1\ and\ \alpha\in W_2\}$
- W_1与W_2的**和空间**为：$W_1+W_2=\{\alpha|\alpha=\alpha_1+\alpha_2,\ \alpha_1\in W_1,\ \alpha_2\in W_2\}$

> **两个重要的维数公式**

- $\dim(W_1\cap W_2)\leqslant\dim W_i\leqslant\dim(W_1+W_2)\leqslant\dim V$
- $\dim W_1+\dim W_2=\dim(W_1+W_2)+\dim(W_1\cap W_2)$

> **直和子空间：如果$W=W_1+W_2$，并且$W_1\cap W_2=\{0\}$，那么称W是W_1与W_2的直和子空间，表示为$W=W_1\oplus W_2$。**

> **直和补子空间：对n维空间V的任何子空间W，设$\alpha_1,\dots,\alpha_r$为W的基，$r<n$，把它们扩充为V的基$\{\alpha_1,\dots,\alpha_r;\beta_{r+1},\dots,\beta_n\}$，$\quad U=L\{\beta_{r+1},\dots,\beta_n\}$，有$V=W\oplus U$成立，则称$U$是$W$的直和补子空间。**

---

(8) 若$(\alpha_1\ \alpha_2\ \dots\ \alpha_n)$是线性空间$V_n(F)$的一组基，则

$$V_n(F)=L\{\alpha_1,\alpha_2,..,\alpha_n\}$$

对一个矩阵$A\in F^{m\times n}$，可以得到两个与A相关的子空间：

$$N(A)=\{X|AX=0\}\subseteq F^n$$

$$R(A)=L\{A_1,A_2,\dots,A_n\}\subseteq F^m$$

其中$N(A)$称为矩阵A的**零空间**，$R(A)$称为矩阵A的**列空间**。

---

(9) 内积：

- **欧氏空间的内积**：$(\alpha,\beta)=\alpha^T\beta;\quad(A,B)=tr(AB^T)$
- **酉空间的内积**：$(\alpha,\beta)=\beta^H\alpha;\quad(A,B)=tr(B^HA)$

> 柯西不等式：$|(\alpha,\beta)|^2\leqslant(\alpha,\alpha)(\beta,\beta)$

> **正交补子空间：设U为内积空间$V_n(F)$的一个子空间，定义$V_n(F)$上的一个子集$U^\perp=\{\alpha\mid\alpha\in V_n(F),\ \forall\beta\in U,\ (\alpha,\beta)=0\}$称为$U$的正交补子空间，有$V_n(F)=U\oplus U^\perp$。**

---

(10) 设T是线性空间$V_n(F)$上的线性变换，则满足$T(k_1\alpha_1+k_2\alpha_2)=k_1T(\alpha_1)+k_2T(\alpha_2)$，则有：

> **像空间**：$R(T)=\{\beta|\ \exists\alpha\in V_n(F),s.t.\ \beta=T(\alpha)\}$**是$V_n(F)$上的子空间，称为$T$的像空间；$\dim R(T)$称为$T$的秩。**

> **零空间**：$N(T)=\{\alpha|\ T(\alpha)=0\}$**是$V_n(F)$上的子空间，称为$T$的零空间；$\dim N(T)$称为$T$的零度。**

（11）设T为$V_n(F)$上的线性变换，$\{\alpha_1,\alpha_2,\dots,\alpha_n\}$是$V_n(F)$的基，若存在$n$阶方阵$A$，有：

$$T(\alpha_1\ \alpha_2\ \dots\ \alpha_n)=(\alpha_1\ \alpha_2\ \dots\ \alpha_n)A$$

称A为T在基$\{\alpha_1,\alpha_2,\dots,\alpha_n\}$下的矩阵。

- 设α与$T(\alpha)$在基$\{\alpha_1,\alpha_2,\dots,\alpha_n\}$下的坐标分别是$X$与$Y$，则有：$Y=AX$。
- 设$\{\alpha_1,\alpha_2,\dots,\alpha_n\}$和$\{\beta_1,\beta_2,\dots,\beta_n\}$是$V_n(F)$的两组基，且有$(\beta_1\ \beta_2\ \dots\ \beta_n)=(\alpha_1\ \alpha_2\ \dots\ \alpha_n)C$；$T$在两组基下的变换矩阵分别是$A$与$B$，则$B=C^{-1}AC$。

