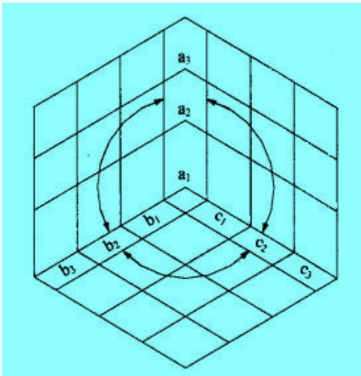
矩阵的相关性质

Lin Wenjie

Nanjing Normal University

版本：1.0

更新：2025 年 10 月 13 日



1 Hermite 矩阵的酉对角化（谱定理）

1.1 基本定义

定义 1.1 (Hermite 矩阵) 设 $A \in \mathbb{C}^{n\times n}$ 为复矩阵，若其共轭转置等于自身，即

$$AA^\dagger = A$$

其中 $A^\dagger = \overline{A}^T$（$\bar{\cdot}$ 表示复共轭，$\cdot^T$ 表示转置），则称 A 为 Hermite 矩阵。

定义 1.2 (酉矩阵) 设 $U \in \mathbb{C}^{n\times n}$ 为复矩阵，若满足

$$U^\dagger U = I$$

其中 I 为 $n \times n$ 单位矩阵，则称 U 为酉矩阵。酉矩阵全体记为 $\mathcal{U}_n$ 酉矩阵的列向量构成 $\mathbb{C}^n$ 的标准正交基（即列向量间内积为 0，自身内积为 1）。

注 1.1 设 $U \in \mathbb{C}^{m\times n}$，若满足

$$U^\dagger U = I$$

我们记满足上述性质的全体 $U \in \mathbb{C}^{m\times n}$ 为 $\mathrm{St}(m, n)$(Stiefel 流形), 其中 $m \geq n$，相当于一种非方阵的酉矩阵. 注意这里一定要 $m \geq n$，否则 $U^\dagger U = I$ 不可能成立。

1.2 关键性质

引理 1.1 (Hermite 矩阵的特征值为实数) 设 A 为 Hermite 矩阵，λ 为其任一特征值，$x \neq 0$ 为对应的特征向量（即 $Ax = \lambda x$），则 $\lambda \in \mathbb{R}$。

证明. 对等式 $Ax = \lambda x$ 两边取共轭转置，得：

$$x^\dagger A^\dagger = \overline{\lambda} x^\dagger$$

因 A 是 Hermite 矩阵（$A^\dagger = A$），故：

$$x^\dagger A = \overline{\lambda} x^\dagger$$

两边右乘 x，结合 $Ax = \lambda x$，得：

$$x^\dagger A x = \overline{\lambda} x^\dagger x \quad \text{且} \quad x^\dagger A x = \lambda x^\dagger x$$

因此 $\lambda x^\dagger x = \overline{\lambda} x^\dagger x$。由于 $x \neq 0$，内积 $x^\dagger x = |x_1|^2 + \cdots + |x_n|^2 > 0$，故 $\lambda = \overline{\lambda}$，即 λ 为实数。

引理 1.2 (不同特征值的特征向量正交) 设 A 为 Hermite 矩阵，$\lambda \neq \mu$ 为其两个互异特征值，x, y 为对应的特征向量（$Ax = \lambda x$，$Ay = \mu y$），则 x 与 y 正交，即 $x^\dagger y = 0$。

证明. 由 $Ax = \lambda x$，两边左乘 $y^\dagger$ 得：

$$y^\dagger A x = \lambda y^\dagger x$$

左边可改写为 $(Ay)^\dagger x$（因 $y^\dagger A = (Ay)^\dagger$），结合 $Ay = \mu y$ 及 $\mu \in \mathbb{R}$（引理 1），得：

$$(Ay)^\dagger x = (\mu y)^\dagger x = \mu y^\dagger x$$

因此 $\mu y^\dagger x = \lambda y^\dagger x$，即 $(\mu - \lambda) y^\dagger x = 0$。因 $\lambda \neq \mu$，故 $y^\dagger x = 0$，即 x 与 y 正交。

