矩阵分析 定理证明补充

说明：本注解涉及课件中未标明过程的证明内容，以朱萍萍老师（周一、周三7-8节）课件为准.大部分内容为笔者自行推理（也存在部分为课堂板书），因此可能存在过于"复杂"或"冗余"的证明过程.若有不严谨之处还请指出.

本文中∂表示多项式次数或矩阵次数(课件或板书中∂或$\deg$写法均存在)；称两行/列交换、一行/列乘以常数、一行/列乘常数或多项式并加到另一行/列的操作为"第一/二/三初等变换".

第一章 线性空间与线性变换II

1.4 线性映射

定理（用坐标计算线性映射）

设线性映射 $\sigma: V_1 \to V_2$ 在入口基 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \dots, \boldsymbol{\alpha}_n$ 和出口基 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_m$ 下的矩阵表示为 $\boldsymbol{A}$． 又设 $\boldsymbol{\alpha} \in V_1$ 在入口基下的坐标为 $\boldsymbol{x} \in \mathbb{F}^n$， 则 $\sigma(\boldsymbol{\alpha}) \in V_2$ 在出口基下的坐标为 $\boldsymbol{Ax} \in \mathbb{F}^m$．

令$\alpha = \sum_i x_i\alpha_i$，则$\sigma(\alpha) = \sum_i \sigma(\alpha_i) = \sum_i x_i(\sum_j a_{ji}\beta_j) = \sum_i(\sum_j x_j a_{ij})\beta_i = \begin{bmatrix}\beta_1 & \dots & \beta_m\end{bmatrix} Ax$.

1.5 矩阵的等价与相似

定义 （方阵的不变子空间）

设 $\boldsymbol{A}\in\mathbb{F}^{n\times n}$，子空间 $W\in\mathbb{F}^{n}$称为$\boldsymbol{A}$的不变子空间，如果

$$\boldsymbol{A}(W)\subseteq W,\quad \text{其中}\boldsymbol{A}(W)=\{\boldsymbol{Ax}:\boldsymbol{x}\in W\}$$

例 $\{\mathbf{0}\}$，$\mathbb{F}^{n}$，$\ker\boldsymbol{A}$，$\mathrm{im}\boldsymbol{A}$都是$\boldsymbol{A}$的不变子空间

若$x\in\ker A$，$Ax=0$，$A(Ax)=0$，$Ax\in\ker A$.
若$x\in\mathrm{im}A$，$\exists u$使得$Au=x$.此时$Ax=A(Au)$，$Ax\in\mathrm{im}A$.

定理 （方阵的不变子空间与相似三角化的等同性）

$$\boldsymbol{A}[\boldsymbol{P}_1\quad \boldsymbol{P}_2]=[\boldsymbol{P}_1\quad \boldsymbol{P}_2]\begin{bmatrix}\boldsymbol{B}_{11} & \boldsymbol{B}_{12}\\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22}\end{bmatrix}$$

(1) $\boldsymbol{B}_{21}=0$ 等价于 $\mathrm{im}\boldsymbol{P}_1$是$\boldsymbol{A}$的不变子空间；

(2) $\boldsymbol{B}_{12}=0$ 等价于 $\mathrm{im}\boldsymbol{P}_2$是$\boldsymbol{A}$的不变子空间；

(3) $\boldsymbol{B}_{21}=0$，$\boldsymbol{B}_{12}=0$ 等价于 $\mathrm{im}(\boldsymbol{P}_1)$是$\boldsymbol{A}$的不变子空间且 $\mathrm{im}(\boldsymbol{P}_2)$是$\boldsymbol{A}$的不变子空间。

1.$B_{21}=0$即$AP_1=P_1B_{11}$.
则$\exists u$使$x=P_1u,x\in\mathrm{im}P_1$，$Ax=AP_1u=P_1B_{11}u\in\mathrm{im}P_1$.
则$\mathrm{im}P_1$是A的不变子空间.
2.$AP_1=P_1B_{11}+P_2B_{21}$.
P为相似变换，则P的所有列向量线性无关.
若$\mathrm{im}P_1$是A的不变子空间，则对$x\in\mathrm{im}P_1,x=P_1u$，$AP_1x=P_1B_{11}u+P_2B_{21}u\in\mathrm{im}P_1$.
此时$P_2B_{21}u=0$，必有$B_{21}=0$.
(2)(3)同理.

定理 （相似对角化的条件）

$$A[p_1 \quad p_2 \quad \cdots \quad p_n]=[p_1 \quad p_2 \quad \cdots \quad p_n]\begin{bmatrix} b_{11} & 0 & \dots & 0 \\ 0 & b_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & b_{nn} \end{bmatrix}$$

$$\Updownarrow$$

$\mathrm{im}\boldsymbol{p}_j$是 A 的不变子空间，$j=1,2,\ldots,n$

由于p为向量，因此下文中u,v均为1*1（标量）.

1.$Ap_j=b_{jj}p_j$，对$x\in\mathrm{im}p_j, x=p_ju$，$Ax=Ap_ju=b_{ii}p_ju\in\mathrm{im}p_j$，$\mathrm{im}p_j$是$A$的不变子空间.

