第一章　　矩阵代数

1.1 分块矩阵

1.2 广义逆

1.3 拉直运算和**Kronecker**积

1.4 矩阵的微商

§1.1 分块矩阵

定义1.1.1

若$A=\left(a_{ij}\right)$为$p\times q$阶矩阵，分成四块，使 得

$$A_{11}:\qquad k\times l,\quad A_{12}:\qquad k\times\left(q-l\right),$$

$$A_{21}:\left(p-k\right)\times l,\quad A_{22}:\left(p-k\right)\times\left(q-l\right),$$

则$A=\begin{pmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{pmatrix}$称为矩阵$A$的分块表示形式.

性质**1**

若A和B有相同的分块，则

$$A + B = \begin{pmatrix} A_{11} + B_{11} & A_{12} + B_{12} \\ A_{21} + B_{21} & A_{22} + B_{22} \end{pmatrix}$$

性质2

若C为$q\times r$矩阵，它分为$C=\begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}$，其中

$C_{11}: l\times m$，$C_{12}: l\times(r-m)$，$C_{21}:(q-l)\times m$，

$C_{22}:(q-l)\times(r-m)$，则

$$AC=\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}\begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}$$

$$=\begin{pmatrix} A_{11}C_{11}+A_{12}C_{21} & A_{11}C_{12}+A_{12}C_{22} \\ A_{21}C_{11}+A_{22}C_{21} & A_{21}C_{12}+A_{22}C_{22} \end{pmatrix}$$

性质3

若A为方阵，A_{11}也为方阵，

(1) 若$|A_{11}| \neq 0$，则 $|A| = |A_{11}||A_{22\cdot 1}|$

其中 $A_{22\cdot 1} = A_{22} - A_{21}A_{11}^{-1}A_{12}$

(2) 若$|A_{22}| \neq 0$，则 $|A| = |A_{11\cdot 2}||A_{22}|$

其中 $A_{11\cdot 2} = A_{11} - A_{12}A_{22}^{-1}A_{21}$

证明：(1) $|A_{11}| \neq 0$, **利用**

$$\begin{pmatrix} I & 0 \\ -A_{21}A_{11}^{-1} & I \end{pmatrix}\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}\begin{pmatrix} I & -A_{11}^{-1}A_{12} \\ 0 & I \end{pmatrix}$$

$$=\begin{pmatrix} A_{11} & 0 \\ 0 & A_{22\cdot 1} \end{pmatrix} \qquad (1.1)$$

两边同时取行列式即可.

(2) $\left|A_{22}\right| \neq 0$, 利用

$$\begin{pmatrix} I & -A_{12}A_{22}^{-1} \\ 0 & I \end{pmatrix}\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}\begin{pmatrix} I & 0 \\ -A_{22}^{-1}A_{21} & I \end{pmatrix}$$

$$= \begin{pmatrix} A_{11\cdot 2} & 0 \\ 0 & A_{22} \end{pmatrix} \qquad (1.2)$$

两边同时取行列式即可.

性质4

若A为可逆方阵，A_{11}和A_{22}均为方阵

(1) 若$\left|A_{11}\right| \neq 0$，则

$$A^{-1}=\begin{pmatrix} A_{11}^{-1}+A_{11}^{-1}A_{12}A_{22\cdot 1}^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22\cdot 1}^{-1} \\ -A_{22\cdot 1}^{-1}A_{21}A_{11}^{-1} & A_{22\cdot 1}^{-1} \end{pmatrix}$$

(2) 若$\left|A_{22}\right| \neq 0$，则

$$A^{-1}=\begin{pmatrix} A_{11\cdot 2}^{-1} & -A_{11\cdot 2}^{-1}A_{12}A_{22}^{-1} \\ -A_{22}^{-1}A_{21}A_{11\cdot 2}^{-1} & A_{22}^{-1}+A_{22}^{-1}A_{21}A_{11\cdot 2}^{-1}A_{12}A_{22}^{-1} \end{pmatrix}$$

性质4（续）

(3) 若 $\left|A_{11}\right| \neq 0$, $\left|A_{22}\right| \neq 0$, 则

$$A^{-1} = \begin{pmatrix} A_{11\cdot 2}^{-1} & -A_{11}^{-1}A_{12}A_{22\cdot 1}^{-1} \\ -A_{22}^{-1}A_{21}A_{11\cdot 2}^{-1} & A_{22\cdot 1}^{-1} \end{pmatrix}$$

证明: (1)对(1.1)式两边求逆

$$\begin{pmatrix} I & -A_{11}^{-1}A_{12} \\ 0 & I \end{pmatrix}^{-1} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}^{-1} \begin{pmatrix} I & 0 \\ -A_{21}A_{11}^{-1} & I \end{pmatrix}^{-1} = \begin{pmatrix} A_{11}^{-1} & 0 \\ 0 & A_{22\cdot1}^{-1} \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}^{-1} = \begin{pmatrix} I & -A_{11}^{-1}A_{12} \\ 0 & I \end{pmatrix} \begin{pmatrix} A_{11}^{-1} & 0 \\ 0 & A_{22\cdot1}^{-1} \end{pmatrix} \begin{pmatrix} I & 0 \\ -A_{21}A_{11}^{-1} & I \end{pmatrix}$$

$$= \begin{pmatrix} A_{11}^{-1} + A_{11}^{-1}A_{12}A_{22\cdot1}^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22\cdot1}^{-1} \\ A_{22\cdot1}^{-1}A_{21}A_{11}^{-1} & A_{22\cdot1}^{-1} \end{pmatrix}$$

§1.2 矩阵的广义逆

一、减号逆

定义1.2.1 设A为$n\times p$阶矩阵，若存在一个$p\times n$阶矩阵X，使得 $AXA=A$，则称X为A的广义逆或A的减号逆，记作$X=A^{-}$.

性质1 任何矩阵的广义逆一定存在，但可能不唯一。

证明：设$rank(A)=r$，$\underset{n\times p}{A}$可分解为

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

其中P和Q分别为n和p阶非奇异方阵，于是有

$$AXA=A\Leftrightarrow$$

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

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} QXP \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

记

$$QXP = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \quad \text{其中} T_{11} : r \times r$$

于是

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

$$\Leftrightarrow T_{11} = I_r$$

所以$QXP=\begin{pmatrix} I_r & T_{12} \\ T_{21} & T_{22} \end{pmatrix}$

即$X=Q^{-1}\begin{pmatrix} I_r & T_{12} \\ T_{21} & T_{22} \end{pmatrix}P^{-1}$

其中T_{12}，T_{21}，T_{22}可任意选择，X就是A的广义逆A^{-}，

这表明A^{-}一定存在，但可能不唯一.

