基本知识

第 0 章 复习与引申

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

初等变换：若 A 的秩为 r，则存在可逆阵 P 和 Q 使得

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

同时将 I_r 拆成 e_i 形式可以得到

$$A = A_1 + \cdots + A_r$$

对于可逆阵，则可分解为一些初等矩阵的乘积。

若 A 为方阵，存在可逆阵 B 和幂等阵 C 使得

$$A = P\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}Q = PQQ^{-1}\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}Q = BC$$

如果 n 阶方阵 A 的每行元素和为 a，则有 $Ae = ae$，其中 $e = (1, \cdots, 1)^T$

令 M 为循环阵（见第0章习题11），首先定义下面这个矩阵

$$A^k = \begin{bmatrix} O & I_{n-k} \\ I_k & O \end{bmatrix}$$

则有 $M = a_0 I + a_1 A + a_2 A^2 + \cdots + a_{n-1}A^{n-1}$

满秩分解求解，首先对 A 进行行变换，化成**行最简型** B，通过行最简型 B 可以从 A 中获取对应的列 C，则 $A = CB$

$\det(A+B) = \det(A^T + B^T)$

> 已知 $A^2 - A = 2I$，且 $A + I$ 的秩为 r，求 A 的可能形式

$$A^2-A-2I=0\Rightarrow(A-2I)(A+I)=0$$

所以 A 的特征值只能为 2 或者 -1，同时由于最小多项式没有重根，所以能相似于对角阵，对角线只能为 -1 或2，由于 $A+I$ 的秩为 r，所以 A 相似于

$$\begin{pmatrix}2I_{n-r}&O\\O&-I_r\end{pmatrix}$$

即存在可逆阵 P，使得

$$A=P\begin{pmatrix}2I_{n-r}&O\\O&-I_r\end{pmatrix}P^{-1}$$

证明 $A^2=A$ 的充要条件为 $r(A)+r(I_n-A)=0$

必要性：若 $A^2=A$，则有 $A(I-A)=0$，所以有

$$r(A(I-A))\geq r(A)+r(I_n-A)-n\Rightarrow r(A)+r(I_n-A)\leq n$$

又 $r(A)+r(I_n-A)\geq r(I_n-A+A)=n$，所以 $r(A)+r(I_n-A)=0$

充分性：假设 ξ 是 $AX=0$ 的基础解系，η 为 $AX=X$ 的基础解系，则 ξ 和 η 线性无关，且为 $A(I-A)X=0$ 的解。所以有下面不等式

$$n-r(A)+n-r(I_n-A)\leq n-r(A(I-A))\Rightarrow r(A(I-A))\leq 0$$

所以 $A(I-A)=0$，即 $A^2=A$。

设 A，B 为 n 阶对合阵，且 $\det AB<0$，证明：存在非零列向量 X，使得 $BAX+X=0$

要证明存在非零列向量 X，使得等式成立，相当于证明 $BA+I$ 秩不为 n，即证明 $\det(BA+I)=0$，下面进行证明：

$$\det(BA+I)=\det(BA+B^2)=\det B(A+B)=\det B\det(A+BA^2)=\det A\det B\det(I+BA)$$
$$\Rightarrow(1-\det AB)\det(I+BA)=0\Rightarrow\det(I+BA)=0$$

如果 e_i 为第 i 位为 1 的单位向量，则有 Ae_i 为 A 的第 i 列，e_i^TA 为 A 的第 i 行

第一章 线性空间与线性变换

证明子空间，首先需要证明子空间非空，同时证明线性运算封闭（若 $A,B\in V$，则 $kA+lB\in V$）

证明直和，可以通过下面几种方法（已经能够确定 $V=V_1+V_2$）

1. $V_1\cap V_2=\{0\}$
2. 零分解式唯一，即若 $\alpha_i\in V_i$，且 $\sum\limits_i^n\alpha_i=0$，能推出 $\alpha_i=0$
3. $\dim V_1+\dim V_2=\dim(V_1+V_2)$（确定是有限维的情况下使用）

下面介绍如何确定 $V=V_1+V_2$

假设 $\alpha\in V$，对 α 进行一些等价代换如 $\alpha=\alpha-f(\alpha)+f(\alpha)$（具体根据题目来定），使得 $\alpha-f(\alpha)\in V_1$，$f(\alpha)\in V_2$

记 V(F) 到 U(F) 的一切线性映射之集合为 Hom(V, U)

零映射：将 V 中每一向量都映射为 U 的零向量，恒等变换：将 V 中每个向量映射为本身

设 f 在基偶 $\{\alpha_i\}_1^n$ 和 $\{\beta_i\}_i^n$ 下的矩阵为 A，表示为 $f(\alpha_i)=\beta_iA$

一个线性变换的核空间和值域存在下面这一关系：

$$\dim R(f)+\dim K(f)=\dim V=n$$

注意该等式不能说明 $R(f)\cap K(f)=0$，判断 $R(f)+K(f)$ 为直和的方法有：（1）将两者的基展开排列，进行初等变换，判断矩阵的秩是否等于空间维度；（2）令 $\alpha\in R(f)\cap K(f)$，$\beta\in V$，有 $\alpha=f(\beta)$，接下来根据题中所给条件证明 $\alpha=0$，一个示例为假设 $f^2=f$，有 $\alpha=f(\beta)=f^2(\beta)=f(\alpha)=0$。

> 证明 $R(f)=V$ 的充分必要条件为 $K(f)=\{0\}$

因为 $\dim R(f)+\dim K(f)=\dim V=n$，所以
$R(f)=V\Leftrightarrow\dim R(f)=n\Leftrightarrow\dim K(f)=0\Leftrightarrow K(f)=\{0\}$

同理 $V=R(f)+K(f)$ 的充分必要条件为 $R(f)\cap K(f)=\{0\}$

$V=R(f)+K(f)\Leftrightarrow\dim V=\dim(R(f)+K(f))\Leftrightarrow\dim(R(f)\cap K(f))=0$

> 证明两个空间相同

1. 证明两个空间相互包含
2. 证明一个空间包含另一个，且维数相等

线性变换 f 的特征子空间：$V_\lambda = \{x|f(x) = \lambda x, x \in V\}$

> 已知一组基，如何找到线性无关的一组基
>
> 如果基向量是列向量，则将其按列排列，同时做列变换
>
> 如果基向量是行向量，则将其按行排列，同时做行变换
>
> 已知 V_1 和 V_2，求 $V_1 + V_2$，$V_1 \cap V_2$
>
> $V_1 + V_2$：将 V_1 和 V_2 的基排在一起，通过行（列）变换找到线性无关的一组基
>
> $V_1 \cap V_2$：假设 $V_1 = \{a_1, a_2\}$，$V_2 = \{b_1, b_2\}$，待定系数法求解 $k_1 * a_1 + k_2 * a_2 = k_3 * b_1 + k_4 * b_2$ 得到基，求出 k_1, k_2, k_3, k_4 后，将 $k_1 * a_1 + k_2 * a_2$ 作为 $V_1 \cap V_2$ 的基

> ⚠ **Warning**
>
> 如果需要求 V_1 和 V_2 的基，注意需要验证基之间线性无关（即 $k_1 a_1 + k2 a_2 = 0$ 能得到 $k_1 = k_2 = 0$ ）