2.若$\mathrm{im}p_j$是A的不变子空间，对$x=p_ju_j$有$Ax=Ap_ju_j\in\mathrm{im}p_j$，因此可令$Ax=p_jv_j$.

令$\frac{v_j}{u_j}=b_{jj}$，以上方法得到n个等式组合即$AP=P\mathrm{diag}\{b_{jj}\}$.

第二章 多项式矩阵与若当标准型

2.1 多项式矩阵及其Smith标准型

定理（单位模阵的行列式刻画）

多项式矩阵 $\boldsymbol{U}(\lambda)\in\mathbb{F}^{n\times n}[\lambda]$ 是单位模阵，当且仅当行列式 $|\boldsymbol{A}(\lambda)|\in\mathbb{F}[\lambda]$ 为非零常值多项式.

1.若U为单位模阵，则$\exists V(\lambda)\in\mathbb{F}^{n\times n}[\lambda]$使$U(\lambda)V(\lambda)=I$，则$|U(\lambda)||V(\lambda)|=1$.

$\partial(|U(\lambda)|)+\partial(|V(\lambda)|)=0$，有$\partial(|U(\lambda)|)=\partial(|V(\lambda)|)=0$.

2.若$\partial(|U(\lambda)|)=0$，取$\frac{U^*(\lambda)}{|U(\lambda)|}=U^{-1}(\lambda)=V(\lambda)\in\mathbb{F}^{n\times n}[\lambda]$，则$U(\lambda)V(\lambda)=V(\lambda)U(\lambda)=I$.

引理（用初等变换将左上角降次）

设多项式矩阵 $\boldsymbol{A}(\lambda)$ 的左上角元素 $a_{11}(\lambda)\neq 0$（不是零多项式，下同），并且 $\boldsymbol{A}(\lambda)$ 中至少有一个元素不能被 $a_{11}(\lambda)$ 整除．那么 $\boldsymbol{A}(\lambda)$ 等价于一个多项式矩阵 $\boldsymbol{B}(\lambda)$，使得 $b_{11}(\lambda)\neq 0$，且 $\deg(b_{11}(\lambda))<\deg(a_{11}(\lambda))$.

考虑左上角元素满足$\partial(a_{11}(\lambda))\leq(a_{ij}(\lambda))$，否则可通过第一初等变换将对应元素移至左上角.

(1)$\exists a_{i1}(\lambda), i\neq 1$使$a_{11}(\lambda)\nmid a_{i1}(\lambda)$，$a_{i1}(\lambda)$可表示为$a_{11}(\lambda)p(\lambda)+q(\lambda)$，其中$\partial(q(\lambda))<$

$\partial(a_{11})(\lambda)$.

$$\begin{bmatrix} a_{11}(\lambda) \\ a_{11}(\lambda)p(\lambda)+q(\lambda) \end{bmatrix} \to \begin{bmatrix} a_{11}(\lambda) \\ q(\lambda) \end{bmatrix} \to \begin{bmatrix} q(\lambda) \\ a_{11}(\lambda) \end{bmatrix}.$$

(2)$\exists a_{1j}(\lambda), j \neq 1$使$a_{11}(\lambda) \nmid a_{1j}(\lambda)$，同理.

(3)$i, j \neq 1$，$\exists a_{ij}(\lambda)$使$a_{11}(\lambda) \nmid a_{ij}(\lambda)$，且$a_{11}(\lambda) \mid a_{i1}(\lambda), a_{11}(\lambda) \mid a_{1j}(\lambda)$（否则即(1)(2)情形）.

$$\begin{bmatrix} a_{11}(\lambda) & a_{11}(\lambda)s(\lambda) \\ a_{11}(\lambda)t(\lambda) & a_{11}(\lambda)p(\lambda)+q(\lambda) \end{bmatrix} \to$$
$$\begin{bmatrix} a_{11}(\lambda) & 0 \\ a_{11}(\lambda)t(\lambda) & a_{11}(\lambda)p(\lambda)+q(\lambda)-a_{11}(\lambda)t(\lambda)s(\lambda) \end{bmatrix} \to$$
$$\begin{bmatrix} a_{11}(\lambda) & 0 \\ a_{11}(\lambda) & a_{11}(\lambda)p(\lambda)+q(\lambda)-a_{11}(\lambda)t(\lambda)s(\lambda) \end{bmatrix} \to \begin{bmatrix} a_{11}(\lambda) & p(\lambda)-t(\lambda)s(\lambda) \\ a_{11}(\lambda) & q(\lambda) \end{bmatrix} \to$$
$$\begin{bmatrix} q(\lambda) & a_{11}(\lambda) \\ p(\lambda)-t(\lambda)s(\lambda) & a_{11}(\lambda) \end{bmatrix}.$$

定理（多项式矩阵的 Smith 标准型）

任意多项式矩阵 $A(\lambda) \in \mathbb{F}^{m\times n}[\lambda]$ 等价于下面的 Smith 标准型

$$\left[\begin{array}{cccc|c} d_1(\lambda) & & & \mathbf{0} & \\ & d_2(\lambda) & & & \mathbf{0} \\ & & \ddots & & \\ \mathbf{0} & & & d_r(\lambda) & \\ \hline & & \mathbf{0} & & \mathbf{0}_{m-r,n-r} \end{array}\right],$$

其中 $d_i(\lambda), i=1,2,\ldots,r$， 为首项系数为 1 的非零多项式， 满足 $d_i(\lambda)$ 整除 $d_{i+1}(\lambda)$， 记为 $d_i(\lambda) \mid d_{i+1}(\lambda)$， $i=1,2,\ldots,r-1$.

