1 行列式

设 $A=[a_{ij}]$ 是 n 阶方阵，按一行或者一列展开：则

$$a_{1j}A_{1k}+a_{2j}A_{2k}+\cdots+a_{nj}A_{nk}=\delta_{jk}|A| \tag{1}$$

其中 $\delta_{jk}=0$ 如果 $k\neq j$, $\delta_{jj}=1$.

应用：设 A 的每行元素求和都为 3, 且 $|A|=8$, 求 $\sum_{i=1}^{n}\sum_{j=1}^{n}A_{ij}$.

定理 1.1 (克拉默法则)**.** *设 $A=[\alpha_1,\alpha_2,\cdots,\alpha_n]$ 为 n 阶方阵，考虑线性方程组 $AX=\beta$，若 $|A|\neq 0$, 则方程组有唯一解*

$$X=A^{-1}\beta$$

或者 $X=[x_1,x_2,\cdots,x_n]^T$ 满足

$$x_1=\frac{D_1}{D},\ x_2=\frac{D_2}{D},\ \cdots,x_n=\frac{D_n}{D}$$

其中 $D=|A|$ 和 $D_i=|[\alpha_1,\alpha_2,\cdots,\alpha_{i-1},\beta,\alpha_{i+1},\cdots,\alpha_n]|$ 对 $i=1,\cdots,n$.

应用：用于解 n 元 n 个方程的方程组的方程。

2 线性方程组

- 消元法解方程：$AX=\beta$
- 用初等行变换化增广矩阵为阶梯型：$\bar{A}=[A,\beta]$。
- 有解的判断方法：有解当且仅当 $r(A)=r(\bar{A})$, 当且仅当 β 可由 A 的列向量线性表示。
- 唯一解：$r(A)=r(\bar{A})=$ 未知量的个数。
- 消元法解方程一定要掌握，在后期的矩阵对角化、找特征向量、把一个向量扩充为一组正交基 P167 例 4.8.10 等。

3 矩阵计算

- 矩阵多项式：$f(x)=a_kx^k+a_{k-1}x^{k-1}+\cdots+a_1x+a_0$, 且 A 为 n 阶，则 $f(A)=a_kA^k+a_{k-1}A^{k-1}+\cdots+a_1A+a_0E$.
- 这要结合矩阵对角化来学习:若存在可逆 P,使得 $P^{-1}AP=diag\{\lambda_1,\cdots,\lambda_n\}$, 则

  $$P^{-1}f(A)P=diag\{f(\lambda_1),\cdots,f(\lambda_n)\}$$

  在计算 $f(A)$ 或者 $|f(A)|$ 时可以直接利用，例如 $|A+E|$ 的行列式。
- 矩阵行列式与矩阵乘法的关系：设 A,B 为 n 阶，则 $|AB|=|A||B|$. 证明可以不看，或者应用分块矩阵给出一个证明。

- 矩阵的迹：A 为 n 阶，定义 $trA=\sum_{i=1}^n a_{ii}$. 则 $tr(AB)=tr(BA)$ 对任意的 $A_{n\times m}$ 和 $B_{m\times n}$ 成立。

  可以用于证明：不存在 n 阶 A,B，使得 $AB-BA=E$.

- 注：$(A+B)^2=(A+B)(A+B)=A^2+BA+AB+B^2$.

- 设 α,β 为两个 n 维列向量，若 $A=\alpha\beta^T$, 则 $A^k=(\beta^T\alpha)^{k-1}A$, 其中 $\beta^T\alpha$ 是一个数。

- 设 $A=diag\{\lambda_1,\cdots,\lambda_n\}$ 且 $\lambda_i\neq\lambda_j$, 则与 A 可交换的矩阵一定是对角矩阵。

  分块矩阵类比：把 λ_i 换成 $\lambda_i E_{r_i}$, 则与 A 可交换的矩阵一定是分块对角矩阵。

- 可逆矩阵：$A^{-1}=\frac{1}{|A|}A^*$, 其中 A^* 为 A 的伴随矩阵。注：$A^*=[b_{ij}]$ 的元素满足 $b_{ij}=A_{ji}$, 而不是 A_{ij}.

  例：$A=\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ 则 $A^*=\begin{bmatrix} d & -b \\ -c & b \end{bmatrix}$ 和 $A^{-1}=\frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & b \end{bmatrix}$.

- 已知矩阵满足某些等式，求矩阵的逆：一般配出单位矩阵 E.
  例：$A^2-2A-3E=0$ 求 A^{-1}.
  解：$A(A-2E)=3E$ 可知 $\frac{1}{3}A(A-2E)=E$, 从而 $A^{-1}=\frac{1}{3}(A-2E)$.

  例：设 $A+B=AB$，求 $A-E$ 的逆。
  解：$A=(A-E)B$, 从而 $E+(A-E)=(A-E)B$, 所以 $E=(A-E)(B-E)$. 因此 $(A-E)^{-1}=B-E$.

- 分块矩阵：掌握 2×2 的分块，比如求 $P=\begin{bmatrix} A & 0 \\ C & B \end{bmatrix}$ 的逆（P96 例 3.3.5）

- 分块矩阵：掌握对角分块矩阵的性质：$A=diag\{A_1,A_2,\cdots,A_t\}$. 乘法，加法，数乘，行列式，求逆。

  应用：求矩阵的逆。例如 P97 例 3.3.6.

- 附录四：对任意矩阵 $A_{n\times m}, B_{m\times n}$, 则 $|E_n+AB|=|E_m+BA|$. 特别地，在计算行列式的时候可以直接应用，在计算特征多项式时也会用到。**例：若 A,B 是 n 阶方阵，证明 AB 和 BA 的特征多项式相等。**

- 初等矩阵：对单位矩阵施行一次初等变换的矩阵。初等矩阵都是可逆矩阵，逆矩阵也是初等矩阵。

  重要的定理：对矩阵施行一次初等行变换等价于左乘一个初等矩阵，施行一次列变换等价于右乘一个初等矩阵。

推论：若 $A_{m\times n}$ 的秩为 $r(A)=r$, 则存在可逆矩阵 P,Q 使得

$$PAQ=\left[\begin{array}{cc} E_r & 0 \\ 0 & 0 \end{array}\right]$$

推论：若 A 可逆，则存在一序列初等矩阵 $P_1,P_2,\cdots,P_k$ 使得

$$A=P_1P_2\cdots P_k.$$

推论:若 A 可逆,则 $A^{-1}C$, 可以通过如下的方式求得:$A^{-1}C=P_k^{-1}\cdots P_2^{-1}P_1^{-1}C$, 等价于：对分块矩阵 $[A,C]$ 做初等行变换把 A 化为 E, 则同时有 C 化为 $A^{-1}C$. P104-105.

应用：用于用初等变换求矩阵的逆或者求 $A^{-1}C$.

对于右乘 CA^{-1}, 则考虑列变换。

推论：若 A,B 可逆，则 $r(AC)=r(C)=r(CB)=r(ACB)$.

4 线性空间

- 线性空间: 有加法和数乘的集合。例如数域 P 上 $m\times n$ 矩阵组成的空间。次数小于 n 次的多项式全体组成的空间。
- 线性无关：一组向量如果表示零向量的表示法唯一，则称该向量组线性无关。
- 要会判断一组向量是线性相关或者线性无关：如果是 P^n 中的向量可以用矩阵的秩来判断。比如
  $\alpha_1,\cdots,\alpha_s\in P^n$ 线性相关当且仅当 $A=[\alpha_1,\cdots,\alpha_s]$ 满足 $r(A)<s$.
  $\alpha_1,\cdots,\alpha_s\in P^n$ 线性无关当且仅当 $A=[\alpha_1,\cdots,\alpha_s]$ 满足 $r(A)=s$.
  要理解这两个性质的证明，在具体的计算中，如果 $\alpha_1,\cdots,\alpha_s$ 是线性相关的，要找出它的一个极大无关组 (例如找一组向量的生成子空间的一组基)，此时可以对矩阵 $A=[\alpha_1,\cdots,\alpha_s]$ 施行初等**行变换**, 化为阶梯型 $B=[\alpha_1',\cdots,\alpha_s']$. 此时只要确定 B 中列的一个极大无关组，对应回 A 中的列就是向量组 $\alpha_1,\cdots,\alpha_s$ 的一个极大无关组。具体例子 P124 例 4.2.1, P127 9.

  对于一般的 n 维线性空间中的向量组 $\alpha_1,\cdots,\alpha_s$, 只要取定一组基 $e_1,\cdots,e_n$, 然后考虑向量在这组基下的坐标 $X_1,\cdots,X_s\in P^n$, 此时判断 $\alpha_1,\cdots,\alpha_s$ 的线性关系 (无关、相关)，等价于判断 $X_1,\cdots,X_s$ 的线性关系，因此要熟练掌握 P^n 中如何判断一组向量的线性关系。
- 4.3 节向量极大无关组：这一节的证明都需要掌握，特别是定理 4.3.2, 定理 4.3.3

- 极大无关组：向量组的一个线性无关的部分向量组，且与原向量组等价。要求：掌握如何求一个向量组的极大无关组。P130 例 4.3.1, P131 1.
- 极大无关组与原向量组等价。
- 定理：向量组 $\alpha_1, \cdots, \alpha_r$ 可以由 $\beta_1, \beta_2, \cdots, \beta_s$ 线性表示，若 $\alpha_1, \cdots, \alpha_r$ 线性无关，则 $r \leq s$.