性质2 若A非退化，则A^-唯一，且$A^- = A^{-1}$.

证明：因为$AA^{-1}A = A$，所以A^{-1}是A的一个广义逆. 若X也是A的广义逆，则 $AXA = A$ $\Leftrightarrow X = A^{-1}$，唯一性得证.

性质3 $\quad rk(A^-) \geq rk(A)$

证明 由性质1中知，若 $rank(A)=r$，$\underset{n\times p}{A}$ 可分解为

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

其中P和Q分别为n和p阶非奇异方阵，于是有

$$A^- = Q^{-1}\begin{pmatrix} I_r & T_{12} \\ T_{21} & T_{22} \end{pmatrix}P^{-1}$$

所以 $rk(A^-) \geq rk(A)$.

性质4

$$rk(A) = rk(AA^{-}) = rk(A^{-}A)$$

$$= tr(AA^{-}) = tr(A^{-}A)$$

证明 若$rank(A)=r$，$\underset{n\times p}{A}$可分解为

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

其中P和Q分别为n和p阶非奇异方阵，

于是有

$$A^{-}=Q^{-1}\begin{pmatrix} I_r & T_{12} \\ T_{21} & T_{22} \end{pmatrix}P^{-1}$$

$$\Rightarrow AA^{-} = P\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} QQ^{-1} \begin{pmatrix} I_r & T_{12} \\ T_{21} & T_{22} \end{pmatrix} P^{-1}$$

$$= P\begin{pmatrix} I_r & T_{12} \\ 0 & 0 \end{pmatrix} P^{-1}$$

$$\Rightarrow (AA^{-})^2 = AA^{-}$$

所以AA^{-}为幂等阵

同理

$$A^{-}A = Q^{-1}\begin{pmatrix} I_r & T_{12} \\ T_{21} & T_{22} \end{pmatrix} P^{-1}P\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} Q$$

$$= Q^{-1}\begin{pmatrix} I_r & 0 \\ T_{21} & 0 \end{pmatrix} Q$$

$$\Rightarrow (A^{-}A)^2 = A^{-}A$$

所以AA^{-}和$A^{-}A$均为幂等阵

特别，

$$若\ rk(\underset{n\times p}{A}) = p，则\ A^{-}A = I_p$$

$$若\ rk(\underset{n\times p}{A}) = n，则\ AA^{-} = I_n$$

性质5

对任意矩阵A，有

$$A'A(A'A)^{-}A' = A' \quad A(A'A)^{-}A'A = A$$

$$A'A(A'A)^{-}A'A = A'A$$

证明

$$(1)Ax=0 \Leftrightarrow A'Ax=0$$

事实上,

$$Ax=0 \Rightarrow A'Ax=0$$

$$\Rightarrow x'A'Ax=0 \Rightarrow Ax=0$$

(2) $Ax = Ay \Leftrightarrow A'Ax = A'Ay$

(3)对$\forall x \in R^p$，有

$A'A(A'A)^- A'Ax = A'Ax$，利用(2)

$\Rightarrow$ 对$\forall x \in R^p$，有$A(A'A)^- A'Ax = Ax$，

$\Rightarrow A(A'A)^- A'A = A$

同理可证 $A'A(A'A)^- A' = A'$

等号两边同时消去一个相同矩阵的一般结论：

定理

(1) $ABC = 0$ 和 $BC = 0$ 等价的充分必要条件是 $rk(AB) = rk(B)$.

(2) $CAB = 0$ 和 $CA = 0$ 等价的充分必要条件是 $rk(AB) = rk(A)$.

性质6

$A(A'A)^{-}A'$是一个投影矩阵，

且与$(A'A)^{-}$的取法无关.

证明

(1) $A(A'A)^{-}A'$与$(A'A)^{-}$的取法无关.

设$rk(\underset{n\times p}{A})=r$，$(A'A)_1^{-}$和$(A'A)_2^{-}$是$A'A$的两个广义

逆, 则有$A'A(A'A)_1^{-}A'A = A'A(A'A)_2^{-}A'A = A'A$

利用性质5$\Rightarrow A'A(A'A)_1^{-}A' = A'A(A'A)_2^{-}A' = A'$

$\Rightarrow A(A'A)_1^{-}A' = A(A'A)_2^{-}A'$

(2) $A(A'A)^{-}A'$一定是对称阵.

由于$A'A$为对称阵，存在正交阵P，使得

$$A'A = P'diag(\lambda_1^2,\ \lambda_2^2,\cdots\lambda_r^2,0,\cdots,0)P$$

$$= P'diag(\lambda_1,\ \lambda_2,\cdots\lambda_r,1,\cdots,1)\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}\underbrace{diag(\lambda_1,\ \lambda_2,\cdots\lambda_r,1,\cdots,1)P}_{Q}$$

$$= Q'\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}Q \text{ 其中}\lambda_i \neq 0, \text{取}(A'A)^{-} = Q^{-1}\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}\left(Q^{-1}\right)',$$

它是对称阵. 所以 $A(A'A)^{-}A'$一定是对称阵.

(3)$A(A'A)^{-}A'$是幂等阵.

$$\left(A(A'A)^{-}A'\right)^{2}=\underline{A(A'A)^{-}A'A}(A'A)^{-}A'$$

$$=A(A'A)^{-}A'$$

故$A(A'A)^{-}A'$是一个投影矩阵

性质7

$A(A'A)^{-}A'$是到$R(A)$空间的投影矩阵，

记作$P_A = A(A'A)^{-}A'$

(1)$\forall x \in R(A)$，有$P_A x = x$.

(2)对$\forall x \in R^n, P_A x \in R(A)$.

证明

(1)$\forall x \in R(A)$，有$P_A x = x$.

事实上

$\forall x \in R(A)$，存在$y \in R^p$，使得$x = Ay$.

$$P_A x = A(A'A)^- A'Ay = Ay = x$$

(2)对$\forall x \in R^n, P_A x \in R(A)$.