经以上引理降次后，设矩阵所有元素均能被$a_{11}(\lambda)$整除，则矩阵第1行、第1列其他元素均化为0.

即$A(\lambda) = \begin{bmatrix} d_1(\lambda) & a_{12}^{(1)}(\lambda)d_1(\lambda) & \dots \\ a_{21}^{(1)}(\lambda)d_1(\lambda) & \ddots & \\ \vdots & & \ddots \end{bmatrix} \to \begin{bmatrix} d_1(\lambda) & \\ & A_1(\lambda) \end{bmatrix} \to \cdots \to$

$$\begin{bmatrix} d_1(\lambda) & & & 0 \\ & d_2(\lambda) & & \\ & & \ddots & \\ 0 & & & d_r(\lambda) \\ 0 & & & 0 \end{bmatrix}.$$

（初等变换不改变矩阵的秩，而每次变换时$r(A_i(\lambda)) = 1 + r(A_{i+1}(\lambda))$，因此能进行$r-1$次该步骤而$A_{r-1}(\lambda)$为$1\times 1$，即对角线上有$r$个非0元素）

定理（初等行、列变换不改变多项式矩阵的行列式因子）

设两个多项式矩阵 $\boldsymbol{A}(\lambda), \boldsymbol{B}(\lambda) \in \mathbb{F}^{m\times n}[\lambda]$ 等价，则它们的各阶行列式因子分别相同.

第一、二类初等变换显然不改变行列式因子.
考虑第三类初等行变换$a_i \to a_i + p(\lambda) a_j$:
若i, j行均在子式中，该子式行列式不变.
若i行不在子式中，该子式不变.
若j行不在子式中且i行在子式中，不妨记变换前行列式因子为$D(\lambda)$，原子式为$\begin{bmatrix} r_i(\lambda) \\ \vdots \end{bmatrix}$，则变换后子式为$\begin{bmatrix} r_i(\lambda) \\ \vdots \end{bmatrix} + p(\lambda) \begin{bmatrix} r_j(\lambda) \\ \vdots \end{bmatrix}$.
其中，$\begin{bmatrix} r_j(\lambda) \\ \vdots \end{bmatrix}$为矩阵中的另一子式，其行列式同样含有因式$D(\lambda)$，因此变换后行列式仍含有因式$D(\lambda)$，即行列式因子不变.

问题（Smith 型的唯一性）

Smith 型是否唯一？确切的说，若同一多项式矩阵 $\boldsymbol{A}(\lambda) \in \mathbb{F}^{m\times n}[\lambda]$ 分别被两个不同的初等变换序列化成 Smith 标准型

$$\left[\begin{array}{cccc:c} d_1(\lambda) & & & 0 & \\ & d_2(\lambda) & & & \mathbf{0} \\ & & \ddots & & \\ 0 & & & d_r(\lambda) & \\ \hdashline & \mathbf{0} & & & \mathbf{0}_{m-r,n-r} \end{array}\right] \text{ 及 } \left[\begin{array}{cccc:c} e_1(\lambda) & & & 0 & \\ & e_2(\lambda) & & & \mathbf{0} \\ & & \ddots & & \\ 0 & & & e_t(\lambda) & \\ \hdashline & \mathbf{0} & & & \mathbf{0}_{m-t,n-t} \end{array}\right]$$

那么，是否有 $r = t$？若 $r = t$，是否 $d_i(\lambda) = e_i(\lambda)$，$i = 1,2,\ldots,r$？

$r = t$已在上述Smith标准型推导中论证.
由Smith型元素间的整除关系，Smith型中$D_1(\lambda) = d_1(\lambda)$，$D_i(\lambda) = \prod_1^i d_j(\lambda)$.
又多项式因子不变，依次可得$d_i(\lambda) = e_i(\lambda)$.
以上即可说明Smith型，行列式因子，不变因子相互决定.

定理（幺模阵写为初等矩阵的乘积）

（1） 幺模阵的 Smith 标准型为单位矩阵.
（2） 幺模阵可写为有限个初等矩阵的乘积.

(1)单位模阵行列式为常数，则$d_n(\lambda)=1$.
$\partial(d_i(\lambda))\le\partial(d_n(\lambda)),i=1,...,n-1$，则所有不变因子均为1，Smith标准型为单位阵.
(2)单位模阵可经过有限次初等变换变为Smith型，即可写为有限初等矩阵乘积.