> ⓘ **Note**
>
> $R(A)$ 和 $K(A)$ 的基向量的维度可以不同，这里维度不同指的是 $R(A)$ 的基可以用若干个 5 维向量组成，$K(A)$ 的基可以用若干个 3 维向量组成。

> 若已知 f 在一组基 α 下的矩阵 A，求 f 在另外一组基 β 下的矩阵 B
>
> 先求过渡矩阵 $\beta = \alpha P$，再由 $B = P^{-1}AP$ 求出 B

同构：设 V 与 U 是数域 F 上线性空间，如果存在 V 到 U 的一个双射 σ，且 σ 又是线性映射，则称 V 与 U 同构，σ 是 V 到 U 的一个同构映射。

设 σ 为 V 到 U 的同构映射，则 V 中向量 $\alpha_1, \cdots, \alpha_k$ 线性无关的充要条件为 $\sigma(\alpha_1), \cdots, \sigma(\alpha_k)$ 线性无关

两个有限维线性空间 V 与 U 同构的充要条件为 $\dim V = \dim U$

> 如何证明 V 中 f 的矩阵必相似于某个矩阵

首先根据题目所给信息定义两个空间 V_1 和 V_2，证明这两个空间是 V 的子空间，再证明 V 是这两个子空间的直和，证明完成后，假设这两个子空间的基，然后通过一开始定义的空间的表达形式给出 f 的矩阵（f 在不同的基下的矩阵是相似）

第二章 内积空间和等距变换

如何证明内积空间：（1）$\langle\beta,\alpha\rangle=\overline{\langle\alpha,\beta\rangle}$（2）$\langle\alpha+\beta,\gamma\rangle=\langle\alpha,\gamma\rangle+\langle\beta,\gamma\rangle$（3）$\langle k\alpha,\beta\rangle=k\langle\alpha,\beta\rangle$（4）$\langle\alpha,\alpha\rangle\geq 0$

上三角矩阵的逆矩阵仍为上三角矩阵，若 A 可逆，则 $A\beta=0\Rightarrow\beta=0$（A 的列向量组线性无关，所以 β 只能为 0）

假设对角分块矩阵 A 为

$$A=\begin{pmatrix}A_1 & 0 & 0\\ 0 & A_2 & 0\\ 0 & 0 & A_3\end{pmatrix}$$

则行列式等于对角子阵行列式的乘积 $|A|=|A_1|\,|A_2|\,|A_3|$

线性变换 $f(x)$ 在某个基下的矩阵的特征值 λ_i，求解对应的特征子空间的基 X_i，则有

$$f(X_1,\cdots,X_n)=(X_1,\cdots,X_n)\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}$$

正定阵是实对称的。

一个不等式：$\left|\int_a^b f_i(x)g_i(x)dx\right|\leq\int_a^b|f_i(x)g_i(x)|dx$

$tr(A^T)=tr(A)$

证明内积空间 `<A B>` 是欧式空间或者酉空间，则需要证明 `<A B>` 满足四个条件：

- $\langle B,A\rangle=\overline{\langle A,B\rangle}$
- $\langle A_1+A_2,B\rangle=\langle A_1,B\rangle+\langle A_2,B\rangle$
- $\langle kA,B\rangle=k\langle A,B\rangle$
- $\langle A,A\rangle\geq 0$，当且仅当 A=0 时取等

α 的长度为 $\sqrt{\langle\alpha,\alpha\rangle}$，记为 $||\alpha||$

称 $d(\alpha-\beta)=\|\alpha-\beta\|$ 为 α 和 β 的距离

如果矩阵 A 为酉矩阵，则其列向量组为酉空间 C^n 的标准正交基，反之也对（A^HA的 i 行 j 列元素正是 A 的第 j 列与第 i 列这两个向量的内积）

一个不等式：$|\langle X,Y\rangle|^2 \leq \langle X,X\rangle \langle Y,Y\rangle$（用量纲理解，左边是单位的4次方，右边也应该为单位的4次方）

一个不等式：$\|X+Y\| \leq \|X\|+\|Y\|$

度量矩阵 G：$g_{ij}=\langle \alpha_i,\alpha_j\rangle$，$\alpha$ 在 V 的一组基 $\{\alpha_i\}_1^n$ 的坐标为 X，β 的坐标为 Y。则有

$$\langle \alpha,\beta\rangle = X^TG\bar{Y} = Y^HG^TX$$

可逆阵的 UT 分解：设 A 为可逆阵，可以分解为一个酉矩阵 U 和主对角元恒正的上三角阵 T，A = UT，且分解唯一。具体求解方法为：先用schmidt正交化 A 的列向量组求 U，再用 $T=U^{-1}A$ 求 T，注意 $U^{-1}=U^H$

> ⓘ **Note**
>
> 事实上，任意方阵 $A\in C^{n\times n}$ 都可以分解为酉矩阵 U 与主对角元非负的上三角阵 T 的乘积，即 A = UT

若 U_2 和 U_1 均为酉矩阵，则 $U_2^HU_1$ 也是酉矩阵 $\left(U_2^HU_1\right)^H\left(U_2^HU_1\right)=I$

α 与 β 正交：$\alpha\perp\beta \Leftrightarrow \cos\phi=0 \Leftrightarrow \langle\alpha,\beta\rangle=0 \Leftrightarrow \frac{\langle\alpha,\beta\rangle}{\|\alpha\|+\|\beta\|}=0$

> 证明 f 是正交变换

设 f 为内积空间 V 的线性变换，V 为欧式空间，证明

$$\langle f(\alpha),f(\beta)\rangle = \langle\alpha,\beta\rangle$$

若 A 为正交阵，则 Y = AX 为正交变换，此外还有 $A^T=A^{-1}$，f 是正交变换的充要条件为 $\|f(\alpha)\| = \|\alpha\|$

> ⓘ **Note**
>
> 若 V 为酉空间，则称 f 是酉变换

若 $\beta\in W$，使得 $d(\beta,\alpha)=\min\{d(\xi,\alpha)|\xi\in W\}$，则称 β 为 α 在子空间 W 上的正投影，$\|\beta-\alpha\|$ 为 α 到 W 的最短距离，若 $\varepsilon_1,\cdots,\varepsilon_r$ 为 W 的一组标准正交基，则 $\beta=\langle\alpha,\varepsilon_1\rangle\varepsilon_1+\cdots+\langle\alpha,\varepsilon_r\rangle\varepsilon_r$

> 镜像变换的矩阵相似于 $diag(-1,1,\cdots,1)$，下面给出证明
>
> 已知内积空间 V 中的镜像变换为 $f(\xi)=\xi-2\langle\xi,\omega\rangle\omega$，设 w 为 V 中的单位基
>
> 则有 $f(\omega)=\omega-2\langle\omega,\omega\rangle\omega=-\omega$
>
> 现在将 w 扩充为标准正交基 $(\omega,\xi_1,\cdots,\xi_{n-1})$，则有 $f(\xi_i)=\xi_i$，可得

$$f(\omega, \xi_1, \cdots, \xi_{n-1}) = (\omega, \xi_1, \cdots, \xi_{n-1}) \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}$$

