第三章

内积空间

引导问题

在线性空间中, 向量之间的基本运算只有**线性运算**.
几何中向量有度量性质,向量的度量性质在许多问题中有着特殊的地位,而在线性空间中不能体现.

因此有必要引入度量的概念.

解析几何中,向量的度量性质通过内积来表示,
而内积有明显的代数性质,
内积作为引入度量的基本概念.

提纲

欧式空间、酉空间

标准正交基、Gram—Schmidt方法

酉变换、正交变换

正规矩阵

Hermite矩阵

正定二次齐式、正定Hermite 矩阵

【定义】：设 V 是实数域 R 上一线性空间, 在 V 上定义了一个二元实函数, 称为**内积**, 记作 (α,β), 它具有以下性质:

1) $(\alpha,\beta)=(\beta,\alpha)$;

2) $(k\alpha,\beta)=k(\alpha,\beta)$;

3) $(\alpha+\beta,\gamma)=(\alpha,\gamma)+(\beta,\gamma)$;

4) $(\alpha,\alpha)\geq 0$, 当且仅当 $\alpha=0$ 时 $(\alpha,\alpha)=0$.

其中 α,β,γ 是 V 中任意的向量, k 是任意实数, 这样的线性空间 V 称为**欧几里得空间**. 简称**欧氏空间**

在欧氏空间的定义中，对线性空间的维数并无要求，可以是有限维的，也可以是无限维的.

几何中向量的内积

$$<\alpha,\beta>=|\alpha||\beta|\cos(\alpha,\beta)$$

满足定义中的四条性质，所以几何空间是一个欧氏空间。

欧氏空间举例

例 在线性空间 R^n 中, 对于向量

$$\alpha = [a_1, a_2, \ldots, a_n]^T, \beta = [b_1, b_2, \ldots, b_n]^T,$$

定义内积

$$(\alpha, \beta) = a_1 b_1 + a_2 b_2 + \ldots + a_n b_n = \alpha^T \beta.$$

R^n就成为一个欧氏空间

以后用 R^n 来表示欧氏空间时，若没有特别说明，内积就用上式定义

在n = 3时, 上式就是几何空间中向量的内积在直角坐标系中的坐标表达式.

例 在闭区间 $[a,b]$ 上的所有实连续函数构成的空间 $C[a,b]$ 中, 对于函数 $f(x)$, $g(x)$ 定义内积

$$(f,g)=\int_a^b f(x)g(x)\mathrm{d}x\ .$$

可以验证, $C[a,b]$构成一欧氏空间.

同样地, 线性空间 $P[x]_n$ 按照上述内积也构成欧氏空间.

定义(酉空间) 设 V 是复数域上的线性空间, 在 V上定义了一个二元复函数, 称为内积, 记作 (α, β), 它具有以下性质:

1) $(\alpha, \beta) = \overline{(\beta, \alpha)}$, 这里 $\overline{(\beta, \alpha)}$ 是 (β, α) 的共轭复数;

2) $(k\alpha, \beta) = k(\alpha, \beta)$;

3) $(\alpha + \beta, \gamma) = (\alpha, \gamma) + (\beta, \gamma)$;

4) (α, α) 是非负实数, 且 $(\alpha, \alpha) = 0$ 当且仅当 $\alpha = 0$.

α, β, γ 是 V 中任意的向量, k 为任意复数,

这样的线性空间称为复欧氏空间, 或称酉空间。

欧氏空间与酉空间统称为内积空间。

例　在线性空间 C^n 中, 对向量

$$\alpha = [a_1, a_2, \ldots, a_n]^{\mathrm{T}}, \beta = [b_1, b_2, \ldots, b_n]^{\mathrm{T}},$$

定义内积为

$$(\alpha, \beta) = a_1\overline{b_1} + a_2\overline{b_2} + \ldots + a_n\overline{b_n} = \alpha^T\overline{\beta}$$

C^n 成为一个酉空间.

以后用 C^n 来表示酉空间时，若没有特别说明，内积就用上式定义

欧氏空间的性质：

(1) $(\alpha, k\beta) = k(\alpha, \beta)$

(2) $(\alpha, \beta + \gamma) = (\alpha, \beta) + (\alpha, \gamma)$

(3) $(\sum_{i=1}^{t} k_i\alpha_i, \beta) = \sum_{i=1}^{t} k_i(\alpha_i, \beta)$

(4) $(\sum_{i=1}^{t} \alpha, k_i\beta_i) = \sum_{i=1}^{t} k_i(\alpha, \beta_i)$

酉空间的性质：

$$\boxed{(1)\ (\alpha, k\beta) = \bar{k}(\alpha, \beta)}$$

$$(2)\ (\alpha, \beta+\gamma) = (\alpha, \beta) + (\alpha, \gamma)$$

$$(3)\ \left(\sum_{i=1}^{t} k_i\alpha_i, \beta\right) = \sum_{i=1}^{t} k_i(\alpha_i, \beta)$$

$$\boxed{(4)\ \left(\sum_{i=1}^{t} \alpha, k_i\beta_i\right) = \sum_{i=1}^{t} \bar{k}_i(\alpha, \beta_i)}$$

北京理工大学 BEIJING INSTITUTE OF TECHNOLOGY

内积空间的度量

定义： 设V为内积空间, 向量$\alpha\in V$的**长度**定义为非负实数

$$\|\alpha\|=\sqrt{(\alpha,\alpha)}$$

内积空间C^n中任意向量$\alpha=(a_1, a_2, \ldots, a_n)$，其长度为

$$\|\alpha\|=\sqrt{\sum_{i=1}^{n}|a_i|^2}$$

例： 在内积空间C^4 中求下列向量的长度

(1) $\alpha=(1+2i,-i,3,2+\sqrt{2}i)^T$

(2) $\beta=(1,-2,3,4)^T$

解：

$$\|\alpha\|=\sqrt{(1^2+2^2)+(1^2)+(3^2)+(2^2+\sqrt{2}^2)}$$

$$=\sqrt{5+1+9+6}=\sqrt{21}$$

$$\|\beta\|=\sqrt{1^2+2^2+3^2+4^2}$$

$$=\sqrt{1+4+9+16}=\sqrt{30}$$

定义：设V是n维内积空间, $\{\alpha_i\}$为其一组基底, 对于V 中的任意两个向量

选取合适的基底，使得度量矩阵$G=E$

$$\alpha=\sum_{i=1}^{n} x_i\alpha_i\,,\quad \beta=\sum_{j=1}^{n} y_j\alpha_j$$

那么α与β的内积

$$(\alpha,\beta)=\left(\sum_{i=1}^{n} x_i\alpha_i\,,\sum_{j=1}^{n} y_j\alpha_j\right)=\sum_{i,j=1}^{n} x_i\overline{y_j}(\alpha_i,\alpha_j)$$