  要掌握这个定理的各种推论：

  - 向量组 $\alpha_1, \cdots, \alpha_r$ 可以由 $\beta_1, \beta_2, \cdots, \beta_s$ 线性表示, 若 $r > s$, 则向量组 $\alpha_1, \cdots, \alpha_r$ 线性相关。例如：若 α_1, α_2 可以由 β 线性表示，则必有 α_1, α_2 线性相关。
  - 向量组 (I) 可由向量组 (II) 线性表示，则 秩$(I) \leq$ 秩(II)。
  - 线性无关组 $\alpha_1, \cdots, \alpha_r$ 与 $\beta_1, \beta_2, \cdots, \beta_s$ 等价，则 $r = s$.
  - 一个向量组的极大无关组的向量个数是唯一确定的。
  - 若线性无关向量 $\alpha_1, \cdots, \alpha_r$ 可以被 $\beta_1, \beta_2, \cdots, \beta_r$ 线性表示, 则这两组向量等价。
  - 若 $\alpha_1, \cdots, \alpha_n$ 是线性空间 V 的一组基,且 $\alpha_1, \cdots, \alpha_n$ 可以被 $\beta_1, \beta_2, \cdots, \beta_n$ 线性表示，则 $\beta_1, \beta_2, \cdots, \beta_n$ 为 V 的一组基。
- 证明一组向量 $\alpha_1, \cdots, \alpha_s$ 是线性无关的，一般是利用反证法：假设存在一组不全为零的数 k_i 使得 $k_1\alpha_1 + \cdots + k_s\alpha_s = \theta$. 然后利用题目已知的条件带入得到一个关于 $k_1, \cdots, k_s$ 的方程组，然后判断这组方程只有零解。
  或者直接假设某个 $k_i \neq 0$, 利用已知条件推出 $k_i = 0$, 得出矛盾。

  具体例子可以参考 P125 例 4.2.2, P127 10，或证明正交向量组是线性无关的。另外可以对比理解 P128 定理 4.3.3 的证明 (这个证明要求掌握。)
- 属于 P 上的线性空间的基：$e_1, \cdots, e_n \in V$ 在数域 P 上线性无关，且 V 中任一向量可由这组向量表示。称数域 P 上的线性空间的维数为 n.
  - $\mathbb{C}$ 是实数域上的 2 维线性空间，基为：$i, 1$.
  - $\mathbb{C}$ 是复数域上的 1 维线性空间，基为：1.
  - $\mathbb{R}_+ = \{x > 0, x \in \mathbb{R}\}$ 按如下运算是实数域上的线性空间，维数为 1.
    加法：$a \otimes b = ab$. 数乘：$k \cdot a = a^k$, 其中 $a, b \in \mathbb{R}_+$, $k \in \mathbb{R}$.
    该线性空间的零元素为 1, 任一不等于 1 的正实数都是该线性空间的基。
- 坐标表示：给定一组基 $e_1, \cdots, e_n$, 对 $\alpha \in V$ 且 $\alpha = x_1e_1 + \cdots + x_ne_n$，则称 $X = [x_1, \cdots, x_n]^T$ 是 α 在基 $e_1, \cdots, e_n$ 的坐标。**给定一组基后，线性空间向量组的线性无关、相关性可以通过相应的坐标在 P^n 中的无关与相关性来判断**

- 定理：给定 V 上一组基 $e_1, \cdots, e_n$, 设 $\alpha_1, \cdots, \alpha_s$ 为 V 上的一组向量，对应的坐标记为

$$X_i = [a_{1i}, a_{2i}, \cdots, a_{ni}]^T,\ i = 1, 2, \cdots, s$$

以坐标列向量构成 $n \times s$ 矩阵 $A = [X_1, X_2, \cdots, X_s]$，则

  - $\alpha_1, \cdots, \alpha_s$ 线性相关当且仅当 $r(A) < s$.
  - $\alpha_1, \cdots, \alpha_s$ 线性无关当且仅当 $r(A) = s$.

要熟练运用该定理，特别是判断向量组无关与相关问题上。比如如下的各种推论：

  - n 维线性空间的 $n+1$ 个向量必定线性相关。
  - 线性空间 V 中的向量组 $\beta_1, \cdots, \beta_s$ 线性无关，则 $\dim V \geq s$.
  - n 维线性空间的 n 个无关的向量是一组基。
  - 两组基的过渡矩阵是可逆矩阵。

- 过渡矩阵：设 (I): $e_1, e_2, \cdots, e_n$ 与 (II): $e'_1, e'_2, \cdots, e'_n$ 是 n 维线性空间 V 的两组基，$e'_i (1 \leq i \leq n)$ 经基 (I) 线性表示为

$$\begin{cases} e'_1 &= m_{11}e_1 + m_{21}e_2 + \cdots + m_{n1}e_n, \\ e'_2 &= m_{12}e_1 + m_{22}e_2 + \cdots + m_{n2}e_n, \\ \vdots & \qquad\qquad \vdots \\ e'_n &= m_{1n}e_1 + m_{2n}e_2 + \cdots + m_{nn}e_n. \end{cases} \tag{2}$$

以 $e'_i (1 \leq i \leq n)$ 在基（I）下的坐标作为列向量构成的矩阵

$$M = \begin{bmatrix} m_{11} & m_{12} & \cdots & m_{1n} \\ m_{21} & m_{22} & \cdots & m_{2n} \\ \vdots & \vdots & & \vdots \\ m_{n1} & m_{n2} & \cdots & m_{nn} \end{bmatrix} \tag{3}$$

称为基 (I) 到基 (II) 的**过渡矩阵**.

- **基变换公式:**

$$[e'_1, e'_2, \cdots, e'_n] = [e_1, e_2, \cdots, e_n]M. \tag{4}$$

- 坐标变换公式：α 是 V 中一个向量，α 在基 (I) 基 (II) 下的坐标分别为

$$X = [x_1, x_2, \cdots, x_n]^T,\ X' = [x'_1, x'_2, \cdots, x'_n]^T.$$

则

$$X = MX',\ X' = M^{-1}X. \tag{5}$$

要掌握如何计算过渡矩阵（注意区分行和列），如何运用过渡矩阵求不同基下的坐标（坐标变换公式），例：P136，P137。

- 子空间：线性空间 V 的**子集** W, 关于 V 中的加法和数乘封闭。
- 生成子空间：$\alpha_1, \cdots, \alpha_s \in V$, 则它们生成的子空间

$$L(\alpha_1, \cdots, \alpha_s) = \{k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s \mid k_1, \cdots, k_s \in P\}$$

注：$\alpha_1, \cdots, \alpha_s$ 不一定线性无关。

  - 设 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 与 $\beta_1, \cdots, \beta_r$ 为 V 的两组向量，则
    1. $L(\alpha_1, \alpha_2, \cdots, \alpha_s) = L(\beta_1, \beta_2, \cdots, \beta_r)$ 当且仅当 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 与 $\beta_1, \cdots, \beta_r$ 等价。
    2. $\dim L(\alpha_1, \alpha_2, \cdots, \alpha_s) =$ 秩$\{\alpha_1, \alpha_2, \cdots, \alpha_s\}$, 且 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 的极大无关组均可作为 $L(\alpha_1, \alpha_2, \cdots, \alpha_s)$ 的一组基。
  - 若 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 线性无关，则 $L(\alpha_1, \alpha_2, \cdots, \alpha_s)$ 是 s 维的线性空间。
  - $\dim L(\alpha_1, \alpha_2, \cdots, \alpha_s) \leq s$.
  - 若 $\beta_1, \cdots, \beta_r \in V$ 是一个向量组, 如果 $\beta_1, \cdots, \beta_r \in L(\alpha_1, \alpha_2, \cdots, \alpha_s)$, 则

    $$L(\beta_1, \cdots, \beta_r) \subset L(\alpha_1, \alpha_2, \cdots, \alpha_s).$$

    在证明中经常用到这些性质, 要求熟练掌握。
    掌握如果计算生成子空间的基和维数：等价于寻找向量组的极大无关组。P140
  - 解空间与基础解系：$m \times n$ 矩阵 A 的秩为 $r(A) = r$, 则线性方程组 $AX = 0$ 的解的全体记为 $W \subset P^n$(称为该线性方程组的解空间), 维数满足 $\dim W = n - r$, W 的一组基称为该方程组的一个基础解系。
    * 掌握计算基础解系：解方程，然后找出通解，计算出来的解要代回原方程去验证，P145。
    * 要证明一组向量是一个方程组的基础解系主要分三步：一，判断矩阵 A 的秩，然后确定解空间的维数。二，证明该向量组每个元素都是方程组的解。三，证明该向量组线性无关且个数等于解空间的维数。P149 5，6.
    * 寻找一个方程组的基础解系：1，判断解空间的维数。2，找出线性无关的解。3，线性无关的解的个数等于维数。
    * 基础解系的定理的应用：判断矩阵的秩，例 P 148，P150 7，8，9，14，15.
      1. 如果要判断两个矩阵 A, B 的秩相等，只需证明 $AX = 0$ 与 $BX = 0$ 同解，例如证明 $r(A) = r(A^TA)$.
      2. 如果判断两个矩阵 A, B 的秩 $r(A) \leq r(B)$, 只需证明 $BX = 0$ 的解都是 $AX = 0$ 的解，例如 $r(AB) \leq r(B)$.
      3. 证明矩阵秩的不等式问题可以利用上一章的分块矩阵，初等变换等方法。