V 为齐次线性方程组 $AX = 0$ 的解空间，求其正交补空间 $V^\perp$ 的标准正交基

因为 $V \in K(A)$，所以 $V^\perp \in R(A^H)$，因此只需要将 A^H 的列向量化为标准正交基即可（需要单位化）。有时只需要求一组基，那么只需找出 A^H 中线性无关的一组基即可（通过初等变换）

若已知 $\langle f(\alpha), f(\beta) \rangle = \langle \alpha, \beta \rangle$，证明 f 是线性变换

证明 $\langle f(k\alpha + l\beta) - kf(\alpha) - lf(\beta), f(k\alpha + l\beta) - kf(\alpha) - lf(\beta) \rangle = 0$ 即可

第三章 矩阵的相似标准形

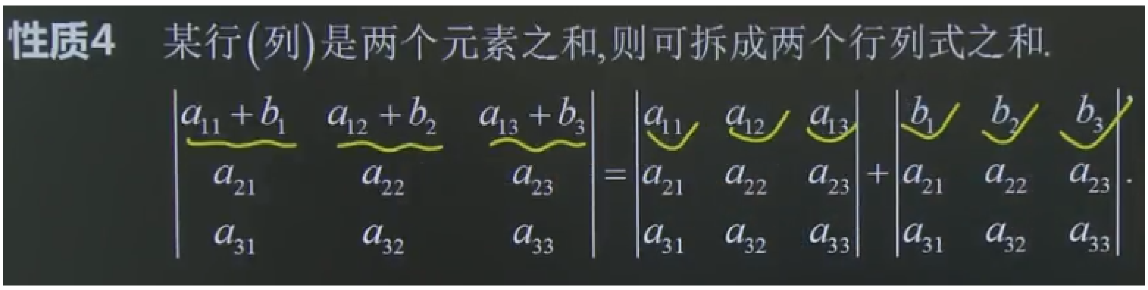
A 的特征多项式为 $C(\lambda) = \det(\lambda I - A)$，且

$$\det(\lambda I - A) = \lambda^n + \sum_{k=1}^{n} (-1)^k b_k \lambda^{n-k}$$

其中 b_k 为 A 的所有 k 阶主子式之和，特别有 $b_1 = trA$，$b_n = \det A$

以 $A = (A_1, A_2, \cdots, A_n)$ 为例，b_k 表示在 A 中挑选 $n-k$ 列替换成 e_i 之后计算行列式，将所有的可能加起来，若 $A = (A_1, A_2, A_3)$，则 $b_2 = |(e_1, A_2, A_3)| + |(A_1, e_2, A_3)| + |(A_1, A_2, e_3)|$（随机挑选 $3-2$ 列替换为 e_i）

特别的，有 $C(A) = C(\lambda_1)C(\lambda_2)\cdots C(\lambda_n) = 0$（Hamiltom-Cayley定理）

> **Caution**
>
> 行列式具有可拆性，即某一行（列）可以被拆开计算行列式（一次只能8一行（列））
>
> 
>

多项式 $\lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n$ 的友阵 F 为

$$F=\begin{bmatrix}0 & 0 & 0 & \cdots & 0 & -a_n \\ 1 & 0 & 0 & \cdots & 0 & -a_{n-1} \\ 0 & 1 & 0 & \cdots & 0 & -a_{n-2} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -a_1\end{bmatrix}$$

若 A，B 为 n 阶方阵，$\det(\lambda I-AB)=\det(\lambda I-BA)$

若 A、B 分别为 s × n 和 n × s 矩阵，则 $\lambda^n\det(\lambda I_s-AB)=\lambda^s\det(\lambda I_n-BA)$

如果遇到求解 $\det\left(\lambda A-\alpha\beta^T\right)=0$，可以采用以下方法

$$\begin{aligned}&\det\left(\lambda A-\alpha\beta^T\right)=\det\left(\lambda I_n-\alpha\beta^TA^{-1}\right)\det\left(A\right)=0\\&\det\left(\lambda I_n-\alpha\beta^TA^{-1}\right)=\lambda^n-tr\left(\alpha\beta^TA^{-1}\right)\lambda^{n-1}=\lambda^n-A^{-T}\beta\alpha^T\lambda^{n-1}\end{aligned}$$

> 求解矩阵方程 $X^2-X-20I=0$
>
> $$X^2-X-20I=\left(X-5I\right)\left(X+4I\right)=0$$
>
> 所以 X 的特征值为 5 或 -4，则有 X 相似于 $diag(5I_r,4I_{n-r})$，所以 $X=Pdiag(5I_r,4I_{n-r})P^{-1}$

Schur 引理：任一 n 阶复方阵 A 必酉相似于上三角阵，即存在酉矩阵 U 使得 $U^HAU=T$

如果特征值全为实数，则 A 正交相似于实上三角阵。

$|-A|=(-1)^n|A|$

若有多项式 $\varphi(x)$ 使 $\varphi(x)=0$，则称 $\varphi(x)$ 为化零多项式，A 的化零多项式中，次数最小并且最高次系数为 1 的叫做 A 的最小多项式。

如果 A 和 B 的最小多项式不同，则必定不相似。

$AP=B$ 这种情况**不是相似**

若 A 和 B 相似，则有

1. A 和 B 的秩、行列式，迹相等
2. 特征值相同（特征向量不一定相同）
3. 特征多项式相同，最小多项式相同

特征子空间：$V_{\lambda_0}=\{\xi|f(\xi)=\lambda_0\xi,\ \xi\in V\}$ 或者 $V_{\lambda_0}=\{X|AX=\lambda_0X,\ X\in C^n\}$

若 V 的某组基下 f 的矩阵为 A，则 f 的特征值便是 A 的特征值。

几何重数 s_i 小于等于代数重数 c_i，几何重数为特征子空间的维数，或者说是$(A-\lambda I)X=0$ 中解空间的维数，即 $n-r(A-\lambda I)$，代数重数等于特征多项式中对应特征值的幂次。

若代数重数之和为空间 V 的维度，下面三条命题等价：

（1）几何重数和代数重数相等
（2）f 的矩阵可以相似于对角阵
（3）V 为特征子空间的直和。

f 可以相似对角化的充要条件是**最小多项式无重因式。**

f 可以相似对角化，即 f 在一组基下的矩阵 A 可以相似对角化，即存在可逆矩阵 P，$P^{-1}AP=\Lambda$

对应 λ 的特征子空间的维度等于 $(A-\lambda I)X=0$ 的维度，相当于 $n-(A-\lambda I)$

Jordan 幂零块：对角线元素为 0 的Jordan块

对于方阵 A，则必存在**可逆阵** P 使得 $P^{-1}AP=J$

A B 相似当且仅当有相同的 Jordan标准形

通过 J 求 P 时，如果遇到 A 过于复杂，可以先进行行化简，记得需要记下行化简的步骤，方便后续操作

A 的特征值的最大模为 A 的谱半径 $\rho(A)$，$\rho(A)\le\rho_1$，且 $\rho(A)\le\rho_2$，其中ρ_1 为按行取绝对值相加后的最大值，ρ_2 为按列取绝对值相加后的最大值

