Ch4 Hermite二次型

§4.1 Hermite阵、正规阵

实二次型：$f(x_1,\cdots,x_n)=\sum_{i,j} a_{ij}x_ix_j = X^TAX=\langle AX,X\rangle$

其中 $A^T=A\in\mathbb{R}^{n\times n}$

推广到复数的情况：$X\in\mathbb{C}^n$

$f(X)=\langle AX,X\rangle = X^HAX=\sum_{i,j} a_{ij}\bar{x}_i x_j \in\mathbb{R}$

$\Leftrightarrow f^H=f$
$\Leftrightarrow X^H(A-A^H)X=0$
$\Leftrightarrow a_{ij}=\bar{a}_{ji}$
$\Leftrightarrow A^H=A$

$\Leftrightarrow A^H=A$ —— Hermite阵或H阵

Hermite二次型

Schur引理：设 $A\in\mathbb{C}^{n\times n}$，则∃酉阵使 $U^HAU=\begin{bmatrix}\lambda_1 & & * \\ & \ddots & \\ & & \lambda_n\end{bmatrix}$.

定理4.1.1 设A是H阵，则∃酉阵U，使

$$U^HAU=\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix},\ (\lambda_i\in\mathbb{R},\ i=1,\cdots,n).$$

H阵的性质：(1) 特征值全为实数；
(2) 属于不同特征值的特征向量正交.
(3) H阵是正规阵

定理4.1.2 方阵A酉相似于对角阵 ⇔ A是正规阵
(即 $A^HA=AA^H$)

§4.2 Hermite二次型

设Hermite二次型 $f(X)=\langle AX,X\rangle = X^HAX$

$\xrightarrow[C\text{可逆}]{\text{作}X=CY} Y^H(C^HAC)Y$

B仍是H阵

则称A与B共轭合同 $\xrightarrow{\text{记}B=C^HAC} Y^HBY$ ⟶ Hermite二次型

$\xrightarrow{\text{取}X=UY} \lambda_1|y_1|^2+\cdots+\lambda_n|y_n|^2$ ⟶ 标准形

求标准形的方法. ① 正交变换法 ② 初等变换法

例1. 求可逆线性变换化Hermite二次型为标准形.

$$f(x_1,x_2,x_3)=|x_1|^2+(1-i)\bar{x}_1x_2+(2+i)\bar{x}_1x_3+(1+i)x_1\bar{x}_2 +3|x_2|^2+(2-i)x_1\bar{x}_3+2|x_3|^2.$$

解：$\begin{bmatrix} A \\ I \end{bmatrix} = \begin{bmatrix} 1 & 1-i & 2+i \\ 1+i & 3 & 0 \\ 2-i & 0 & 2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow[\substack{r_2-(1+i)c_1 \\ r_3-(2-i)c_1}]{\substack{c_2-(1-i)c_1 \\ c_3-(2+i)c_1}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1-3i \\ 0 & -1+3i & -3 \\ 1 & -1+i & -2-i \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

$\rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -13 \\ 1 & -1+i & -6-3i \\ 0 & 1 & 1+3i \\ 0 & 0 & 1 \end{bmatrix}$ 令 $X = \begin{bmatrix} 1 & -1+i & -6-3i \\ 0 & 1 & 1+3i \\ 0 & 0 & 1 \end{bmatrix} Y$，

则 $f = |y_1|^2 + |y_2|^2 - 13|y_3|^2$

再令 $y_1 = z_1,\ y_2 = z_2,\ y_3 = \frac{1}{\sqrt{13}} z_3$

则 $f = |z_1|^2 + |z_2|^2 - |z_3|^2$. #

(惯性定理) Hermite二次型的标准形中，系数为正的项数 正惯性指数 及系数为负的项数 负惯性指数 唯一.