- 矩阵的行秩 = 矩阵行向量组的秩 = 矩阵的秩 = 矩阵的列秩 = 矩阵的列向量组的秩。**一般情况矩阵的行向量组和列向量组不在同一个空间。在遇到矩阵的秩的问题时，如果用上一章的结论无法证明，可以考虑用矩阵的行秩或者列秩来考虑，这需要掌握向量组的线性无关性与相关性，例 P148**

- 欧氏空间：有内积的线性空间。可以理解为线性空间中引进的向量与向量的乘法，满足交换律，分配率等，可以与 $\mathbb{R}$ 上的乘法类比。

- $\mathbb{R}^n$ 上的标准内积：$\alpha = [a_1, \cdots, a_n]^T$, $\beta = [b_1, \cdots, b_n]^T$, 则内积

  $$(\alpha, \beta) = a_1 b_1 + \cdots + a_n b_n = \alpha^T \beta$$

  其中后一个等号是矩阵的乘法。

- 任一 n 阶正定矩阵 A 可以定义 $\mathbb{R}^n$ 的一个内积：$(\alpha, \beta) = \alpha^T A \beta$. 特别地，标准内积是由 $A = E$ 定义的。

- $\mathbb{R}^n$ 中任一内积对应一个正定矩阵 A, 这个正定矩阵就是内积在一组基下的度量矩阵。

- 内积满足线性性：$(\sum_{i=1}^m k_i \alpha_i, \sum_{j=1}^n t_j \beta_j) = \sum_{i=1}^m \sum_{j=1}^n (\alpha_i, \beta_j)$.

- $|(\alpha, \beta)| \leq ||\alpha|| \cdot ||\beta||$, 等号成立当且仅当 α, β 相关（$\beta = k\alpha$, 或者 $\alpha = k\beta$.）
  - 可以用来证明: 若 A 正定, 且 $X, Y \in \mathbb{R}^n$, 则有 $|X^T A Y|^2 \leq |X^T A X| \cdot |Y^T A Y|$. 该结论对半正定矩阵也对, 可以考虑 $A + \epsilon E$, 然后另 $\epsilon \to 0$.

- $(\alpha, \beta) = 0$ 称 α 与 β 正交或者相互垂直。

- $||\alpha + \beta|| \leq ||\alpha|| + ||\beta||$

- 若 $\alpha \perp \beta$, 则 $||\alpha + \beta||^2 = ||\alpha||^2 + ||\beta||^2$

- 度量矩阵: 设 $e_1, \cdots, e_n$ 为 V 的一组基, 矩阵 $A = [a_{ij}]_{n \times n} = [(e_i, e_j)]_{n \times n}$ 称为基 $e_1, \cdots, e_n$ 下的度量矩阵。**理解度量矩阵的定义，例 P162.**
  - 度量矩阵是对称矩阵。
  - 若 α, β 的坐标为 $X, Y \in \mathbb{R}^n$, 则 $(\alpha, \beta) = X^T A Y$, 由内积的正定性，可知 A 是正定矩阵。
  - 计算两向量的内积，可以通过度量矩阵来计算 $(\alpha, \beta) = X^T A Y$. 例: P162.
  - 设 $\alpha_1, \cdots, \alpha_n \in \mathbb{R}^n$, 记 $B = [\alpha_1, \cdots, \alpha_n]$, 则该基下的度量矩阵为 $A = [(\alpha_i, \alpha_j)]_{n \times n} = [\alpha_i^T \alpha_j]_{n \times n} = B^T B$. 由于 B 可逆，可知 $A = B^T B$ 正定。

- 正交向量组：两两正交的非零向量构成的向量组。
  - 正交向量组线性无关。证明利用反证法和正交性。
  - 正交基：n 维欧氏空间 n 个两两正交的向量组成的向量组。

- 标准正交基：正交基且每个向量都是单位长度。**从正交基到标准正交基，只需做单位化**：$\beta = \alpha/||\alpha||$.
- 由度量矩阵的定义可知：正交基对应的度量矩阵为对角矩阵，标准正交基对应的度量矩阵为单位矩阵。特别地，若 $\alpha_1, \cdots, \alpha_n \in \mathbb{R}^n$ 为一组标准正交基，则矩阵 $B = [\alpha_1, \cdots, \alpha_n]$ 为正交矩阵：$B^TB = E$.
- 取定标准正交基后，计算向量的内积比较容易。设 $e_1, \cdots, e_n$ 是一个标准正交基，α, β 的坐标为 $X = [x_1, \cdots, x_n]^T, Y = [y_1, \cdots, y_n]^n$, 则
  1. $(\alpha, \beta) = X^TY$
  2. $x_i = (\alpha, e_i)$.
- 给定一组标准正交基，要计算一个向量的坐标，只需计算内积。$\alpha = (\alpha, e_1)e_1 + \cdots + (\alpha, e_n)e_n$. 例 P165, P171 11. **这个结论只对标准正交基成立。**

• 任一正交向量组都可以扩充为一组正交基。

• 施密特正交化：任一一组基 $\alpha_1, \cdots, \alpha_n$，都存在上三角矩阵 B 且 $b_{ii} > 0$ 使得向量组 $[\beta_1, \cdots, \beta_n] = [\alpha_1, \cdots, \alpha_n]B$ 为标准正交基。

  - 要**熟练**掌握施密特正交化的证明，从一组基变为正交基时用到。P167 例 4.8.10, P172 13, 15.
  - QR 分解：任意可逆的实矩阵 A 都存在正交矩阵 Q 和上三角矩阵 R 且 $r_{ii} > 0$, 使得 $A = QR$, 并且这样的 Q, R 是唯一。

    Proof. 设 $A = [\alpha_1, \cdots, \alpha_n]$, 则列向量 $\alpha_1, \cdots, \alpha_n \in \mathbb{R}^n$ 为一组基，由施密特正交化存在一个上三角矩阵 S 且 $s_{ii} > 0$，和标准正交基 $\beta_1, \cdots, \beta_n$, 满足

    $$[\beta_1, \cdots, \beta_n] = [\alpha_1, \cdots, \alpha_n]S.$$

    记矩阵 $Q = [\beta_1, \cdots, \beta_n]$, 由于 $\beta_1, \cdots, \beta_n$ 是标准正交矩阵，所以 Q 是正交矩阵。因此 $Q = AS$. 从而 $A = QS^{-1}$, 记 $R = S^{-1}$ 即得分解公式。整个证明还需下面两个结论：

    * 上三角矩阵的逆矩阵还是上三角矩阵。**按照伴随矩阵的定义可证明。**
    * 唯一性：上三角型的正交矩阵必是对角矩阵且对角元为 1 或 -1, 如果对角元素都大于零，则该矩阵一定是单位矩阵。**利用：正交矩阵的每一行、每一列作为向量，向量的长度为 1.**

    □

  - Cholesky 分解：任一正定矩阵 A 存在唯一的上三角矩阵 R 且 $r_{ii} > 0$, 满足 $A = R^TR$.

    Proof. 由于 A 正定，所以存在可逆矩阵 B 使得 $A = B^TB$. 对 B 应用 QR 分解，有上三角矩阵 R 和正交矩阵 Q 使得

    $$B = QR$$

且 $r_{ii} > 0$. 则 $A = B^T B = R^T Q^T Q R = R^T R$.