注 1.2 一般的复矩阵只能得到不同特征值对应的特征向量线性无关，而 Hermite 矩阵可以得到相互正交，这是一个更强的结论 (因为正交一定线性无关)。

引理 1.3 (特征向量的正交补空间是不变子空间) 设 A 为 Hermite 矩阵，λ 为其特征值，x 为单位特征向量（$x^\dagger x = 1$），记 $M = \mathrm{span}\{x\}$，其正交补空间为：

$$M^\perp = \{y \in \mathbb{C}^n \mid y^\dagger x = 0\}$$

则 $M^\perp$ 是 A 的不变子空间（即 $y \in M^\perp \implies Ay \in M^\perp$）。

证明. 对任意 $y \in M^\perp$，需证 $(Ay)^\dagger x = 0$。由 $Ax = \lambda x$ 及 $\lambda \in \mathbb{R}$，得 $A^\dagger x = x^\dagger A^\dagger = \overline{\lambda} x^\dagger = \lambda x^\dagger$（因 $A^\dagger = A$）。因此：

$$(Ay)^\dagger x = y^\dagger A^\dagger x = y^\dagger (\lambda x^\dagger) = \lambda y^\dagger x = 0 \quad (\text{因} y \in M^\perp)$$

故 $Ay \in M^\perp$，即 $M^\perp$ 是不变子空间。

1.3 谱定理的证明

定理 1.4 (Hermite 矩阵的谱定理 (Hermite 矩阵可以对角化)) 对任意 $n \times n$ Hermite 矩阵 A，存在 $n \times n$ 酉矩阵 U 和实对角矩阵 Λ，使得：

$$U^\dagger AU = \Lambda$$

其中 Λ 的对角元为 A 的特征值，U 的列向量为对应的特征向量（标准正交）。

证明. 用数学归纳法证明：

- 基础步（$n = 1$）：
  1 阶 Hermite 矩阵为实数 a（满足 $\overline{a} = a$）。取酉矩阵 $U = [1]$，则 $U^\dagger AU = [a]$ 为对角矩阵，结论成立。
- 归纳步（假设 $n-1$ 阶成立，证明 n 阶成立）：
  设 A 为 $n \times n$ Hermite 矩阵，由代数基本定理，A 存在特征值 $\lambda_1 \in \mathbb{R}$（定理 1.1）和单位特征向量 $x_1 \in \mathbb{C}^n$（${x_1}^\dagger x_1 = 1$）。
  记 $M = \text{span}\{x_1\}$，其正交补 $M^\perp$ 为 $n-1$ 维子空间（定理 1.3）。A **在** $M^\perp$ **上的限制** $A|_{M^\perp}$ **是** $n-1$ **阶 Hermite 矩阵（因** $M^\perp$ **是不变子空间且** A **对称）**。
  由归纳假设，存在 $n-1$ 阶酉矩阵 V（列向量为 $M^\perp$ 的标准正交基），使得：

$$V^\dagger (A|_{M^\perp})V = \Lambda' \quad (\Lambda'\text{为 } n-1 \times n-1 \text{实对角矩阵})$$

  构造 $n \times n$ 矩阵 $U = [x_1 \quad V']$，其中 V' 是将 V 的列嵌入 $\mathbb{C}^n$ 得到的矩阵（列向量属于 $M^\perp$）。因 x_1 与 V' 的列正交且均为单位向量，故 U 是酉矩阵（$U^\dagger U = I$）。
  计算 $U^\dagger AU$：

$$U^\dagger AU = \begin{pmatrix} {x_1}^\dagger \\ V'^\dagger \end{pmatrix} A \begin{pmatrix} x_1 & V' \end{pmatrix} = \begin{pmatrix} {x_1}^\dagger Ax_1 & {x_1}^\dagger AV' \\ V'^\dagger Ax_1 & V'^\dagger AV' \end{pmatrix}$$

  - 左上角：${x_1}^\dagger Ax_1 = \lambda_1 {x_1}^\dagger x_1 = \lambda_1$；- 右上角与左下角：因 $AV' \subset M^\perp$，故 ${x_1}^\dagger AV' = 0$，同理 $V'^\dagger Ax_1 = 0$；- 右下角：$V'^\dagger AV' = V^\dagger (A|_{M^\perp})V = \Lambda'$。
  因此：

$$U^\dagger AU = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \Lambda' \end{pmatrix} = \Lambda$$