> 设 α 和 β 为 n 维列向量，则 $\alpha\beta^H$ 的 Jordan 标准形为
>
> 首先有 $tr(\alpha\beta^H)=tr(\beta^H\alpha)=\beta^H\alpha$，再由 $r(\alpha\beta^H)=1$，所以有特征多项式
>
> $$\left|\lambda I-\alpha\beta^H\right|=\lambda^n-\beta^H\alpha\lambda^{n-1}$$
>
> 所以特征值为 0 （n-1重），$\beta^H\alpha$
>
> （1）$\beta^H\alpha\neq0$，则有
>
> $$J=\begin{pmatrix}\beta^H\alpha & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0\end{pmatrix}$$
>
> （2）$\beta^H\alpha=0$，则有

$$J=\begin{pmatrix}0 & 1 & & \\ & 0 & & \\ & & \ddots & \\ & & & 0\end{pmatrix}$$

根子空间：$V_i=\{\xi|(f-\lambda_i I)^{c_i}\xi=0,\xi\in V\}$，事实上根子空间是 $(f-\lambda_i I)^{c_i}$ 的核，也是 f 的不变子空间，且 V 为根子空间的直和。

设 g 为线性空间 W 的线性变换，k 为正整数，$\xi\in W$，若 W 的子空间 $\operatorname{span}\left\{\xi,g(\xi),\cdots,g^{k-1}(\xi)\right\}$，则称其为 g 的**循环不变子空间**。充分必要条件是 $g^k(\xi)=0$

在一个盖尔圆内，特征值的轨迹 $\lambda_i(t)$ 从 $\lambda_i(0)=a_{ii}$ 连续地变化到 $\lambda_i(1)=\lambda_i$。

若（实）矩阵 A 的盖尔圆全为 1 区（指盖尔圆之间不相交），则 A 可相似于（实）对角阵。

若 A 为对角占优矩阵，则 $\det A\neq 0$，则 A 可逆，且 $\rho(A)<\max_i\{2|a_{ii}|\}$；若 A 的主对角元都是正实数且为对角占优矩阵，则 A 的特征值全在右半平面。

第四章 Hermite 二次型

若 $X^H=X$，则称其为 Hermite 阵，Hermite 阵必酉相似于一个实对角阵，（正规阵（A 和 A^H 可交换）可以酉相似于一个对角阵，而普通的方阵只能酉相似于一个上三角阵）

A 为正规阵的充要条件为 $\|AX\|=\left\|A^HX\right\|$，证明如下

必要性：$\|AX\|=\sqrt{X^HA^HAX}=\sqrt{X^HAA^HX}=\left\|A^HX\right\|$
充分性：先证 AA^H 和 A^HA 的对角线元素相等，令 $X=e_k$

$$\|AX\|=\|Ae_k\|=\sqrt{e_k{}^HA^HAe_k}=\sqrt{e_k{}^HAA^He_k}=\left\|A^He_k\right\|$$

由上可得对角线元素相等，再证非对角线元素，令 $X=e_i+e_j(i\neq j)$

$$\|A\left(e_i+e_j\right)\|=\sqrt{\left(e_i+e_j\right)^HA^HA\left(e_i+e_j\right)}=\sqrt{\left(e_i+e_j\right)^HAA^H\left(e_i+e_j\right)}$$

可得 $e_i^H\left(A^HA-AA^H\right)e_j+e_j{}^H\left(A^HA-AA^H\right)e_i=0$，但从这一个方程还无法得出结论，再令 $X=e_i+ie_j(i\neq j)$，有

$$ie_i^H\left(A^HA-AA^H\right)e_j-ie_j{}^H\left(A^HA-AA^H\right)e_i=0$$

两个方程结合起来可得 $AA^H = A^H A$

> 对于正规阵 A 而言，$AX = \lambda X$ 的充分必要条件为 $A^H X = \bar{\lambda} X$

首先易得 $A - \lambda I$ 为正规阵，所以有

$$
\begin{aligned}
&(A-\lambda I)X = 0 \Rightarrow X^H\left(A^H - \bar{\lambda} I\right)(A-\lambda I)X = 0 \\
&\Rightarrow X^H(A-\lambda I)\left(A^H - \bar{\lambda} I\right)X = 0 \Rightarrow \left(A^H - \bar{\lambda} I\right)X = 0
\end{aligned}
$$

==上三角的正规阵必是对角阵==，设 T 为上三角的正规阵（归纳法，设 n-1 的正规阵 T_1 为对角阵）：

$$
T = \begin{pmatrix} r & \alpha \\ O & T_1 \end{pmatrix}
$$

由 $T^H T = TT^H$ 可得

$$
\begin{pmatrix} r & \alpha \\ O & T_1 \end{pmatrix}\begin{pmatrix} r & O \\ \alpha^H & T_1 \end{pmatrix} = \begin{pmatrix} r^2 + \alpha\alpha^H & \\ & \end{pmatrix} = \begin{pmatrix} r^2 & \\ & \end{pmatrix} = \begin{pmatrix} r & O \\ \alpha^H & T_1 \end{pmatrix}\begin{pmatrix} r & \alpha \\ O & T_1 \end{pmatrix}
$$

即 $\alpha\alpha^H = O$，所以 $\alpha = O$，T 为对角阵。

设 $A \in C^{n\times n}$，C 为 n 阶可逆阵，则称 A 与 $C^H AC$ 共轭合同

正定阵：对于一切非零列向量 X，有 $f(X) = X^H AX > 0$，其中 A 为 Hermite 阵

如果 A 为正定阵，则有 $A = P^H P$，P 为可逆阵，此外存在正定阵 S，使得 $A = S^2$

n 阶 Hermite 阵，A 正定的充要条件是 A 的 n 个顺序主子式都大于0。

若 X_0 为 A 对应特征值 λ_0 的特征向量，则有 $R(X_0) = \frac{X_0^H AX_0}{X_0^H X_0} = \lambda_0$

$A^H A$ **为半正定阵**，若 B 为半正定阵，则 $Q^H BQ$ 也是半正定阵（Q 为可逆阵）

奇异值分解：对秩为 r 的矩阵 $A \in C^{s\times n}$，必定存在 s 阶和 n 阶的酉矩阵 U 与 V 使得

$$
U^H AV = \begin{pmatrix} D & O \\ O & O \end{pmatrix}_{s\times n}
$$

其中 $D = \operatorname{diag}\left(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \cdots, \sqrt{\lambda_r}\right)$，而 $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0$。

极分解：设 $A \in C^{n\times n}$，则存在酉阵 U 及半正定阵 M，使得 $A = UM$

瑞利商：设 A 为 Herimite 阵，最大特征值为 λ_1，最小特征值为 λ_n

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

$\lambda_1 = \max R(X) \quad \lambda_n = \min R(X)$，X 为向量。

注意瑞利商分母与正定阵的相似之处，可通过 $X = P^{-1}Y$ 在分母处构造一个正定阵，此外对于分子，若 $P^{-1}BP = A$，P 为可逆阵，则有

$$R(X) = \frac{X^H AX}{X^H X} = \frac{X^H BX}{X^H X}$$

Courant 极大极小原理：

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

一个等式：$X^2 = X^H X$

> 若 A 为正定阵，则有 $C^H AC = I$，如果 B 为Hermite阵，则有 $C^H BC = \Lambda$，两者的 C 相同，证明如下：

因为 A 为正定阵，所以存在可逆矩阵 P 使得 $P^H AP = I$（注意这里的 P 不是酉矩阵，不满足 $P^H P = I$），并且 $P^H BP$ 仍为 Hermite 阵（直接证明 $A^H = A$ 即可），因此存在酉矩阵 U 使得 $U^H P^H BPU = \Lambda$，而 $P^H AP = I \Rightarrow U^H P^H APU = I$，因此取 C = PU 即可。