故$P_A = A(A'A)^- A'$是到$R(A)$空间的投影矩阵

二、加号逆

定义1.2.2 设A是一个$n\times p$阶矩阵，若存在一个$p\times n$阶矩阵X，使得

$$\begin{cases} AXA = A & XAX = X \\ (AX)' = AX & (XA)' = XA \end{cases}$$

则称X是A的加号逆，也称$Moore-Penrose$逆，记作$X = A^{+}$.

性质1

对任何矩阵$\boldsymbol{A}$，$\boldsymbol{A}^+$存在且唯一.

(1) 存在性

(2) 唯一性

证明 (存在性) 令$rk(\underset{n\times p}{A}) = r$,

(1)当$r = 0$时，必有$A = 0$，这时取$A^{+} = 0$即可.

(2)当$r > 0$时，则存在可逆阵P_1和Q_1使得

$$A = P_1\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} Q_1 = P_1\begin{pmatrix} I_r \\ 0 \end{pmatrix}\begin{pmatrix} I_r & 0 \end{pmatrix} Q_1 \triangleq PQ'$$

其中$P = P_1\begin{pmatrix} I_r \\ 0 \end{pmatrix}$，$Q' = \begin{pmatrix} I_r & 0 \end{pmatrix} Q_1$，且$P$和$Q$均为列满秩矩阵. 所以$P'P$和$Q'Q$均可逆.

令$X = Q(Q'Q)^{-1}(P'P)^{-1}P'$，我们来验证$X$满足加号逆的四个条件.

$$(1)\ AXA = PQ'Q(Q'Q)^{-1}(P'P)^{-1}P'PQ' = PQ' = A$$

$$(2)\ XAX = Q(Q'Q)^{-1}(P'P)^{-1}P'PQ'Q(Q'Q)^{-1}(P'P)^{-1}P'$$

$$= Q(Q'Q)^{-1}(P'P)^{-1}P' = X$$

$$(3)\ AX = PQ'Q(Q'Q)^{-1}(P'P)^{-1}P' = P(P'P)^{-1}P' \quad \text{对称}$$

$$(4)\ XA = Q(Q'Q)^{-1}(P'P)^{-1}P'PQ' = Q(Q'Q)^{-1}Q' \quad \text{对称}$$

故X是A的加号逆，存在性得证.

(唯一性)

若A_1^+和A_2^+是A的两个加号逆，则

$$A_1^+ = A_1^+ A A_1^+ = A_1^+ (A A_1^+)' = A_1^+ (A_1^+)' A' = A_1^+ (A_1^+)' (A A_2^+ A)'$$

$$= A_1^+ (A_1^+)' A' (A A_2^+)' = A_1^+ (A A_1^+)' (A A_2^+)' = A_1^+ A A_1^+ A A_2^+$$

$$= A_1^+ A A_2^+$$

$$A_2^+ = A_2^+ A A_2^+ = (A_2^+ A)' A_2^+ = A' (A_2^+)' A_2^+ = (A A_1^+ A)' (A_2^+)' A_2^+$$

$$= A' (A_1^+)' A' (A_2^+)' A_2^+ = (A_1^+ A)' (A_2^+ A)' A_2^+ = A_1^+ A A_2^+ A A_2^+$$

$$= A_1^+ A A_2^+$$

唯一性得证.

性质2

若A行满秩，则$A^{+}=A'(AA')^{-1}$.

若A列满秩，则$A^{+}=(A'A)^{-1}A^{T}$.

若A可逆，则$A^{+}=A^{-1}$.

性质3 $(A^{+})^{+} = A$

性质4 $A^{+} = A'(AA')^{+} = (A'A)^{+} A'$

性质5 $(A'A)^{+} = A^{+}(A^{+})'$

性质6 $(A')^{+} = (A^{+})'$

由此，当A为对称阵时，A^{+}也是对称矩阵

性质7

记$A = PQ'$，其中

$$rk(\underset{n\times p}{A}) = rk(\underset{n\times r}{P}) = rk(\underset{p\times r}{Q}) = r$$

则$A^+ = (Q^+)'P^+$.

注意：

一般，对任意两个矩阵

$$(AB)^{+} = B^{+}A^{+}$$

未必成立.

性质8

若A是投影矩阵，则 $A^+ = A$

性质9

若A为n阶对称方阵，有分解 $A = H'\Lambda H$，

H为正交矩阵， $\Lambda = diag(\lambda_1, \lambda_2, \cdots, \lambda_n)$, 令

$$\lambda^+ = \begin{cases} \lambda^{-1} & \text{若}\lambda \neq 0 \\ 0 & \text{若}\lambda = 0 \end{cases}$$

则 $A^+ = H' diag(\lambda_1{}^+, \lambda_2{}^+, \cdots \lambda_n{}^+) H$.

性质**10**

AA^{+}与$A^{+}A$均为投影矩阵，

记 $P_A = AA^{+}$

$$P_{A'} = A^{+}A$$

§ 1.3 矩阵的拉直运算和Kronecker积

$$\underset{n\times p}{A} = (a_1, a_2, \cdots, a_p)$$

$$\vec{A} = vec(A) = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{pmatrix}$$

定义1.3.1 设$\mathbf{A}=(a_{ij})=(a_1,a_2,\cdots,a_p)=\begin{pmatrix} a'_{(1)} \\ a'_{(2)} \\ \vdots \\ a'_{(n)} \end{pmatrix}$

为$n\times p$阶矩阵，则 $vec(A)=\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{pmatrix}$称为矩阵 A的

按列拉直运算，记作 $vec(A)$或$\vec{A}$.