2.2 数域上矩阵的特征矩阵

定理 （矩阵相似的充要条件是特征矩阵等价）

两个数字矩阵$\boldsymbol{A},\boldsymbol{B}\in\mathbb{F}^{n\times n}$相似

$$\Updownarrow$$

$\boldsymbol{\lambda I}-\boldsymbol{A},\boldsymbol{\lambda I}-\boldsymbol{B}\in\mathbb{F}^{n\times n}[\boldsymbol{\lambda}]$作为多项式矩阵等价

(1)A,B相似则存在可逆变换P使$PAP^{-1}=B$.
$P(\lambda I-A)P^{-1}=\lambda I-B$，且$P$为数值矩阵，因此取单位模阵$Q(\lambda)=P,R(\lambda)=P^{-1}$，则$\lambda I-A,\lambda I-B$等价.
(2)取单位模阵$U(\lambda),V(\lambda)$，$U(\lambda)(\lambda I-A)V(\lambda)=\lambda I-B$.
令$U(\lambda)=(\lambda I-B)Q(\lambda)+R$，其中由$\partial(\lambda I-B)=1$得$R\in\mathbb{F}^{n\times n}$.
若$R=0$，两边取行列式次数即$|\lambda I-B||Q(\lambda)|=|U(\lambda)|$，即$n+0=0$，不成立，即$R\neq0$.
此时$U(\lambda)(\lambda I-A)=(\lambda I-B)V^{-1}(\lambda),R(\lambda I-A)=(\lambda I-B)[V^{-1}(\lambda)-Q(\lambda)(\lambda I-A)]$.
两边取次数：$0+1=1+0$，即$S=V^{-1}(\lambda)-Q(\lambda)(\lambda I-A)\in\mathbb{F}^{n\times n}$.
比较λ次数，$R=S,RA=BS$.
接下来考虑R可逆：
令$U^{-1}(\lambda)=(\lambda I-A)\tilde{Q}(\lambda)+\tilde{R}$.
$I=U(\lambda)U^{-1}(\lambda)$
$=[(\lambda I-B)Q(\lambda)+R][(\lambda I-A)\tilde{Q}(\lambda)+\tilde{R}]$
$=(\lambda I-B)Q(\lambda)U^{-1}(\lambda)+R(\lambda I-A)\tilde{Q}(\lambda)+R\tilde{R}$
$=(\lambda I-B)Q(\lambda)U^{-1}(\lambda)+(\lambda I-B)S\tilde{Q}(\lambda)+R\tilde{R}$.
$R\tilde{R}-I=-(\lambda I-B)[Q(\lambda)U^{-1}(\lambda)+S\tilde{Q}(\lambda)]$，若左边不为0，则两边取次数时左边为0而右边不小于1.
因此$R\tilde{R}-I=0$，即R可逆.
令$P=R^{-1}$，$P^{-1}AP=B$.

定理（矩阵相似的各种刻画）

给定两个矩阵 $A, B \in \mathbb{F}^{n\times n}$. 下列条件等价:

（1） A 与 B 相似.

（2） $\lambda I - A$ 与 $\lambda I - B$ 作为多项式矩阵等价.

（3） $\lambda I - A$ 与 $\lambda I - B$ 作为多项式矩阵有相同的 Smith 标准型.

（4） $\lambda I - A$ 与 $\lambda I - B$ 作为多项式矩阵有相同的各阶行列式因子.

（5） $\lambda I - A$ 与 $\lambda I - B$ 作为多项式矩阵有相同的各阶不变因子.

（6） $\lambda I - A$ 与 $\lambda I - B$ 作为多项式矩阵有相同的初等因子组.

(1)(2)等价已证明，且(2-5)等价已证明.接下来证明(5)(6)等价.

(5)→(6)显然.

(6)→(5)：假设初等因子：$p_1^{t_{11}}, p_1^{t_{12}}, \ldots, p_1^{t_{1m_1}}, \ldots, p_l^{t_{l1}}, \ldots, p_1^{t_{lm_l}}$，其中每个最简多项式$p_i(\lambda)$的次数$t_{ij}$按降序排列.

则依次取$d_{n-k+1}(\lambda) = \prod_i p_i^{t_{ik}}$，其中若$k \geq m_i$即初等因子组中$p_i(\lambda)$的个数小于$k$，则$t_{in} = 0$.此时满足$d(k) \mid d(k+1)$，即作为各阶不变因子.

例（任一矩阵与其转置相似）

设矩阵 $A \in \mathbb{F}^{n\times n}$. 则 A 的转置 A^{T} 与 A 相似.

在$\lambda I - A$中取任一子式，均可在$\lambda I - A^T$中取其转置，因此$\lambda I - A, \lambda I - A^T$各阶行列式因子相等，$A, A^T$相似.