第五章 范数及矩阵函数

范数 ν 需要满足

1. 正定性：对于非零向量 α，$\nu(\alpha) > 0$
2. 齐次性：对于一切数k，有 $\nu(k\alpha) = |k|\nu(\alpha)$
3. 三角不等式：$\nu(\alpha + \beta) \le \nu(\alpha) + \nu(\beta)$

p-范数计算：$\|X\|_p = \left(\sum_i |x_i|^p \right)^{1/p}$，注意 X 为向量

一个不等式 $\|X + Y\|_2 \le \|X\|_2 + \|Y\|_2$，下面给出证明

$$
\begin{aligned}
&\|X+Y\|_2^2=(X+Y)^H(X+Y)=\langle X+Y,X+Y\rangle\\
&=\langle X,X\rangle+\langle X,Y\rangle+\langle Y,X\rangle+\langle Y,Y\rangle=\langle X,X\rangle+2\mathrm{Re}\langle X,Y\rangle+\langle Y,Y\rangle\\
&\leq\langle X,X\rangle+2\left|\langle X,Y\rangle\right|+\langle Y,Y\rangle\leq\|X\|_2^2+2\|X\|_2\|Y\|_2+\|Y\|_2^2\\
&=\left(\|X\|_2+\|Y\|_2\right)^2
\end{aligned}
$$

对于 2-范数，向量形式：$\left(\sum_i |a_i|^p\right)^{1/p}=\left(\alpha^H\alpha\right)^{1/2}$，矩阵形式：$\left(\sum_{i,j}|a_{ij}|^p\right)^{1/p}=\sqrt{tr\left(A^HA\right)}=\sqrt{tr\left(AA^H\right)}$

矩阵范数的相容性：$\|AB\|_{m_1}\leq\|A\|_{m_1}\|B\|_{m_1}$，$\|\ \|_{m_1}$ 和 $\|\ \|_F$ 都是相容的，$\|\ \|_\infty$ 不是相容的

矩阵范数还有一种特殊形式—算子范数

$$
\begin{aligned}
\|A\|_1&=\max_{\|Y\|_1=1}\|AY\|_1=\max_j\left\{\sum_{i=1}^s|a_{ij}|\right\}\\
\|A\|_2&=\max_{\|Y\|_2=1}\|AY\|_2=\sqrt{\rho(A^HA)}\\
\|A\|_\infty&=\max_{\|Y\|_\infty=1}\|AY\|_\infty=\max_i\left\{\sum_{j=1}^n|a_{ij}|\right\}
\end{aligned}
$$

> **Caution**
>
> 算子范数与向量范数的表达形式基本相同，但是注意算子范数针对的是矩阵，而不是向量。

设 $A\in C^{n\times n}$，则 $\lim_{k\to\infty}A^k=O$ 的充要条件为 A 的谱半径 $\rho(A)<1$

对于任意一类相容的矩阵范数有 $\rho(A)<\|A\|$

> 已知 A，求 e^A
>
> 先求 Jordan 标准形，$P^{-1}AP=J$
>
> 根据公式求解 e^J，$e^A=Pe^JP^{-1}$
>
> 或者根据 Jordan 标准形求出最小多项式，使用待定系数法求出 e^A 的线性表示

设 A，B 均为 C 中的矩阵，O 为 n 阶零矩阵，则

$$
\begin{aligned}
&e^{O}=I \\
&\text{if } AB=BA, \text{ then } e^{A+B}=e^{A}e^{B}=e^{B}e^{A} \\
&(e^{A})^{-1}=e^{-A} \\
&\det(e^{A})=e^{trA}
\end{aligned}
$$

> 求解 $\det(e^{At})$ 的几种方法

1. 直接求出 e^{At}，再求 $\det(e^{At})$
2. 先求出 A 的特征值 λ_i，再求 e^{At} 的特征值 $e^{\lambda_i t}$，再求 $\det(e^{At})$
3. 先求 At 的特征值 $\lambda_i t$，再通过 $\det(e^{At})=e^{tr(At)}$ 求解

例题

内积和投影

定义内积为 $\langle\varphi(x),\psi(x)\rangle=\int_{-1}^{1}\varphi(x)\psi(x)dx$，令 $\alpha=1$，$\beta=x$，$\eta=x^2$，$W=L(\alpha,\beta)$，求 η 在 W 中的正投影

> 该问题等同于求 $\eta_0\in W$，使得 $\eta-\eta_0=\min\eta_i\in W||\eta-\eta_i||$，令 $\eta_0=a+bx$
>
> 解法1：求正投影即求 $\eta_0\in W$，使得 $\langle\eta-\eta_0,\alpha\rangle=\langle\eta-\eta_0,\beta\rangle=0$，两个等式解两个未知量即可
>
> 该解法的核心是在投影的空间内找到具有代表性的矢量，即空间的基，本题已经直接给出基（α 和 β ），如果没有给出，就需要自己寻找基，基的数量和未知数的个数相同，以保证可以解出结果。
>
> 解法2：直接求 $\langle\eta-\eta_0,\eta-\eta_0\rangle$，使其最小
>
> $$
> \begin{aligned}
> \langle\eta-\eta_0,\eta-\eta_0\rangle&=\int_{-1}^{1}\left(x^2-a-bx\right)\left(x^2-a-bx\right)dx \\
> &=\frac{2}{5}-\frac{4}{3}a+2a^2+\frac{2}{3}b^2=2\left(a-\frac{1}{3}\right)^2+\frac{2}{3}b^2+\frac{8}{45}
> \end{aligned}
> $$
>
> 最小即令 a=1/3，b=0 这种方法可能很难得到结果，许多时候还是解法1更适用
>
> 解法3：与解法1类似，先找到投影空间的基向量，再对基向量进行施密特正交化得到η_i，通过类似傅里叶变换的拟合的方法直接解出
>
> $$
> \eta_0=\sum_{i=1}^{n}\langle\eta,\eta_i\rangle\eta_i
> $$
>
> 以上题为例，先进行正交化（本身已经正交，只需单位化）得到 $\eta_1=\frac{\sqrt{2}}{2}$ $\quad\eta_2=\frac{\sqrt{6}}{2}x$

> $$\begin{aligned}\eta_0 &= \langle\eta,\eta_1\rangle\eta_1 + \langle\eta,\eta_2\rangle\eta_2 \\ &= \frac{\sqrt{2}}{2}\int_{-1}^{1}\frac{\sqrt{2}}{2}x^2dx + \frac{\sqrt{6}x}{2}\int_{-1}^{1}\frac{\sqrt{6}}{2}x^3dx = \frac{1}{3}\end{aligned}$$