  其中 Λ 为实对角矩阵。由归纳法，结论对所有 $n \geq 1$ 成立。

2 复矩阵的奇异值分解（SVD）

定理 2.1 设 $A \in \mathbb{C}_r^{m\times n}$，则存在 $U \in \mathcal{U}_m$ 和 $V \in \mathcal{U}_n$，使得

$$U^\dagger AV = \begin{pmatrix} \Sigma_r & 0 \\ 0 & 0 \end{pmatrix}$$

其中 $\Sigma_r = \mathrm{diag}(\sigma_1, \ldots, \sigma_r)$，$\sigma_1 \geq \cdots \geq \sigma_r > 0$。

证明. 由于 $A^\dagger A$ 是半正定的 Hermite 矩阵，且 $\mathrm{rank}(A^\dagger A) = \mathrm{rank}(A)$，则根据定理 1.4, 有

$$V^\dagger(A^\dagger A)V = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_n^2)$$

其中 $V = [v_1, \ldots, v_n] \in \mathcal{U}_n$，$\sigma_1 \geq \cdots \geq \sigma_r > 0$，$\sigma_{r+1} = \cdots = \sigma_n = 0$。

令 $V_1 = [v_1, \ldots, v_r]$，$V_2 = [v_{r+1}, \ldots, v_n]$，$\Sigma_r = \mathrm{diag}(\sigma_1, \ldots, \sigma_r)$。则

$$\begin{aligned} {V_1}^\dagger(A^\dagger A)V_1 &= \Sigma_r^2, \\ {V_2}^\dagger(A^\dagger A)V_2 &= 0, \end{aligned}$$

上式表明 AV_1 的列是相互正交的, 而下式表明 AV_2 的列都是零向量，即 $AV_2 = 0$。因此，令

$$U_1 = AV_1\Sigma_r^{-1}$$

则 ${U_1}^\dagger U_1 = I_r$。再取 $U_2 \in \mathbb{C}^{m\times(m-r)}$ 使 $U = [U_1, U_2] \in \mathcal{U}_m$（$\mathbb{C}^m$ 中的任何一组正交向量都可扩展成全空间的一组正交基），则有

$$U^\dagger AV = \begin{pmatrix} U_1^\dagger AV_1 & U_1^\dagger AV_2 \\ U_2^\dagger AV_1 & U_2^\dagger AV_2 \end{pmatrix} = \begin{pmatrix} \Sigma_r & 0 \\ 0 & 0 \end{pmatrix}.$$

注 2.1 考察一下奇异值的几何意义。考虑一个简单的例子。设

$$A = [u_1, u_2] \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} [v_1, v_2]^T,$$

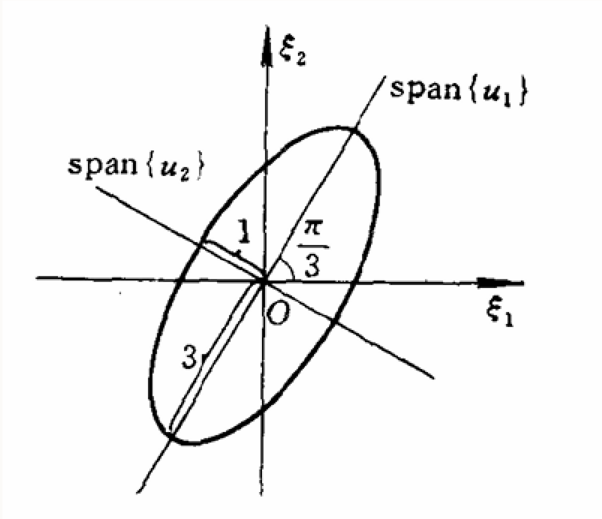
其中 $u_1 = \left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right)^T, u_2 = \left(\frac{\sqrt{3}}{2}, -\frac{1}{2}\right)^T, v_1 = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)^T, v_2 = \left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right)^T$。那么对任意的 $x = \xi_1 v_1 + \xi_2 v_2 \in \mathbb{R}^2$，有