令$g_{ij}=(a_i,a_j)$　$i,j=1,2,\ldots,n$

$$G=[g_{ij}]_{n\times n}$$

称G为基底$\{\alpha_i\}$的**度量矩阵**

$$(\alpha,\beta)=X^TG\overline{Y}$$

北京理工大学 BEIJING INSTITUTE OF TECHNOLOGY

例： 在内积空间$P[x]_2$中，已知基底向量1，x，x^2之间的内积如右表：

	1	x	x^2
1	1	1	1
x	1	1	1
x^2	1	1	1

计算：

(1) $\|2+3x+4x^2\|$

(2) $(3+2x,\ 1+4x+2x^2)$

解： 在基底$\{1,x,x^2\}$下的度量矩阵

$$G=\begin{bmatrix}1&1&1\\1&1&1\\1&1&1\end{bmatrix}$$

$$\|2+3x+4x^2\|=\sqrt{[2,3,4]\begin{bmatrix}1&1&1\\1&1&1\\1&1&1\end{bmatrix}\begin{bmatrix}2\\3\\4\end{bmatrix}}$$

$$=\sqrt{81}=9$$

$(3+2x,\ 1+4x+2x^2)$

$$=[3,2,0]\begin{bmatrix}1&1&1\\1&1&1\\1&1&1\end{bmatrix}\begin{bmatrix}1\\4\\2\end{bmatrix}$$

$$=35$$

北京理工大学 BEIJING INSTITUTE OF TECHNOLOGY

定理：向量长度具有如下性质

(1) $\|\alpha\| \geq 0$, $\|\alpha\|=0$当且仅当$\alpha=0$

(2) $\|k\alpha\|=|k|\|\alpha\|$

(3) $\|\alpha+\beta\| \leq \|\alpha\|+\|\beta\|$

三角不等式

(4) $|(\alpha,\beta)| \leq \|\alpha\|\,\|\beta\|$

Cauchy-Schwarz不等式

定义：设 V为欧氏空间,向量α和β之间的**距离**：

$$d(\alpha, \beta)=\|\alpha - \beta\|$$

两个非零向量 α, β的**夹角**θ:

$$\theta=\arccos\frac{(\alpha, \beta)}{\|\alpha\|\|\beta\|}, \quad 0\le\theta\le\pi$$

$$\cos\theta=\frac{(\alpha,\beta)}{\|\alpha\|\,\|\beta\|}$$

显然

$$\theta=\frac{\pi}{2}\Leftrightarrow(\alpha, \beta)=0$$

几何意义垂直
代数定义正交

定义： 长度为 **1** 的向量称为**单位向量**。

对于任何一个非零的向量α,

向量$\dfrac{\alpha}{\|\alpha\|}$是单位向量

向量单位化

提纲

欧式空间、酉空间

标准正交基、Gram—Schmidt方法

酉变换、正交变换

正规矩阵

Hermite矩阵

正定二次齐式、正定Hermite 矩阵

定义：在内积空间V 中, 如果 $(\alpha, \beta)=0$,

则称α与β正交, 记为$\alpha \perp \beta$.

0向量与任何一个向量都是正交的.

定义：设$\{\alpha_i\}$为一组不含有零向量的向量组,

如果$\{\alpha_i\}$内的任意两个向量彼此正交,

则称其为正交向量组.

定义：如果一个正交向量组中,

任何一个向量都是单位向量,

则称此向量组为标准正交向量组.

例　在C^3中向量组

$$\alpha_1=\frac{1}{3}\begin{pmatrix}-2\\-1\\2\end{pmatrix},\ \alpha_2=\frac{1}{3}\begin{pmatrix}2\\-2\\1\end{pmatrix},\ \alpha_3=\frac{1}{3}\begin{pmatrix}1\\2\\2\end{pmatrix}.$$

是标准正交向量组.

向量组

$$\beta_1=\begin{pmatrix}-\cos\theta\\0\\i\sin\theta\end{pmatrix},\beta_2=\begin{pmatrix}0\\1\\0\end{pmatrix},\beta_3=\begin{pmatrix}i\sin\theta\\0\\-\cos\theta\end{pmatrix}.$$

也是标准正交向量组.

定理： 正交向量组(不含零向量)是线性无关向量组

定理说明：

n维内积空间中, 两两正交的非零向量组中不能包含超过n个向量。

几何意义：

平面上不存在三个两两垂直的非零向量,

三维空间中不存在四个两两垂直的非零向量.

定义：在n维内积空间中，

由n个正交向量组成的基底称为正交基底；

由 n个标准正交向量组成的基底称为标准正交基底.

注意：标准正交基底不唯一.

定理：向量组$\{\alpha_i\}$为正交向量组的充要条件是

$$(\alpha_i,\alpha_j)=0,\quad \forall i\neq j$$

定理：向量组$\{\alpha_i\}$为标准正交向量组的充要条件是

$$(\alpha_i,\alpha_j)=\delta_{ij}=\begin{cases}1, & \forall i=j\\ 0, & \forall i\neq j\end{cases}.$$

● n个向量$\{\alpha_i\}$是标准正交基的充要条件为它的度量矩阵G为单位矩阵.

定理：由一个线性无关的向量组出发可以构造一个正交向量组, 甚至是一个标准正交向量组.