性质1

若c,d是实数，A和B是两个大小相同的矩阵，则

$$vec(c\cdot A+d\cdot B)=c\cdot vec(A)+d\cdot vec(B)$$

性质2

$$\underset{n\times p}{A}=\sum_{i,j}a_{ij}E_{ij}=\sum_{i,j}a_{ij}e_i e_j'=\sum_j a_j e_j'=\sum_i e_i a_{(i)}'$$

(矩阵A的不同表示形式)

性质3

$$a_j = Ae_j = \sum_i a_{ij} e_i$$

$$a_{(i)} = A' e_i = \sum_j a_{ij} e_j$$

性质4

$$a_{ij} = e_i' A e_j$$

性质5 $$tr(E'_{rs}A) = a_{rs}$$

事实上，

$$tr(E'_{rs}A) = tr(e_s e'_r A)$$

$$= tr(e'_r A e_s) = e'_r A e_s = a_{rs}$$

性质6

$$tr(AB)=\sum_{i,j} a_{ij} b_{ji} = [vec(A')]'[vec(B)]$$

定义1.3.2 设$A=A'$为p阶方阵，令

$$svec(A)=(a_{11},a_{21},\cdots,a_{p1},a_{22},a_{32},\cdots,a_{p2}\cdots a_{pp})'$$

为$\dfrac{p(p+1)}{2}$维向量，称$svec(A)$为对称矩阵的拉直运算

定义1.3.3 设$A=(a_{ij})$为$n\times p$阶矩阵，B为$m\times q$阶矩阵，则A和B的kronecker积定义为

$$A\otimes B=(a_{ij}B)=\begin{pmatrix} a_{11}B & a_{12}B\cdots a_{1p}B \\ a_{21}B & a_{22}B\cdots a_{2p}B \\ \cdots & \cdots \qquad \cdots \\ a_{n1}B & a_{n2}B\cdots a_{np}B \end{pmatrix}$$

它是$nm\times pq$阶矩阵。

性质1

若α是任一实数，则

$$(\alpha A)\otimes B = A\otimes(\alpha B) = \alpha(A\otimes B)$$

性质2（分配律）

$$A \otimes (B + C) = A \otimes B + A \otimes C$$

$$(B + C) \otimes A = B \otimes A + C \otimes A$$

性质3（结合律）

$$(A\otimes B)\otimes C = A\otimes(B\otimes C)$$

性质4

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

性质5

$$(A \otimes B)' = A' \otimes B'$$

性质6

$$(A \otimes B)(C \otimes D) = AC \otimes BD$$

证明：我们用$[A\otimes B]_{\alpha\beta}$表示$A\otimes B$的$(\alpha,\beta)$块

$$[(A\otimes B)(C\otimes D)]_{\alpha\beta}=\sum_{\gamma=1}^{p}[A\otimes B]_{\alpha\gamma}[C\otimes D]_{\gamma\beta}$$

$$=\sum_{\gamma=1}^{p}a_{\alpha\gamma}B\cdot c_{\gamma\beta}D=\sum_{\gamma=1}^{p}a_{\alpha\gamma}c_{\gamma\beta}BD$$

$$=(AC)_{\alpha\beta}BD=[AC\otimes BD]_{\alpha\beta}$$

所以$(A\otimes B)(C\otimes D)=AC\otimes BD$

性质7

若$\boldsymbol{A}$和$\boldsymbol{B}$为非退化方阵，则

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

性质8

设$A: n\times m$，$B: p\times q$，$X: m\times p$，则

$$vec(AXB) = (B' \otimes A)vec(X)$$

注：这个性质很重要，它建立了拉直运算和Kronecker积之间的联系。

证明：令$(A)_i$和$(A)_{(j)}$分别表示A的第i列和第j行，则

$$(AXB)_k = AXBe_k = A\left(\sum_{j=1}^{p}(X)_j e_j'\right)Be_k$$

$$= \sum_{j=1}^{p} A(X)_j e_j' Be_k = \sum_{j=1}^{p} b_{jk} A(X)_j$$

$$= \begin{pmatrix} b_{1k}A & b_{2k}A & \cdots & b_{pk}A \end{pmatrix} \begin{pmatrix} (X)_1 \\ (X)_2 \\ \vdots \\ (X)_p \end{pmatrix} = \left((B)_k' \otimes A\right) vec(X).$$

$$vec(AXB) = \begin{pmatrix} (AXB)_1 \\ (AXB)_2 \\ \vdots \\ (AXB)_q \end{pmatrix} = \begin{pmatrix} (B)'_1 \otimes A \\ (B)'_2 \otimes A \\ \vdots \\ (B)'_q \otimes A \end{pmatrix} vec(X)$$

$$= (B' \otimes A) vec(X)$$

性质9

若A和B均为方阵，则

$$tr(A \otimes B) = tr(A) \cdot tr(B)$$

性质10

若x和y为列向量，则

$$xy' = x \otimes y' = y' \otimes x$$

性质11

若A为n阶方阵，其特征值为$\{\lambda_i, i=1,2,\cdots,n\}$，相应的特征向量为$\{x_i, i=1,2,\cdots,n\}$，$B$为$m$阶方阵，其特征值为$\{\mu_i, i=1,2,\cdots,m\}$，相应的特征向量$\{y_i, i=1,2,\cdots,m\}$，则$A\otimes B$的特征值为$\{\lambda_i\mu_j, i=1,2,\cdots,n, j=1,2,\cdots m\}$，相应的特征向量为$\{x_i\otimes y_j, i=1,2,\cdots,n, j=1,2,\cdots,m\}$.

证明：利用$Jordan$标准形，存在可逆矩阵 P和Q，

使得 $$P^{-1}AP=\begin{pmatrix} D_1 & 0 & \cdots 0 \\ 0 & D_2 & \cdots 0 \\ \cdots & \cdots & \cdots \\ 0 & 0 & \cdots D_k \end{pmatrix}$$

其中$$D_i=\begin{pmatrix} \lambda_i & 1 & 0 & \cdots 0 \\ 0 & \lambda_i & 1 & \cdots 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots \lambda_i \end{pmatrix}$$

$$Q^{-1}BQ=\begin{pmatrix} C_1 & 0 & \cdots 0 \\ 0 & C_2 & \cdots 0 \\ \cdots & \cdots & \cdots \\ 0 & 0 & \cdots C_k \end{pmatrix}$$

$$\text{其中} C_i=\begin{pmatrix} \mu_i & 1 & 0 & \cdots 0 \\ 0 & \mu_i & 1 & \cdots 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots \mu_i \end{pmatrix}$$

$$(P^{-1}\otimes Q^{-1})(A\otimes B)(P\otimes Q)$$

$$=\begin{pmatrix} D_1 & 0 & \cdots 0 \\ 0 & D_2 & \cdots 0 \\ \cdots & \cdots & \cdots \\ 0 & 0 & \cdots D_k \end{pmatrix}\otimes\begin{pmatrix} C_1 & 0 & \cdots 0 \\ 0 & C_2 & \cdots 0 \\ \cdots & \cdots & \cdots \\ 0 & 0 & \cdots C_k \end{pmatrix}$$