唯一性：若存在上三角矩阵 P 且对角元素 $p_{ii} > 0$ 满足 $A = P^T P = R^T R$. 则 $(R^T)^{-1} P^T = RP^{-1}$, 而前者是下三角，后者是上三角，因此 $RP^{-1}, (R^T)^{-1} P^T$ 是对角矩阵。特别地，$R^{-1}P = (R^T)^{-1} P^T = RP^{-1}$, 即 RP^{-1} 的逆等于本身且是对角矩阵，因此 $RP^{-1} = E$, 从而 $R = P$. □

5 矩阵对角化

这一章最核心的问题就是研究矩阵如何对角化，在什么条件下可对角化，对角化的矩阵有哪些应用。**不是所有矩阵都可对角化**。

- 特征值/特征向量：A 是数域 P 上的 n 阶方阵，若存在非零向量 ξ 和 $\lambda_0 \in P$ 满足 $A\xi = \lambda_0 \xi$，则称 λ_0 为 A 的特征值（或特征根），ξ 为对应 λ_0 的特征向量。
- 特征多项式：$f(\lambda) = |\lambda E - A|$.
  - A 的特征值一定是特征多项式 $f(\lambda) = 0$ 的解，特征多项式 $f(\lambda) = 0$ 的解一定是特征值。
  - 求特征值和特征向量的一般步骤：通过特征多项式 $|\lambda E - A| = 0$ 解出特征值 $\lambda_1, \cdots, \lambda_n$, 然后解方程 $(\lambda_i E - A)X = 0$, 对应的解就是特征向量。
  - 复数域上一定有 n 个特征值（计算重数）。
- A 可对角化当且仅当存在 n 个线性无关的特征向量。设线性无关的特征向量为 $\xi_1, \cdots, \xi_n$, 对应的特征值为 $\lambda_1, \cdots, \lambda_n$, 则矩阵 $P = [\xi_1, \cdots, \xi_n]$ 满足：$P^{-1}AP = diag\{\lambda_1, \cdots, \lambda_n\}$. **理解这个定理的证明。**
- 相似：设 A 与 B 为 n 阶方阵，若存在可逆矩阵 P 使得 $P^{-1}AP = B$，则称 A 与 B 相似。
- 设 A 的全体互异的特征值为 $\lambda_1, \cdots, \lambda_s$, 则 A 可对角化
  - 当且仅当 $r(\lambda_i E - A) = s_i$ 满足 $\sum_{i=1}^{s} (n - s_i) = n$.
  - 当且仅当 $n - s_i = n_i$，$i = 1, \cdots s$，其中 λ_i 的重数为 n_i。

  主要用于判断一个矩阵是否可以对角化，这是充要条件。例 P206，P207 3，6 P 208 7，8. 证明满足 $A^2 = E$ 的矩阵可对角化，满足 $A^2 = A$ 的矩阵 A 可对角化。
- 设 A 的所有特征值为 $\lambda_1, \cdots, \lambda_n$(**在复数域下总存在 n 个特征值**), 则
  - $tr(A) = \sum_{i=1}^{n} a_{ii} = \sum_{i=1}^{n} \lambda_i$
  - $|A| = \lambda_1 \lambda_2 \cdots \lambda_n$ **这两个结论不需要 A 可对角化。**
  - 应用：P215 9，P 212 例 5.3.4.

- $h(x) = a_k x^k + \cdots + a_1 x + a_0$ 是多项式，若 λ_0 是 A 的特征值，则 $h(\lambda_0)$ 是 $h(A)$ 的特征值。
- A 可逆当且仅当每个特征值都非零，且 A^{-1} 的特征值是 A 的特征值的倒数。
- 相似矩阵有相同的特征多项式。具有相同特征多项式的矩阵不一定相似。推论：相似矩阵有相同的迹 $tr(A) = tr(B)$ 和相同的行列式 $|A| = |B|$. **可以用于判断两矩阵是否相似，P212 例 5.3.4，P215 9.**
- A 与 B 相似，则 A^k 与 B^k 相似，$h(A)$ 与 $h(B)$ 相似。
  - A 与 B 相似，则 $|h(A)| = |h(B)|$.
  - 若 A 与对角矩阵相似，即 $A = P^{-1} diag\{\lambda_1, \cdots, \lambda_n\} P$, 则 $h(A) = P^{-1} diag\{h(\lambda_1), \cdots, h(\lambda_n)\} P$. **这可以用于计算一个矩阵 A^k, 首先是对角化，找出特征值，找出特征向量，然后构造 P.**
  - 若 A 可对角化，则 $|h(A)| = h(\lambda_1) h(\lambda_2) \cdots h(\lambda_n), tr(h(A)) = \sum_{i=1}^{n} h(\lambda_i)$. **该结论对于 A 不可对角化也成立。**
  - 具体应用：P214 例 5.3.5, P215 5.
- 实对称矩阵的特征值都是实数 (**即复数域解特征多项式得到的 n 个解都是实数。**)，特征向量也是实向量。
- 实对称矩阵不同特征值的特征向量相互正交。
- 实对称矩阵一定可对角化，即总存在 n 个线性无关的特征向量。**只需熟记该结论，在遇到实对称矩阵对角化问题，直接解特征值，然后解特征向量，得到线性无关的特征向量一定是 n 个。**
- 实对称矩阵 A，存在正交矩阵 U 使得 $U^T AU$ 为对角矩阵。U 的构造方法
  - 利用特征多项式 $|\lambda E - A| = 0$ 解出所有特征值。
  - 解线性方程组 $(\lambda_i E - A) X = 0$, 得到一个基础解系，利用施密特正交化得到一组单位正交的基础解系。
  - 上面得到 n 个特征向量作为列向量构成的矩阵就是所求的正交矩阵。
  - 注：施密特正交化只需**分别**对每一个特征值对应的特征向量做。如果是计算实对称矩阵的对角化问题，建议利用以上方法构造正交矩阵，这样的好处是正交矩阵的逆只需通过转置即可得到。P 217 例 5.4.1, P218 例 5.4.2.
  - 注：一般的矩阵不存在正交矩阵使得它是对角。存在正交矩阵使得矩阵对角化当且仅当该矩阵是实对称的。

6 二次型

二次型 $f(x_1, \cdots, x_n) = X^TAX$, 其中 $A = [a_{ij}]_{n\times n}$ 对称矩阵, $X = [x_1, \cdots, x_n]^T$. 本章主要考察如何利用非退化线性替换 $X = CY$ 化二次型为标准型，等价于把对称矩阵合同到对角矩阵。

- 任意一个数域 P 上的二次型都存在非退化的线性替换化为标准形。
- 数域 P 上的对称矩阵 A，总存在 P 上的可逆矩阵 C 使得 C^TAC 为对角矩阵。
- 对于实对称矩阵 A，则存在正交矩阵 U 使得 $U^TAU = diag\{\lambda_1, \cdots, \lambda_n\}$ 其中 λ_i 是 A 的特征值。**化实二次型为标准形，等价于求正交矩阵把实对角矩阵的对角化。P238.**
- A 与 B 合同，则 $r(A) = r(B)$. $|A| > 0$ 当且仅当 $|B| > 0$. 一般情况不会有 $|A| = |B|$. **对比矩阵的相似和矩阵的等价来学习。**
- 二次型的秩：标准形中不为零的项的个数 $=r(A)$.
- 复对称矩阵 A 与 B 合同当且仅当 $r(A) = r(B)$.
- 实二次型的规范形 $y_1^2 + \cdots + y_p^2 - y_{p+1}^2 - \cdots - y_r^2$ 是唯一确定的，即 p, r 是唯一确定。
  - $r = r(A)$
  - p 称为正惯性指数 $=A$ 的正特征值的个数。
  - $r - p$ 称为负惯性指数 $=A$ 的负特征值的个数。
  - 实对称矩阵 A 与 B 在实数域上合同当且仅当 $r(A) = r(B)$ 和正惯性指数相同。
  - 计算正负惯性指数：设 $A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}$ 为 n 阶可逆矩阵，其中 A_1,A_2 为 $p \times n$, $(n-p)\times n$ 矩阵，计算二次型 $f = X^T(A_1^TA_1 - A_2^TA_2)X$ 的正负惯性指数。
  - 计算正负惯性指数：设 $\alpha \in \mathbb{R}^n$ 且 $\alpha^T\alpha = 1$, 求 $E - 2\alpha\alpha^T$ 的的正负惯性指数。**利用** $|E_n + AB| = |E_m + BA|$
- 正定二次型：实二次型，对任意非零实向量 X 有 $X^TAX > 0$.
- 半正定二次型：实二次型，对任意非零实向量 X 有 $X^TAX \geq 0$.
- A 实对称，则如下相互等价
  1. A 正定。
  2. A 的特征值均大于零。
  3. 存在可逆矩阵 B 使得 $A = B^TB$，**这个结论只有在证明中才用到**
  4. A 与单位矩阵合同

5. A 的顺序主子式都大于零。

判断一个实对称矩阵是否正定：（1）通过定义 $X^TAX > 0$ 来判断（例如：A, B 正定，证明 $A+B$ 是正定。例：A 正定，C 可逆，证明 C^TAC 正定。例：P254 10. ）。（2）证明 A 的特征值都大于零（例 P254 11.）。(3) 计算 A 的顺序主子式，P253 中的例 6.4.3，需要熟练掌握。