Gram—Schmidt正交化与单位化方法:

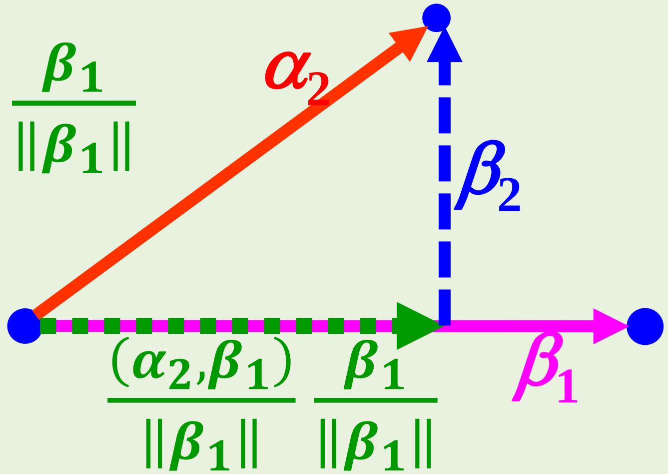
$\{\alpha_1, \alpha_2, \ldots, \alpha_r\}$为 n 维内积空间 V 中的 r 个线性无关的向量,

利用这 r 个向量可以构造一个标准正交向量组,

构造的标准正交向量组是 $\mathrm{span}\{\alpha_1, \alpha_2, \ldots, \alpha_r\}$空间的一个标准正交基.

第一步：正交化

$$\beta_2 = \alpha_2 - \frac{(\alpha_2,\beta_1)}{\|\beta_1\|}\frac{\beta_1}{\|\beta_1\|}$$



$$\beta_1 = \alpha_1$$

$$\beta_2 = \alpha_2 - \frac{(\alpha_2,\beta_1)}{(\beta_1,\beta_1)}\beta_1$$

………

$$\beta_r = \alpha_r - \frac{(\alpha_r,\beta_1)}{(\beta_1,\beta_1)}\beta_1 - \cdots - \frac{(\alpha_r,\beta_{r-1})}{(\beta_{r-1},\beta_{r-1})}\beta_{r-1}$$

$\{\beta_1, \beta_2, \dots, \beta_r\}$是一个正交向量组

第二步：单位化

$$\eta_1 = \frac{\beta_1}{\|\beta_1\|},\quad \eta_2 = \frac{\beta_2}{\|\beta_2\|},\cdots,\eta_r = \frac{\beta_r}{\|\beta_r\|}$$

$\{\eta_1, \eta_2, \eta_3, \dots, \eta_r\}$就是一个标准的正交向量组.

例： 已知C^3中三个线性无关的向量

$$\alpha_1=[1,i,0]^T,\quad \alpha_2=[1,0,i]^T,\quad \alpha_3=[0,0,1]^T$$

用Gram—Schmidt正交化方法将其标准正交化.

解：正交化

$$\beta_1=\alpha_1=\begin{bmatrix}1\\ i\\ 0\end{bmatrix}$$

$(\alpha_2,\beta_1)=1$

$(\beta_1,\beta_1)=2$

$$\beta_2=\alpha_2-\frac{(\alpha_2,\beta_1)}{(\beta_1,\beta_1)}\beta_1=\begin{bmatrix}1\\ 0\\ i\end{bmatrix}-\frac{1}{2}\begin{bmatrix}1\\ i\\ 0\end{bmatrix}=\frac{1}{2}\begin{bmatrix}1\\ -i\\ 2i\end{bmatrix}$$

$$\beta_3=\alpha_3-\frac{(\alpha_3,\beta_1)}{(\beta_1,\beta_1)}\beta_1-\frac{(\alpha_3,\beta_2)}{(\beta_2,\beta_2)}\beta_2$$

$(\alpha_3,\beta_1)=0$

$(\alpha_3,\beta_2)=-i$

$$=\begin{bmatrix}0\\ 0\\ 1\end{bmatrix}-0\begin{bmatrix}1\\ i\\ 0\end{bmatrix}-\frac{-i}{\frac{3}{2}}\times\frac{1}{2}\begin{bmatrix}1\\ -i\\ 2i\end{bmatrix}=\frac{1}{3}\begin{bmatrix}i\\ 1\\ 1\end{bmatrix}$$

$(\beta_2,\beta_2)=\frac{3}{2}$

$(\beta_3,\beta_3)=\frac{1}{3}$

单位化：

$$\eta_1=\frac{\beta_1}{\|\beta_1\|}=\frac{\sqrt{2}}{2}\begin{bmatrix}1\\ i\\ 0\end{bmatrix}$$

$$\eta_2=\frac{\beta_2}{\|\beta_2\|}=\frac{\sqrt{6}}{6}\begin{bmatrix}1\\ -i\\ 2i\end{bmatrix}$$

$$\eta_3=\frac{\beta_3}{\|\beta_3\|}=\frac{\sqrt{3}}{3}\begin{bmatrix}i\\ 1\\ 1\end{bmatrix}$$

$\{\eta_1,\eta_2,\eta_3\}$即为所求的标准正交向量组

提纲

欧式空间、酉空间

标准正交基、Gram—Schmidt方法

酉变换、正交变换

正规矩阵

Hermite矩阵

正定二次齐式、正定Hermite 矩阵

定义：设$A\in C^{m\times n}$, 用$\overline{A}$ 表示以A的元素的共轭复数为元素组成的矩阵,记$A^H=(\overline{A})^T$

则称A^H 为 A的共轭转置矩阵.

共轭转置矩阵的性质：

(1) $(A+B)^H=A^H+B^H$

(2) $(kA)^H=\overline{k}A^H, \forall k\in \mathrm{C}$

(3) $(AB)^H=B^HA^H$

(4) $(A^k)^H=(A^H)^k, \forall k\in \mathrm{Z}^+$

(5) $(A^H)^H=A$

(6) $|A^H|=\overline{|A|}$

(7) 若A可逆，$(A^H)^{-1}=(A^{-1})^H$

正交矩阵$Q^{n\times n}$:

n 阶实矩阵Q 满足 $Q^TQ=QQ^T=I$

酉矩阵$U^{n\times n}$:

n 阶复矩阵U 满足 $U^HU=UU^H=I$

例:

(1)设$\alpha\in C^n$且 $\alpha^H\alpha=1$ ($(\alpha, \alpha)=1$或$\|\alpha\|=1$), 记

$$H=I-2\alpha\alpha^H$$

则H是一个酉矩阵.