仍为上三角形矩阵，对 角线元为$\lambda_i\mu_j, i=1,2,\cdots,n,$

$j=1,2,\cdots m$，所以$A\otimes B$的特征值为

$\{\lambda_i\mu_j, i=1,2,\cdots,n, j=1,2,\cdots m\}.$

若$Ax=\lambda x$，$By=\mu y$，

$$\Rightarrow (A\otimes B)(x\otimes y)=(Ax)\otimes(By)$$

$$=(\lambda x)\otimes(\mu y)=\lambda\mu(x\otimes y),$$

所以

$$\{x_i\otimes y_j, i=1,2,\cdots,n, j=1,2,\cdots,m\}$$

为相应的特征向量.

性质12

条件同性质11，则

$$\left|\underset{n\times n}{A}\otimes\underset{m\times m}{B}\right|=\left|\underset{n\times n}{A}\right|^{m}\left|\underset{m\times m}{B}\right|^{n}$$

§1.4 矩阵的微商和变换的雅可比

矩阵微商是通常微商的推广，是求极大似然估计和最小二乘估计的工具，利用它也可以方便地求多变量积分变换的雅可比行列式。

一、矩阵对标量的微商

二、矩阵的标量函数对矩阵的微商

三、向量对向量的微商

四、矩阵对矩阵的微商

一、矩阵对标量的微商

定义1.4.1 设$Y=(y_{ij}(t))$是$p\times q$矩阵，它的元素是t的函数，

$$\frac{\partial\{Y\}}{\partial t}=\left(\frac{\partial y_{ij}(t)}{\partial t}\right)$$

称为Y对t的微商.

$$\frac{\partial\{Y\}}{\partial t} = \begin{pmatrix} \frac{\partial y_{11}(t)}{\partial t} & \frac{\partial y_{12}(t)}{\partial t} & \cdots & \frac{\partial y_{1q}(t)}{\partial t} \\ \frac{\partial y_{21}(t)}{\partial t} & \frac{\partial y_{22}(t)}{\partial t} & \cdots & \frac{\partial y_{2q}(t)}{\partial t} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial y_{p1}(t)}{\partial t} & \frac{\partial y_{p2}(t)}{\partial t} & \cdots & \frac{\partial y_{pq}(t)}{\partial t} \end{pmatrix}$$

性质**1**

$$\frac{\partial\{\boldsymbol{X}+\boldsymbol{Y}\}}{\partial \boldsymbol{t}}=\frac{\partial\{\boldsymbol{X}\}}{\partial \boldsymbol{t}}+\frac{\partial\{\boldsymbol{Y}\}}{\partial \boldsymbol{t}}$$

性质2

$$\frac{\partial\{\boldsymbol{XY}\}}{\partial \boldsymbol{t}}=\frac{\partial\{\boldsymbol{X}\}}{\partial \boldsymbol{t}}\boldsymbol{Y}+\boldsymbol{X}\frac{\partial\{\boldsymbol{Y}\}}{\partial \boldsymbol{t}}$$

性质3

$$\frac{\partial\{\boldsymbol{X}\otimes\boldsymbol{Y}\}}{\partial \boldsymbol{t}}=\frac{\partial\{\boldsymbol{X}\}}{\partial \boldsymbol{t}}\otimes\boldsymbol{Y}+\boldsymbol{X}\otimes\frac{\partial\{\boldsymbol{Y}\}}{\partial \boldsymbol{t}}$$

性质4

$$\left(\frac{\partial\{\boldsymbol{X}\}}{\partial \boldsymbol{t}}\right)' = \frac{\partial\{\boldsymbol{X}'\}}{\partial \boldsymbol{t}}$$

性质5

$$\frac{\partial\{\boldsymbol{X}\}}{\partial x_{ij}} = \boldsymbol{E}_{ij}$$

性质6

$$\frac{\partial\{\boldsymbol{AXB}\}}{\partial \boldsymbol{x}_{ij}} = \boldsymbol{AE}_{ij}\boldsymbol{B}$$

性质7

$$\frac{\partial\{X^{-1}\}}{\partial t} = -X^{-1}\frac{\partial\{X\}}{\partial t}X^{-1}$$

(作业)

提示：利用 $XX^{-1} = I$ 和性质2.

二、矩阵的标量函数对矩阵的微商

定义1.4.2 设$y = f(X)$是$m\times n$阶矩阵X的函数，y对X的微商定义为

$$\frac{\partial y}{\partial \{X\}} = \left(\frac{\partial y}{\partial x_{ij}}\right)$$

$$\frac{\partial y}{\partial \{X\}} = \begin{pmatrix} \frac{\partial y}{\partial x_{11}} & \frac{\partial y}{\partial x_{12}} & \cdots & \frac{\partial y}{\partial x_{1n}} \\ \frac{\partial y}{\partial x_{21}} & \frac{\partial y}{\partial x_{22}} & \cdots & \frac{\partial y}{\partial x_{2n}} \\ \cdots & \cdots & & \cdots \\ \frac{\partial y}{\partial x_{m1}} & \frac{\partial y}{\partial x_{m2}} & \cdots & \frac{\partial y}{\partial x_{mn}} \end{pmatrix}$$

性质1

$$\left(\frac{\partial f(\boldsymbol{X})}{\partial\{\boldsymbol{X}\}}\right)' = \frac{\partial f(\boldsymbol{X})}{\partial\{\boldsymbol{X}'\}}$$

性质2

若X为n阶方阵，则

$$\frac{\partial tr(X)}{\partial \{X\}} = I_n$$

性质3

$$\frac{\partial tr(AXB)}{\partial\{X\}} = A'B',$$

特别 $$\frac{\partial tr(AX)}{\partial\{X\}} = A'$$

证明 $tr(AXB)=\sum_i e_i' AXBe_i$

$$\frac{\partial tr(AXB)}{\partial x_{kl}}=\sum_i e_i' AE_{kl}Be_i=\sum_i \underline{e_i' Ae_k}\,\underline{e_l' Be_i}$$

$$=\sum_i a_{ik}b_{li}=\sum_i b_{li}a_{ik}=(BA)_{lk}=(A'B')_{kl}$$

所以$\dfrac{\partial tr(AXB)}{\partial\{X\}}=A'B'$

性质4

$$\frac{\partial tr(AX)}{\partial\{X\}}$$

$$=\begin{cases} A' \\ A+A'-diag(a_{11},a_{22},\cdots,a_{nn}) & 若X=X' \end{cases}$$

证明 仅对X为对称矩阵时给出证明即可.