(证明：略)

Hermite二次型的分类：设 $f(X) = X^H A X$ 是Hermite二次型

$\forall X \neq 0$, $\begin{cases} 若\ f(X) > 0\ (<0)，则称f为正定(负定)二次型，A为正定(负定)阵； \\ 若\ f(X) \geq 0\ (\leq 0)，则称f为半正定(半负定)二次型，A为半正定(半负定)阵； \end{cases}$

正定性判定准则：设A是n阶Hermite阵，则以下等价：

绝招 ⟶ 1° A是正定阵（$\forall X \neq 0$，则 $f(X) = X^H A X > 0$）；

2° 与A共轭合同的是正定阵（$C^H A C$正定，$\forall C$可逆）；

3° A的正惯性指数 $= n$；

第二招 ⟶ 4° A的特征值全大于0；（$\Rightarrow$ A可逆）

5° $A = S^2$，其中S正定；

第三招 ⟶ 6° $A = P^H P$，其中P可逆；

第一招 ⟶ 7° A的顺序主子式全大于0.

证明：$2^\circ \Leftrightarrow 1^\circ \Rightarrow 3^\circ \Rightarrow 4^\circ \Rightarrow 5^\circ \Rightarrow 6^\circ \Rightarrow 1^\circ$.

$1^\circ \Rightarrow 2^\circ$. 设 $C^HAC=B$，C可逆．则对 $\forall x \neq 0$，

$x^HBx=(Cx)^HA(Cx)>0$ （$\because Cx\neq 0$）．

$2^\circ \Rightarrow 1^\circ$. $A\simeq A$.

$1^\circ \Rightarrow 3^\circ$. $\because A$是H阵，$\therefore \exists$可逆阵C，使 $C^HAC=\begin{bmatrix}d_1&&\\&\ddots&\\&&d_n\end{bmatrix}$.

作 $x=Cy$，则 $f(x)=y^H(C^HAC)y=d_1|y_1|^2+\cdots+d_n|y_n|^2$

取 $x=Ce_k$，则 $d_k=f(Ce_k)>0$ （$\because Ce_k\neq 0$），$k=1,\cdots,n$

$3^\circ \Rightarrow 4^\circ$. $\because A$是H阵，$\therefore \exists$酉阵U使 $U^HAU=\begin{bmatrix}\lambda_1&&\\&\ddots&\\&&\lambda_n\end{bmatrix}$

作 $x=Uy$，有 $f(x)=\lambda_1|y_1|^2+\cdots+\lambda_n|y_n|^2$

由3°知 $\lambda_i>0$，$i=1,\cdots,n$.

$4^\circ \Rightarrow 5^\circ$. $\because U^HAU=\begin{bmatrix}\lambda_1&&\\&\ddots&\\&&\lambda_n\end{bmatrix}$ （$\lambda_i>0, i=1,\cdots,n$）

$\therefore A=U\begin{bmatrix}\lambda_1&&\\&\ddots&\\&&\lambda_n\end{bmatrix}U^H=\left(U\begin{bmatrix}\sqrt{\lambda_1}&&\\&\ddots&\\&&\sqrt{\lambda_n}\end{bmatrix}U^H\right)^2\triangleq S^2$，且$S$正定.

$5^\circ \Rightarrow 6^\circ$. 设 $A=S^2$，$\because S$正定 $\therefore S$可逆且$S^H=S$，故$A=S^HS$.

$6^\circ \Rightarrow 1^\circ$ $\forall x\neq 0$，则 $Px\neq 0$，故 $x^HAx=(Px)^H(Px)>0$. #

例1. 设H阵A正定且为酉阵，证明：$A=I$.

证明：设 $A=U\begin{bmatrix}\lambda_1&&\\&\ddots&\\&&\lambda_n\end{bmatrix}U^H$ $\because A$正定 $\therefore \lambda_i>0$, $i=1,\cdots,n$

又A为酉阵，则 $A=A^H=A^{-1}=U\begin{bmatrix}\frac{1}{\lambda_1}&&\\&\ddots&\\&&\frac{1}{\lambda_n}\end{bmatrix}U^H$

$\therefore \lambda_i=\frac{1}{\lambda_i}$，即 $\lambda_i=1$ 故 $A=I$. #

例2. 设n阶H阵A正定，又B是n阶方阵，且$A-B^HAB$也正定，证明：B的谱半径$\rho(A)<1$.

