第 **VI** 章

克罗内克（**Kronecker**）积及其应用

6．1　Kronecker 积

6．1．1 Kronecker 积的概念

定义 1—1　设 A=$(a_{ij}) \in c^{m\times n}, B=(b_{ij}) \in c^{p\times q}$，则称如下的分块矩阵

$$A\otimes B=\begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{bmatrix} \in C^{mp\times nq}$$

为 A 的克罗内克（**Kronecker**）积，或称 A 与 B 的直积，或张量积，简记为 $A\otimes B=(a_{ij}B)_{m\times n}$，即 $A\otimes B$ 是一个 $m\times n$ 块的分块矩阵，最后是一个 $mp\times nq$ 阶的矩阵。

例 **1—1**　设 $A=\begin{bmatrix} a & b \\ c & d \end{bmatrix}, B=\begin{bmatrix} x \\ y \end{bmatrix}$，那么

$$A\otimes B=\begin{bmatrix} aB & bB \\ cB & dB \end{bmatrix}=\begin{bmatrix} ax & bx \\ ay & by \\ cx & dx \\ cy & dy \end{bmatrix}_{4\times 2}$$

$$B\otimes A=\begin{bmatrix} xA \\ yA \end{bmatrix}=\begin{bmatrix} xa & xb \\ xc & xd \\ ya & yb \\ yc & yd \end{bmatrix}=\begin{bmatrix} ac & bx \\ cx & dx \\ ay & by \\ cy & dy \end{bmatrix}_{4\times 2}$$

由这个例子可以看出，$A\otimes B$ 与 $B\otimes A$ 一般不是同一矩阵，即 Kronecker 积不满足交换律，但它们的阶数是相同的。

对单位矩阵，有

$$I_n\otimes I_m=I_m\otimes I_n=I_{mn}$$

6．1．　2　Kronecker 积的性质

不难验证，矩阵的 Kronecker 积满足下列运算律：

1. $k(A\otimes B)=kA\otimes B=A\otimes kB,\qquad k\in c$；

2. 分配律 $(A+B)\otimes C=A\otimes C+B\otimes C$；

3. 结合律 $(A\otimes B)\otimes C=A\otimes(B\otimes C)$.

下面我们来研究Kronecker积的另一个重要性质，这条性质对进一步研究Kronecker积有着重要的作用。

定理 1—1 设 $A=(a_{ij})_{m\times n},B=(b_{ij})_{s\times r},C=(c_{ij})_{n\times p},D=(d_{ij})_{r\times t}$，则

$$(A\otimes B)(C\otimes D)=AC\otimes BD \tag{1—1}$$

证 因为

$$\begin{aligned}(A\otimes B)(C\otimes D)&=(a_{ij}B)(c_{ij}D)\\&=(\sum_{k=1}^{n}a_{ik}c_{kj}BD)=((AC)_{ij}BD)\\&=AC\otimes BD\end{aligned}$$

式中（AC）ij 是矩阵 AC 中第 i 行第 j 列的元素。 证毕

推论 若 $A=(a_{ij})_{m\times m},B=(b_{ij})_{n\times n}$，则

$$A\otimes B=(A\otimes I_n)(I_m\otimes B)=(I_m\otimes B)(A\otimes I_n)$$

定理 1—2 设 $A=(a_{ij})_{m\times n},B=(b_{ij})_{p\times q}$，则

$$(A\otimes B)^T=A^T\otimes B^T \tag{1 — 2}$$

$$(A\otimes B)^H=A^H\otimes B^H \tag{1 — 3}$$

证 因为

$$(A\otimes B)^T=(a_{ij}B)^T=\begin{bmatrix}a_{11}B & \cdots & a_{1nB}\\ \vdots & & \vdots\\ a_{m1} & \cdots & a_{mn}B\end{bmatrix}^T$$

$$=\begin{bmatrix}a_{11}B^T & \cdots & a_{m1}B^T\\ \vdots & & \vdots\\ a_{1n}B^T & \cdots & a_{mn}B^T\end{bmatrix}=A^T\otimes B^T$$

同理可证 $(A\otimes B)^H=A^H\otimes B^H$。 证毕

定理 1—3　设 A，B 分别为 m 阶和 n 阶可逆矩阵，则 $A \otimes B$ 也为可逆矩阵，且

$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1} \qquad (1—4)$$

证　由式（1—1）有

$$\begin{aligned}&(A \otimes B)(A^{-1} \otimes B^{-1}) = (AA^{-1} \otimes BB^{-1}) \\ &= I_m \otimes I_n = I_{mn}\end{aligned}$$

即 $$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$$ 证毕

由式（1—2）、(1—4）可见，对于 Kronecker 积，转置和求逆的反序法则不再成立，这也是与通常的矩阵乘法的主要区别之一。

定理 1—4　设 $A=(a_{ij})_{m\times n}, B=(b_{ij})_{p\times q}$，则

$$rank(A \otimes B) = rank(A)rank(B) \qquad (1—5)$$

证　设 A 与 B 的标准形为 A_1 与 B_1，即

$$MAN = A_1 \ , \qquad PBQ = B_1 \qquad (1—6)$$

其中M、N、P、Q 分别为m阶、n 阶、p 阶和 q 阶非奇异矩阵，且

$$A_1 = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix}, \quad B_1 = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix}$$