2.3 复数域上矩阵的Jordan标准型

引理　考虑复数域$\mathbb{F} = \mathbb{C}$，设$\lambda I - A$的非常数不变因子

$h_i(\lambda) = (\lambda - c_1)^{r_1}(\lambda - c_2)^{r_2}\cdots(\lambda - c_k)^{r_k}$, 则有

$$\begin{bmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & h_i(\lambda) \end{bmatrix}_{n_i\times n_i} \sim \begin{bmatrix} \begin{bmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & (\lambda - c_1)^{r_1} \end{bmatrix}_{r_1\times r_1} & & \\ & \ddots & \\ & & \begin{bmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & (\lambda - c_k)^{r_k} \end{bmatrix}_{r_k\times r_k} \end{bmatrix}_{n_i\times n_i}$$

左边：$D_{n_i}(\lambda) = h_i(\lambda), D_k(\lambda) = 1(k < n_i)$.

右边：显然$D_{n_i}(\lambda) = h_i(\lambda)$.

考虑n_i-1阶，存在n_i-1阶子式行列式为$\frac{h_i(\lambda)}{(\lambda_k-c_k)^{r_i}}$且分别唯一，公因式为1.
因此$D_{n_i-1}(\lambda)=1$，更低阶行列式因子也为1.

引理 （若当块的Smith型）

称r_i阶矩阵

$$J_i=\begin{bmatrix} c_i & 1 & & \\ & c_i & 1 & \\ & & \cdots & \cdots \\ & & & c_i \end{bmatrix}$$

为Jordan块，其特征矩阵的Smith标准形为

$$\begin{bmatrix} I_{r_i-1} & \\ & (\lambda-c_i)^{r_i} \end{bmatrix}$$

考虑行列式因子相等证等价.
$\lambda I-J_i$可按以下方式选取$1\sim r_i-1$阶下三角矩阵，行列式为常数，因此$1\sim r_i-1$阶行列式因子为1.

$$\begin{bmatrix} \lambda-c_i & -1 & & \\ & \lambda-c_i & -1 & \\ & & \ddots & -1 \\ & & & \lambda-c_i \end{bmatrix} \quad \begin{bmatrix} \lambda-c_i & -1 & & \\ & \lambda-c_i & -1 & \\ & & \ddots & -1 \\ & & & \lambda-c_i \end{bmatrix}$$

r_i阶行列式因子为$(\lambda-c_i)^{r_i}$.

定理 （复数域上矩阵的若当标准型）

$\boldsymbol{A} \in \mathbb{C}^{n\times n}$，$\lambda\boldsymbol{I}$-$\boldsymbol{A}$的初等因子组为

$$(\lambda-c_1)^{r_1},(\lambda-c_2)^{r_2},\ldots,(\lambda-c_q)^{r_q}$$

取 $\boldsymbol{J}_i = \begin{bmatrix} c_i & 1 & & \\ & c_i & 1 & \\ & & \cdots & \cdots \\ & & & c_i \end{bmatrix}_{r_i\times r_i}$，则$\boldsymbol{A}$相似于 $\boldsymbol{J} = \begin{bmatrix} \boldsymbol{J}_1 & & & \\ & \boldsymbol{J}_2 & & \\ & & \cdots & \\ & & & \boldsymbol{J}_q \end{bmatrix}$

称$\boldsymbol{J}$为$\boldsymbol{A}$的Jordan标准型。

$\lambda I - A$的Smith型即前文引理的左侧矩阵$diag[1,...,1,h(\lambda)]$.按照前文引理，可分解的各Smith标准型均可变为Jordan块的特征矩阵.

$$\begin{bmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & h_i(\lambda) \end{bmatrix}_{n_i\times n_i} \sim \begin{bmatrix} \begin{bmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & (\lambda-c_1)^{r_1} \end{bmatrix}_{r_1\times r_1} \backsim \begin{bmatrix} \lambda-c_1 & 1 & & \\ & \lambda-c_1 & 1 & \\ & & \cdots & \cdots \\ & & & \lambda-c_1 \end{bmatrix} = \lambda-J_1 & & \\ & \ddots & \\ \lambda-J_n \backsim & & \begin{bmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & (\lambda-c_k)^{r_k} \end{bmatrix}_{r_k\times r_k} \end{bmatrix}_{n_i\times n_i}$$

实际上，此处c_i可重复.

第三章 内积空间

3.1 内积与Gram矩阵

定理 （Gram矩阵的性质）

(1)Hermite性： $G^{\mathrm{H}}=\overline{G}^{\mathrm{T}}=G$

(2) 非负定性： $\forall z\in\mathbb{C}^s$， $z^{\mathrm{H}}Gz\geq 0$

(3) G正定 等价于 $\{\alpha_i\}$线性无关

(4)rank(G)=rank($\{\alpha_1, \alpha_2, ..., \alpha_s\}$)

(1)对Gram矩阵，$g_{ij}=\langle\alpha_i,\alpha_j\rangle=\overline{\langle\alpha_j,\alpha_i\rangle}=\overline{g_{ji}}$.
(2)$z^HGz=\sum_{i=1}^s\sum_{j=1}^s\overline{z_i}\langle\alpha_i,\alpha_j\rangle z_j=\sum_{i=1}^s\sum_{j=1}^s\overline{z_i}\alpha_i^H\alpha_jz_j=$
$(\sum_{i=1}^s\overline{z_i}\alpha_i^H)(\sum_{j=1}^s\alpha_jz_j)=\langle u,u\rangle\geq 0$.
即z^HGz可表示为向量u的内积.
(3)若$\{a_i\}$线性无关，等价于$u=0$当且仅当$z=0$，等价于G正定.
(4)若其中一向量可由其它向量表示，则G中该向量对应行、列可由第三类初等变换化为0.