若$X = X'$，则

$$(1)\ \frac{\partial tr(AX)}{\partial x_{kk}} = \frac{\partial \sum_i e_i' AXe_i}{\partial x_{kk}} = \sum_i e_i' AE_{kk} e_i$$

$$= \sum_i e_i' Ae_k \underline{\underline{e_k' e_i}} = a_{kk}$$

(2)若$k \neq l$，则

$$\frac{\partial tr(AX)}{\partial x_{kl}} = \frac{\partial \sum_i e_i' AXe_i}{\partial x_{kl}} = \sum_i e_i' AE_{kl} e_i + \sum_i e_i' AE_{lk} e_i$$

$$= \sum_i e_i' Ae_k \underline{\underline{e_l' e_i}} + \sum_i e_i' Ae_l \underline{\underline{e_k' e_i}} = a_{lk} + a_{kl}$$

$$\text{所以} \frac{\partial tr(AX)}{\partial \{X\}} = A + A' - diag(a_{11}, a_{22}, \cdots, a_{nn})$$

性质5

$$\frac{\partial x'Ax}{\partial\{x\}}=(A+A')x$$

$$\frac{\partial tr(X'AX)}{\partial\{X\}}=(A+A')X$$

作业

证明 $x'Ax = tr(x'Ax) = tr(Axx') = \sum_i e_i' Axx' e_i$

$$\frac{\partial x'Ax}{\partial x_l} = \sum_i \left(\frac{\partial e_i' Ax}{\partial x_l} \right) x' e_i + \sum_i e_i' Ax \left(\frac{\partial x' e_i}{\partial x_l} \right)$$

$$= \sum_i e_i' A e_l x' e_i + \sum_i e_i' Ax e_l' e_i = \sum_i a_{il} x_i + e_l' Ax$$

$$\frac{\partial \{XY\}}{\partial t} = \frac{\partial \{X\}}{\partial t} Y + X \frac{\partial \{Y\}}{\partial t}$$

$$\sum_i a_{il} x_i + e_l' A x$$

$$= e_l' A' x + e_l' A x$$

$$= e_l' (A' + A) x = \left((A' + A) x\right)_l$$

所以 $$\frac{\partial x' A x}{\partial \{x\}} = (A + A') x$$

性质6

$$\frac{\partial a'x}{\partial \{x\}} = a$$

性质7

$$\frac{\partial \det(X)}{\partial\{X\}} = \left(X'\right)^{-1} \det(X)$$

证明 $\det(X)=\sum_i x_{ij}X_{ij}$，其中$X_{ij}$是$x_{ij}$的代数余子式. 易见$\dfrac{\partial\det(X)}{\partial x_{ij}}=X_{ij}$

$$\therefore \frac{\partial \det(X)}{\partial\{X\}}=\begin{pmatrix} X_{11} & X_{12} & \cdots & X_{1n} \\ X_{21} & X_{22} & \cdots & X_{2n} \\ \cdots & \cdots & & \cdots \\ X_{n1} & X_{n2} & \cdots & X_{nn} \end{pmatrix}$$

又 $$X^{-1}=\frac{1}{\det(X)}\begin{pmatrix} X_{11} & X_{21} & \cdots & X_{n1} \\ X_{12} & X_{22} & \cdots & X_{n2} \\ \cdots & \cdots & & \cdots \\ X_{1n} & X_{2n} & \cdots & X_{nn} \end{pmatrix}$$

所以 $$\frac{\partial \det(X)}{\partial\{X\}}=\left(X'\right)^{-1}\det(X)$$

三、向量对向量的微商

定义1.4.3 设x为n维向量，$y=\left(y_1,y_2,\cdots,y_m\right)'$

$$=\left(f_1(x),f_2(x),\cdots,f_m(x)\right)'$$

则y对x的微商定义为

$$\frac{\partial y'}{\partial x}=\left(\frac{\partial y_j}{\partial x_i}\right)$$

$$\frac{\partial \boldsymbol{y}'}{\partial \boldsymbol{x}} = \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_2}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_1} \\ \frac{\partial y_1}{\partial x_2} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_m}{\partial x_2} \\ \cdots & \cdots & \cdots & \\ \frac{\partial y_1}{\partial x_n} & \frac{\partial y_2}{\partial x_n} & \cdots & \frac{\partial y_m}{\partial x_n} \end{pmatrix}$$

性质1

$$\text{若}\boldsymbol{y}=\boldsymbol{A}\boldsymbol{x}\text{，则}\frac{\partial \boldsymbol{y}'}{\partial \boldsymbol{x}}=\boldsymbol{A}'$$

证明：$y_j=\sum_k a_{jk}x_k \qquad j=1,2,\cdots,m.$

$$\frac{\partial y_j}{\partial x_i}=a_{ji}=\left(A'\right)_{ij}$$

所以
$$\frac{\partial y'}{\partial x}=A'$$

性质2

若$\boldsymbol{Y}=\boldsymbol{AXB}$，则

$$\frac{\partial\left(vec(\boldsymbol{Y})\right)'}{\partial vec(\boldsymbol{X})}=\boldsymbol{B}\otimes\boldsymbol{A}'$$

证明

$$vec(Y) = (B' \otimes A)vec(X)$$

由性质1立即得证.