正交变换

> 假设 $\|\eta\| = \sqrt{2}$，定义 V 上的线性变换 f 如下：$f(\mathbf{x}) = \mathbf{x} - k\langle\mathbf{x},\eta\rangle\eta$，问实数 k 取什么值时，f 为正交变换

解法1：直接通过求解 $\langle f(x), f(x)\rangle$ 获得

解法2：将 η 扩充为一组标准正交基，具体步骤如下

(1) $\varepsilon_1 = \frac{\eta}{\|\eta\|} = \frac{\sqrt{2}}{2}\eta$
(2) 将 ε_1 扩充为一组标准正交基
(3) $f(\varepsilon_1) = \varepsilon_1 - k\langle\varepsilon_1,\eta\rangle\eta = \varepsilon_1 - k\left\langle\varepsilon_1,\sqrt{2}\varepsilon_1\right\rangle\sqrt{2}\varepsilon_1 = (1-2k)\varepsilon_1$
(4) $f(\varepsilon_i) = \varepsilon_i - k\langle\varepsilon_i,\eta\rangle\eta = \varepsilon_i \quad i \geq 2$

所以 f 在 V 的标准正交基下的矩阵为 diag(1-2k,1,...,1)，正交变换等价于 A 为正交矩阵，则 1-2k 为 1 或者 -1，所以 k=0 或 1。

> 假设 V 是有限维欧式空间，$w \in V$ 是单位向量，V 上的线性变换 f 定义如下：对任意 $\eta \in V, f(\eta) = \eta - 2\langle\eta,w\rangle w$，定义内积为 $\langle\varphi(x),\psi(x)\rangle = \int_0^1 \varphi(x)\psi(x)dx$，求 $\alpha = 1$ 及 $\beta = x$ 的长度，并求正实数 k 及单位向量 $w \in R[x]_3$，使得如下的正交变换 f 将 α 变成 $k\beta$

$\|\alpha\| = \sqrt{\langle\alpha,\alpha\rangle} = 1 \quad \|\beta\| = \sqrt{\langle\beta,\beta\rangle} = \frac{1}{\sqrt{3}}$

因为 $\|f(\alpha)\| = \|\alpha\| = \|k\beta\|$，则有 k=$\sqrt{3}$，下面求 w

$$\begin{aligned}&f(\alpha) = \alpha - 2\langle\alpha,w\rangle w = 1 - 2\langle 1,w\rangle w = \sqrt{3}x \\ &\Rightarrow 1 - \sqrt{3}x = 2\langle 1,w\rangle w\end{aligned}$$

所以 w 为

$$w = \frac{1-\sqrt{3}x}{\|w\|} = \frac{1-\sqrt{3}x}{\sqrt{2-\sqrt{3}}} = \frac{2\left(1-\sqrt{3}x\right)}{\sqrt{6}-\sqrt{2}} = \frac{\left(\sqrt{6}+\sqrt{2}\right)\left(1-\sqrt{3}x\right)}{2}$$

> 证明 f 是 V 上的正交变换当且仅当 k=1 （假设 k 非零），下面给出充分性证明

若 f 是 V 上的正交变换，取 $\alpha = w$，则有

$$
\begin{aligned}
&2(1-2k)^2 = \langle f(w), f(w) \rangle = \langle w, w \rangle = 2 \\
&(1-2k)^2 = 1 \Rightarrow k = 1 (k \neq 0)
\end{aligned}
$$

> 证明若 f 是 V 上的正交变换 $f(\alpha) = \alpha - 2\langle \alpha, w \rangle w$，$w \neq 0$，则 $||w|| = 1$

$$
\begin{aligned}
&f(w) = w - 2\langle w, w \rangle w = \left(1 - 2 \| w \|^2\right) w \\
&\|f(w)\| = \left|1 - 2 \| w \|^2\right| \| w \| \Rightarrow \left|1 - 2 \| w \|^2\right| = 1 \Rightarrow \| w \| = 1
\end{aligned}
$$

当 $||w|| = 1$ 时，对于任意 $\alpha \in V$，有

$$
\begin{aligned}
&\|f(\alpha)\|^2 = \langle \alpha - 2\langle \alpha, w \rangle w, \alpha - 2\langle \alpha, w \rangle w \rangle \\
&= \langle \alpha, \alpha \rangle - 2\langle \alpha, w \rangle \langle \alpha, w \rangle - 2\langle \alpha, w \rangle \langle \alpha, w \rangle + 4\langle \alpha, w \rangle^2 \langle w, w \rangle \\
&= \langle \alpha, \alpha \rangle = \| \alpha \|^2
\end{aligned}
$$

因此 f 确实是正交变换。

线性变换

> 定义线性变换 f 为

$$
f(X) = \begin{pmatrix} a & 2a \\ 3a & 4a \end{pmatrix} \quad a = tr(X)
$$

1. 求 f 在 $E_{11}, E_{12}, E_{21}, E_{22}$ 下的矩阵 M

代入求解即可，解得

$$M=\begin{pmatrix}1&0&0&1\\2&0&0&2\\3&0&0&3\\4&0&0&4\end{pmatrix}$$

2. 求解 f 的值域 R(f) 和核空间的基及维数

对于值域，根据第 1 问，选择 M 中的列向量组，可以得到基为

$$\begin{pmatrix}1&2\\3&4\end{pmatrix}$$

维数为 1。

对于核空间，则转为求解 MX=0（必要的情况下可以进行**行变换**），显然存在下面的解

$$\xi_1=(1,0,0,-1)^T\quad\xi_2=(0,1,0,0)^T\quad\xi_3=(0,0,1,0)^T$$

当然这个解不是我们所需要的，需要写成矩阵的形式，即 $E_{11}-E_{22}$，E_{12}，E_{21}，维数为3。

3. 求解 f 的特征值及相应的特征子空间的基

直接求 $|\lambda I-M|$ 即可，需要注意的是**不要对进行初等行变换后的 M 求特征值**

$$|\lambda I-M|=\begin{vmatrix}\lambda-1&0&0&-1\\-2&\lambda&0&-2\\-3&0&\lambda&-3\\-4&0&0&\lambda-4\end{vmatrix}=\lambda^3(\lambda-5)$$

所以特征值为 0（3重）、5

再求特征子空间，先对特征值为 0 求子空间，即求 $MX=\lambda X=0$，由 2 可得基为 $E_{11}-E_{22}$，E_{12}，E_{21}（注意是矩阵形式）

对特征值为 5 求子空间，即求 $(M-5I)X=0$，解得 $\xi=(1,2,3,4)^T$，所以基为 $E_{11}+2E_{12}+3E_{21}+4E_{22}$

> **Caution**
>
> 在求解特征值时特征值上的幂次为代数重数，**不代表**特征子空间的基的维度（几何重数），几何重数代表了最小多项式的特征值幂次

4. 是否存在 $C^{2\times 2}$ 的基, 使得 f 在这组基下的矩阵为对角阵? 为什么?