（12）设T是线性空间$V_n(F)$上的线性变换，W是$V_n(F)$的子空间，如果$\forall\alpha\in W$，$T(\alpha)\in W$，即值域$T(W)\subseteq W$，则称W是T的**不变子空间**。

> **重要例题**
>
> **设T是欧式空间R^3上的线性变换，对R^3中单位矢量u，$\forall x\in R^3$，$T(x)=x-(1-k)(x,u)u$，问：T的不变子空间的直和分解以及相应的矩阵分解。**

> 答：对向量u有 $T(u)=u-(1-k)(u,u)u=u-(1-k)u=ku$ 所以以u为基向量的空间是不变子空间，表示为$L\{u\}$;
> 同理，对于u的正交补子空间$u^{\perp}$，对于任意向量$X\in u^{\perp}$，有
> $T(X)=X-(1-k)(X,u)u=X-0=X$ 于是另一个不变子空间为$u^{\perp}$；即$R^3=L\{u\}\oplus u^{\perp}$。
>
> 显然有$L\{u\}$是一维空间，特征值k对应的特征向量是$u_1=u$；那么$u^{\perp}$就是二维空间，特征值1对应两个线性无关的特征向量，可以找到两个单位正交特征向量u_2,u_3，所以相应的矩阵分解为$\begin{bmatrix}k&&\\&1&\\&&1\end{bmatrix}$，对应的特征向量组 $\{u_1,u_2,u_3\}$为标准正交基。

（13）**正交变换（酉变换）：线性变换T不改变向量内积，即$(T(\alpha),T(\beta))=(\alpha,\beta)$。**

- 正交变换T关于任一标准正交基的矩阵C满足$C^TC=CC^T=I$；酉变换关于任一标准正交基的矩阵U满足$U^HU=UU^H=I$。
- 正交矩阵的行列式为±1；酉矩阵的行列式的模长为1。

（14）**常见的正交变换**

- R^2**上绕原点逆时针旋转θ角的线性变换T_θ称为*正交变换***，在标准正交基下对应的变换矩阵$\begin{bmatrix}\cos\theta&-\sin\theta\\\sin\theta&\cos\theta\end{bmatrix}$是**正交矩阵**。
- 空间R^3**上绕过原点的直线l旋转θ角的变化T_{L_θ}为*正交变换***，在标准正交基下对应的变换矩阵$\begin{bmatrix}1&&\\&\cos\theta&-\sin\theta\\&\sin\theta&\cos\theta\end{bmatrix}$是**正交矩阵**。

---

2. Jordan标准形

（1）若有$T(\xi)=\lambda\xi$，称λ为T的特征值，ξ为T的特征向量。如果A是线性变换T对应的矩阵，那么，λ和ξ也是A的特征值和特征向量。

（2）设$\lambda_1,\lambda_2,\ldots,\lambda_s$是$V_n(F)$上线性变换$T$的$s$个互异特征值，$V_{\lambda_i}$是$\lambda_i$的**特征子空间**，其中$i=1,2,\ldots,s$，则：

- V_{λ_i}是T的**不变子空间**；
- $\lambda_i\neq\lambda_j$时，$V_{\lambda_i}\cap V_{\lambda_j}=0$；
- 若λ_i是k_i重（*代数重数*）的，$\dim V_{\lambda_i}$是*几何重数*，则有$\dim V_{\lambda_i}\leqslant k_i$。

（3）线性变换T有对角矩阵表示的充分必要条件是 **T有n个线性无关的特征向量。**

> **幂等矩阵：** $A^2=A$，**A相似于对角矩阵**$\begin{bmatrix} I_r & \\ & 0 \end{bmatrix}$**，其中$r$为矩阵$A$的秩。**

> **乘方矩阵：** $A^2=I$，**A相似于对角阵**$\begin{bmatrix} I_s & \\ & I_t \end{bmatrix}$**，其中$s+t=n$。**

（4）关于秩的不等式：

$rank(A\pm B)\leqslant rank(A)+rank(B)$

$rank(A)+rank(B)-n\leqslant rank(A_{m\times n}B_{n\times m})\leqslant \min(r(A),r(B))$

$if\ A_{m\times n}B_{n\times m}=0,\quad rank(A)+rank(B)\leqslant n$

（5）形如$J(\lambda)=\begin{bmatrix} \lambda & 1 & & \\ & \lambda & 1 & \\ & & \ldots & 1 \\ & & & \lambda \end{bmatrix}$，称为**Jordan块**。Jordan块呈上三角，主对角线是它的全部特征值，特点是**主对角线上元素相等，紧邻上方元素$a_{i,i+1}=1$，其余元素为0。**