- 若 A 正定，则
  - A^{-1} 正定。
  - A^* 正定
  - 若 C 可逆，则 C^TAC 正定。
- A 正定，则有 $|A| > 0$. 若实对称矩阵 A 满足 $|A| < 0$, 则存在非零实向量 X 使得 $X^TAX < 0$.
- A 实对称，则如下相互等价:
  1. A 半正定。
  2. A 的特征值均大于等于零。
  3. 存在矩阵 B 使得 $A = B^TB$。
  4. A 的所有主子式都大于等于零。
- 如果已知一个矩阵 A 是正定矩阵或者证明 A 正定，去证明相应的结论：
  - 用正交矩阵对角化。例：A, B 正定，且 $AB = BA$，证明：AB 也正定。**一般的结论：A, B 可对角化，且 $AB = BA$, 则 A, B 可同时对角化。**
  - 利用特征值大于零：例：A 实对称，证明存在 $k > 0$ 使得 $kE+A$ 正定。例：A 实对称，证明存在 k_0, k_1 使得 $k_0X^TX \le X^TAX \le k_1X^TX$. 例：$A$ 半正定且可逆，则 A 正定。**该结论可以用来证明 P254 8,9.**
  - 若结论对于 $A = E$ 很容易证明，此时可以利用分解 $A = B^TB$ 来证明，然后提出 B^T 和 B. 例：A, B 正定，且 $AB = BA$，证明：AB 也正定。例：A, B $A-B$ 都正定，证明 $B^{-1} - A^{-1}$ 正定。例：$A = [a_{ij}]$, $B = [b_{ij}]$ 都正定，证明 $D = [a_{ij}b_{ij}]$ 也正定。

7 扰动技巧

- 若要证明一个结论对于半正定矩阵 A 成立，如果对于正定矩阵可以证明，则利用正定矩阵 $xE + A$ 其中 $(x > 0)$，得到一个结论，然后令 $x \to 0$ 可得半正定矩阵的结论。
- 若要证明一个结论对于一般矩阵成立, 如果对于可逆矩阵可以证明，则当 x 不是特征值时，$A - xE$ 是可逆矩阵，验证这个矩阵也满足题目的假设，由于 $A - xE$ 可逆，可以利用已经证明的可逆的结论，然后令 $x \to 0$.
- 只要是代数结论，不涉及求逆的，都可以尝试应用上面的技巧。

- 例：n 阶矩阵 A, B, 若 $A^k = 0$ 和 $AB = BA$, 证明 $|A+B| = |B|$.

  Proof. 1）若 B 可逆，要证 $|A+B| = |B|$，只需证 $|AB^{-1}+E| = 1$. 由 $AB = BA$, 可得 $B^{-1}A = AB^{-1}$. 从而
  $$(AB^{-1})^k = A^k(B^{-1})^k = 0.$$
  设 AB^{-1} 的特征值为 $\lambda_1, \cdots, \lambda_n$, 则 λ_i^k 是 $(AB^{-1})^k$ 的特征值，特别地，$\lambda_i = 0, i = 1, 2, \cdots, n$.

  假设 $\lambda_1', \cdots, \lambda_n'$ 为矩阵 $AB^{-1}+E$ 的特征值，对应的特征向量为 $\xi_1', \cdots, \xi_n'$, 则对任意的 $i = 1, 2, \cdots, n$, 有
  $$(AB^{-1}+E)\xi_i' = \lambda_i'\xi_i.$$
  可知 $AB^{-1}\xi_i' = (\lambda_i' - 1)\xi_i'$. 即 $\lambda_i' - 1$ 为 AB^{-1} 的特征值（等于零）。因此 $\lambda_i' = 1, i = 1, \cdots, n$. 因此 $|AB^{-1}+E| = \lambda_1'\lambda_2'\cdots\lambda_n' = 1$. 对于 B 可逆，结论已证明。

  2）对于 B 不可逆，假设 B 的特征值为 $\lambda_1, \cdots, \lambda_n$, 则当 $\lambda \neq \lambda_i$ 时，$D_\lambda = B - \lambda E$ 是可逆的，且满足 $AD = DA$. 利用 1）证明的结论，可得对任意的 $\lambda \neq \lambda_i$, 都有
  $$|A+B-\lambda E| = |B - \lambda E|.$$
  而上式等式两边都是多项式，在有限个点外相等，则它一定是一个恒等式，特别地 $\lambda = 0$ 也是等式，即 $|A+B| = |B|$. **另一个看法：由于等式两边都是多项式，是关于 λ 的连续函数，因此令 $\lambda \to 0$ 可得 $|A+B| = |B|$. (因为有限个点外的值不影响极限。)**

  □

- 例：n 阶矩阵 A, B, C, D, 若 $AC = CA$, 证明：
  $$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |AD - CB|.$$

  Proof. 1) 若 A 可逆，则
  $$\begin{aligned} \begin{vmatrix} A & B \\ C & D \end{vmatrix} &= \begin{vmatrix} E & 0 \\ -CA^{-1} & E \end{vmatrix} \begin{vmatrix} A & B \\ C & D \end{vmatrix} && (6) \\ &= \begin{vmatrix} A & B \\ 0 & D - CA^{-1}B \end{vmatrix} && (7) \\ &= |A||D - CA^{-1}B| && (8) \\ &= |A||D - A^{-1}CB| = |AD - CB|. && (9) \end{aligned}$$
  2) 若 A 不可逆，则考虑 $A' = A - \lambda E$ 当 λ 不是 A 的特征值时，A' 时可逆的，且满足 $A'C = CA'$. 应用 1) 的结论可得
  $$\begin{vmatrix} A - \lambda E & B \\ C & D \end{vmatrix} = |(A - \lambda E)D - CB|.$$

由于等式两边都是关于 λ 的多项式，等式表明两个多项式在有限个点外恒等，则两个多项式处处相等，特别地, 取 $\lambda=0$, 得

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |AD-CB|.$$

□

8 例题

- 例：A 实 n 阶方阵，若 $AA^T=A^2$, 证明 $A=A^T$.

  Proof. 令 $B=A-A^T$, 则 $B^T=-B$, 要证明 $B=0$, 只需证明 $tr(BB^T)=0$ 即 $tr(B^2)=0$.

  $$tr(B^2)=tr(A^2+(A^T)^2-AA^T-A^TA)=tr((A^T)^2-A^TA)=tr(A^T(A^T-A)) \tag{10}$$

  由于 $tr(CD)=tr(DC)$, 可得

  $$tr(B^2)=tr((A^T-A)A^T)=tr((A^T)^2-AA^T)$$

  对 $AA^T=A^2$ 取转置得 $(A^T)^2=AA^T$, 因此

  $$tr(B^2)=0$$

  . □

- 例：A 实对称，如果存在非零列向量 α 使得 $\alpha^TA\alpha>0$, 则 A 至少有一个特征值大于零。

  Proof. 反证法，若 A 的所有特征值都小于等于零。则 $-A$ 的所有特征值都大于等于零，特别地，$-A$ 是半正定的，因此 $\alpha^T(-A)\alpha\geq 0$ 与 $\alpha^TA\alpha>0$ 矛盾。 □

- 例：A,B 为 n 阶方阵，A 可逆，且存在常数 λ, 满足 $A=(A-\lambda E)B$, 求证：$AB=BA$.

  Proof. 1）若 $\lambda=0$, 则 $A=AB$, 由 A 可逆可知 $B=E$, 因此结论显然成立。

  2）若 $\lambda\neq 0$, 则 (**凑 E**) $\lambda E+A-\lambda E=(A-\lambda E)B$. 因而

  $$\lambda E=(A-\lambda E)(B-E).$$

  取行列式可得 $|A-\lambda E||B-E|=\lambda^n\neq 0$. 因此 $A-\lambda E$ 可逆。从而 $B=(A-\lambda E)^{-1}A$. 要证明 $AB=BA$, 只需证明 $A(A-\lambda E)^{-1}A=$

$(A-\lambda E)^{-1}A^2$. 若我们可以证明 $A(A-\lambda E)^{-1}=(A-\lambda E)^{-1}A$, 则显然有 $A(A-\lambda E)^{-1}A=(A-\lambda E)^{-1}A^2$.

而要证明 $A(A-\lambda E)^{-1}=(A-\lambda E)^{-1}A$，等式两边左乘、右乘可逆矩阵 $A-\lambda E$, 只需证明 $(A-\lambda E)A=A(A-\lambda E)$, 展开后可知这是成立。因此证明了 $A(A-\lambda E)^{-1}=(A-\lambda E)^{-1}A$，由上面的分析可得 $AB=BA$. □

- 例：设 A,B 实对称，且 A 的特征值在区间 $[a,b]$ 上，B 的特征值在区间 $[c,d]$ 上，则 $A+B$ 的特征值在区间 $[a+c,b+d]$ 上。

  Proof. 由于 $A-aE$, $B-cE$ 的特征值都大于等于零，所以 $A-aE, B-cE$ 是半正定矩阵。从而 $A+B-(a+c)E$ 也是半正定矩阵，特别地，它的特征值都大于等于零，即 $A+B$ 的特征值都大于等于 $a+c$.