A_1 中数 1 的个数为 rank（A），B_1 中数 1 的个数为 rank（B）。

由式（1—6）有

$$A = M^{-1}A_1N^{-1} \ , \quad B = P^{-1}B_1Q^{-1}$$

于是，由式（1—1）有

$$\begin{aligned}A \otimes B &= (M^{-1}A_1N^{-1}) \otimes (P^{-1}B_1Q^{-1}) \\ &= (M^{-1} \otimes P^{-1})(A_1 \otimes B_1)(N^{-1} \otimes Q^{-1})\end{aligned}$$

由定理 1—3 知，$M^{-1} \otimes P^{-1}, N^{-1} \otimes Q^{-1}$ 均为非奇异矩阵，故

$$rank(A \otimes B) = rank(A_1 \otimes B_1)$$

而 $A_1 \otimes B_1$ 的秩为 $rank(A)rank(B)$，于是

$$rank(A \otimes B) = rank(A)rank(B)$$ 证毕

定理 1—5　设 $\lambda_1, \lambda_2 \cdots \lambda_m$ 是 $A_{m\times n}$ 的 m 个特征值，$\mu_1, \mu_2, \cdots \mu_p$ 是 $B_{p\times p}$ 的 p 个特征值，那么 $A \otimes B$ 的 mp 个特征值为 $\lambda_i \mu_j (i = 1,2\cdots, m; j = 1,2.\cdots, p)$.

证　由第三章§2 知，A 与 B 一定与 Jordan 标准形相似，即存在可逆矩阵 P 与 Q，使得

$$P^{-1}AP = J_1 = \begin{bmatrix} \lambda & & * \\ & \ddots & \\ 0 & & \lambda_m \end{bmatrix}, \qquad Q^{-1}BQ = J_2 = \begin{bmatrix} \mu_1 & & * \\ & \ddots & \\ 0 & & \mu_p \end{bmatrix}$$

即有

$$A = .P\begin{bmatrix} \lambda & & * \\ & \ddots & \\ 0 & & \lambda_m \end{bmatrix}P^{-1}, \qquad B = Q\begin{bmatrix} \mu_1 & & * \\ & \ddots & \\ 0 & & \mu_p \end{bmatrix}Q^{-1}$$

从而由式（1—1）有

$$A \otimes B = (P \otimes Q)(\begin{bmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_m \end{bmatrix} \otimes \begin{bmatrix} \mu_1 & & * \\ & \ddots & \\ 0 & & \mu_p \end{bmatrix})(P^{-1} \otimes Q^{-1})$$

$$= (P \otimes Q)\begin{bmatrix} \lambda_1\begin{bmatrix} \mu_1 & & * \\ & \ddots & \\ 0 & & \mu_p \end{bmatrix} & & * \\ & \ddots & \\ 0 & & \lambda_m\begin{bmatrix} \mu_1 & & * \\ & \ddots & \\ 0 & & \mu_p \end{bmatrix} \end{bmatrix}(P \otimes Q)^{-1}$$

即有

$$A \otimes B \sim \begin{bmatrix} \lambda_1\mu_1 & & & & & & * \\ & \ddots & & & & & \\ & & \lambda_1\mu_p & & & & \\ & & & \ddots & & & \\ & & & & \lambda_m\mu_1 & & \\ & & & & & \ddots & \\ & & & & & & \lambda_m\mu_p \end{bmatrix}$$

从而 $A \otimes B$ 的 mp 个特征值为 $\lambda_i \mu_j (i = 1, \cdots, m; j = 1, \cdots, p)$， 证毕

定理 1—6 设 A 为 m 阶矩阵，B 为 p 阶矩阵，则

$$\det(A \otimes B) = (\det(A))^p (\det(B))^m \qquad (1—7)$$

证 设 A 与 B 的 Jordan 标准形分别为 J_1 和 J_2，于是存在非奇异矩阵 P 与 Q，有

$$P^{-1}AP = J_1, \quad Q^{-1}BQ = J_2$$

由式（1—1），有

$$A \otimes B = \left(PJ_1OP^{-1}\right) \otimes \left(QJ_2Q^{-1}\right)$$

$$= (P \otimes Q)(J_1 \otimes J_2)(P \otimes Q)^{-1}$$

于是

$$\det(A \otimes B) = \det(J_1 \otimes J_2)$$

显然，当 J_1, J_2 均为下（上）三角矩阵时，$(J_1 \otimes J_2)$ 也为下（上）三角矩阵，故有

$$\det(A \otimes B) = \det(J_1 \otimes J_2) = \prod_{j=1}^{p}(\lambda_1\mu_j)\prod_{j=1}^{p}(\lambda_2\mu_j)\cdots\prod_{j=1}^{p}(\lambda_m\mu_j)$$

$$= \prod_{j=1}^{m}(\lambda_j)\prod_{j=1}^{p}(\mu_j)$$

$$= (\det(A))^p (\det(B))^m$$

其中 $\lambda_1, \lambda_2 \cdots, \lambda_m$ 为 A 的特征值，$\mu_1, \mu_2, \cdots \mu_p$ 是 B 的特征值， 证毕