$$y = Ax = \eta_1 u_1 + \eta_2 u_2,$$

其中 $\eta_1 = 3\xi_1$，$\eta_2 = \xi_2$。因此，如果 $\|x\|_2 = 1$，即 $\xi_1^2 + \xi_2^2 = 1$，则对应的 $y = \eta_1 u_1 + \eta_2 u_2$ 满足

$$\frac{\eta_1^2}{3^2} + \eta_2^2 = 1.$$

这表明 A 将 $\mathbb{R}^2$ 中的单位圆 $\{x\in\mathbb{R}^2:\|x\|_2=1\}$ 变成了椭圆 $E_2=\{y=Ax:\|x\|_2=1\}$，而两个奇异值正好是这一椭圆的两个半轴长；长轴所在的直线是 $\text{span}\{u_1\}$，短轴所在的直线是 $\text{span}\{u_2\}$，它们分别是 $\text{span}\{v_1\}$ 和 $\text{span}\{v_2\}$ 的像（如下图所示）。



对于一般的情形，不妨设 $m=n$，我们亦有

$$E_n=\{y\in\mathbb{C}^n:y=Ax,\ x\in\mathbb{C}^n,\ \|x\|_2=1\}$$

是一个超椭球面，它的 n 个半轴长正好是 A 的 n 个奇异值 $\sigma_1\geq\sigma_2\geq\cdots\geq\sigma_s\geq 0$，这些轴所在的直线正好是 A 的左奇异向量所在的直线，它们分别是对应的右奇异向量所在直线的像。

注 2.2 对于 SVD 分解 $A=U\begin{pmatrix}\Sigma_r & 0\\ 0 & 0\end{pmatrix}V^\dagger$, 可以要进一步写成：

$$A=\tilde{U}\Sigma_r\tilde{V^\dagger}(SVD),$$

(即截断的 SVD 分解) 其中 $\tilde{U}\in\text{St}m,r,\tilde{V}\in\text{St}n,r$, 于是 Ax 相当于先对 x 做一个 $\mathbb{C}^n\to\mathbb{C}^r$ 的保范映射，再在 $\mathbb{C}^r$ 中做一个缩放变换 (Σ_r), 最后再做一个 $\mathbb{C}^r\to\mathbb{C}^m$ 的保范映射.

3 复矩阵的 QR 分解

定义 3.1 (QR 分解) 设 A 是一个 $m\times n$ 复矩阵，且 $m\geq n$。如果存在一个 $m\times r$ 酉矩阵 Q 和一个 $r\times n$ 上梯形矩阵 R，使得

$$A=QR$$

则称此分解为 A 的 QR 分解, 其中 $r=\text{rank}(A)$。

定理 3.1 (列满秩矩阵 QR 分解的存在性与唯一性) 任意列满秩复矩阵 $A\in\mathbb{C}^{m\times n}$ 都存在一个 $m\times n$ 酉矩阵 Q 和一个 $n\times n$ 上三角矩阵 R，使得

$$A=QR$$

。如果进一步要求 R 的对角元为正实数，则该分解是唯一的。

证明. **存在性的证明：**

我们使用 Gram-Schmidt 正交化过程来构造证明。设 $A = [a_1, a_2, \ldots, a_n]$，其中 $a_i \in \mathbb{C}^m$ 是 A 的列向量。

(1) 正交化过程

定义：

- $u_1 = a_1$
- $q_1 = \dfrac{u_1}{\|u_1\|}$
- 对于 $k = 2, 3, \ldots, n$：

$$u_k = a_k - \sum_{j=1}^{k-1} \langle a_k, q_j \rangle q_j$$

$$q_k = \frac{u_k}{\|u_k\|} \quad (\text{如果} u_k \neq 0)$$

其中 $\langle x, y \rangle = y^\dagger x$ 是复向量空间中的内积。

(2) 构造 QR 分解

令 $Q = [q_1, q_2, \ldots, q_n]$，则 $Q^\dagger Q = I_n$。

定义上三角矩阵 R 的元素为：

$$r_{ij} = \begin{cases} \langle a_j, q_i \rangle, & \text{如果} i \leq j \\ 0, & \text{如果} i > j \end{cases}$$

特别地，$r_{kk} = \|u_k\|$。

(3) 验证分解

对于每个 $k = 1, 2, \ldots, n$，有：

$$a_k = \sum_{i=1}^{k} r_{ik} q_i = \sum_{i=1}^{k} \langle a_k, q_i \rangle q_i$$