假设一组基 $\beta_1,\beta_2,\beta_3,\beta_4$ 可以使得 f 在这组基下的矩阵为对角阵 Λ，假设 $E_{11},E_{12},E_{21},E_{22}$ 到 $\beta_1,\beta_2,\beta_3,\beta_4$ 的过渡矩阵为 P，即 $(\beta_1,\beta_2,\beta_3,\beta_4)=(E_{11},E_{12},E_{21},E_{22})P$，有

$$
\begin{aligned}
&f\left(\beta_1,\beta_2,\beta_3,\beta_4\right)=f\left(E_{11},E_{12},E_{21},E_{22}\right)P\\
&(\beta_1,\beta_2,\beta_3,\beta_4)\Lambda=(E_{11},E_{12},E_{21},E_{22})MP\\
&(E_{11},E_{12},E_{21},E_{22})P\Lambda=(E_{11},E_{12},E_{21},E_{22})MP\\
&MP=P\Lambda
\end{aligned}
$$

与特征值的矩阵表达形式相对比，可以看到 P 实际为特征函数构成的矩阵，而 Λ 则是对角线元素为对应特征值的对角矩阵。

本题可以转化为 M 是否可以相似对角化，因为 f 的各特征值的几何重数与代数重数相等，所以可以相似对角化。事实上，在$E_{11}-E_{22}$，E_{12}，E_{21}，$E_{11}+2E_{12}+3E_{21}+4E_{22}$ 上的矩阵为 diag(5, 0, 0,0)

如果特征向量组成的矩阵不可逆，假设特征向量的基的维度为3，那么对于 f 的特征向量，假设为 4 维的特征向量，这四个特征向量必定相关（多的能被少的表示，则多的线性相关），所以不会存在基使得 f 在这组基下的矩阵为对角阵。

> **Caution**
>
> 注意不能通过 f 的矩阵是否可逆（或者矩阵的秩）来判断 f 是否可以相似对角化，如果 f 为实对称阵，那么 f 的矩阵可以相似于对角阵

> 证明 $C^{2\times 2}=R(f)+K(f)$

先求出 R(f) 和 K(f) 的基，将基按行排列得到矩阵，通过矩阵判断这些基是否线性无关，如果线性无关，则成立

> 证明 $C^{2\times 2}=R(f)\oplus K(f)$

先判断 R(f) 和 K(f) 的维度之和是否等于 $C^{2\times 2}$ 的维度，再判断 $R(f)$ 和 $K(f)$ 的基是否线性无关

> 定义 $C^{2\times 2}$ 上的线性变换 f 为对任意 $X\in C^{2\times 2}$，$f(X)=AX$，其中 A 为
>
> $$A=\begin{pmatrix}1&2\\3&4\end{pmatrix}$$

给出 $C^{2\times 2}$ 的两个 2 维不变子空间 V_1 和 V_2 使得 $C^{2\times 2}=V_1\oplus V_2$

首先假设本题已经得到如下结论，在基 $E_{11},E_{21},E_{12},E_{22}$ 下的矩阵为

$$\begin{pmatrix} 1 & 2 & 0 & 0 \\ 3 & 4 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 3 & 4 \end{pmatrix}$$

观察该矩阵可以得到 E_{11} 和 E_{21} 张成的空间 V_1 以及 E_{12} 和 E_{22} 张成的空间 V_2 满足条件，假设 $\alpha = aE_{11} + bE_{21} \in V_1$，则显然 $f(\alpha) \in V_1$，同时该矩阵为 4 阶矩阵，AX=0 只有零解，因此 K(f) 为 0，dimR(f)=4，V_1 和 V_2 正好构成了 R(f)。

> 设 f 是 V 上的线性变换，证明：$V = K(f) \oplus K(I - f) \Leftrightarrow f^2 = f$

仅给出 $K(f) \oplus K(I - f) \Rightarrow f^2 = f$ 的证明

对于 $\alpha \in K(f)$ 和 $\beta \in K(I - f)$，有 $\gamma = k\alpha + l\beta \in V$

$$\begin{aligned} f(\gamma) &= f(k\alpha + l\beta) = kf(\alpha) + lf(\beta) = l\beta \\ f^2(\gamma) &= f^2(k\alpha + l\beta) = f(l\beta) = l\beta = f(\gamma) \end{aligned}$$

Jordan 标准形

> 求 $(2I + A)^{100}$ 的 Jordan 标准形

先求出 A 的 Jordan 标准形 J，则问题转换为求 $(2I + J)^{100}$

> 求子空间 $V = \{X|AX = XA\}$ 维数

先求 A 的 Jordan 标准形 J，$P^{-1}AP = J$，
$AX = XA \Rightarrow PJP^{-1}X = XPJP^{-1} \Rightarrow JP^{-1}XP = P^{-1}XPJ$
令 $Y = P^{-1}XP$，根据 Y 求出来的基的维数和 X 求出来的基的维数相等。

与正定相关的证明

> 假设 α 和 β 为两个 n 维相互正交的单位列向量，实数 p，q 均小于 1，证明矩阵 $A = I - p\alpha\alpha^H - q\beta\beta^H$ 正定

首先 α 和 β 是两个相互正交的单位列向量，因此可以将其扩充为一组标准正交基，$\alpha, \beta, w_1, \cdots, w_{n-2}$，

对于任意的 n 维非零列向量 $\mathbf{x}$，设 $\mathbf{x} = x_1\alpha + x_2\beta + x_3w_1 + \cdots + x_nw_{n-2}$

则有 $\mathbf{x}^H A\mathbf{x} = \mathbf{x}^H\mathbf{x} - p|x_1|^2 - q|x_2|^2 > 0$

> 设 A, B 是阶数相同的 Hermite 阵，且 A 是正定的，若 $A^{-1}B$ 的特征值均大于 -1，证明 A+B 正定

因为 A 正定，所以存在可逆阵 P （主对角元恒正的上三角阵）使得 $A=P^HP$，注意到 $A+B=P^HP+B=P^H\left(I+P^{-H}BP^{-1}\right)P$，又

有 $PA^{-1}BP^{-1}=P\left(P^HP\right)^{-1}BP^{-1}=PP^{-1}P^{-H}BP^{-1}=P^{-H}BP^{-1}$，因此 $A^{-1}B$ 与 B 相似，所以有相同

的特征值，所以 B 的特征值均大于 -1，因此 $P^H\left(I+P^{-H}BP^{-1}\right)P$ 的特征值均大于 0，所以 A + B 正定。

> 已知 A，B 都是 n 阶 Hermite 阵，且 A 是正定的，设 AB 的特征值均为 1，证明 $AB=I$

因为 A 是正定的，所以 $A=P^HP$ （P为可逆阵），则有 $P^{-H}ABP^H=P^{-H}P^HPBP^H=PBP^H$，因此 AB 与 PBP^H 相似，因为 PBP^H 为 Hermite 阵，相似于对角阵，AB 也会相似于对角阵，因此 AB 相似于 I，即存在可逆阵 Q 使得 $Q^{-1}(AB)Q=I\Rightarrow AB=I$

> 已知 n 阶 Hermite 阵 A 是正定的，B 是 n×s 矩阵，证明矩阵 B^HAB 是正定的当且仅当 B 的秩为 s

$r(B)=s\Rightarrow x^HB^HABx>0$：因为 B 的秩为 s，所以当 x 不为 0 时，$Bx\neq 0$，$x^HB^HABx>0$

$x^HB^HABx>0\Rightarrow r(B)=s$：因为 B^HAB 正定，所以 $r(B^HAB)=s$，$s=r(B^HAB)\leq r(B)\leq s\Rightarrow r(B)=s$