定理 1—7 （1）若 A，B 均为对角矩阵时，则 $A \otimes B$ 也是对角矩阵；

（2）若 A，B 均为对称矩阵时，则 $A \otimes B$ 也是对称矩阵；

（3）若 A，B 均为 Hermite 矩阵时，则 $A \otimes B$ 也是 Hermite 矩阵；

（4）若 A，B 均为正交（酉）矩阵时，则 $A \otimes B$ 也是正交（酉）矩阵。

定理的证明作为练习。

由例 1—1 我们已看到，Kronecker 积的交换律不成立，即 $A \otimes B$ 一般不等于 $B \otimes A$，但是，我们仍有下面的性质。

定理 1—8　设 A 为 m 阶矩阵，B 为 n 阶矩阵，则有 $A \otimes B$ 相似于 $B \otimes A$。

证　容易验证，对矩阵 $A \otimes I_n$ 进行一系列"相合"变换（对矩阵的行和相应的列进行相同的初等变换，这里是指对调矩阵的第 I 行与第 J 行，然后再对调第 I 列与第 J 列。），可以变成 $I_n \otimes A$，即存在一个 mn 阶置换矩阵（有限个初等矩阵的乘积）P，使

$$P^T\left(A \otimes I_n\right)P = I_n \otimes A$$

同理，对矩阵 $I_m \otimes B$ 也有

$$P^T\left(I_m \otimes B\right)P = B \otimes I_m$$

再由此种初等矩阵的性质知 $P^T P = I$，有

$$P^T\left(A \otimes B\right)P = P^T\left(A \otimes I_n\right)\left(I_m \otimes B\right)P$$

$$= P^T\left(A \otimes I_n\right)PP^T\left(I_m \otimes B\right)P$$

$$= (I_n \otimes A)(B \otimes I_m)$$

$$= B \otimes A$$

证毕

矩阵在 Kronecker 积的意义下也有幂的概念。

定义 1—2　设有矩阵 $A \in C^{m \times n}$，记

$$\underbrace{A \otimes A \otimes \cdots \otimes A}_{k个} = A^{(k)}$$

它是一个 $m^k \times n^k$ 阶矩阵。

定理 1—9 设 $A\in C^{m\times n},B\in C^{n\times p}$，则

$$(AB)^{(k)}=A^{(k)}B^{(k)} \tag{1 — 8}$$

证 用归纳法，当 $k=1$ 时，显然成立，设 $k-1$ 时定理成立，则

$$\begin{aligned}(AB)^{(k)}&=(AB)\otimes(AB)^{(k-1)}\\&=(AB)\otimes A^{(k-1)}B^{(k-1)}\\&=(A\otimes A^{(k-1)})(B\otimes B^{(k-1)})\\&=A^{(k)}B^{(k)}\end{aligned}$$

证毕

关于 Kronecker 积的多项式的特征值问题，我们有下面的结论。

定理 1—10 设 $f(x,y)=\sum_{i,j=0}^{p}a_{ij}x^iy^j$ 是变量 x,y 的复系数多项式，对于 $A\in C^{m\times n},B\in C^{n\times p}$ 定义 $m\ n$ 阶矩阵：

$$f(A;B)=\sum_{i,j=0}^{p}a_{ij}A^i\otimes B^j \tag{1 — 9}$$

如果 A 和 B 的特征值分别是 $\lambda_1,\lambda_2\cdots,\lambda_m$ 和 $\mu_1,\mu_2,\cdots\mu_n$，它们对应的特征向量分别是 $x_1,x_2,\cdots x_m$ 和 $y_1,y_2,\cdots y_n$，则矩阵 $f(A;B)$ 的特征值是 $f(\lambda_r;\mu_s)$，而对应 $f(\lambda_r;\mu_s)$ 的特征向量为 $x_r\otimes y_s\ (r=1,\cdots,m;s=1,\cdots,n)$。