四、矩阵对矩阵的微商

定义1.4.4 设X是$m\times n$阶矩阵，$F(X)$是$p\times q$阶矩阵，记$F(X)=\left(f_{ij}(X)\right)=\left(f_{ij}\right)$，则

$$\frac{\partial F(X)}{\partial X}=\frac{\partial[vec(F(X))]'}{\partial vec(X)}$$

称为$F(X)$对X的微商。

$$
\begin{pmatrix}
\frac{\partial f_{11}}{\partial x_{11}} & \frac{\partial f_{21}}{\partial x_{11}} & \cdots & \frac{\partial f_{p1}}{\partial x_{11}} & \frac{\partial f_{12}}{\partial x_{11}} & \frac{\partial f_{22}}{\partial x_{11}} & \cdots & \frac{\partial f_{p2}}{\partial x_{11}} & \cdots & \frac{\partial f_{1q}}{\partial x_{11}} & \frac{\partial f_{2q}}{\partial x_{11}} & \cdots & \frac{\partial f_{pq}}{\partial x_{11}} \\
\frac{\partial f_{11}}{\partial x_{21}} & \frac{\partial f_{21}}{\partial x_{21}} & \cdots & \frac{\partial f_{p1}}{\partial x_{21}} & \frac{\partial f_{12}}{\partial x_{21}} & \frac{\partial f_{22}}{\partial x_{21}} & \cdots & \frac{\partial f_{p2}}{\partial x_{21}} & \cdots & \frac{\partial f_{1q}}{\partial x_{21}} & \frac{\partial f_{2q}}{\partial x_{21}} & \cdots & \frac{\partial f_{pq}}{\partial x_{21}} \\
\cdots & \cdots & & \cdots & & & & \cdots & & & \cdots & & \\
\frac{\partial f_{11}}{\partial x_{m1}} & \frac{\partial f_{21}}{\partial x_{m1}} & \cdots & \frac{\partial f_{p1}}{\partial x_{m1}} & \frac{\partial f_{12}}{\partial x_{m1}} & \frac{\partial f_{22}}{\partial x_{m1}} & \cdots & \frac{\partial f_{p2}}{\partial x_{m1}} & \cdots & \frac{\partial f_{1q}}{\partial x_{m1}} & \frac{\partial f_{2q}}{\partial x_{m1}} & \cdots & \frac{\partial f_{pq}}{\partial x_{m1}} \\
\frac{\partial f_{11}}{\partial x_{12}} & \frac{\partial f_{21}}{\partial x_{12}} & \cdots & \frac{\partial f_{p1}}{\partial x_{12}} & & & & & \cdots & \frac{\partial f_{1q}}{\partial x_{12}} & \frac{\partial f_{2q}}{\partial x_{12}} & \cdots & \frac{\partial f_{pq}}{\partial x_{12}} \\
\frac{\partial f_{11}}{\partial x_{22}} & \frac{\partial f_{21}}{\partial x_{22}} & \cdots & \frac{\partial f_{p1}}{\partial x_{22}} & & & & & \cdots & \frac{\partial f_{1q}}{\partial x_{22}} & \frac{\partial f_{2q}}{\partial x_{22}} & \cdots & \frac{\partial f_{pq}}{\partial x_{22}} \\
\cdots & \cdots & & \cdots & & & & & & & \cdots & & \\
\frac{\partial f_{11}}{\partial x_{m2}} & \frac{\partial f_{21}}{\partial x_{m2}} & \cdots & \frac{\partial f_{p1}}{\partial x_{m2}} & & & & & \cdots & \frac{\partial f_{1q}}{\partial x_{m2}} & \frac{\partial f_{2q}}{\partial x_{m2}} & \cdots & \frac{\partial f_{pq}}{\partial x_{m2}} \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\frac{\partial f_{11}}{\partial x_{1n}} & \frac{\partial f_{21}}{\partial x_{1n}} & \cdots & \frac{\partial f_{p1}}{\partial x_{1n}} & \frac{\partial f_{12}}{\partial x_{1n}} & \frac{\partial f_{22}}{\partial x_{1n}} & \cdots & \frac{\partial f_{p2}}{\partial x_{1n}} & \cdots & \frac{\partial f_{1q}}{\partial x_{1n}} & \frac{\partial f_{2q}}{\partial x_{1n}} & \cdots & \frac{\partial f_{pq}}{\partial x_{1n}} \\
\frac{\partial f_{11}}{\partial x_{2n}} & \frac{\partial f_{21}}{\partial x_{2n}} & \cdots & \frac{\partial f_{p1}}{\partial x_{2n}} & \frac{\partial f_{12}}{\partial x_{2n}} & \frac{\partial f_{22}}{\partial x_{2n}} & \cdots & \frac{\partial f_{p2}}{\partial x_{2n}} & \cdots & \frac{\partial f_{1q}}{\partial x_{2n}} & \frac{\partial f_{2q}}{\partial x_{2n}} & \cdots & \frac{\partial f_{pq}}{\partial x_{2n}} \\
\cdots & \cdots & & \cdots & & & & \cdots & & & \cdots & & \\
\frac{\partial f_{11}}{\partial x_{mn}} & \frac{\partial f_{21}}{\partial x_{mn}} & \cdots & \frac{\partial f_{p1}}{\partial x_{mn}} & \frac{\partial f_{12}}{\partial x_{mn}} & \frac{\partial f_{22}}{\partial x_{mn}} & \cdots & \frac{\partial f_{p2}}{\partial x_{mn}} & \cdots & \frac{\partial f_{1q}}{\partial x_{mn}} & \frac{\partial f_{2q}}{\partial x_{mn}} & \cdots & \frac{\partial f_{pq}}{\partial x_{mn}}
\end{pmatrix}
$$

(2,2)块
为m行p列
矩阵

$$
\begin{array}{cccc}
\dfrac{\partial f_{11}}{\partial x_{11}} & \dfrac{\partial f_{21}}{\partial x_{11}} & \cdots & \dfrac{\partial f_{p1}}{\partial x_{11}} \\
\dfrac{\partial f_{12}}{\partial x_{11}} & \dfrac{\partial f_{22}}{\partial x_{11}} & \cdots & \dfrac{\partial f_{p2}}{\partial x_{11}} \\
\cdots & \cdots & \cdots & \cdots \\
\dfrac{\partial f_{1q}}{\partial x_{11}} & \dfrac{\partial f_{2q}}{\partial x_{11}} & \cdots & \dfrac{\partial f_{pq}}{\partial x_{11}}
\end{array}
$$