（6）每个n阶方阵A都相似于一个Jordan矩阵，即存在可逆矩阵P，有：

$$P^{-1}AP=J_A=\begin{bmatrix} J_1(\lambda_1) & & & \\ & J_2(\lambda_2) & & \\ & & \ldots & \\ & & & J_s(\lambda_s) \end{bmatrix}$$

其中J_A称为**Jordan标准形。**

（7）**Jordan标准形**的求法：

- 求矩阵A的特征多项式$|\lambda I-A|=(\lambda-\lambda_1)^{k_1}(\lambda-\lambda_2)^{k_2}\ldots(\lambda-\lambda_s)^{k_s}$，其中$k_i$是特征值$\lambda_i$的**代数重数，决定了对角线上特征值$\lambda_i$的个数**；
- 对λ_i，由$(A-\lambda_i I)X=0$，求A的**线性无关的特征向量**$\alpha_1,\alpha_2,\ldots,\alpha_{t_i}$，其中$t_i$是特征值$\lambda_i$的**几何重数，决定了Jordan块的个数**；

- 如果$k_i = t_i$，即**代数重数等于几何重数**，说明λ_i对应的Jordan块是对角阵；
- 如果$t_i < k_i$，就选择合适的特征向量α_j，利用$|A-\lambda_i I| = \alpha_j$**求Jordan链**，确定每一个小Jordan块的阶数。

- 将所有特征值λ_i对应的Jordan块组合起来，形成Jordan矩阵J_A。

（8）**矩阵多项式**可以表示为$g(A) = a_m A^m + a_{m-1}A^{m-1} + \ldots + a_1 A + a_0 I$，由于有$A = PJ_AP^{-1}$，所以有：

$$g(A) = P\begin{bmatrix} g(J_1(\lambda_1)) & & & \\ & g(J_2(\lambda_2)) & & \\ & & \cdots & \\ & & & g(J_s(\lambda_s)) \end{bmatrix}P^{-1}$$

而对于$g(J(\lambda))$则有：

$$g(J(\lambda)) = \begin{bmatrix} g(\lambda) & g'(\lambda) & \ldots & \frac{g^{(r-1)}(\lambda)}{(r-1)!} \\ & g(\lambda) & \ldots & . \\ & & \ldots & . \\ & & & g(\lambda) \end{bmatrix}$$

对于**常用的幂指数形式**有：

$$J^k(\lambda) = \begin{bmatrix} \lambda^k & \frac{(\lambda^k)'}{1!} & \frac{(\lambda^k)''}{2!} & \ldots \\ & \lambda^k & \ldots & . \\ & & \ldots & . \\ & & & \lambda^k \end{bmatrix}$$

（9）使$g(A) = 0$的多项式$g(\lambda)$称为A的**化零多项式**，**特征多项式必是矩阵A的化零多项式。**

> **注：化零多项式的根一定包含了所有的特征值，但不能说化零多项式的根一定是特征值。**

（10）对于**最小多项式**$m_T(\lambda)$

- $m_T(\lambda)$最高项系数为`1`；
- $m_T(\lambda)$是T的一个化零多项式；
- $m_T(\lambda)$是化零多项式中次数最低的那一个。

> **最小多项式$m_T(\lambda)$的根一定包含了所有的特征值λ_i，子式$(\lambda-\lambda_i)^{r_i}$的幂$r_i$等于Jordan标准形中关于特征值$\lambda_i$的Jordan块中的最高阶数**。

> > 比如矩阵A有一个代数重数为`3`的特征值`2`，该特征值对应两个Jordan块，分别是$\begin{bmatrix} 2 & 1 \\ & 2 \end{bmatrix}$以及$[2]$，说明其中其最高阶数为2，那么在最小多项式中对应的子式为$(\lambda-2)^2$。

---

3. 矩阵的分解

（1）**等价标准形**

对于$A \in C^{m\times n}$，存在可逆矩阵$P \in C^{m\times m}, Q \in C^{n\times n}$，使得

$$A = P \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} Q$$

其中r是矩阵A的秩。

---

(2) **相似标准形**

存在可逆矩阵$P \in C^{n\times n}$，有

$$A = P \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \dots & \\ & & & \lambda_n \end{bmatrix} P^{-1}$$