Householder矩阵

(2) $$\begin{bmatrix} 1 & & & & & & & & & & \\ & \ddots & & & & & & & & & \\ & & 1 & & & & & & & & \\ & & & \cos\theta & 0 & \cdots & 0 & \sin\theta & & & \\ & & & 0 & 1 & \cdots & 0 & 0 & & & \\ & & & \vdots & \vdots & \ddots & \vdots & \vdots & & & \\ & & & 0 & 0 & \cdots & 1 & 0 & & & \\ & & & -\sin\theta & 0 & \cdots & 0 & \cos\theta & & & \\ & & & & & & & & 1 & & \\ & & & & & & & & & \ddots & \\ & & & & & & & & & & 1 \end{bmatrix}$$

是正交矩阵.

Givens矩阵.

定理：设U是n阶复方阵，以下命题等价

(a) U 是酉矩阵;

(b) U 是可逆阵并且$U^H = U^{-1}$;

(c) U^H是酉矩阵;

(d) U的n个列向量组是标准正交向量组;

(e) U的n个行向量组是标准正交向量组;

(f) $\forall x \in C^n$，$y=Ux$ 与x的长度相等，即$y^H y = x^H x$.

$$\forall x \in C^n, \quad (Ux, Ux) = (x, x).$$

酉矩阵与正交矩阵的性质

若**U, V都是n阶酉矩阵(或都是正交阵), 则 UV**也是 酉矩阵(正交阵**).**

正交矩阵的行列式等于 **+1** 或 **−1**

若U是n阶酉矩阵，则 **|(det(U))|=1.**

定义：设$\{\alpha_1, \alpha_2, \ldots, \alpha_r\}$为$R^n$空间的标准正交列向量组，则称$n\times r$型矩阵

$$Q_1=[\alpha_1, \alpha_2, \ldots, \alpha_r]$$

为一个次正交矩阵. 一般地将其记为$Q_1\in Q_r^{n\times r}$。

定义：设$\{\alpha_1, \alpha_2, \ldots, \alpha_r\}$为$C^n$空间的标准正交列向量组，则称$n\times r$型矩阵

$$U_1=[\alpha_1, \alpha_2, \ldots, \alpha_r]$$

为一个次酉矩阵. 一般地将其记为$U_1\in U_r^{n\times r}$。

引理：$U_1\in U_r^{n\times r}$的充分必要条件是 $U_1^H U_1=I_r$

$Q_1\in Q_r^{n\times r}$的充分必要条件是 $Q_1^T Q_1=I_r$

定义： 设V是一个n维内积空间, σ是V的一个线性变换, 如果对任意的$\alpha, \beta \in V$都有

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

则称σ是V的一个**酉变换**.

定义： 设V是一个n维欧式空间, σ是V的一个线性变换, 如果对任意的$\alpha, \beta \in V$都有

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

则称σ是V的一个**正交变换**.

定理：设V是一个n维内积空间, σ是V 的一个线性变换, 那么下列陈述等价:

(1) σ是酉变换;

(2) $\|\sigma(\alpha)\|=\|\alpha\|, \quad \forall \alpha\in V$

(3) 将V 的标准正交基底变换成标准正交基底;

(4) 酉变换在标准正交基下的矩阵表示为酉矩阵.

注意：关于正交变换也有类似的性质.

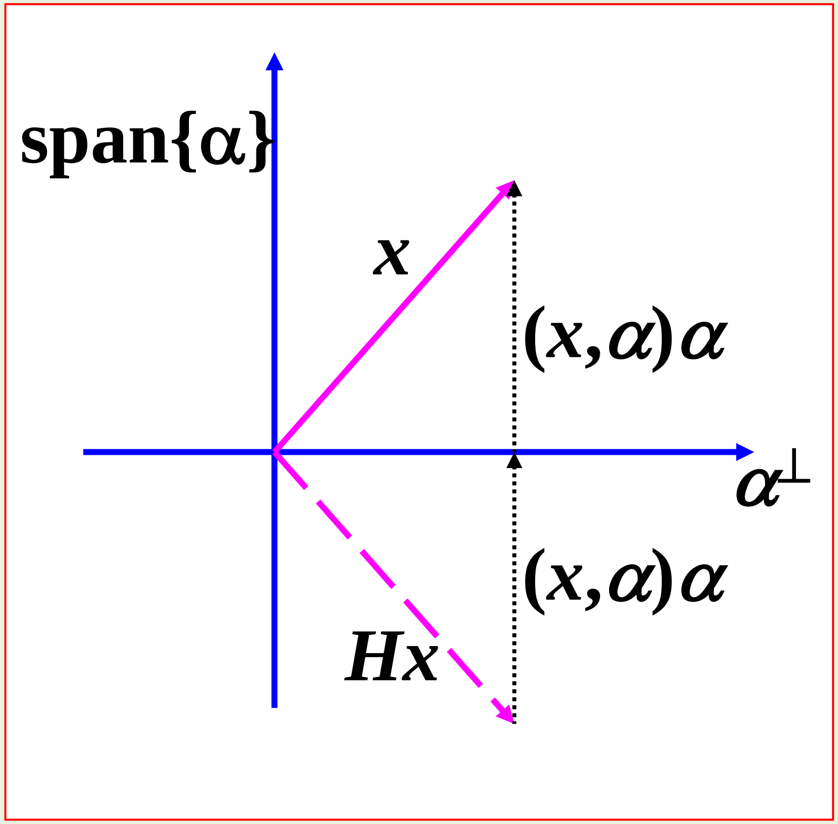
例：Householder矩阵$H=I-2\alpha\alpha^H$ 对应的变换是酉变换

若α是R^n中的一单位向量

$\alpha^{\perp}$表示与向量α垂直的$(n-1)$维超平面

则H保持超平面$\alpha^{\perp}$上的向量不变

而把非超平面$\alpha^{\perp}$中的向量变成关于$\alpha^{\perp}$中超平面的**镜像向量**



镜像变换、反射变换

$\forall x\in R^n$,

$$Hx=Ix-2\alpha\alpha^H x$$
$$=x-2(x,\alpha)\alpha$$

提纲

欧式空间、酉空间

标准正交基、Gram—Schmidt方法

酉变换、正交变换

正规矩阵

Hermite矩阵

正定二次齐式、正定Hermite 矩阵

定义 设A, B为n阶复(或实)方阵, 若存在n阶酉(或正交) 矩阵U, 使得

$$\boxed{U^H AU = U^{-1}AU = B}$$

则称A**酉相似**(或**正交相似**)于B .