3.2 正交投影与最佳逼近

定理（投影定理——有限维情形）

设V是复数域上的内积空间，$\beta\in V$，W是V上的子空间，则

(1)存在唯一的$\alpha^*\in W$，使得

$$d(\beta,\alpha^*)\leq d(\beta,\alpha),\quad \forall\alpha\in W$$

(2)设$\alpha_1, \alpha_2, ..., \alpha_s$是$W$的一组基，记

$$\alpha^*=k_1\alpha_1+k_2\alpha_2+\ldots+k_s\alpha_s=[\alpha_1, \alpha_2, ..., \alpha_s]k$$

则

$$k=G(\{\alpha_i\})^{-1}G(\{\alpha_i\};\beta)$$

(1)考虑$\alpha^* \perp \beta, \alpha = \alpha^* + u, u \in W$，则$d(\beta, \alpha) = \|\beta - \alpha^* - u\| = \|\beta - \alpha^*\| + \|u\|$.
$d(\beta, \alpha)$最小当且仅当$\|u\| = 0$，即$\alpha = \alpha^*$，唯一存在性得证.
(2)若$\beta - \alpha^* \perp \alpha^*$，则$\langle \alpha^*, \alpha^* \rangle = \langle \alpha^*, \beta \rangle$.
此时对$i = 1, ..., s$，$\langle \alpha_i, \alpha^* \rangle = \sum_j k_j \langle \alpha_i, \alpha_j \rangle = \langle \alpha_i, \beta \rangle$.

$$\begin{bmatrix} \langle \alpha_1, \alpha_1 \rangle & ... & \langle \alpha_1, \alpha_s \rangle \\ \vdots & & \vdots \\ \langle \alpha_s, \alpha_1 \rangle & ... & \langle \alpha_s, \alpha_s \rangle \end{bmatrix} k = \begin{bmatrix} \langle \alpha_1, \beta \rangle \\ \vdots \\ \langle \alpha_s, \beta \rangle \end{bmatrix}，即 G(\{\alpha_i\})k = G(\{\alpha_i\}; \beta).$$

α_i为一组基，则Gram矩阵非奇异，则$k = G^{-1}(\{\alpha_i\})G(\{\alpha_i\}; \beta)$.

3.3 标准正交基与酉矩阵

内积空间中标准正交组的好处体现在：

定理（沿正交组展开时坐标的解耦性）

设$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_s$是复内积空间$V$的一个正交组，$\boldsymbol{\beta} \in V$. 若有

$$\boldsymbol{\beta} = \boldsymbol{\alpha}_1 k_1 + \boldsymbol{\alpha}_2 k_2 + \cdots + \boldsymbol{\alpha}_s k_s$$

则

$$k_i = \frac{\langle \boldsymbol{\alpha}_i, \boldsymbol{\beta} \rangle}{\langle \boldsymbol{\alpha}_i, \boldsymbol{\alpha}_i \rangle}, i = 1, 2, \ldots, s.$$

$\langle \alpha_i, \beta \rangle = \sum_j k_i \langle \alpha_i, \alpha_j \rangle$，其中$i \neq 0$时$\langle \alpha_i, \alpha_j \rangle = 0$.
即$\langle \alpha_i, \beta \rangle = k_i \langle \alpha_i, \alpha_i \rangle$.

定理 从n维内积空间的任一组基$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, ..., \boldsymbol{\alpha}_n$出发，可通过Schmidt正交化方法构造出一个标准正交基。

（1）正交化：

$$\boldsymbol{\beta}_1 = \boldsymbol{\alpha}_1,$$

$$\boldsymbol{\beta}_j = \boldsymbol{\alpha}_j - \left(\boldsymbol{\beta}_1 k_{1,j} + \boldsymbol{\beta}_2 k_{2,j} + \cdots + \boldsymbol{\beta}_{j-1} k_{j-1,j}\right), j = 2, 3, \ldots, s.$$

其中，$k_{ij} = \dfrac{\langle \boldsymbol{\beta}_i, \boldsymbol{\alpha}_j \rangle}{\langle \boldsymbol{\beta}_i, \boldsymbol{\beta}_i \rangle}, i < j.$

（2）标准化：$\tilde{\boldsymbol{\beta}}_j = \boldsymbol{\beta}_j \dfrac{1}{\|\boldsymbol{\beta}_j\|}, j = 1, 2, \cdots, s$

正交化$j-1$个向量时,$\{\beta_i\}$为一组标准正交基.
令$\alpha_j = w_j + h_j$，其中$w_j \in \mathrm{span}(\beta_1, ..., \beta_{j-1})$，$h_j \perp \mathrm{span}(\beta_1, ..., \beta_{j-1})$.