---

(3) **LU分解**

> 定义：L**是下三角矩阵，**U**是上三角矩阵**，$A = LU$。

> 求法：

- 对于$(A \mid I_n)$，**只用第**i**行乘数**k**加到第**j**行（**$i < j$**）型初等变换将**A**化为上三角形**U**，可以得到**$(U \mid P)$；
- **可知**$PA = U$**，于是有**$L = P^{-1}$**，则**$A = LU$**。**

---

(4) **LDV分解**

> 定义：L, V分别是**对角线元素为1**的下三角矩阵和上三角矩阵，D为对角矩阵，$A = LDV$。

> 求法：

> > **方法一：**

- 由LU分解得到$A = LU$；
- 通过每行除以对应的对角线上元素的值，**将**U**的对角线元素化为1**，得到$U = DV$；
- 有$A = LDV$。

> > **方法二：**

- 取矩阵A对角线第一个元素，得到矩阵$A_1 = [a_{11}]$，则有$A_1 = L_1 D_1 V_1 = [1][a_{11}][1]$；
- 取包含对角线前两个元素的二阶矩阵$A_2 = \begin{bmatrix} A_1 & \alpha \\ \beta & a_{22} \end{bmatrix}$，则有矩阵$A_2 = L_2 D_2 V_2$，其中$L_2 = \begin{bmatrix} L_1 & 0 \\ x & 1 \end{bmatrix}$，$D_2 = \begin{bmatrix} D_1 & 0 \\ 0 & d_2 \end{bmatrix}$，$V_2 = \begin{bmatrix} V_1 & y \\ 0 & 1 \end{bmatrix}$，求得未知量$x, d_2, y$；
- 以此类推，最终得到$A = L_n D_n V_n$。

---

(5) **满秩分解**

> 定义：对于$rank(A) = r$的矩阵A，若存在秩为r的矩阵$B \in F^{m\times r}$，$C \in F^{r\times n}$，有$A = BC$，称为矩阵A的满秩分解。

> 求法：方法较多，一般只用最简单的第3种。

- 用**行初等变换**把A化为Hermite标准形；
- 依Hermite标准形中向量e_i所在的列的位置第j_i列，相应地取出A的第j_i列a_{ji}，得到 A的**列向量极大无关组**$\{a_{j_1}, a_{j_2}, \ldots, a_{j_r}\}$，$B = (a_{j_1}, a_{j_2}, \ldots, a_{j_r})$；
- A的Hermite矩阵中的**非零行**构成矩阵C，得到满秩分解$A = BC$。

> 举个例子：
>
> > 求矩阵$A = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 2 & 2 \\ 1 & 0 & 1 \end{bmatrix}$的满秩分解。
>
> > 答：用行初等变换化A为Hermite标准形：

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

> > 可知$rank(A) = 2$，A的前两列线性无关，取出构成B；取出A的Hermite标准形的前两行作为C，有：

$$B = \begin{bmatrix} 1 & 1 \\ 0 & 2 \\ 1 & 0 \end{bmatrix}, C = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}, A = BC$$

（6）**谱分解**

> 定义：矩阵A互异的特征值$\{\lambda_1, \lambda_2, \ldots, \lambda_s\}$称为矩阵$A$的**谱**。**可相似对角化是可以谱分解的充要条件。**

> 求法：

- 通过求特征值和特征向量得到$A = P \begin{bmatrix} \lambda_1 & & & & & & & & & \\ & \ldots & & & & & & & & \\ & & \lambda_1 & & & & & & & \\ & & & \lambda_2 & & & & & & \\ & & & & \ldots & & & & & \\ & & & & & \lambda_2 & & & & \\ & & & & & & \ldots & & & \\ & & & & & & & \lambda_s & & \\ & & & & & & & & \ldots & \\ & & & & & & & & & \lambda_s \end{bmatrix} P^{-1}$；
- 对角阵$\Lambda = \lambda_1 \begin{bmatrix} I_{r_1} & & & \\ & 0 & & \\ & & \ldots & \\ & & & 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 0 & & & \\ & I_{r_2} & & \\ & & 0 & \\ & & & 0 \end{bmatrix} + \ldots + \lambda_s \begin{bmatrix} 0 & & & \\ & 0 & & \\ & & 0 & \\ & & & I_{r_s} \end{bmatrix}$，令$Q_i = \begin{bmatrix} 0 & & & \\ & \ldots & & \\ & & I_{r_i} & \\ & & & \ldots \end{bmatrix}$；
- 得到$A = \sum_{i=1}^{s} = \lambda_i P_i$，其中$P_i = PQ_iP^{-1}$。