定理(Schur引理、Schur分解) 设A为n 阶复矩阵, 则存在一个n阶酉矩阵U, 使得

$$U^H A U = T$$

其中T是上三角矩阵, 且n个对角线元素为A的n个特征值$\lambda_1, \lambda_2, \ldots, \lambda_n$, 而且对角元可以按要求的次序排列。

任何矩阵相似Jordan矩阵

若A为实矩阵, 此时是否能找到实正交矩阵Q, 使得$Q^TAQ=T$ 为实上三角矩阵呢?

一般情形是不对的, 由于T的全部对角线元素为A的所有特征值, 而实矩阵A的特征值可以不是实数. 上三角矩阵T的对角线元素皆为实数又都是A的特征值(可以不是实数), 两者矛盾.

但若实矩阵A的所有特征值皆为实数, 则此时一定存在实正交矩阵Q使得$Q^TAQ=T$为实上三角矩阵

正规矩阵: 方阵A 满足$\underline{AA^H=A^HA}$

实正规矩阵: 方阵A 满足$\underline{AA^T=A^TA}$

Examples:

(a) 若U是<u>酉</u>矩阵, 则 $U^HU=E=UU^H$, 故所有酉矩阵都是正规矩阵.

正规矩阵的性质与结构定理

引理1 设A是一个正规矩阵，则与A酉相似的矩阵一定是正规矩阵.

证(思路)：根据酉相似与正规矩阵的定义即可证明.

引理2 设A是一个正规矩阵，且又是上(下)三角矩阵，则 A 必为对角矩阵.

证(思路)：根据正规矩阵的定义，
比较AA^H和A^HA对角线上的元素
即可证明.

定理(正规矩阵的结构定理)　设A为n阶复矩阵，则A是正规矩阵⇔存在一个n阶酉矩阵U，使得

$$U^H AU = D = \begin{bmatrix} \lambda_1 & & & \mathbf{0} \\ & \lambda_2 & & \\ & & \ddots & \\ \mathbf{0} & & & \lambda_n \end{bmatrix}$$

其中$\lambda_1, \lambda_2, \ldots, \lambda_n$是矩阵$A$的特征值.

证明思路：⇒：用Schur分解定理，A酉相似于上三角矩阵D，D既是上三角矩阵又是正规矩阵，所以D只能是对角矩阵.

⇐：直接根据正规矩阵定义进行验证.

给定n阶正规矩阵A, 存在酉矩阵U, 使得

$$U^H A U = D.$$

- 酉矩阵U：U的n个列向量是A的n个两两正交的单位特征向量,
- 对角矩阵D：A的所有特征值构成的对角矩阵.

$$\Downarrow$$

推论1 n 阶正规矩阵有n个线性无关的特征向量 .

推论2 正规矩阵属于不同特征值的特征向量彼此正交.

举例说明：可对角化的矩阵不一定可酉对角化.

设 X, Y 是两个线性无关但是不正交的向量，比如

取 $P=[X\ \ Y]=\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$, $D=\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$,

则 $A=PDP^{-1}=\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}$

$$=\begin{bmatrix} 1 & 0 \\ 4 & -1 \end{bmatrix}$$

A可对角化，但不能酉对角化

如何求酉矩阵U以及对角矩阵D，使得$U^H AU=D$?

步骤

- 求出A的所有特征值
- 对每个特征值λ, 求特征子空间V_λ以及它的标准正交基:
  解特征矩阵$(\lambda E-A)x=0$求得它的一个基础解系,
  用Gram—Schmidt方法把基础解系单位正交化,
  得到特征子空间V_λ的一组标准正交基
- 把所有特征子空间的标准正交基并在一起构成C^n的一组标准正交基, 以这些基向量为列构造出来的酉矩阵U就是要求的相似变换酉矩阵
- 对角矩阵D的第k个对角元就是酉矩阵U的第k个列向量对应的特征值

【例】： 已知 $A=\begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix}$

验证A是正规矩阵, 且求酉矩阵U, 使U^HAU为对角阵.

解：

$$AA^H=\begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix}\begin{bmatrix}1 & 1\\ -1 & 1\end{bmatrix}=\begin{bmatrix}2 & 0\\ 0 & 2\end{bmatrix}$$

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

$\therefore A$是正规矩阵

$$|\lambda I-A|=\begin{vmatrix}\lambda-1 & 1\\ -1 & \lambda-1\end{vmatrix}=(\lambda-1)^2+1$$

A的特征值：$\lambda_1=1+i,\ \lambda_2=1-i$.

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

$$\boxed{\lambda_1=1+i,\quad \lambda_2=1-i}$$

当λ=1+i时, 特征矩阵

$$\lambda I\text{-}A=\begin{bmatrix}i & 1\\ -1 & i\end{bmatrix}\Rightarrow\begin{bmatrix}1 & -i\\ 0 & 0\end{bmatrix}$$

特征值1+i的单位特征向量

$$\alpha_1=\frac{1}{\sqrt{2}}\begin{bmatrix}i\\ 1\end{bmatrix}$$

当λ=1–i时, 特征矩阵

$$\lambda I\text{-}A=\begin{bmatrix}-i & 1\\ -1 & -i\end{bmatrix}\Rightarrow\begin{bmatrix}1 & i\\ 0 & 0\end{bmatrix}$$

特征值1–i的单位特征向量

$$\alpha_2=\frac{1}{\sqrt{2}}\begin{bmatrix}-i\\ 1\end{bmatrix}$$

令

$$U=[\alpha_1,\alpha_2]=\frac{1}{\sqrt{2}}\begin{bmatrix}i & -i\\ 1 & 1\end{bmatrix}$$

U是酉矩阵，且满足

$$\boxed{U^HAU=\begin{bmatrix}1+i & 0\\ 0 & 1-i\end{bmatrix}}$$

定理 设A是正规矩阵，则

A是酉矩阵$\Leftrightarrow$ A的所有特征值的模长均为1.