因此：

$$A = [a_1, a_2, \ldots, a_n] = [q_1, q_2, \ldots, q_n] R = QR$$

唯一性的证明：

假设有两个不同的分解 $A = Q_1 R_1 = Q_2 R_2$，那么：

$$Q_2^* Q_1 = R_2 R_1^{-1}$$

- 左边是酉矩阵的乘积，仍然是酉矩阵;
- 右边是两个上三角矩阵的乘积，仍然是上三角矩阵;

- 一个既是酉矩阵又是上三角矩阵的矩阵必须是对角矩阵;
- 由于 R_1, R_2 对角线都是正实数，且乘积为单位矩阵，这个对角矩阵只能是单位矩阵

因此 $Q_1 = Q_2, R_1 = R_2$

定理 3.2 (一般的复矩阵的 QR 分解) 对于任意复矩阵 $A \in \mathbb{C}^{m\times n}$，存在一个 $m \times r$ 酉矩阵 Q , 一个 $r \times n$ 上梯形矩阵 R, 一个置换矩阵 $P \in \mathbb{C}^{n\times n}$，使得

$$AP = QR,$$

其中，$r = \operatorname{rank}(A) \leq n$.

证明. 我们用置换矩阵 $P \in \mathbb{C}^{n\times n}$ 作用于 A 的列向量组，使得 AP 的前 r 个向量线性无关。不妨设 $AP = [a_1, a_2, \ldots, a_n]$, 则 $a_1, \ldots, a_r$ 线性无关。

设 $[a_1, \ldots, a_r] = A_r \in \mathbb{C}^{m\times r}$ 列满秩，则根据上一个定理，存在一个 $m \times r$ 酉矩阵 Q_r 和一个 $r \times r$ 上三角矩阵 R_r，使得

$$A_r = Q_r R_r$$

对于 $k = r + 1, \ldots, n$, 存在 $x_k \in \mathbb{C}^r$ 使得 $a_k = A_r x_k = Q_r R_r x_k$, 记 $y_k = R_r x_k \in \mathbb{C}^r$, 令 $R'_r = [y_{r+1}, \ldots, y_n] \in \mathbb{C}^{r\times(n-r)}$, 则

$$AP = Q_r \begin{bmatrix} R_r & R'_r \end{bmatrix} := QR,$$

其中 R 为上梯形矩阵。

注 3.1 上面的定理可以进一步扩充：Q 的列向量可以扩充为一组单位正交基得到 $\overline{Q} = \begin{bmatrix} Q & Q_\perp \end{bmatrix}$ 为 $m \times m$ 酉矩阵，于是相应的 R 变成了

$$\overline{R} = \begin{bmatrix} R \\ 0 \end{bmatrix}$$

因此，

$$A = \begin{bmatrix} Q & Q_\perp \end{bmatrix} \begin{bmatrix} R \\ 0 \end{bmatrix},$$

这成为完全 QR 分解，但是这种分解是不唯一的：

例 3.1 (秩亏缺矩阵的 QR 分解) 考虑秩为 1 的矩阵：

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix}$$

一种可能的 QR 分解为：

$$Q=\begin{bmatrix}\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}}\\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}}\\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}}\end{bmatrix},\quad R=\begin{bmatrix}\sqrt{3} & 2\sqrt{3} & 3\sqrt{3}\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}$$

另一种分解为：

$$Q=\begin{bmatrix}\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & 0\\ \frac{1}{\sqrt{3}} & 0 & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}\end{bmatrix},\quad R=\begin{bmatrix}\sqrt{3} & 2\sqrt{3} & 3\sqrt{3}\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}$$