（7）**Schur分解**

> 定义：对可逆矩阵A，存在**酉矩阵**U和**主对角线上元素都为正的上三角矩阵**R，使$A = UR$。

> 求法：

- 取矩阵$A=(A_1,A_2,...,A_n)$的列向量，进行**施密特正交化**，得到$u_1,u_2,...,u_n$，有$U=(u_1,u_2,...,u_n)$；
- 再由$R=U^HA$得到R，于是$A=UR$。

（8）几种特殊矩阵：

- **正规矩阵**：$A^HA=AA^H$ （**正规矩阵酉相似于对角阵**）
- **酉矩阵**：$A^HA=AA^H=I$
- **Hermite矩阵**：$A^H=A$

（9）**奇异值分解（SVD分解）**

> **奇异值**：对$rank(A)=r$的矩阵A，矩阵A^HA的**非零特征值**有$\lambda_1\geqslant\lambda_2\geqslant...\geqslant\lambda_r>0$，则称正数$\sigma_i=\sqrt{\lambda_i}$为矩阵$A$的奇异值。

> 定义：对$rank(A)=r$的矩阵$A\in C^{m\times n}$，奇异值有$\sigma_1\geqslant\sigma_2\geqslant...\geqslant\sigma_r>0$，则存在**酉矩阵**$U\in C^{m\times m}$，$V\in C^{n\times n}$，分块矩阵$\Sigma=\begin{bmatrix}\Delta & 0\\ 0 & 0\end{bmatrix}$，有$A=U\Sigma V^H$，其中$\Delta=\begin{bmatrix}\sigma_1 & & & \\ & \sigma_2 & & \\ & & ... & \\ & & & \sigma_r\end{bmatrix}$。

> 求解：

- 由特征多项式$|\lambda I-A^HA|=0$求得特征值$\lambda_1\geqslant\lambda_2\geqslant..\geqslant\lambda_n$，（**务必按照从大到小排列**），以及每个特征值对应的特征向量$\alpha_1,\alpha_2,...,\alpha_n$；
- 对特征向量进行施密特正交化和单位化（一般只需要单位化），得到单位正交向量组$v_1,v_2,..,v_n$，则$V=(v_1,v_2,...,v_n)$；
- 对于非零特征值$\lambda_1,...,\lambda_r$对应奇异值$\sigma_1,...,\sigma_r$，于是有$u_i=\frac{1}{\sigma_i}Av_i$，这样得到了$r$个列向量，剩余的设为$\beta$，通过正交的特性$u_i^T\beta=0$即可求得，
- 于是得到$A=U\Sigma V^H$

（10）**极分解**

> 定义：对于$rank(A)=r$的矩阵$A\in C^{n\times n}$，可以被分解为$A=PQ$，其中P为半正定矩阵，Q为酉矩阵。

> 求法：

- 对A进行奇异值分解，得到$A=U\Sigma V^H$；
- 可以得到$A=(U\Sigma U^H)(UV^H)$，于是$P=U\Sigma U^H$，$Q=UV^H$，$A=PQ$。

---

4. 矩阵的广义逆

（1）设$A\in C^{m\times n},B\in C^{n\times m}$，若有$BA=I_n$，则称$B$是$A$的一个**左逆**。

> 等价条件：

- A的零空间$N(A)=0$
- $m\geqslant n$，$rank(A)=n$，即A是**列满秩的**

- A^HA可逆

---

(2) 设$A \in C^{m\times n}, C \in C^{n\times m}$，有$AC = I_m$，则称$C$是$A$的一个**右逆**。

> 等价条件:

- A的列空间$R(A) = C^m$
- $m \leqslant n$，$rank(A) = m$，即A是**行满秩的**
- AA^H可逆

---

(3) 对于$A \in C^{m\times n}$，$G \in C^{n\times m}$，有$AGA = A$，称G是A的一个**减号广义逆**。

> 求法:

- 对$rank(A) = r$的矩阵A，有矩阵$\begin{bmatrix} A & I_m \\ I_n & 0 \end{bmatrix}$进行**初等变换，对$A$行变换时$I_m$保持同步，对$A$列变换时，$I_n$保持同步，将$A$化为最简形**，得到$\begin{bmatrix} I_r & 0 & P \\ 0 & 0 & \\ & & \\ Q & & 0 \end{bmatrix}$；
- 有$G = Q\begin{bmatrix} I_r & U \\ V & W \end{bmatrix}P$，其中$U, V, W$是满足固定阶次的任意矩阵。

---

(4) **加号广义逆（M-P逆）**

> **定义**：对于矩阵$A \in C^{m\times n}$，$G \in C^{n\times m}$，满足4条

- $AGA = A$
- $GAG = G$
- $(AG)^H = AG$
- $(GA)^H = GA$

> 称G为A的M-P逆。

> **求法**:

> > **方法一**:

- 对矩阵A进行**满秩分解**，得到$A = BC$;
- 则$A^+ = C^H(CC^H)^{-1}(B^HB)^{-1}B^H$，也就是等于`C的右逆 x B的左逆`。

> > **方法二**:

- 对矩阵A进行**奇异值分解**，得到$A = U\begin{bmatrix} \Delta & 0 \\ 0 & 0 \end{bmatrix}V^H$;
- 则$A^+ = V\begin{bmatrix} \Delta^{-1} & 0 \\ 0 & 0 \end{bmatrix}U^H$。

> **性质**

- $rank(A) = rank(A^+)$
- $rank(A^+A) = rank(AA^+) = rank(A)$

---

(5) **投影变换**

> **定义**: $C^n = L \oplus M, \quad x = y + z, \quad y \in L, z \in M$, 投影变换$\sigma$就是把$C^n$映射成子空间$L$, 称$\sigma$是从$C^n$沿子空间$M$到子空间$L$的投影变换, 在一组基下对应的矩阵称为**投影矩阵**, 子空间L称为**投影子空间**。显然有, 子空间L就是σ的像空间$R(\sigma)$, M就是σ的核空间$N(\sigma)$, 于是$C^n = R(\sigma) \oplus N(\sigma)$。

> > **σ是投影变换的充要条件是σ关于某组基下的矩阵A是幂等矩阵, 即$A^2 = A$。**

> **求法**

- 找出像空间L的一组基$y_1, y_2, \ldots, y_r$, 得到矩阵$B = (y_1 \ y_2 \ \ldots \ y_r)$; 找出$M$的一组基$z_{r+1}, \ldots, z_n$, 得到矩阵$C = (z_{r+1} \ \ldots \ z_n)$;
- 于是有投影矩阵$A = (B|0)(B|C)^{-1}$。

---

(6) **正交投影变换**

> **定义**: 若$C^n = R(\sigma) \oplus N(\sigma)$, $R(\sigma)$的**正交补空间**是$R(\sigma)^\perp = N(\sigma)$, 称$\sigma$是**正交投影变换**, 其在**标准正交基**下对应的矩阵称为正交投影矩阵。

> > **σ是正交投影变换的充要条件是A是幂等Hermite矩阵, 即$A^2 = A$, $A^H = A$。**

> **求法**

$$A = (B|0)(B|C)^{-1} = (B|0)((B|C)^H(B|C))^{-1}(B|C)^H = B(B^HB)^{-1}B^H$$

---

(7) **最佳最小二乘解**

$A \in C^{m\times n}$, $b \in C^m$, 则$x_0 = A^+b$是线性方程组$Ax = b$的最佳最小二乘解。

$A \in C^{m\times n}$, $B \in C^{m\times k}$, 则$X_0 = A^+B$是$AX = B$的最佳最小二乘解。

---

5. 矩阵分析

---

(1) **向量范数**满足**正定性、齐次性和三角不等式**, 定义了范数的内积空间称为**赋范空间。**

---

(2) 重要的**向量范数**:

对于复向量$x = (x_1 \ \ x_2 \ \ \ldots \ \ x_n)$, 有:

- **2-范数:** $|x|| = \sqrt{|x_1|^2 + |x_2|^2 + \ldots + |x_n|^2}$
- **1-范数:** $||x||_1 = |x_1| + |x_2| + \ldots + |x_n|$
- **∞范数:** $||x||_\infty = \max_i |x_i|$