证明：A正规，存在酉矩阵使得 $U^HAU=D$为对角阵.

A是酉矩阵$\Leftrightarrow$ D是酉矩阵$\Leftrightarrow$ D的所有特征值的模长均为1 $\Leftrightarrow$ A的所有特征值的模长均为1.

定理(Schur 不等式) 设A为n阶复方阵, $\lambda_1, \lambda_2, \ldots, \lambda_n$ 为矩阵A 的特征值，则

$$\sum_{i=1}^{n}|\lambda_i|^2 \le \sum_{i,j=1}^{n}|a_{ij}|^2$$

其中等号成立 $\Leftrightarrow$ A酉相似于对角矩阵.

证：设存在酉矩阵U, 使得$U^HAU=R$为上三角矩阵,

$$R^H=(U^HAU)^H=U^HA^HU, \quad RR^H=U^HAA^HU,$$

$$\begin{array}{ccc} \mathrm{tr}(AA^H) & = & \mathrm{tr}(RR^H) \\ \| & & \| \\ \sum_{i,j=1}^{n}|a_{ij}|^2 & & \sum_{i=1}^{n}|\lambda_i|^2+\sum_{i<j}|r_{ij}|^2 \end{array}$$

提纲

欧式空间、酉空间

标准正交基、Gram—Schmidt方法

酉变换、正交变换

正规矩阵

Hermite矩阵

正定二次齐式、正定Hermite 矩阵

R域对称矩阵

R域反对称矩阵

定义: 设$A\in C^{n\times n}$, 如果 $A^H=A$,

那么称A为Hermite矩阵; 简称H阵

如果$A^H=-A$,

那么称A为反Hermite矩阵.简称反H阵

$$A^H=A \Leftrightarrow a_{ij}=\overline{a_{ji}},\quad \forall i,j=1,2,\ldots,n$$

$$\Leftrightarrow \boxed{\mathrm{Re}(a_{ij})=\mathrm{Re}(a_{ji}),\quad \mathrm{Im}(a_{ij})=-\mathrm{Im}(a_{ji})}$$

对角线上元素虚部为0

$$A^H=-A \Leftrightarrow -a_{ij}=\overline{a_{ji}},\quad \forall i,j=1,2,\ldots,n$$

$$\Leftrightarrow \boxed{\mathrm{Re}(a_{ij})=-\mathrm{Re}(a_{ji}),\quad \mathrm{Im}(a_{ij})=\mathrm{Im}(a_{ji})}$$

对角线上元素实部为0

例 判断下列矩阵是H阵还是反H阵.

$$(1)\begin{bmatrix} 4i & 2+i & 4+2i \\ -2+i & i & 1 \\ -4+2i & -1 & -2i \end{bmatrix}$$

反H阵

$$(2)\begin{bmatrix} 6 & 1+2i & 3i \\ 1-2i & 9 & 1-i \\ -3i & 1+i & -7 \end{bmatrix}$$

H阵

$$(3)\begin{bmatrix} 0 & 1-i & 8i \\ -1-i & 0 & 4-i \\ 8i & -4-i & 0 \end{bmatrix}$$

反H阵

$$(4)\begin{bmatrix} 3 & 1+3i & 2i \\ 1-3i & 4 & 1+5i \\ -2i & 1-5i & 5 \end{bmatrix}$$

H阵

(5) 实对称矩阵

H阵

(6)反实对称矩阵

反H阵

H-矩阵的基本性质

反H-矩阵的基本性质

引理：设A为n阶矩阵, 则

(1) $A+A^H$, A^HA, AA^H都是H-阵.

(2) $A-A^H$是反 H-阵.

(3) 若A是H-阵，则A^k也是H-阵，k为任意正整数.

(4) 若A是可逆的 H-阵， 则A^{-1}也是可逆的H-阵.

(5) 若A是 H-阵(反H-阵)，则iA 是反H-矩阵 (H-阵).

(6)若A, B都是 H-阵，则$\forall k, l\in$R，$kA+lB$也是H-阵。

(7) 若A, B都是 H-阵，则AB, BA也是H-阵 $\Leftrightarrow$ $AB=BA$

思路：根据H阵的定义

定理: 设A为n阶复矩阵，则

(1) A是H-阵$\Leftrightarrow \forall X\in C^n$, X^HAX 是实数.

(2) A是H-阵$\Leftrightarrow \forall n$ 阶方阵B, B^HAB为H-阵.

证明(1)：$\Rightarrow$: A是H阵,$(X^HAX)^H= X^HAX$, X^HAX是实数

$\Leftarrow$: $\forall X\in C^n$, $\forall Y\in C^n$, $(X+Y)^HA(X+Y)$是实数,

$(X+Y)^HA(X+Y)= X^HAX+ X^HAY+Y^HAX +Y^HAY$

取$X=ie_j$, $Y=e_k$, $X^HAY+Y^HAX=-ia_{jk}+ia_{kj}$是实数

取$X=e_j$, $Y=e_k$, $X^HAY+Y^HAX=a_{jk}+a_{kj}$是实数

$\mathrm{Re}(a_{jk})=\mathrm{Re}(a_{kj})$

$\mathrm{Im}(a_{jk})=-\mathrm{Im}(a_{kj})$

证明(2)： $\Rightarrow$根据H阵定义,

$\Leftarrow$取$B=E$

北京理工大学 BEIJING INSTITUTE OF TECHNOLOGY

H-阵的结构定理

定理(H-阵酉相似于实对角矩阵) 设A为n阶复矩阵, 则A是 H-阵 $\Leftrightarrow$ 存在一个n阶酉矩阵U, 使得

$$U^H A U = \begin{bmatrix} \lambda_1 & & & \mathbf{0} \\ & \lambda_2 & & \\ & & \ddots & \\ \mathbf{0} & & & \lambda_n \end{bmatrix}$$

其中$\lambda_1, \lambda_2, \ldots, \lambda_n \in R$.