> 已知 A，B 是 n 阶正定矩阵，并且矩阵方程 AX + XA = B 有唯一解。如果矩阵 C 是这个矩阵方程的解，证明：C 也是正定矩阵。

因为 A，B 是正定矩阵，则有 $A^H=A$，$B^H=B$，由 $AX+XA=B\Rightarrow X^HA^H+A^HX^H=B^H$ 可得

$$X^HA+AX^H=B$$

因为是唯一解，且 C 是矩阵方程的解，所以有 $C^H=C$，令 η 为对应 λ 的特征向量

$$\begin{aligned}&\eta^HB\eta=\eta^H(AC+CA)\eta=\eta^HAC\eta+\eta^HC^HA\eta\\&=\lambda\eta^HA\eta+\lambda\eta^HA\eta=2\lambda\eta^HA\eta>0\end{aligned}$$

其中 $\eta^HA\eta>0$，所以 $\lambda>0$，因此 C 的特征值均大于 0，即 C 是正定矩阵。

> 设 A 是 n 阶 Hermite 矩阵，α 是 n 维列向量，且 $\alpha\alpha^H < 1$。证明 $I - \alpha\alpha^H$ 是正定阵，且 $A - A\alpha\alpha^H$ 相似于实对角阵。

首先易证 $I - \alpha\alpha^H$ 是正定阵，所以 $I - \alpha\alpha^H = P^H P$，P 为可逆阵，则有

$$A - A\alpha\alpha^H \sim PA(I - \alpha\alpha^H)P^{-1} = PAP^H$$

由 $\left(PAP^H\right)^H = PAP^H$ 可得 PAP^H 为 Hermite 阵，所以 $A - A\alpha\alpha^H$ 相似于实对角阵。

存在唯一

> 设 Hermite 矩阵 A 是正定的， m 是正整数，证明存在唯一正定矩阵 B 使得 $A = B^m$

因为 A 为 n 阶 Hermite 阵，存在酉矩阵 U，使得 $U^H AU = diag(\lambda_1, \lambda_2, \cdots, \lambda_n)$

（存在性）设 $B = Udiag\left(\sqrt[m]{\lambda_1}, \sqrt[m]{\lambda_2}, \cdots, \sqrt[m]{\lambda_n}\right)U^H$，$B^m = A$

（唯一性）设正定矩阵 B 和 C 满足 $A = B^m = C^m$，下面给出自己的证明

令 $B = U_1\Lambda_1U_1^H \quad C = U_2\Lambda_2U_2^H$，下证两者相等，由 $B^m = U_1\Lambda_1^mU_1^H = U_2\Lambda_2^mU_2^H = C^m$ 可得 $U_2^HU_1\Lambda_1^m = \Lambda_2^mU_2^HU_1$，设 $U_2^HU_1 = (u_{ij})$，可得 $\Lambda_1 = \Lambda_2$，进而有 $B = U_1\Lambda_1U_1^H = U_2\Lambda_2U_2^H = C$

（下面这个证明似乎存在问题，不过有一些启发性）
下证 $U_1 = U_2$：由 $U_1U_1^H = U_2U_2^H$ 可得 $U_1^HU_2 = U_1^HU_2$，首先 $U_1^HU_2$ 仍为酉矩阵，所以存在酉矩阵 P 使得 $P^HU_1^HU_2P = diag(\lambda_1, \cdots, \lambda_n)$，这里的 λ_i 不为0（因为 $U_1^HAU_1 = \Lambda_1^m \Rightarrow AU_1U_2^H = U_1\Lambda_1^mU_2^H$，因此 $U_1U_2^H$ 的行列式不为 0），因此有

$$P^H\left(U_1^HU_2\right)^{-1}P = diag\left(\frac{1}{\lambda_1}, \cdots, \frac{1}{\lambda_n}\right) = P^HU_2^HU_1P = diag(\overline{\lambda_1}, \cdots, \overline{\lambda_n})$$

因此 $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 1$，因此 $U_1^HU_2 = I$，即 $U_1 = U_2$。

正规阵

> 已知 A 是 n 阶正规矩阵，并且 A 有 n 个互异的特征值。如果矩阵 B 与 A 可交换，即 AB = BA，证明：B 也是正规矩阵。

首先 A 是 n 阶正规矩阵，则存在酉矩阵 U 使得 $U^HAU = \Lambda$，Λ 的对角线的元素不同。由 AB = BA 可得

$$AB = U\Lambda U^HB = BU\Lambda U^H = BA \Rightarrow \Lambda U^HBU = U^HBU\Lambda$$

令 $U^HBU = (c_{ij})$，则有

$$\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

即 $\lambda_i c_{ij} = \lambda_j c_{ij}$，由于 $\lambda_i \neq \lambda_j (i \neq j)$，所以 $c_{ij}(i \neq j) = 0$，所以 $U^H BU = \Lambda_2$，即 B 也是正规矩阵。

与正交向量组相关的证明

> 设 A 是 n 阶 Hermite 阵，λ_0 是 A 的最大特征值，集合 $S = \{x \in C^n | \parallel x \parallel = 1\}$，其中 $||x||$ 表示通常内积下向量 x 的长度，证明：$\lambda_0 = \max_{x \in S} x^H Ax$

设 A 是 n 阶 Hermite 阵，所以 A 的特征值全为实数，设 A 的 n 个特征值由大到小为 $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$，$\xi_1, \xi_2, \cdots, \xi_n$ 为标准正交向量组，而且 $A\xi_1 = \lambda_1 \xi_1, \cdots, A\xi_n = \lambda_n \xi_n$

对于任意 $x \in S$，设 $x = k_1\xi_1 + \cdots + k_n\xi_n$，则 $|k_1|^2 + \cdots + |k_n|^2 = x^H x = \parallel x \parallel = 1$

$$\begin{aligned} & x^H Ax = k_1 x^H A\xi_1 + \cdots + k_n x^H A\xi_n \\ & = \lambda_1 k_1 x^H \xi_1 + \cdots + \lambda_n k_n x^H \xi_n = \lambda_1 |k_1|^2 + \cdots + \lambda_n |k_n|^2 \\ & \leq \lambda_1 \left(|k_1|^2 + \cdots + |k_n|^2 \right) = \lambda_1 = \lambda_0 \end{aligned}$$

所以，$\lambda_0 = \max_{x \in S} x^H Ax$

范数相关

> 设 A是n阶非零方阵，证明：$\|A\|_F = \|A\|_2$ 的充分必要条件是 $r(A) = 1$

因为 A 是 n 阶非零方阵，所以 $x^H A^H Ax = (Ax)^H Ax \geq 0$，即 $A^H A$ 半正定，所以 $A^H A$ 的特征值只能大于或等于 0，因为 $\|A\|_F = \|A\|_2 \Rightarrow \sqrt{tr(A^H A)} = \sqrt{\rho(A^H A)}$，所以只能有一个非零特征值，所以 AA^H 的秩为 1，$r(A) = r(AA^H) = 1$（$AX = 0$ 和 $A^H AX = 0$ 同解）。