证 由

$$Ax_r=\lambda_rx_r\quad,\quad By_s=\mu_sy_s$$

有

$$A^ix_r=\lambda^i{}_rx_r\quad,\quad B^jy_s=\mu_s^j\quad y_s$$

于是

$$\begin{aligned}f(A;B)x_r\otimes y_s&=(\sum_{i,j=0}^{p}a_{ij}A^i\otimes B^j)(x_r\otimes y_s)\\&=\sum_{i,j=0}^{p}a_{ij}(A^i\otimes B^j)(x_r\otimes y_s)\\&=\sum_{i,j=0}^{p}a_{ij}(A^ix_r\otimes B^jy_s)\end{aligned}$$

$$= \sum_{i,j=0}^{p} a_{ij} \lambda_r^i \ \mu_s^j \ x_r \otimes \ y_s$$

$$= f(\lambda_r, \mu_s) x_r \otimes y_s \qquad \text{证毕}$$

特别地，若取 $f(x,y) = xy$，则有

$$f(A;B) \ = A \otimes B$$

应用本定理，便有定理 1—5 的结论，即

推论 1　$A \otimes B$ 的特征值为 $m \ n$ 个数 $\lambda_r \mu_s \ (r = 1, \cdots, m; s = 1, \cdots, n)$，且对应 $\lambda_r \mu_s$ 的特征向量为 $x_r \otimes y_s$。

若取 $f(x,y) = x + y$，即 $f(x,y) = xy^0 \ + x^0 y$，则

$$f(A;B) \ = A \otimes I_n + I_m \otimes B$$

应用本定理，便有

推论 2　$A \otimes I_n + I_m \otimes B$ 的特征值是 $\lambda_r + \mu_s$，其对应的特征向量是 $x_r \otimes y_s \ (r = 1, \cdots, m; s = 1, \cdots, n)$。

矩阵 $A \otimes I_n + I_m \otimes B$ 称为 A 与 B 的 **Kronecker** 和。

最后，我们还要介绍一个在数理统计中很有用的矩阵。

定义 1—3　元素为 1 或-1 的方阵 $H \in R^{n \times m}$，若有

$$H H^T = nI_n \qquad (1 — 10)$$

则称 H 为 n 阶哈达马矩阵。

定理 1—11　设 H_m 与 H_n 均为哈达马矩阵，则矩阵 $H_m \otimes H_n$ 为 $m \ n$ 阶的哈达马矩阵。

证　因为

$$(H_m \otimes H_n)(H_m \otimes H_n)^T = (H_m \otimes H_n)(H_m{}^T \otimes H_n{}^T)$$

$$
\begin{aligned}
&= (H_m H_m^{\ T}) \otimes (H_n H_n^{\ T}) = (mI_m) \ \otimes (nI_n) \\
&= mnI_{mn}
\end{aligned}
$$

故按定义，$H_m \otimes H_n$ 为 $m \ n$ 阶的哈达马矩阵。　　　　　　　　证毕

本节讨论的 Kronecker 积，特别是哈达马矩阵在数理统计中应用很广。

6．2 Kronecker 积应用举例

6．2．1 线性矩方程

利用矩阵 Kronecker 积的性质，能够方便地研究一般线性矩方程

$$A_1XB_1 + A_2XB_2 + \cdots + A_pXB_p = C \qquad （2 —1）$$

的相容性及其解法等问题，这里 $A_i \in C^{m\times m}, B_i \in C^{n\times n}, C \in C^{m\times n}$ 为已知矩阵，$X \in C^{m\times n}$ 是未知矩阵。

对于矩阵方程（2—1）可以转化为通常的线性方程组

$$Gx = c \qquad （2 — 2）$$

来讨论，其中系数矩阵 G 与 A、B 有头，向量 x 与矩阵 X 有头，向量 c 与矩阵 C 有头，为此，先引入下面矩阵拉直的概念。

矩阵的拉直

定义 2—1　设 $A = (a_{ij})_{m\times n}$，将 A 的各行依次按列纵排得到的 $m \ n$ 维列向量，这种运算称为 A 的拉直，记为 $\vec{A}$，即

$$\vec{A} = (a_{11}, a_{12}, \cdots, a_{1n}, a_{21}, a_{22}, \cdots, a_{2n}, \cdots, a_{m1}, a_{m2}, \cdots, a_{mn})^T \qquad （2 — 3）$$

从定义 2—1 可看出，A 是 $mn \times 1$ 阶矩阵，即为一个列向量，这个列向量先把 A 的第一行按顺序写在前面，依次再写第二行，…，最后写第 m 行。

例 2—1　设 $A = \begin{bmatrix} 1 & -1 \\ 3 & 1 \end{bmatrix}$，　　则　$\vec{A} = (1,-1,3,1)^T$

定理 2—1　拉直算子是线性的，即

$$\overrightarrow{A+B} = \vec{A} + \vec{B}, \qquad \overrightarrow{kA} = k\vec{A}$$

这些都是显然的。

定理 2—2

1. $\overrightarrow{xy^T} = x \otimes y$, 其中 x, y 为 n 维列向量；

2. $E_{ij} = e_i e_j^T$, 其中 E_{ij} 表示 (i, j) 元素为 1，其余元素为 0 的 $m \times n$ 阶矩阵，e_i 表示第 i 个元素为 1，其余元素为 0 的列向量；

4. $\vec{E}_{ij} = e_i \otimes e_j$.