证明：由H-阵为正规矩阵、正规矩阵的结构定理以及H-阵对角线上元素都是实数即得

推论 实对称矩阵正交相似于实对角矩阵.

Hermite二次型 (Hermite二次齐次多项式)

定义: n个复变量$x_1, x_2, \ldots, x_n$，系数为复数的二次齐次多项式

$$f(x_1, x_2, \cdots, x_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\overline{x_i}x_j$$

若$a_{ij}=\overline{a_{ji}}$

则称为 Hermite二次型。

记$X=(x_1, x_2, \ldots, x_n)^T \in C^n$,

$$A=\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

则Hermite 二次型可以记为

$$f(x_1, x_2, \ldots, x_n)=X^H AX$$

H-阵A称为Hermite二次型对应的矩阵，

并称A的秩为Hermite二次型的秩.

对于Hermite二次型作可逆的线性替换

$$X=CY \quad (C\text{可逆})$$

则

$$f(x_1, x_2, \ldots, x_n)=X^HAX$$

$$=Y^HC^HACY \qquad \text{令}B=C^HAC$$

$$=Y^HBY$$

复合同

Hermite 二次型中最简单的一种是只含有纯平方项无交叉项的二次型

$$f(y_1, y_2, \cdots, y_n) = \lambda_1 \overline{y_1} y_1 + \lambda_2 \overline{y_2} y_2 + \cdots + \lambda_n \overline{y_n} y_n$$

称这种形状的 **Hermite** 二次型为**标准形的 Hermite二次型.**

定理　对于任意一个 Hermite 二次型

$$f(x_1, x_2, \cdots, x_n) = X^H AX$$

必存在酉线性替换 $X=UY$ (U为酉矩阵)

可以将 Hermite 二次型 $f(X)$ 化为标准形

$$f(y_1, y_2, \cdots, y_n) = \lambda_1 \overline{y_1} y_1 + \lambda_2 \overline{y_2} y_2 + \cdots + \lambda_n \overline{y_n} y_n$$

其中$\lambda_1, \lambda_2, \ldots, \lambda_n$是 H-阵$A$的特征值.

证明：把矩阵的结论翻译成Hermite二次型的结论即可

例: 写出下面Hermite二次型的矩阵表达式，并用酉线性替换将其化为标准形.

$$f(x_1,x_2,x_3)=i\overline{x_1}x_2+\overline{x_1}x_3-ix_1\overline{x_2}+x_1\overline{x_3}$$

解:

令 $X=(x_1, x_2, x_3)^T$

$$f(x_1, x_2, x_3)=X^HAX$$

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

求A的特征值，得

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

$$\lambda_1=0,\ \lambda_2=\sqrt{2},\ \lambda_3=-\sqrt{2}.$$

分别求它们的特征子空间的标准正交基

$$u_1=\frac{1}{\sqrt{2}}\begin{bmatrix}0\\ 1\\ -i\end{bmatrix}, u_2=\frac{1}{2}\begin{bmatrix}\sqrt{2}i\\ 1\\ i\end{bmatrix}, u_3=\frac{1}{2}\begin{bmatrix}-\sqrt{2}i\\ 1\\ i\end{bmatrix}.$$

构造酉矩阵$U=\begin{bmatrix}0 & \frac{\sqrt{2}i}{2} & -\frac{\sqrt{2}i}{2}\\ \frac{1}{\sqrt{2}} & \frac{1}{2} & \frac{1}{2}\\ -\frac{i}{\sqrt{2}} & \frac{i}{2} & \frac{i}{2}\end{bmatrix}$, 令$X=UY$, 则$f$ 化为

$$f=\sqrt{2}\cdot\overline{y_2}y_2-\sqrt{2}\cdot\overline{y_3}y_3.$$

北京理工大学 BEIJING INSTITUTE OF TECHNOLOGY

提纲

欧式空间、酉空间

标准正交基、Gram—Schmidt方法

酉变换、正交变换

正规矩阵

Hermite矩阵

正定二次齐式、正定Hermite 矩阵

Hermite 二次型分类

定义 对于给定的 **Hermite** 二次型

$$f(X)=f(x_1,x_2,\cdots,x_n)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}\overline{x_i}x_j=X^H AX$$

如果对于任意一组不全为零的复数$x_1, x_2, \ldots, x_n$,

都有 $$f(x_1,x_2,\cdots,x_n)>0(\geq 0)$$

则称该 **Hermite** 二次型为正定的(半正定的)，并称

相应的H-阵A为正定的(半正定的) .

如果对于任意一组不全为零复数$x_1, x_2, \ldots, x_n$,

都有 $$f(x_1,x_2,\cdots,x_n)<0(\leq 0)$$

则称该 **Hermite** 二次型为负定的(半负定的)，并称

相应的H-阵A为负定的(半负定的) .

例 判断下列 Hermite 二次型的类别

(1) $f(x_1,x_2,x_3)=4\overline{x_1}x_1+8\overline{x_2}x_2+3\overline{x_3}x_3$ 正定的

(2) $f(x_1,x_2,x_3)=12\overline{x_2}x_2+9\overline{x_3}x_3$ 半正定的

(3) $f(x_1,x_2,x_3)=-7\overline{x_1}x_1+6\overline{x_2}x_2+\overline{x_3}x_3$ 不定的

(4) $f(x_1,x_2,x_3)=-\overline{x_1}x_1-4\overline{x_2}x_2-3\overline{x_3}x_3$ 负定的

(5) $f(x_1,x_2,x_3)=-6\overline{x_1}x_1-13\overline{x_3}x_3$ 半负定的

定理： 对于给定的 Hermite 二次型$f(X) = X^HAX$

下列叙述是等价的.