> **有限维线性空间的任意两种向量范数都是等价的**。

---

(3) **矩阵范数**满足**正定性、齐次性、三角不等式以及相容性。**

---

(4) 重要的**矩阵范数**和**诱导范数**

- **F范数**：$||A||_F = [tr(A^H A)]^{\frac{1}{2}}$
- **列和范数**：$||A||_1 = \max\limits_j(\sum_{i=1}^n |a_{ij}|)$，即**每一列各元素模相加其中的最大值**
- **谱范数**：$||A||_2 = \sqrt{\lambda_1}$，其中$\lambda_1$是$A^H A$的**最大特征值**
- **行和范数**：$||A||_\infty = \max\limits_i(\sum_{j=1}^n |a_{ij}|)$，即**每一行各元素模相加其中的最大值**

（5）**向量收敛和矩阵收敛**必须其中的每一个元素都收敛。

> 向量按分量收敛的充要条件是**它按任意一个向量范数收敛**。

> 当$k \to \infty$时，$||A^{(k)} - A|| \to 0$，称**矩阵序列按矩阵范数收敛于A**

（6）**谱半径**

> **定义**：$\lambda_1, \lambda_2, \ldots, \lambda_n$是矩阵$A \in C^{n\times n}$的全部特征值，称$\rho(A) = \max\limits_i |\lambda_i|$为$A$的**谱半径**。

> $A^k \to 0(k \to \infty)$**的充要条件是**$\rho(A) < 1$。

> A**的谱半径是**A**的任意一种矩阵范数的下确界**。

（7）**矩阵幂级数**

若复变量z的幂级数$\sum_{k=0}^\infty a_k z^k$的**收敛半径为**$R$，而方阵$A \in C^{n\times n}$的**谱半径为**$\rho(A)$，则

- 当$\rho(A) < R$时，矩阵幂级数$\sum_{k=0}^\infty a_k A^k$**收敛**；
- 当$\rho(A) > R$时，矩阵幂级数$\sum_{k=0}^\infty a_k A^k$**发散**

> 当求解A的特征值比较困难时，由于A的每个范数都是谱半径$\rho(A)$的上界，只需要找到一种特殊的矩阵范数$||A||$，使得$||A|| < R$，就能说明矩阵幂级数收敛。（**优先考虑行和、列和范数**）

（8）常用的幂级数

收敛域是整个复平面的幂级数

$e^A = \sum_{k=0}^\infty \frac{1}{k!} A^k$

$\cos A = \sum_{k=0}^\infty \frac{(-1)^k}{(2k)!} A^{2k}$

$\sin A = \sum_{k=0}^\infty \frac{(-1)^k}{(2k+1)!} A^{2k+1}$

收敛域为复平面$|z| < 1$的幂级数

$(I - A)^{-1} = \sum_{k=0}^\infty A^k, \quad \rho(A) < 1$

$\ln(I + A) = \sum_{k=1}^\infty \frac{(-1)^{k-1}}{k} A^k, \quad \rho(A) < 1$

（9）矩阵函数的两种求法

> **方法一：Jordan标准形法**

- 求矩阵A的Jordan标准形J_A，得到$A = PJ_AP^{-1}$

- 设解析函数为$f(z)$，则对每一个Jordan块有$f(J_i)=\begin{bmatrix} f(\lambda_i) & \frac{f'(\lambda_i)}{1!} & \frac{f''(\lambda_i)}{2!} & \dots \\ & f(\lambda_i) & \frac{f'(\lambda_i)}{1!} & \dots \\ & & \dots & \\ & & & f(\lambda_i) \end{bmatrix}$，得到$f(J_A)$
- 最后得到$f(A)=Pf(J_A)P^{-1}$

> 这种方法的难点在于**需要求Jordan链**，过程中可以会遇到麻烦。**如果不同特征值个数较多，建议使用第一种；而如果特征值比较单一，并且 *代数重数 - 几何重数* > 2，建议使用第二种**。

> **方法二：最小多项式法**

- 先计算A的Jordan标准形，由此得到最小多项式$m_A(\lambda)=(\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\dots(\lambda-\lambda_s)^{n_s}$，其中幂次和有$\sum_{i=1}^{s} n_i=m$；
- 得到$g(\lambda)=c_0+c_1\lambda+\dots+c_{m-1}\lambda^{m-1}$，并令$g^{(j)}(\lambda_i)=f^{(j)}(\lambda_i)$，解得系数$c_0,c_1,\dots,c_{m-1}$；
- 最后得到$f(A)=c_0I+c_1A+\dots+c_{m-1}A^{m-1}$