  另一方面，$bE-A$, $dE-B$ 的特征值都大于等于零，所以 $bE-A$, $dE-B$ 是半正定矩阵，从而 $(b+d)E-(A+B)$ 也是半正定矩阵，特征地，它的特征值都大于等于零，即 $A+B$ 的特征值都小于等于 $b+d$. □

- 例：A 实对称，证明存在 k_0,k_1, 使得对任意的实向量 X 有 $k_0X^TX\le X^TAX\le k_1X^TX$.

  Proof. 由于 A 实对称，所以特征值都是实数。设 $\lambda_1,\cdots,\lambda_n$ 为 A 的特征值，不妨设 λ_1 为最小的特征值，λ_n 为最大的特征值。令 $k_0=\lambda_1-1$, $k_1=\lambda_n+1$. 则 $A-k_0E$, k_1E-A 的特征值都大于零，因此它们都是正定矩阵，特别地，有

  $$X^T(A-k_0E)X\ge 0$$
  $$X^T(k_1E-A)X\ge 0$$

  因此 $k_1X^TX\le X^TAX\le k_1X^TX$. □

- 例：求二次型 $f(x_1,\cdots,x_n)=\sum_{j=1}^n x_j^2+2a\sum_{1\le j<k\le n}x_jx_k$ 的秩和符号差，这里 a 是实数。

  Proof. 二次型的矩阵为 $A=[a_{ij}]$, 其中 $a_{ii}=1, i=1,\cdots,n$. 当 $i\ne j$ 有 $a_{ij}=a$. 则 $A=(1-a)E+a\alpha\alpha^T$, 这里 $\alpha=[1,1,\cdots,1]^T$. 计算 A 的秩和符号差只需计算出 A 的特征值。

  A 的特征多项式

  $$h(\lambda)=|(\lambda+a-1)E-a\alpha\alpha^T|=(\lambda+a-1)^{n-1}(\lambda+a-1-a\alpha^T\alpha)=(\lambda+a-1)^{n-1}(\lambda-1-(n-1)a).$$

  因此 A 的所有特征值为 $1-a,1-a,\cdots,1-a,1+a(n-1)$

  1) 当 $n=1$ 时，二次型的秩为 1, 符号差为 1.
  2) 当 $a=1,n\ge 2$ 时，二次型的秩为 1, 符号差为 1.
  3) 当 $a=-1/(n-1),n\ge 2$ 时，二次型的秩为 $n-1$, 符号差为 $n-1$.
  4) 当 $-1/(n-1)<a<1,n\ge 2$ 时，二次型的秩为 n, 符号差为 n.
  5) 当 $a>1$ 时，二次型的秩为 n, 符号差为 $2-n$.
  6) 当 $a<-1/(n-1)$, 二次型的秩为 n, 符号差为 $n-2$. □

- 例：设 $\alpha_1, \alpha_2, \alpha_3$ 是 4 个未知量的非齐次方程组 $AX = b$ 的 3 个解，且 $\alpha_1 + \alpha_2 = (2, 2, 4, 6)^T, \alpha_1 + 2\alpha_3 = (0, 3, 0, 6)^T, r(A) = 3$, 求线性方程组 $AX = b$ 的通解。

  Proof. 由于 $r(A) = 3$, 则齐次方程组 $AX = 0$ 的解空间维数为 1，由于

  $$\beta = \alpha_1 + 2\alpha_3 - \frac{3}{2}(\alpha_1 + \alpha_2) = \frac{1}{2}(\alpha_3 - \alpha_1) + \frac{3}{2}(\alpha_3 - \alpha_2) = (-3, 0, -6, -3)^T.$$

  是齐次方程组 $AX = 0$ 的解，且非零，所以是它的基础解系。而

  $$\eta = \frac{1}{2}(\alpha_1 + \alpha_2) = (1, 1, 2, 3)^T$$

  满足 $A\eta = \frac{1}{2}(A\alpha_1 + A\alpha_2) = b$. 因此 η 是 $AX = b$ 的一个特解，因此 $AX = b$ 的通解为 $k\beta + \eta$.

  □

- 例：$A_{n\times m}$ 为实矩阵，若 $tr(A^TA) = 0$, 则 $A = 0$.

  Proof. 设 $B = A^TA$, 则 $b_{ii} = \sum_{k=1}^n a_{ki}^2$. 因此 $0 = tr(B) = \sum_{i=1}^m b_{ii} = \sum_{i=1}^m \sum_{k=1}^n a_{ki}^2$ 可得任意的 k, i 有 $a_{ki} = 0$, 即 $A = 0$. □

- 例：A, B 为实的 n 阶方阵，若 $AB = (B - A^T)A$, 证明 $A = 0$.

  Proof. 由于 $A^TA = BA - AB$. 因此 $tr(A^TA) = tr(BA - AB) = tr(BA) - tr(AB) = 0$. 由上一题结论可知 $A = 0$. □

- 例：$A = diag\{\lambda_1, \cdots, \lambda_n\}$ 且 $\lambda_i \neq \lambda_j$, 则任意与 A 可交换的矩阵都是对角矩阵。

  Proof. 设 $B = [b_{ij}]$ 且 $AB = BA$. 则 $(AB)_{ij} = \lambda_i b_{ij}$, $(BA)_{ij} = \lambda_j b_{ij}$. 因此

  $$\lambda_i b_{ij} = \lambda_j b_{ij}$$

  若 $i \neq j$, 则可得 $b_{ij} = 0$. 因此 B 是对角矩阵。 □

- 例：$A = diag\{\lambda_1 E_{r_1}, \cdots, \lambda_s E_{r_s}\}$ 且 $\lambda_i \neq \lambda_j$, 则任意与 A 可交换的矩阵一定是分块对角矩阵。

  Proof. 设 B 与 A 可交换，把 B 分块为 $B = [B_{ij}]$, 则由 $AB = BA$, 可得 $(AB)_{ij} = \lambda_i B_{ij}$, $(BA)_{ij} = \lambda_j B_{ij}$. 因此

  $$\lambda_i B_{ij} = \lambda_j B_{ij}$$

  若 $i \neq j$, 则可得 $B_{ij} = 0$. 因此 B 是分块对角矩阵。 □

- 例：若 $B = diag\{B_1, B_2, \cdots, B_s\}$ 为分块对角矩阵，则 B 可对角化当且仅当每个 B_i 可对角化。

  Proof. 1）若每个 B_i 都可对角化，则存在可逆矩阵 $P_1, P_2, \cdots, P_s$ 使得 $P_1^{-1}B_1P_1, P_2^{-1}B_2P_2, \cdots, P_s^{-1}B_sP_s$ 为对角矩阵。令 $P = diag\{P_1, P_2, \cdots, P_s\}$，则有 $P^{-1}BP$ 为对角矩阵。

  2) 若 B 可对角化，则存在 n 个线性无关的特征向量 $\xi_1, \xi_2, \cdots, \xi_n$ 对应的特征值为 $\lambda_1, \cdots, \lambda_n$. 按矩阵 B 的分块对每个向量分块

  $$\xi_i = \begin{pmatrix} \xi_{i1} \\ \xi_{i2} \\ \vdots \\ \xi_{is} \end{pmatrix}$$

  由 $B\xi_i = \lambda_i\xi_i$ 可得 $B_j\xi_{ij} = \lambda_i\xi_{ij}$. 若 $\xi_{ij} \neq 0$, 则 ξ_{ij} 为 B_j 的特征向量。设 B_j 的阶数为 n_j.

  要证明 B_j 可对角化，只需说明 $\xi_{j1}, \xi_{j2}, \cdots, \xi_{jn}$ 的秩为 n_j，即存在 n_j 个线性无关的特征向量。

  反证法：若 $\xi_{j1}, \xi_{j2}, \cdots, \xi_{jn}$ 的秩小于 n_j，我们将证明 $\xi_1, \cdots, \xi_n$ 线性相关，得出矛盾。
  若 $\xi_{j1}, \xi_{j2}, \cdots, \xi_{jn}$ 的秩 $= r < n_j$, 不妨设 $j = 1$, $\xi_{11}, \xi_{12}, \cdots, \xi_{1r}$ 线性无关。则可逆矩阵 $P = [\xi_1, \cdots, \xi_n]$ 经过初等列变换可化为

  $$P \to \begin{pmatrix} A_1 & 0_{n_1 \times (n-r)} \\ A_2 & A_3 \end{pmatrix}$$

  而 A_1 是列满秩矩阵 $r(A_1) = r$，经过初等行变换可以把 A_2 消去

  $$P \to \begin{pmatrix} A_1 & 0_{n_1 \times (n-r)} \\ A_2 & A_3 \end{pmatrix} \to \begin{pmatrix} A_1 & 0_{n_1 \times (n-r)} \\ 0_{(n-n_1) \times r} & A_3 \end{pmatrix}$$

  因此 $r(P) = r(A_1) + r(A_3) = r + r(A_3)$. 由于 A_3 是 $(n - n_1) \times (n - r)$ 矩阵, 则 $r(A_3) \leq n - n_1 < n - r$. 故

  $$r(P) = r + r(A_3) \leq n + r - n_1 < n$$

  与 P 可逆矛盾。综上可得 $\xi_{j1}, \xi_{j2}, \cdots, \xi_{jn}$ 的秩为 n_j, 即 B_j 存在 n_j 个线性无关的特征向量，因此 B_j 可对角化。

  □

- 例：设 A, B 都可对角化，证明：$AB = BA$ 当且仅当 A, B 可同时对角化，即存在可逆矩阵 P 使得 $P^{-1}AP, P^{-1}BP$ 都是对角矩阵。

Proof. 1) 若 A, B 可同时对角化，即存在可逆矩阵 P 使得 $P^{-1}AP, P^{-1}BP$ 为对角矩阵，所以 $P^{-1}APP^{-1}BP = P^{-1}BPP^{-1}AP$. 从而 $P^{-1}ABP = P^{-1}BAP$. 由于 P 可逆，可得 $AB = BA$.