4. $Ae_i = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{bmatrix}$ ；

5. $\overrightarrow{e_j A} = (a_{j1}, a_{j2}, \cdots, a_{jn})^T$ ；

定理 2—3　设 $A = (a_{ij})_{m \times n}, B = (b_{ij})_{n \times p}, C = (c_{ij})_{p \times q}$，则

$$\overrightarrow{ABC} = (A \otimes C^T)\vec{B} \qquad (2—4)$$

证　证明分两步，先证

$$\overrightarrow{AE_{ij}C} = (A \otimes C^T)\vec{E}_{ij} \qquad (2—5)$$

其中 E_{ij} 为 $n \times p$ 阶矩阵。

事实上，

$$\overrightarrow{AE_{ij}C} = \overrightarrow{Ae_i e_j^T C} = \overrightarrow{Ae_j (C^T e_j)^T} = Ae_i \otimes C^T e_j$$

另一方面，有

$$(A \otimes C^T)\vec{E}_{ij} = (A \otimes C^T)(e_i \otimes e_j) = Ae_i \otimes C^T e_j$$

即证明了式（2—5），下面再证明式（2—4），由于

$$B = (b_{ij}) = \sum_{i=1}^{n}\sum_{j=1}^{p} b_{ij} E_{ij}$$

所以

$$\overrightarrow{ABC} = \overrightarrow{A(\sum_{i=1}^{n}\sum_{j=1}^{p} b_{ij}E_{ij})C}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{p} b_{ij}\,\overrightarrow{AE_{ij}C}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{p} b_{ij}\,(A \otimes C^{T})\vec{E}_{ij}$$

$$= (A \otimes C^{T})\sum_{i=1}^{n}\sum_{j=1}^{p} b_{ij}\vec{E}_{ij}$$

$$= (A \otimes C^{T})\vec{B}$$

证毕

推论　设 $A = A_{m\times m}, B = B_{n\times n}, X = X_{m\times n}$，则

1. $\overrightarrow{AX} = (A \otimes I_n)\vec{X}$

2. $\overrightarrow{XB} = (I_m \otimes B^{T})\vec{X}$

3. $\overrightarrow{AX + XB} = (A \otimes I_n + I_m \otimes B^{T})\vec{X}$

线性矩阵方程的解

定理2—4　矩阵 $X \in C^{m\times n}$ 是矩阵方程（2—1）的解的充分必要条件是 $x = \vec{X}$ 为通常的线性方程组

$$Gx = c \qquad (2-6)$$

的解，其中 $G = \sum_{i=1}^{p} A_i \otimes B_i^{T}$ ，$c = \vec{C}$ 。

证　对矩阵方程（2—1）两端拉直，有

$$\vec{C} = \overrightarrow{\sum_{i=1}^{p} A_i X B_i} = \sum_{i=1}^{p} \overrightarrow{A_i X B_i}$$

$$= \sum_{i=1}^{p} (A_i \otimes B_i^{T})\vec{X}$$

$$= G\vec{X}$$

即 $Gx = c$ ，故矩阵方程（2—1）的解与通常的线性方程组（2—6）的相同，证毕。

这样，欲求矩阵方程（2—1）的解，只要将它转化为通常的线性方程组（2—6）求解就行了。

推论 1　矩阵方程（2—1）有解（相容）的充要条件是

$$\mathrm{rank}（G|c）=\mathrm{rank}（G）$$

推论 2　矩阵方程（2—1）有唯一解的充要条件是 G 为非奇异的。

下面我们来讨论矩阵方程（2—1）两个重要的特殊情况。

1.　设 $A \in C^{m\times m}, B \in C^{n\times n}, C \in C^{m\times n}$，方程

$$AX+XB=C \qquad (2-7)$$

定理 2—5　矩阵方程（2—7）有唯一解 $X \in C^{m\times n}$ 的充要条件有 A 和-B 没有相同的特征值，即

$$\lambda_i + \mu_j \neq 0 \quad (i=1,\cdots,m; j=1,\cdots,n) \qquad (2-8)$$

证　将矩阵方程（2—7）两端拉直，并利用定理 2—3 推论（3）的结论知，方程（2—7）等价于

$$(A\otimes I_n + I_m \otimes B^T)\vec{X} = \vec{C} \qquad (2-9)$$

再由定理 2—4 的推论 2 知，方程（2—7）有唯一解的充要条件是矩阵 $A\otimes I_n + I_m \otimes B^T$ 是非奇异的，即矩阵 $A\otimes I_n + I_m \otimes B^T$ 没有零特征值。

如果设 A 的特征值为 $\lambda_1,\lambda_2\cdots,\lambda_m$，B（或 B^T）的特征值为 $\mu_1,\mu_2,\cdots\mu_n$，则由定理 1—10 知，矩阵 $A\otimes I_n + I_m \otimes B^T$ 的特征值为 $\lambda_i + \mu_j \quad (i=1,\cdots,m; j=1,\cdots,n)$，于是方程（2—7）有唯一解的充要条件是 $\lambda_i + \mu_j \neq 0$，即 A 与-B 没有相同的特征值。