(1)　$f(X)$是正定的,

(2)　$\forall n$阶可逆矩阵P都有 P^HAP为正定矩阵,

(3)　A的n个特征值都大于零,

(4)　存在n阶可逆矩阵P使得$P^HAP=E$,

(5)　存在n阶可逆矩阵Q使得$A=Q^HQ$,

(6)　存在正线上三角矩阵R使得$A=R^HR$, 且此分解是唯一的. (Cholesky分解)

证明: (1)⇒(2)⇒(3)⇒(4)⇒(5)⇒(6)⇒(1)

要点: (a)正定的定义, (b) H-阵酉相似于对角阵

定理 对于给定的 Hermite 二次型 $f(X) = X^H AX$ 下列叙述是等价的.

(1) $f(X)$ 是正定的,

(2) $\forall n$ 阶可逆矩阵 P 都有 $P^H AP$ 为正定矩阵,

(1)$\Rightarrow$(2)

$f(X)$ 是正定的 $\Leftrightarrow \forall X \neq 0, X^H AX > 0$

$\forall X \neq 0$, P 可逆 $\Rightarrow PX \neq 0$,

$\Rightarrow (PX)^H A(PX) > 0$

$\Leftrightarrow X^H P^H APX > 0$

因此 $P^H AP$ 正定.

定理　对于给定的 Hermite 二次型 $f(X) = X^H AX$ 下列叙述是等价的.

(2) $\forall n$ 阶可逆矩阵 P 都有 $P^H AP$ 为正定矩阵,

(3) A 的 n 个特征值都大于零,

(2) $\Rightarrow$ (3)

A 是H-阵,

存在酉矩阵 U，$U^H AU=D=\text{diag}(\lambda_1, \lambda_2,\ldots, \lambda_n)$,

D 是正定矩阵,

D 的对角线元素都是正的,

A 的所有特征值都是正的.

定理 对于给定的 Hermite 二次型$f(X) = X^HAX$ 下列叙述是等价的.

(3) A的n个特征值都大于零,

(4) 存在n阶可逆矩阵P使得，$P^HAP=E$

(3)⇒(4)

存在酉矩阵U，$U^HAU=D=\mathrm{diag}(\lambda_1, \lambda_2,\ldots, \lambda_n)$,

且所有$\lambda_i>0$,

令$P_1=\mathrm{diag}(\frac{1}{\sqrt{\lambda_1}}, \frac{1}{\sqrt{\lambda_2}},\ldots, \frac{1}{\sqrt{\lambda_n}})$

$P_1^{\mathrm{H}}DP_1=E$

$P_1^{\mathrm{H}}\, U^HAUP_1=E$

取$P=UP_1$, $P^HAP=E$

定理　对于给定的 Hermite 二次型$f(X) = X^H AX$ 下列叙述是等价的.

(4)　存在n阶可逆矩阵P使得，$P^H AP=E$,

(5)　存在n阶可逆矩阵Q使得$A=Q^H Q$,

(4)⇒(5)

$$P^H AP=E$$

$$A=(P^H)^{-1}P^{-1}=(P^{-1})^H P^{-1}$$

取$Q=P^{-1}$即可.

定理　对于给定的 Hermite 二次型$f(X) = X^H AX$下列叙述是等价的.

(5) 存在n阶可逆矩阵Q使得$A=Q^H Q$,

(6) 存在正线上三角矩阵R使得$A=R^H R$，且此分解是唯一的. (Cholesky分解)

(5)$\Rightarrow$(6)

矩阵Q有唯一的正交三角分解，(第四章定理),

$Q=UR$, 其中U是酉矩阵，R是正线上三角矩阵

故$A=Q^H Q=(UR)^H UR=R^H U^H UR=R^H R$.

定理　对于给定的 Hermite 二次型$f(X) = X^H AX$ 下列叙述是等价的.

(1)　$f(X)$ 是正定的,

(6)　存在正线上三角矩阵R使得$A=R^H R$，且此分解是唯一的. (Cholesky分解)

(6)⇒(1)

$\forall X\neq 0 \Rightarrow RX\neq 0$,

$X^H AX=X^H R^H RX =(RX)^H(RX)=(RX,RX)>0$

$f(X)$是正定的.

定理： 对于给定的 Hermite 二次型 $f(X) = X^H A X$ 下列叙述是等价的.

(1) $f(X)$ 是正定的,

(2) $\forall n$阶可逆矩阵P都有 $P^H A P$为正定矩阵,

(3) A的n个特征值都大于零,

(4) 存在n阶可逆矩阵P使得$P^H A P = E$,

(5) 存在n阶可逆矩阵Q使得$A = Q^H Q$,

(6) 存在正线上三角矩阵R使得$A = R^H R$, 且此分解是唯一的. (Cholesky分解)

定理　n阶Hermite(实对称)矩阵 $A=(a_{ij})$是正定的$\Leftrightarrow A$的n个顺序主子式全大于零, 即

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{vmatrix} > 0, \quad k = 1, 2, \cdots, n.$$

定理　对于给定的 Hermite 二次型$f(X) = X^H AX$ 下列叙述是等价的:

(1) $f(X)$ 是半正定的,

(2) $\forall n$ 阶可逆矩阵 P 都有$P^H AP$为半正定矩阵,

(3) A 的 n个特征值全是非负的,

(4) 存在n阶可逆矩阵P 使得

$$P^H AP = \begin{bmatrix} I_r & O \\ O & O \end{bmatrix}$$

(5) 存在秩为r 的n 阶矩阵Q 使得

$$A=Q^H Q$$

证明：参考正定矩阵的证明方法 + 半正定定义

【例】： 设A是正定(半正定) H- 矩阵，证明存在正定(半正定) H- 矩阵H 使得$A=H^2$.

证明： 以正定矩阵为例.

A正定,

存在酉矩阵U, 使得

$$U^H A U = D = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$$

所有的 λ_i 均大于零.

取H为

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

可验证H是正定H-矩阵并且$A=H^2$.

几种特殊矩阵	等价性质
正规矩阵A $A^HA=AA^H$	酉相似对角矩阵
酉矩阵A $A^HA=I$	酉相似对角线元素模为1的对角矩阵
H-阵A $A^H=A$	酉相似对角线元素为实数的对角矩阵 $\forall X\in C^n$, X^HAX 是实数
反H-阵A $A^H=-A$	酉相似对角线元素为虚数或0的对角矩阵 $\forall X\in C^n$, X^HAX 是虚数或0