2) 若 $AB = BA$, 由于 A 可对角化，存在可逆矩阵 P_1 使得 $P_1^{-1}AP_1 = diag\{\lambda_1 E_{r_1}, \lambda_2 E_{r_2}, \cdots, \lambda_s E_{r_s}\}$, 其中 $\lambda_i \neq \lambda_j$. 则

$$P_1^{-1}AP_1P_1^{-1}BP_1 = P_1^{-1}ABP_1 = P_1^{-1}BAP_1 = P_1^{-1}BP_1P_1^{-1}AP_1$$

即 $P_1^{-1}BP_1$ 与对角分块矩阵 $diag\{\lambda_1 E_{r_1}, \lambda_2 E_{r_2}, \cdots, \lambda_s E_{r_s}\}$ 可交换。由于 $\lambda_i \neq \lambda_j$, 可得 $P_1^{-1}BP_1$ 也是对角分块矩阵 $P_1^{-1}BP_1 = diag\{B_1, B_2, \cdots, B_s\}$. 由于 B 可对角化，所以 $P_1^{-1}BP_1$ 也可对角化，由上一例题可知每个 B_i 可对角化，即存在可逆矩阵 $C_1, C_2, \cdots, C_s$ 使得 $C_1^{-1}B_1C_1, C_2^{-1}B_2C_2, \cdots, C_s^{-1}B_sC_s$ 都是对角矩阵。令

$$Q = diag\{C_1, C_2, \cdots, C_s\}.$$

则 Q 可逆，且

$$Q^{-1}P_1^{-1}BP_1Q = diag\{C_1^{-1}B_1C_1, C_2^{-1}B_2C_2, \cdots, C_s^{-1}B_sC_s\} = \text{对角矩阵。}$$

令 $P = P_1Q$, 则 $P^{-1}BP$ 为对角矩阵。另一方面，

$$P^{-1}AP = Q^{-1}P_1^{-1}AP_1Q = Q^{-1}diag\{\lambda_1 E_{r_1}, \lambda_2 E_{r_2}, \cdots, \lambda_s E_{r_s}\}Q = diag\{\lambda_1 E_{r_1}, \lambda_2 E_{r_2}, \cdots, \lambda_s E_{r_s}\}$$

也是对角矩阵，所以 A, B 可同时对角化。 □

- 例：A, B 正定矩阵，$AB = BA$, 证明 AB 是正定矩阵。

  Proof. 证法一：利用上一题的结论，A, B 可同时对角化，即存在可逆矩阵 P 使得 $P^{-1}AP, P^{-1}BP$ 都为对角矩阵且对角元素都是相应的特征值，都大于零，因此 $P^{-1}ABP$ 也是对角矩阵且对角元素是大于零。因此 AB 的特征值都大于零，故 AB 正定。

  证法二：设 AB 的特征值为 λ, 对应的特征向量为 ξ, 则 $AB\xi = \lambda\xi$. 两边同时左乘 $\xi^T B$ 可得

  $$\xi^T BAB\xi = \lambda\xi^T B\xi.$$

  由于 A 正定，B 也正定，可得 $\xi^T B\xi > 0$, $\xi^T BAB\xi > 0$. 故 $\lambda > 0$. 因此 AB 的特征值都大于零，因而 AB 正定。

  证法三：由于 A 正定，则存在可逆矩阵 C 使得 $A = C^T C$. 因此 $AB = C^T CB = C^T(CBC^{-1})C$. 要证 AB 正定，只需证明 CBC^{-1} 正定，由于 CBC^{-1} 与 B 相似，所以特征值都大于零，故 CBC^{-1} 正定。因而 AB 也正定。 □

- 例：$A, B, A - B$ 都正定，证明 $B^{-1} - A^{-1}$ 正定。

Proof. 思路：当 $A=E$ 时，$E-B$ 正定，所以 B 的特征值都小于一。而 B^{-1} 的特征值都大于 1. 所以 $B^{-1}-E$ 正定。一般情形利用 A 正定，存在可逆 C 使得 $A=C^TC$.

由于 A 正定，所以存在可逆 C 使得 $A=C^TC$. 而

$$A-B=C^TC-B=C^T(E-(C^T)^{-1}BC^{-1})C.$$

$$B^{-1}-A^{-1}=B^{-1}-C^{-1}(C^{-1})^T=C^{-1}(CB^{-1}C^T-E)(C^{-1})^T$$

记 $D=(C^T)^{-1}BC^{-1}$, 则 $D^{-1}=CB^{-1}C^T$. 由于 $A-B$ 正定，可知 $E-D$ 正定，所以 D 的特征值都小于 1. 从而 $D^{-1}-E$ 也正定。由于 $B^{-1}-A^{-1}=C^{-1}(D^{-1}-E)(C^{-1})^T$，所以 $B^{-1}-A^{-1}$ 也正定。 □

- 例: 设 $\alpha_1,\cdots,\alpha_n$ 是欧氏空间 V 的一组基, 若 $\beta\in V$ 满足 $(\beta,\alpha_i)=0, i=1,2,\cdots,n$, 则 $\beta=\theta$.

  Proof. 由施密特正交化，存在标准正交基 $e_1,\cdots,e_n$ 和过渡矩阵 B, 使得 $[e_1,\cdots,e_n]=[\alpha_1,\cdots,\alpha_n]B$, 从而可得 e_i 可被向量组 $\alpha_1,\cdots,\alpha_n$ 线性表示。另一方面，由于对任意的 $i=1,2,\cdots,n$ 有 $(\beta,\alpha_i)=0$，则 $(\beta,e_i)=0$. 若 $\beta\neq\theta$, 则 $e_1,e_2,\cdots,e_n,\beta$ 构成一个正交向量组。由于正交向量组线性无关，所以存在 $n+1$ 个线性无关的向量，矛盾！因此 $\beta=\theta$. □

- 例：$A=[a_{ij}], B=[b_{ij}]$ 正定，证明 $D=[a_{ij}b_{ij}]$ 也正定。

  Proof. 由于 B 正定，则存在可逆矩阵 C 使得 $B=C^TC$. 因此 $b_{ij}=\sum_{k=1}^n c_{ki}c_{kj}$. 要证明 D 是正定矩阵，只需对任意非零向量 $X=[x_1,\cdots,x_n]^T$ 证明 $X^TDX>0$. 事实上

$$\begin{aligned}
X^TDX &= \sum_{i=1}^n\sum_{j=1}^n d_{ij}x_ix_j=\sum_{i=1}^n\sum_{j=1}^n a_{ij}b_{ij}x_ix_j && (11)\\
&= \sum_{i=1}^n\sum_{j=1}^n\sum_{k=1}^n a_{ij}c_{ki}c_{kj}x_ix_j && (12)\\
&= \sum_{k=1}^n\sum_{i=1}^n\sum_{j=1}^n a_{ij}(c_{ki}x_i)(c_{kj}x_j). && (13)
\end{aligned}$$

  令 $Y_k=[c_{k1}x_1,\cdots,c_{kn}x_n]^T$, 则

$$X^TDX=\sum_{k=1}^n Y_k^TAY_k.$$

  由于 A 正定，要说明 $X^TDX>0$ 只需说明存在某个 $Y_k\neq 0$ 即可。由于

$$[Y_1,\cdots,Y_k]=Cdiag\{x_1,x_2,\cdots,x_n\}$$

而 C 可逆，$X \neq 0$, 则 $[Y_1, \cdots, Y_k]$ 不为零矩阵，所以存在某个 $Y_k \neq 0$. 故 $X^T DX > 0$, 从而 D 正定。

另一种看法说明存在 $Y_k \neq 0$：若所有的 $Y_k = 0$，按列向量分块 $C = [\alpha_1, \cdots, \alpha_n]$，则对任意的 i 都有 $(X, \alpha_i) = 0$，而 $\alpha_1, \cdots, \alpha_n$ 是 $\mathbb{R}^n$ 的一组基，由上一例题的结论可知 $X = 0$，矛盾！ □

- 例：设实二次型 f 的秩大于 1，则它可以分成两个一次齐次多项式的乘积当且仅当它的秩为 2 且符号差为零。

  Proof. 假设 f 可以表示为两个一次齐次多项式的乘积：

  $$f(x_1, \cdots, x_n) = (a_1x_1 + a_2x_2 + \cdots + a_nx_n)(b_1x_1 + \cdots + b_nx_n)$$