推论　设 $A \in C^{m\times m}$ 则矩阵方程 $AX - XA = 0 \quad (即 AX = XA)$ 必要非零解 $X \in C^{m\times n}$。

2.　设 $A \in C^{m\times m}, B \in C^{n\times n}, C \in C^{m\times n}$，方程

$$X + AXB = C \qquad (2-10)$$

定理 2—6　矩阵方程（2—10）有唯一解 $X \in C^{m\times n}$ 的充要条件是

$\lambda_i\mu_j \neq -1(i=1,2,...,m; j=1,2,...,n)$，$\lambda_i$和$\mu_j$分别为 A 与 B 的特征值。

证　把方程（2—10）两端拉直，有

$$\vec{C} = \overrightarrow{I_m X I_n + AXB} = (I_m \otimes I_n + A \otimes B^T)\vec{X}$$

于是方程（2—10）有唯一解的充要条件是矩阵$_m \otimes I_n + A \otimes B^T$的特征值全不为零，由定理 1—10 知

$$1+\lambda_i\mu_j \neq 0$$

证毕

6．2．2 矩阵函数积的导数

令$X=\left(x_{pq}\right)_{m\times n}, F(X)=\left(f_{ij}(X)\right)_{s\times t}, f_{ij}(X)$为 X 的函数，则 F 关于矩阵 X 的导数为

$$\frac{dF}{dX}=\left(\frac{\partial F}{\partial x_{pq}}\right)_{m\times n},$$

其中

$$\frac{\partial F}{\partial x_{pq}}=\begin{pmatrix} \frac{\partial f_{11}}{\partial x_{pq}} & \cdots & \frac{\partial f_{1t}}{\partial x_{pq}} \\ \vdots & & \vdots \\ \frac{\partial f_{s1}}{\partial x_{pq}} & & \frac{\partial f_{st}}{\partial x_{pq}} \end{pmatrix}=\left(\frac{\partial f_{ij}}{\partial x_{pq}}\right)_{s\times t}$$

定理 2—4 设$X=\left(x_{pq}\right)_{m\times n}, F(X)=\left(f_{ij}(X)\right)_{s\times r}, G(X)=\left(g_{ij}(X)\right)_{r\times t}$，则

$$\frac{d[FG]}{dX}=\frac{dF}{dX}(I_n\otimes G)+(I_m\times F)\frac{dG}{dX} \qquad (2\text{----}11)$$

证 令$A(X)=F(X)\,G(X)$，则

$$a_{ij}(X)=\sum_{k=1}^{r} f_{ik}(X)g_{kj}(X), i=1,2,...,s; j=1,2,...,t$$

由导数定义，我们有

$$\frac{dA}{dX}=\left(\frac{\partial A}{\partial x_{pq}}\right)_{m\times n}$$

$$\frac{\partial A}{\partial x_{pq}}=\left(\frac{\partial a_{ij}}{\partial x_{pq}}\right)_{m\times n}=\left(\sum_{k=1}^{r}\frac{\partial(f_{ik}g_{kj})}{\partial x_{pq}}\right)_{s\times t}=\left(\sum_{k=1}^{r}\frac{\partial f_{ik}}{\partial x_{pq}}\cdot g_{kj}\right)_{s\times t}+\left(\sum_{k=1}^{r}f_{ik}\cdot\frac{\partial g_{kj}}{\partial x_{pq}}\right)_{s\times t}$$

$$= \frac{\partial F}{\partial x_{pq}} \cdot G + F \frac{\partial G}{\partial x_{pq}}$$

所以

$$\frac{dA}{dX} = \left(\frac{\partial F}{\partial x_{pq}} \cdot G \right)_{m\times n} + \left(F \cdot \frac{\partial G}{\partial x_{pq}} \right)_{m\times n}$$

$$= \begin{pmatrix} \frac{\partial F}{\partial x_{11}} \cdot G & \cdots & \frac{\partial F}{\partial x_{1n}} \cdot G \\ \vdots & & \vdots \\ \frac{\partial F}{\partial x_{m1}} \cdot G & \cdots & \frac{\partial F}{\partial x_{mn}} \cdot G \end{pmatrix} + \begin{pmatrix} F \cdot \frac{\partial G}{\partial x_{11}} & \cdots & F \cdot \frac{\partial G}{\partial x_{1n}} \\ \vdots & & \vdots \\ F \cdot \frac{\partial G}{\partial x_{m1}} & \cdots & F \cdot \frac{\partial G}{\partial x_{mn}} \end{pmatrix}$$

$$= \begin{pmatrix} \frac{\partial F}{\partial x_{11}} & \cdots & \frac{\partial F}{\partial x_{1n}} \\ \vdots & & \vdots \\ \frac{\partial F}{\partial x_{m1}} & \cdots & \frac{\partial F}{\partial x_{mn}} \end{pmatrix} \begin{pmatrix} G & & \\ & \ddots & \\ & & G \end{pmatrix} + \begin{pmatrix} F & & \\ & \ddots & \\ & & F \end{pmatrix} \begin{pmatrix} \frac{\partial G}{\partial x_{11}} & \cdots & \frac{\partial G}{\partial x_{1n}} \\ \vdots & & \vdots \\ \frac{\partial G}{\partial x_{m1}} & \cdots & \frac{\partial G}{\partial x_{mn}} \end{pmatrix}$$

$$= \frac{dF}{dX}(I_n \otimes G) + (I_m \otimes F)\frac{dG}{dX}$$

因此

$$\frac{d[FG]}{dX} = \frac{dF}{dX}(I_n \otimes G) + (I_m \times F)\frac{dG}{dX}$$