证明：设 $B\xi=\lambda\xi$，$\xi\neq 0$ 则

$\xi^HB^HAB\xi=(B\xi)^HA(B\xi)=\bar{\lambda}\xi^HA\lambda\xi=|\lambda|^2\xi^HA\xi$

$\therefore 1-|\lambda|^2=1-\dfrac{\xi^H(B^HAB)\xi}{\xi^HA\xi}=\dfrac{\xi^H(A-B^HAB)\xi}{\xi^HA\xi}>0$. （$\because A, A-B^HAB$都正定）

故$|\lambda|<1$ 即 $\rho(A)<1$. #

奇值分解：设 $r(A_{s\times n})=r$，则 $r(A^HA)=r(A)=r$ 且 A^HA 半正定，设其非零特征值为 $\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_r>0$。即 ∃ 酉阵 $V=(X_1,\cdots,X_n)$ 使 $V^H(A^HA)V=\begin{bmatrix}\lambda_1 & & & & \\ & \ddots & & & \\ & & \lambda_r & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0\end{bmatrix}$ 即 $A^HAV=V\begin{bmatrix}\lambda_1 & & & & \\ & \ddots & & & \\ & & \lambda_r & & \\ & & & \ddots & \\ & & & & 0\end{bmatrix}$

$\therefore A^HAX_i=\begin{cases}\lambda_iX_i, & i=1,\cdots,r\\ \theta, & i>r\end{cases}$

即 $\langle AX_i,AX_j\rangle=X_j^HA^HAX_i=\begin{cases}\lambda_i, & 1\leq i=j\leq r\\ 0, & 其余\end{cases}$

取 $Y_i=\frac{AX_i}{\sqrt{\lambda_i}}$ $(i=1,\cdots,r)$，则 $Y_1,\cdots,Y_r$ 为标准正交向量组

将其扩充为 $\mathbb{C}^s$ 的标基 $Y_1,\cdots,Y_s$，记 $U=(Y_1,\cdots,Y_s)$ 为酉阵

故 $A(X_1,\cdots,X_n)=(Y_1,\cdots,Y_s)\begin{bmatrix}D&0\\0&0\end{bmatrix}$，$D=\begin{bmatrix}\sqrt{\lambda_1}& & \\ & \ddots & \\ & & \sqrt{\lambda_r}\end{bmatrix}$

即 $U^HAV=\begin{bmatrix}D&0\\0&0\end{bmatrix}$. #

定理 4.2.6（奇值分解）对秩为 r 的矩阵 $A\in\mathbb{C}^{s\times n}$，∃ 酉阵 U 与 V，使 $U^HAV=\begin{bmatrix}D&0\\0&0\end{bmatrix}_{s\times n}$，$D=\begin{bmatrix}\sqrt{\lambda_1}& & \\ & \ddots & \\ & & \sqrt{\lambda_r}\end{bmatrix}$

且 $\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_r>0$ 为 A^HA 的非零特征值.

作业：$P_{143-145}$ 1, 2, 3, 6, 17.

§4.3 Rayleigh 商

设 $A\in H_n$，若 $Ax=\lambda x$，$x\neq\theta$ 则 $x^HAx=\lambda x^Hx$

$$\therefore \lambda=\frac{x^HAx}{x^Hx}$$

定义 4.3.1 $R(x)=\dfrac{x^HAx}{x^Hx}$，$x\neq\theta$

—— A的Rayleigh商

将A的特征值 $\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_n$

相应的标准正交特征向量组：$X_1,X_2,\cdots,X_n$（为 $\mathbb{C}^n$ 的标基）

$\forall X(\neq\theta)\in\mathbb{C}^n$，则 $X=\sum_{i=1}^n a_iX_i=(X_1,\cdots,X_n)\begin{bmatrix}a_1\\ \vdots\\ a_n\end{bmatrix}$，于是

$$AX=\sum_{i=1}^n a_i\lambda_iX_i=(X_1,\cdots,X_n)\begin{bmatrix}\lambda_1a_1\\ \vdots\\ \lambda_na_n\end{bmatrix}$$

$$R(X)=\frac{X^HAX}{X^HX}=\frac{\langle AX,X\rangle}{\langle X,X\rangle}=\frac{\sum_{i=1}^n|a_i|^2\lambda_i}{\sum_{i=1}^n|a_i|^2}$$