两者都满足 $A=QR$，说明分解不唯一。

注 3.2 完全 QR 分解不唯一的根本原因是：$Q_\perp$ 不唯一，其列向量可以随意换位置。

4 矩阵的 DR 分解

定义 4.1 (DR 分解) 设 A 是一个 $n\times n$ 复矩阵。如果存在一个对角矩阵 D 和一个酉矩阵 R，使得

$$A=DR$$

其中 D 的对角元都是正实数，R 满足 $R^\dagger R=I$，则称此分解为 A 的 DR 分解。

定理 4.1 (DR 分解的存在条件) 复矩阵 A 存在 DR 分解的充分必要条件是 $A^\dagger A$ 是对角矩阵。

证明. 必要性：若 $A=DR$，则

$$A^\dagger A=(DR)^\dagger(DR)=R^\dagger D^\dagger DR=R^\dagger D^2R$$

由于 D 是实对角矩阵，$D^\dagger=D$，且 R 是酉矩阵，故 $A^\dagger A$ 与对角矩阵 D^2 酉相似，因此 $A^\dagger A$ 是对角矩阵。

充分性：若 $A^\dagger A$ 是对角矩阵，记为 Λ，则 Λ 是半正定埃尔米特矩阵。令 $D=\sqrt{\Lambda}$，即 D 的对角元是 Λ 对角元的算术平方根。定义

$$R=AD^{-1}$$

验证 R 是酉矩阵：

$$R^\dagger R=(AD^{-1})^\dagger(AD^{-1})=D^{-1}A^\dagger AD^{-1}=D^{-1}\Lambda D^{-1}=D^{-1}D^2D^{-1}=I$$

因此 $A=DR$ 是 DR 分解。

定理 4.2 (DR 分解与奇异值分解的关系) 若 A 有 DR 分解 $A = DR$，则 D 的对角元是 A 的奇异值，R 的列是 A 的右奇异向量。

证明. 由 $A = DR$ 和 R 是酉矩阵，得 $A^\dagger A = R^\dagger D^2 R$，这表明 D^2 的特征值是 $A^\dagger A$ 的特征值，即 D 的对角元是 A 的奇异值，而 R 的列是 $A^\dagger A$ 的特征向量，即右奇异向量。

5 正交投影

定义 5.1 设 $\mathcal{X} \subset \mathbb{C}^n$ 是一个子空间。如果 $P \in \mathbb{C}^{n\times n}$ 满足：

(1) $\mathcal{R}(P) = \mathcal{X}$,

(2) $P^2 = P$,

(3) $P^\dagger = P$,

则称 P 是映射到 $\mathcal{X}$ 上的正交投影。

注 5.1 从定义可知：正交投影是一个 Hermite 幂等矩阵；对任意的 $x \in \mathbb{C}^n$，有 $Px \in \mathcal{X}$，$(I - P)x \in \mathcal{X}^\perp$；$(I - P)$ 是映射到 $\mathcal{X}^\perp$ 上的正交投影。

定理 5.1 对于任意的子空间 $\mathcal{X} \subset \mathbb{C}^n$，在 $\mathbb{C}^{n\times n}$ 中有且仅有一个映射到 $\mathcal{X}$ 上的正交投影。

证明.

- **存在性**
  设 $\mathcal{X} = \mathcal{R}(V)$，其中 $V \in \mathbb{C}^{n\times l}$ 满足 $V^*V = I$，即 V 的列构成 $\mathcal{X}$ 的一组标准正交基。令
  $$P = VV^\dagger$$
  则容易验证 P 是映射到 $\mathcal{X}$ 上的正交投影。
- **唯一性**
  设 P_1 也是映射到 $\mathcal{X}$ 上的正交投影。则对任意的 $x \in \mathbb{C}^n$，有
  $$\begin{aligned}\|Px - P_1x\|_2^2 &= x^\dagger P^\dagger Px - x^\dagger P^\dagger P_1x - x^\dagger P_1^\dagger Px + x^\dagger P_1^\dagger P_1x \\ &= x^\dagger Px - x^\dagger PP_1x - x^\dagger P_1Px + x^\dagger P_1x \\ &= x^\dagger(I - P_1^\dagger)Px + x^\dagger(I - P)P_1x \\ &= ((I - P_1)x)^\dagger Px + ((I - P)x)^\dagger P_1x \\ &= 0\end{aligned}$$
  从而 $P = P_1$.

注 5.2 对于给定的子空间 $\mathcal{X}$，通常用 $P_{\mathcal{X}}$ 来表示映射到它上的唯一的正交投影。尽管对一个子空间来说，它的标准正交基是不唯一的，但按上述定义的矩阵却是唯一的，与正交基的选取无关。