例 **2-1** 令 $f(\boldsymbol{x}) = \boldsymbol{x}^T A \boldsymbol{x}$，其中 $\boldsymbol{x} = (x_1, x_2, \ldots x_n)^T, A = \left(a_{ij}\right)_{n\times n}$，求 $\frac{df}{d\boldsymbol{x}}$。

解：$\frac{df}{d\boldsymbol{x}} = \frac{d\boldsymbol{x}^T}{d\boldsymbol{x}}(I_1 \otimes A\boldsymbol{x}) + (I_n \otimes \boldsymbol{x}^T)\frac{dA\boldsymbol{x}}{d\boldsymbol{x}}$

$$\frac{d\boldsymbol{x}^T}{d\boldsymbol{x}} = \begin{pmatrix} \frac{\partial \boldsymbol{x}^T}{\partial x_1} \\ \vdots \\ \frac{\partial \boldsymbol{x}^T}{\partial x_n} \end{pmatrix} = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} = I_n，(I_1 \otimes A\boldsymbol{x}) = A\boldsymbol{x}$$

$$\frac{d\boldsymbol{x}^T}{d\boldsymbol{x}}(I_1 \otimes A\boldsymbol{x}) = A\boldsymbol{x}$$

$$(I_n \otimes \boldsymbol{x}^T) = \begin{pmatrix} \boldsymbol{x}^T & & \\ & \ddots & \\ & & \boldsymbol{x}^T \end{pmatrix}，\quad \frac{dA\boldsymbol{x}}{d\boldsymbol{x}} = \begin{pmatrix} \frac{\partial A\boldsymbol{x}}{\partial x_1} \\ \vdots \\ \frac{\partial A\boldsymbol{x}}{\partial x_n} \end{pmatrix} = \begin{pmatrix} \frac{\partial \sum_{j=1}^{n} a_{1j} x_j}{\partial x_1} \\ \vdots \\ \frac{\partial \sum_{j=1}^{n} a_{nj} x_j}{\partial x_1} \\ \vdots \\ \frac{\partial \sum_{j=1}^{n} a_{1j} x_j}{\partial x_n} \\ \vdots \\ \frac{\partial \sum_{j=1}^{n} a_{nj} x_j}{\partial x_n} \end{pmatrix} = \begin{pmatrix} \boldsymbol{a}_1 \\ \vdots \\ \boldsymbol{a}_n \end{pmatrix}$$

$$(I_n \otimes \boldsymbol{x}^T)\frac{dA\boldsymbol{x}}{d\boldsymbol{x}} = \begin{pmatrix} \boldsymbol{x}^T & & \\ & \ddots & \\ & & \boldsymbol{x}^T \end{pmatrix} \begin{pmatrix} \boldsymbol{a}_1 \\ \vdots \\ \boldsymbol{a}_n \end{pmatrix} = \begin{pmatrix} \boldsymbol{x}^T \boldsymbol{a}_1 \\ \vdots \\ \boldsymbol{x}^T \boldsymbol{a}_n \end{pmatrix} = A^T \boldsymbol{x}$$

所以

$$\frac{df}{d\boldsymbol{x}} = \frac{d\boldsymbol{x}^T}{d\boldsymbol{x}}(I_1 \otimes A\boldsymbol{x}) + (I_n \otimes \boldsymbol{x}^T)\frac{dA\boldsymbol{x}}{d\boldsymbol{x}} = (A + A^T)\boldsymbol{x}$$

例 **2-2** 令 $f(A) = \boldsymbol{x}^T A\boldsymbol{x}$，其中 $\boldsymbol{x} = (x_1, x_2, \ldots x_n)^T, A = (a_{ij})_{n\times n}$，求 $\frac{df}{dA}$。