故 $\lambda_n\leq R(X)\leq\lambda_1$，且 $R(X_1)=\lambda_1$，$R(X_n)=\lambda_n$

于是我们有：

定理 4.3.1 设 $A\in H_n$，则 $\lambda_n\leq R(X)\leq\lambda_1$ 且

$$\lambda_n=\min_{X\neq\theta}R(X),\quad \lambda_1=\max_{X\neq\theta}R(X).$$

记 $S_1^\perp=(L[X_1])^\perp=L[X_2,\cdots,X_n]$，则 $\lambda_2=\max\limits_{X\in S_1^\perp}R(X)$

$T_1^\perp=(L[X_n])^\perp$，则 $\lambda_{n-1}=\min\limits_{X\in T_1^\perp}R(X)$

$=L[X_1,\cdots,X_{n-1}]$

一般地，$\lambda_k=\max\limits_{X\in S_{k-1}^\perp}R(X)$，$S_{k-1}^\perp=L[X_k,\cdots,X_n]$

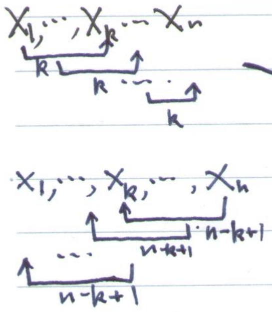
或 $\lambda_k=\min\limits_{X\in T_{n-k}^\perp}R(X)$，$T_{n-k}^\perp=L[X_1,\cdots,X_k]$

注. 这里的 λ_k 依赖于 $X_1,\cdots,X_{k-1}$ 或 $X_{k+1},\cdots,X_n$.

Courant 极大极小原理. 设 $A\in H_n$ 且 $\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_n$，

则 $$\lambda_k=\min_{\dim S=n-k+1}\left\{\max_{X\in S}R(X)\right\}$$

$$=\max_{\dim S=k}\left\{\min_{X\in S}R(X)\right\}.$$



$A\in$ 正规矩阵，则 $U^HAU=\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix}$

其中 U 为酉阵.

例1. 若 A 为 n 阶正规矩阵，则 $\displaystyle\max_{0\neq X\in\mathbb{C}^n}\frac{|X^HAX|}{X^HX}=\rho(A)$.

证：设 A 的特征值为 $\lambda_1,\cdots,\lambda_n$，相应的标准正交向量组为 $X_1,\cdots,X_n$，$\forall X=\sum_{i=1}^n a_iX_i$，$AX=\sum_{i=1}^n a_i\lambda_iX_i$

于是 $$\frac{|X^HAX|}{X^HX}=\frac{|\langle AX,X\rangle|}{\langle X,X\rangle}=\frac{\left|\sum_{i=1}^n|a_i|^2\lambda_i\right|}{\sum_{i=1}^n|a_i|^2}\leq\rho(A).\ \#$$

例2. 设 $A,B\in H_n$，且 $\lambda_1(A)\geq\cdots\geq\lambda_n(A)$，证明
$$\lambda_k(A)+\lambda_n(B)\leq\lambda_k(A+B)\leq\lambda_k(A)+\lambda_1(B)$$

证明：$$\frac{X^H(A+B)X}{X^HX}=\frac{X^HAX}{X^HX}+\frac{X^HBX}{X^HX}$$

且 $$\lambda_n(B)\leq\frac{X^HBX}{X^HX}\leq\lambda_1(B)$$

则 $$\lambda_n(B)+\frac{X^HAX}{X^HX}\leq\frac{X^H(A+B)X}{X^HX}\leq\lambda_1(B)+\frac{X^HAX}{X^HX}$$

由 Courant 极大极小原理可得证. #.

作业：习题四. 6, 7, 13, 17

讲解课后习题（包含以上作业）