> 当**不同特征值的个数比较多或者最小多项式幂次较高**时，计算起来比较复杂，建议使用第一种。

（10）两个知识点：

> 重要的**导数**

- $\frac{dA^{-1}(t)}{dt}=-A^{-1}(t)\left(\frac{d}{dt}A(t)\right)A^{-1}(t)$

> 矩阵指数函数的**行列式**

- $|e^A|=e^{trA}$

（11）矩阵函数应用

> **一阶常系数齐次微分方程组**：$\begin{cases}\dot{x}(t)=Ax(t)\\ x(t_0)=C_{n\times 1}\end{cases}$

解为：$x(t)=e^{A(t-t_0)}x(t_0)$

> **一阶线性常系数非齐次线性方程组**：$\begin{cases}\dot{x}(t)=Ax(t)+f(t)\\ x(t_0)=C\end{cases}$

解为：$x(t)=e^{A(t-t_0)}x(t_0)+\int_{t_0}^{t}e^{A(t-\tau)}f(\tau)d\tau$

---

6. K积

（1）对于矩阵$A=(a_{ij})\in C^{m\times n}$，$B=(b_{ij})\in C^{s\times t}$，则**K积**为：

$$A\otimes B=\begin{bmatrix} a_{11}B & a_{12}B & \dots & a_{1n}B \\ a_{21}B & a_{22}B & \dots & a_{2n}B \\ \dots & \dots & & \dots \\ a_{n1}B & a_{n2}B & \dots & a_{nn}B \end{bmatrix}$$

> K积不具有交换律，即$A \otimes B \neq B \otimes A$

（2）重要性质

- $I \otimes I = I$
- $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$
- $(A \otimes B)^k = A^k \otimes B^k$
- $(A \otimes B) = (I_m \otimes B)(A \otimes I_n)$
- $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$
- $|A \otimes B| = |B \otimes A| = |A|^n |B|^m$ （**这里的 `n` 表示 `B` 的阶数， `m` 表示 `A` 的阶数**）
- $rank(A \otimes B) = rank(A)rank(B)$

（3）**K和**：设$A \in F^{m\times m}$，$B \in F^{n\times n}$， $A \oplus B = A \otimes I_n + I_m \otimes B$

（4）若A的特征值是λ_i，相应的特征向量是x_i；B的特征值是μ_i，相应的特征向量为y_i；则：

- $A \otimes B$**的特征值是**$\lambda_i \mu_i$**，对应的特征向量是**$x_i \otimes y_i$
- $A \oplus B$**的特征值是**$\lambda_i + \mu_i$**，对应的特征向量是**$x_i \otimes y_i$

（5）设$f(z)$是解析函数，$A \in F^{n\times n}$，$f(A)$存在，则

- $f(I_m \otimes A) = I_m \otimes f(A)$
- $f(A \otimes I_m) = f(A) \otimes I_m$

（6）设矩阵$A \in F^{m\times n}$， $A = (A_1, A_2, \ldots, A_n)$，则$Vec(A) = \begin{bmatrix} A_1 \\ A_2 \\ \cdots \\ A_n \end{bmatrix} \in F^{nm}$

> $Vec(ABC) = (C^T \otimes A)Vec(B)$

- $Vec(AX) = (I_s \otimes A)Vec(X)$
- $Vec(XC) = (C^T \otimes I_k)Vec(X)$

（7）求解矩阵方程$AX + XB = D$，将两边同时取**向量化算子**，得到$(I_m \otimes A + B^T \otimes I_n)Vec(X) = Vec(D)$，最后通过常规的求非齐次线性方程组的方法求解。

（8）求微分方程：$\begin{cases} \dot{X}(t) = AX(t) + X(t)B \\ X(0) = C \end{cases}$

- 用向量化算子作用在方程两边，得到$Vec(\dot{X}(t)) = (I_n \otimes A + B^T \otimes I_m)Vec(X(t))$和$Vec(X(0)) = Vec(C)$
- 令$Y(t) = Vec(X(t))$， $C_1 = Vec(C)$， $G = I_n \otimes A + B^T \otimes I_m$，通过求解普通微分方程的方法得到$Y(t) = e^{Gt}C_1$；
- 将$Y(t)$，G，C_1带入化简求得$X(t)$。