解：$\frac{df}{dA} = \frac{d\boldsymbol{x}^T}{dA}(I_n \otimes A\boldsymbol{x}) + (I_n \otimes \boldsymbol{x}^T)\frac{dA\boldsymbol{x}}{d\boldsymbol{x}}$

$$\frac{d\boldsymbol{x}^T}{dA} = 0,$$

$$\frac{d\boldsymbol{x}^T}{dA}(I_n \otimes A\boldsymbol{x}) = 0$$

$$\frac{dA\boldsymbol{x}}{dA} = \left(\frac{\partial A\boldsymbol{x}}{\partial a_{ij}}\right)_{n\times n}，\quad \frac{\partial A\boldsymbol{x}}{\partial a_{ij}} = x_j \boldsymbol{e}_i,$$

$$\frac{dA\boldsymbol{x}}{dA} = \begin{pmatrix} x_1\boldsymbol{e}_1 & x_2\boldsymbol{e}_1 & \cdots & x_n\boldsymbol{e}_1 \\ x_1\boldsymbol{e}_2 & x_2\boldsymbol{e}_2 & \cdots & x_n\boldsymbol{e}_2 \\ \vdots & \vdots & & \vdots \\ x_1\boldsymbol{e}_n & x_2\boldsymbol{e}_n & \cdots & x_n\boldsymbol{e}_n \end{pmatrix}$$

$$(I_n \otimes \boldsymbol{x}^T)\frac{dA\boldsymbol{x}}{d\boldsymbol{x}} = \begin{pmatrix} \boldsymbol{x}^T & & & \\ & \boldsymbol{x}^T & & \\ & & \ddots & \\ & & & \boldsymbol{x}^T \end{pmatrix} \begin{pmatrix} x_1\boldsymbol{e}_1 & x_2\boldsymbol{e}_1 & \cdots & x_n\boldsymbol{e}_1 \\ x_1\boldsymbol{e}_2 & x_2\boldsymbol{e}_2 & \cdots & x_n\boldsymbol{e}_2 \\ \vdots & \vdots & & \vdots \\ x_1\boldsymbol{e}_n & x_2\boldsymbol{e}_n & \cdots & x_n\boldsymbol{e}_n \end{pmatrix}$$

$$= \begin{pmatrix} x_1\boldsymbol{x}^T\boldsymbol{e}_1 & x_2\boldsymbol{x}^T\boldsymbol{e}_1 & \cdots & x_n\boldsymbol{x}^T\boldsymbol{e}_1 \\ x_1\boldsymbol{x}^T\boldsymbol{e}_2 & x_2\boldsymbol{x}^T\boldsymbol{e}_2 & \cdots & x_n\boldsymbol{x}^T\boldsymbol{e}_2 \\ \vdots & \vdots & & \vdots \\ x_1\boldsymbol{x}^T\boldsymbol{e}_n & x_2\boldsymbol{x}^T\boldsymbol{e}_n & \cdots & x_n\boldsymbol{x}^T\boldsymbol{e}_n \end{pmatrix} = \begin{pmatrix} x_1\boldsymbol{x} & x_2\boldsymbol{x} & \cdots & x_n\boldsymbol{x} \end{pmatrix} = \boldsymbol{x}\boldsymbol{x}^T$$

所以

$$\frac{df}{dA} = \boldsymbol{x}\boldsymbol{x}^T$$

例 **2-3** 令 $f(A) = A^T A$，其中 $A = \left(a_{ij}\right)_{n\times n}$，求 $\dfrac{df}{dA}$。

解：$\dfrac{df}{dA} = \dfrac{dA^T}{dA}(I_n \otimes A) + (I_n \otimes A^T)\dfrac{dA}{dA}$

$$\frac{dA}{dA} = \left(\frac{\partial A}{\partial a_{ij}}\right) = \left(E_{ij}\right)_{n\times n}$$

$$\frac{dA^T}{dA} = \left(\frac{\partial A^T}{\partial a_{ij}}\right) = \left(E_{ji}\right)_{n\times n},$$

$$\frac{df}{dA} = \left(E_{ji}\right)_{n\times n}(I_n \otimes A) + (I_n \otimes A^T)\left(E_{ij}\right)_{n\times n} = \left(E_{ji}A + A^T E_{ij}\right)_{n\times n}$$

例 **2-4** 求 $\dfrac{dA^{-1}}{dA}$。

解：$A^{-1}A = I$，$\dfrac{dA^{-1}}{dA}(I_n \otimes A) + (I_n \otimes A^{-1})\dfrac{dA}{dA} = 0$

$$\frac{dA^{-1}}{dA} = -(I_n \otimes A^{-1})\frac{dA}{dA}(I_n \otimes A)^{-1} = -(I_n \otimes A^{-1})\left(E_{ij}\right)_{n\times n}(I_n \otimes A^{-1})$$

$$= -\left(A^{-1}E_{ij}A^{-1}\right)_{n\times n}$$

例 **2-5** $f(A)=\boldsymbol{x}^T A^{-1}\boldsymbol{x}$，求 $\dfrac{df}{dA}$。

解 $$\frac{df}{dA}=\frac{dx}{dA}(I\otimes A^{-1}x)+(I\otimes x^T)\frac{dA^{-1}x}{dA}$$

$$=(I\otimes x^T)\left[\frac{dA^{-1}}{dA}(I\otimes x)+(I\times A^{-1})\frac{dx}{dA}\right]$$

$$=(I\otimes x^T)\frac{dA^{-1}}{dA}(I\otimes x)$$