由以上定理，$w_j=\sum_i k_{i,j}\beta_i=\sum_i \frac{\langle\beta_i,w_j\rangle}{\langle\beta_i,\beta_i\rangle}\beta_i=\sum_i \frac{\langle\beta_i,w_j+h_j\rangle}{\langle\beta_i,\beta_i\rangle}\beta_i=\sum_i \frac{\langle\beta_i,\alpha_i\rangle}{\langle\beta_i,\beta_i\rangle}\beta_i.$

（以上能直接在w_j加h_j，因为h_j与任意β_i垂直即内积为0）

取$\beta_j=h_j$即证.

定理（酉矩阵保持度量）

设$U\in\mathbb{C}^{n\times n}$，则下列等价

（1） U 为酉矩阵

（2）（保内积）对任意 $x,y\in\mathbb{C}^n$，

$$\langle Ux,Uy\rangle=\langle x,y\rangle.$$

（3）（保长度）对任意 $x\in\mathbb{C}^n$，

$$\|Ux\|=\|x\|.$$

(1)→(2)：U为酉矩阵，$\langle Ux,Uy\rangle=x^HU^HUy=x^Hy=\langle x,y\rangle.$

(2)→(1)：$\langle Ux,Uy\rangle=x^HU^HUy=x^Hy=\langle x,y\rangle$，则$U^HU=I$，即$U$为酉矩阵.

(2)→(3)：$\langle Ux,Ux\rangle=\langle x,x\rangle$，则$\|Ux\|=\|x\|.$

(3)→(2)：令$\|U(x+y)\|=\|x+y\|$即$U(x+y),U(x+y)\rangle=\langle x+y,x+y\rangle$，则$\langle Ux,Ux\rangle+\langle Uy,Uy\rangle+\langle Ux,Uy\rangle+\langle Uy,Ux\rangle=\langle x,x\rangle+\langle y,y\rangle+\langle x,y\rangle+\langle y,x\rangle.$

化简得$Re(\langle Ux,Uy\rangle)=Re(\langle x,y\rangle).$

类似地,$\|U(x-y)\|=\|x-y\|$可得$Im(\langle Ux,Uy\rangle)=Im(\langle x,y\rangle)$.即$\langle Ux,Uy\rangle=\langle x,y\rangle.$

注：矩阵的正交三角分解是Schmidt标准正交化在标准酉空间中的具体化

证：$A=[\alpha_1,\alpha_2,\ldots,\alpha_n]$

用Schmidt正交化方法将$\alpha_1,\alpha_2,\ldots,\alpha_n$。标准正交化，得

（1）正交化：

$\beta_1=\alpha_1,$

$\beta_j=\alpha_j-\left(\beta_1k_{1,j}+\beta_2k_{2,j}+\cdots+\beta_{j-1}k_{j-1,j}\right),\ j=2,3,\ldots,s.$

其中，$k_{ij}=\dfrac{\langle\beta_i,\alpha_j\rangle}{\langle\beta_i,\beta_i\rangle},i<j.$

（2）标准化：$\tilde{\beta}_j=\beta_j\dfrac{1}{\|\beta_j\|},j=1,2,\cdots,s$

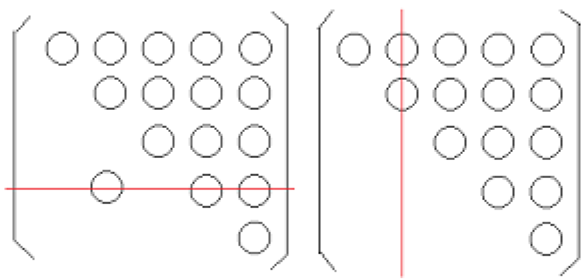
$$\begin{bmatrix}\alpha_1 & \dots & \alpha_n\end{bmatrix} = \begin{bmatrix}1 & \dots & k_{1,n}\\ & \ddots & \vdots \\ & & 1\end{bmatrix}\begin{bmatrix}\beta_1 & \dots & \beta_n\end{bmatrix}.$$

引理1 正线上三角阵的逆还是上三角阵

引理2 两个上三角阵的乘积还是上三角阵

引理3 A既是酉矩阵，又是正线上三角矩阵，则A是单位阵。

1.若$\exists a'_{ij}\neq 0, i>j$，其中$a'_{ij}$为$A^{-1}$中的元素，则$\sum_k a'_{ik}a_{kj}=a'_{ij}a_{jj}$.
若A为正线上三角阵,$a_{jj}\neq 0$，则$A^{-1}A$非对角线元素不为0，矛盾.
(A不为正线上三角阵则$A^{-1}A=I$仍可能成立)



2.考虑上三角阵A,B及$i>j$，$i\leq k, k\leq j$至少有一个不成立，则$\sum_k a_{ik}b_{kj}$每项至少有一个为0，即求和结果（AB下三角元素）为0.

3.A为正线上三角阵, 则A^H为正线下三角阵，A^{-1}为上三角阵.
又A为酉矩阵，$A^{-1}=A^H$，此时A^H为对角阵，即A为对角阵.
A列向量为标准正交基，范数为1可得对角线元素为1，即$A=I$.