  若 $(a_1, a_2, \cdots, a_n)$ 与 $(b_1, b_2, \cdots, b_n)$ 线性相关且 $(b_1, \cdots, b_n) = k(a_1, \cdots, a_n)$，不妨设 $a_1 \neq 0$, 令 $y_1 = a_1x_1 + \cdots + a_nx_n$, $y_2 = x_2, \cdots, y_n = x_n$, 它是一个非退化的线性替换，则有

  $$f = ky_1^2$$

  这与 f 的秩大于 1 矛盾。因此 $(a_1, a_2, \cdots, a_n)$ 与 $(b_1, b_2, \cdots, b_n)$ 线性无关。不妨设

  $$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \neq 0.$$

  令 $y_1 = a_1x_1 + \cdots + a_nx_n$, $y_2 = b_1x_1 + \cdots + b_nx_n$, $y_3 = x_3, \cdots, y_n = x_n$, 则它是一个非退化线性替换，替换后的结果为

  $$f = y_1y_2.$$

  令 $y_1 = z_1 - z_2$, $y_2 = z_1 + z_2$, $y_3 = z_3, \cdots, y_n = z_n$，则

  $$f = z_1^2 - z_2^2.$$

  因此 f 的秩为 2, 符号差为零。

  另一方面：若 f 的秩为 2, 符号差为零，则存在非退化替换 $X = CY$ 使得 $f = y_1^2 - y_2^2 = (y_1 - y_2)(y_1 + y_2)$. 由于 y_1, y_2 都是关于 $x_1, \cdots, x_n$ 的一次齐次多项式，因此 f 可以表示为两个一次齐次多项式的乘积。 □

- 例：A 为 n 阶方阵，α 为 n 维列向量，若 $B = \begin{pmatrix} A & \alpha \\ \alpha^T & 0 \end{pmatrix}$ 的秩 $r(B) = r(A)$, 则 α 可由 A 的列向量表示。

  Proof. 由于 $r(A) = r(B) \geq r(A\ \alpha) \geq r(A)$, 因此不等号都是等号，即 $r(A\ \alpha) = r(A) = r$. 设 A 的列向量组为 $\beta_1, \cdots, \beta_n$, 取它的一个极大无关组 $\beta_{i_1}, \cdots, \beta_{i_r}$. 由于 $r(A\ \alpha) = r(A)$，则 $\beta_{i_1}, \cdots, \beta_{i_r}$ 也是 $\beta_1, \cdots, \beta_n, \alpha$ 的一个极大无关组，特别地，α 可由 $\beta_{i_1}, \cdots, \beta_{i_r}$ 线性表示。 □

- 例：设 A 为正定矩阵，B 实对称，则 $A + B$ 的正惯性指数大于等于 B 的正惯性指数。

Proof. 由于 A 正定，存在可逆矩阵 C 使得 $A=C^TC$. 因此

$$A+B=C^TC+B=C^T(E+(C^{-1})^TBC^{-1})C$$

所以 $A+B$ 与 $E+(C^{-1})^TBC^{-1}$ 合同。故 $A+B$ 的正惯性指数等于 $E+(C^{-1})^TBC^{-1}$ 的正惯性指数。另一方面 B 与 $(C^{-1})^TBC^{-1}$ 合同，所以 B 的正惯性指数等于 $(C^{-1})^TBC^{-1}$ 的正惯性指数。通过计算正特征值的个数可知：$E+(C^{-1})^TBC^{-1}$ 的正惯性指数大于等于 $(C^{-1})^TBC^{-1}$ 的正惯性指数，因此 $A+B$ 的正惯性指数大于等于 B 的正惯性指数。

□

- 例：设 $A=\begin{pmatrix} A_1 \\ A_2 \end{pmatrix}$ 为 n 阶可逆矩阵，其中 A_1,A_2 为 $p\times n$, $(n-p)\times n$ 矩阵，计算二次型 $f=X^T(A_1^TA_1-A_2^TA_2)X$ 的正负惯性指数。

  Proof. 由于 $A^TA=A_1^TA_1+A_2^TA_2$ 为正定矩阵，利用上一例可知

  $A^TA+A_1^TA_1-A_2^TA_2=2A_1^TA_1$ 的正惯性指数大于等于 $A_1^TA_1-A_2^TA_2$ 的正惯性指数。

  而 $A_1^TA_1$ 的秩等于 $r(A_1)=p$, 且 $A_1^TA_1$ 半正定，因此 $A_1^TA_1$ 的正惯性指数等于 p. 从而 $A_1^TA_1-A_2^TA_2$ 的正惯性指数小于等于 p.

  另一方面，同理

  $A^TA+A_2^TA_2-A_1^TA_1=2A_2^TA_2$ 的正惯性指数大于等于 $A_2^TA_2-A_1^TA_1$ 的正惯性指数.

  所以 $A_2^TA_2-A_1^TA_1$ 的正惯性指数小于等于 $n-p$. 因此 $A_1^TA_1-A_2^TA_2$ 的负惯性指数小于等于 $n-p$.

  由于 $D=[A_1^T,-A_2^T]$ 是可逆矩阵，而 $DA=A_1^TA_1-A_2^TA_2$. 因此 $A_1^TA_1-A_2^TA_2$ 可逆。故它的正负惯性指数和为 n，因此 $A_1^TA_1-A_2^TA_2$ 的正负惯性指数分别为 p 和 $n-p$.

  □

- 例：求二次型 $f(x_1,\cdots,x_n)=\sum_{i=1}^k(a_{i1}x_1+\cdots+a_{in}x_n)^2$ 的矩阵。

  Proof. 设 $\alpha_i=[a_{i1},a_{i2},\cdots,a_{in}]^T$, $X=[x_1,\cdots,x_n]^T$. 则

  $$f=\sum_{i=1}^k X^T\alpha_i\alpha_i^TX=X^T(\alpha_1,\cdots,\alpha_k)\begin{pmatrix}\alpha_1^T\\ \vdots \\ \alpha_k^T\end{pmatrix}X.$$

  记 $A=[\alpha_1,\cdots,\alpha_k]$, 则二次型对应的矩阵为 AA^T. □

9 二次型的配方法

对于一般的二次型

$f(x_1,\cdots,x_n)=a_{11}x_1^2+2a_{12}x_1x_2+\cdots+2a_{1n}x_1x_n+a_{22}x_2^2+2a_{23}x_2x_3+\cdots+a_{nn}x_n^2$

利用配方法来化为标准形的主要步骤：归纳的过程，先消去所有 x_1 的交叉项，然后消去所有 x_2 的交叉项，$\cdots$, 消去所有 x_{n-1} 的交叉项。

主要是利用公式:

$$\begin{aligned}&(a_1x_1+a_2x_2+\cdots+a_nx_n)^2 \\ &=a_1^2x_1^2+2a_1a_2x_1x_2+2a_1a_3x_1x_3+\cdots 2a_1a_nx_1x_n+a_2^2x_2^2+2a_2a_3x_2x_3+\cdots+a_n^2x_n^2\end{aligned} \tag{14}$$

即

$$\begin{aligned}&(a_1x_1+a_2x_2+\cdots+a_nx_n)^2-(a_1^2x_1^2+2a_1a_2x_1x_2+2a_1a_3x_1x_3+\cdots 2a_1a_nx_1x_n) \\ &=a_2^2x_2^2+2a_2a_3x_2x_3+\cdots+a_n^2x_n^2.\end{aligned} \tag{15}$$

等式右边不会出现 x_1 的项。

下面我们通过介绍如何消去 x_1 的交叉项来说明配方法:

1. 若 $a_{11}\neq 0$, 不妨设 $a_{11}=a^2>0$ (对于小于零的情形取负号即可)，令

$$\begin{cases} y_1=ax_1+a^{-1}a_{12}x_2+a^{-1}a_{13}x_3+\cdots+a^{-1}a_{1n}x_n \\ y_2=x_2 \\ y_3=x_3 \\ \vdots \\ y_n=x_n \end{cases}$$

   则该线性替换可逆，且替换后

$$f=y_1^2+a_{22}^*y_2^2+a_{23}^*y_2y_3+\cdots+a_{nn}^*y_n^2$$

   没有 y_1 的交叉项，接下来只需依次类推即可消去 y_2 的交叉项,$\cdots$

2. 若 $a_{11}=0$ 且某个 $a_{1i}\neq 0$, 不妨设 $a_{12}\neq 0$. 令

$$\begin{cases} x_1=y_1-y_2 \\ x_2=y_1+y_2 \\ y_3=x_3 \\ \vdots \\ y_n=x_n \end{cases}$$

   则该替换是非退化且替换后

$$f=2a_{12}y_1^2-2a_{12}y_2^2+\cdots$$

   使得 y_1^2 的系数非零，可以利用上面 1 中的情形消去与所有 y_1 的交叉项，依次类推即可化为标准形。