3.4 正规矩阵、Hermite矩阵

定义 设 $A,B\in\mathbb{C}^{n\times n}$，若存在酉矩阵$U$使

$$U^H AU=U^{-1}AU=B,$$

则称A酉相似于B。

定理 （Schur引理）任何一个n阶复矩阵A可酉相似于上三角矩阵。

取$P=[p_1,\dots,p_n]$，$\{p_i\}$为A的特征向量或广义特征向量，则$P^{-1}AP=J$，J为Jordan标准型（上三角矩阵）.
对P正交三角分解$P=UB$，则$B^{-1}U^{-1}AUB=J$即$U^{-1}AU=U^HAU=BJB^{-1}$.
B^{-1}为上三角阵，则BJB^{-1}为上三角阵且与A酉相似.

定义（正规矩阵）一个复方阵若与它的共轭转置矩阵可交换，则称为正规矩阵．即 $A\in\mathbb{C}^{n\times n}$ 称为正规矩阵，若

$$AA^{\mathrm{H}}=A^{\mathrm{H}}A.$$

定理 $A\in\mathbb{C}^{n\times n}$ 可酉相似对角化，当且仅当 A 是正规矩阵.

(1)令$U^HAU=R$，R为上三角阵，$U^HA^HU=R^H$.
$(U^HAU)(U^HA^HU)=U^HAA^HU=RR^H,(U^HA^HU)(U^HAU)=U^HA^HAU=R^HR.$
若A正规，$RR^H=R^HR.$
（课上给的是依次计算左右两边的元素得到R为对角阵，但这里笔者没有验算过）
(2)若A可酉相似对角化即R对角，则$RR^H=R^HR$，$(U^HAU)(U^HA^HU)=(U^HA^HU)(U^HAU)$，得$AA^H=A^HA.$

定理 Hermite矩阵的特征值必是实数。

令$U^HAU=D$，D为对角阵且元素为A特征值.
$U^HA^HU=D^H$，若$A=A^H$则$D=D^H$即D所有元素（A特征值）为实数.

定理 Hermite矩阵A是正定的(非负定的) 等价于 A的n个特征值全为正(非负)实数。

(1)对任意向量x，A正定即$x^HAx>0.$
令$U^HAU=D$，D为对角阵.取$x=Uy$，则$y^TU^TAUy=y^TDy>0.$
上式成立条件为D特征值全为正，即A特征值为正.
(2)若A特征值全为正，则可对角化即$U^HAU=D$，D为正对角阵（正定），即对任意向量y有$y^HDy>0.$
取$y=U^Hx$，则$x^HUDU^Hx=x^HAx>0$，A正定.
非负定证法相同.

定理 $A^{\mathrm{H}}A$ 与 AA^{H} 均为非负定 Hermite 矩阵，它们有相同的非零(正实数)特征值，且若 $A^{\mathrm{H}}Ax=\lambda x$，则 $AA^{\mathrm{H}}Ax=\lambda Ax.$

$(A^HA)^H=A^HA^{H^H}=A^HA$，即A^HA Hermite.
令$U^H(A^HA)U=B^HB=D$，D为对角阵.
B第i列为b_i，则D对角线第i个元素为$b_i^Hb_i\geq0$，即A^HA特征值为非负（A^HA非负定）.

类似地，AA^H Hermite.
$AA^H = AUU^HA^H = BB^H$（此处U, B与上文相同），
A^HA与AA^H有相同特征值.
$A^HAx = \lambda x$，$AA^HAx = \lambda Ax$.

定理（Hermit 矩阵特征值的极值刻画）

设A是Hermite矩阵，则$\lambda_{\max}(A) = \lambda_n = \max\limits_{x \in \mathbb{C}^n, x \neq 0} \dfrac{x^H Ax}{x^H x} = \max\limits_{x \in \mathbb{C}^n, \|x\|=1} x^H Ax$.

以下均限定$\|x\| = 1$.
取A酉特征矩阵$P = [p_1, ..., p_n]$（A可对角化，则有n个不相关的特征向量）.
令$x = Py$，$x^HAx = y^HP^HAPy = y^H \begin{bmatrix} p_1^H \\ \vdots \\ p_n^H \end{bmatrix} \begin{bmatrix} \lambda_1 p_1 & \cdots & \lambda_n p_n \end{bmatrix} y =$
$y^H \begin{bmatrix} \lambda_1 p_1^H p_1 & & \\ & \ddots & \\ & & \lambda_n p_n^H p_n \end{bmatrix} y = \sum_i \overline{y_i} y_i \lambda_i p_i^H p_i = \sum_i \overline{y_i} y_i \lambda_i$.
$\|y\| = 1$，$\max x^HAx = \max \sum_i \overline{y_i} y_i \lambda_i = \max \lambda_i$.

考虑一般情形，若A正规（不一定Hermite），特征值可能为复数，此时无法直接比较大小，但可以得到Rayleigh最大模长即
$\max |x^HAx| = \max |\lambda_i|$.

更一般地，对于一般矩阵A，$\max \|Ax\|_2 = \sqrt{\max x^HA^HAx}$.
由A^HA Hermite，上式即为$\sqrt{\max \lambda(A^HA)}$.（数值分析中矩阵2-范数）