第一行的元素

（依行的先后次序）

$$
\begin{pmatrix}
\dfrac{\partial f_{11}}{\partial x_{11}} & \dfrac{\partial f_{11}}{\partial x_{21}} & \cdots & \dfrac{\partial f_{11}}{\partial x_{m1}} \\
\dfrac{\partial f_{11}}{\partial x_{12}} & \dfrac{\partial f_{11}}{\partial x_{22}} & \cdots & \dfrac{\partial f_{11}}{\partial x_{m2}} \\
\cdots & \cdots & \cdots & \cdots \\
\dfrac{\partial f_{11}}{\partial x_{1n}} & \dfrac{\partial f_{11}}{\partial x_{2n}} & \cdots & \dfrac{\partial f_{11}}{\partial x_{mn}}
\end{pmatrix}
$$

第一列的元素

（依行的先后次序）

$$
\begin{pmatrix}
\dfrac{\partial f_{12}}{\partial x_{12}} & \dfrac{\partial f_{22}}{\partial x_{12}} \cdots & \dfrac{\partial f_{p2}}{\partial x_{12}} \\
\dfrac{\partial f_{12}}{\partial x_{22}} & \dfrac{\partial f_{22}}{\partial x_{22}} \cdots & \dfrac{\partial f_{p2}}{\partial x_{22}} \\
\cdots & \cdots & \cdots \\
\dfrac{\partial f_{12}}{\partial x_{m2}} & \dfrac{\partial f_{22}}{\partial x_{m2}} \cdots & \dfrac{\partial f_{p2}}{\partial x_{m2}}
\end{pmatrix}
$$

若将其分块，每块均为**m**行**p**列的矩阵，则第**(2,2)**个分块矩阵。

性质1

若X是$m\times n$阶矩阵，则

$$\frac{\partial X}{\partial X}=\frac{\partial(vec(X))'}{\partial(vec(X))}=I_{mn}$$

性质2

若$F(X)$是$p\times q$阶矩阵，$G(X)$是$q\times r$阶矩阵，

X是$m\times n$阶矩阵，则

$$\frac{\partial F(X)G(X)}{\partial X}$$

$$=\frac{\partial F(X)}{\partial X}(G(X)\otimes I_p)+\frac{\partial G(X)}{\partial X}(I_r\otimes F'(X))$$

性质3

$$\frac{\partial(AXB)}{\partial X} = B \otimes A'$$

可以先拉直，再求导.

证明：利用性质2

$$\frac{\partial(AXB)}{\partial X}=\frac{\partial(AX)}{\partial X}(B\otimes I)+\frac{\partial B}{\partial X}(I\otimes X'A')$$

$$=\left[\frac{\partial(A)}{\partial X}(X\otimes I)+\frac{\partial X}{\partial X}(I\otimes A')\right](B\otimes I)$$

$$=B\otimes A'$$

性质4

$$\frac{\partial \boldsymbol{X}^{-1}}{\partial \boldsymbol{X}} = -[\boldsymbol{X}^{-1} \otimes (\boldsymbol{X}^{-1})']$$

证明 由$\boldsymbol{XX^{-1}=I}$，及性质2

$$\frac{\partial(\boldsymbol{XX^{-1}})}{\partial \boldsymbol{X}}=\frac{\partial \boldsymbol{X}}{\partial \boldsymbol{X}}(\boldsymbol{X^{-1}\otimes I})+\frac{\partial \boldsymbol{X^{-1}}}{\partial \boldsymbol{X}}(\boldsymbol{I\otimes X'})$$

$$=(\boldsymbol{X^{-1}\otimes I})+\frac{\partial \boldsymbol{X^{-1}}}{\partial \boldsymbol{X}}(\boldsymbol{I\otimes X'})=\boldsymbol{0}$$

所以 $$\frac{\partial X^{-1}}{\partial X}=-\left(X^{-1}\otimes I\right)\left(I\otimes(X^{-1})'\right)$$

$$=-[X^{-1}\otimes(X^{-1})']$$

五、复合函数求导的公式

定理1.4.1(复合函数求导公式) 设$\Psi(X)$是矩阵变量X的数值函数，X的每个元素x_{ij}均为变量t的函数，则

$$\frac{\partial \psi(X)}{\partial t} = tr\left[\frac{\partial\{X\}}{\partial t}\left(\frac{\partial \psi}{\partial\{X\}}\right)'\right]$$

证明：由复合函数求导公式得到

$$\frac{\partial \psi(X)}{\partial t} = \sum_{j=1}^{n}\sum_{i=1}^{m} \frac{\partial \psi(X)}{\partial x_{ij}} \frac{\partial x_{ij}}{\partial t}$$

$$= tr\left[\frac{\partial\{X\}}{\partial t} \left(\frac{\partial \psi}{\partial\{X\}} \right)' \right]$$

定理1.4.2(矩阵的复合函数求导公式)

设$\boldsymbol{F}(\boldsymbol{G}(\boldsymbol{X}))$是两个矩阵函数的复合函数，

则

$$\frac{\partial \boldsymbol{F}(\boldsymbol{G}(\boldsymbol{X}))}{\partial \boldsymbol{X}} = \frac{\partial \boldsymbol{G}(\boldsymbol{X})}{\partial \boldsymbol{X}} \frac{\partial \boldsymbol{F}(\boldsymbol{G})}{\partial \boldsymbol{G}}$$

性质1

$$\frac{\partial \det(X(t))}{\partial t} = \det(X)\, tr\left(X^{-1} \frac{\partial \{X\}}{\partial t} \right)$$

证明：

$$\frac{\partial \det(X(t))}{\partial t} = tr\left[\frac{\partial\{X\}}{\partial t}\left(\frac{\partial \det(X)}{\partial\{X\}}\right)'\right]$$

$$= tr\left[\frac{\partial\{X\}}{\partial t}\left((X')^{-1}\det(X)\right)'\right]$$

$$= tr\left[\frac{\partial\{X\}}{\partial t}\left(X^{-1}\det(X)\right)\right] = \det(X)tr\left(X^{-1}\frac{\partial\{X\}}{\partial t}\right)$$

性质2

$$\frac{\partial \ln \det(X)}{\partial t} = tr\left(X^{-1} \frac{\partial \{X\}}{\partial t} \right)$$

证明：

$$\frac{\partial \ln \det(X)}{\partial t} = \frac{1}{\det(X)} \frac{\partial \det(X)}{\partial t}$$

$$= tr\left(X^{-1} \frac{\partial \{X\}}{\partial t} \right)$$

六、雅可比(Jacobi)行列式的计算

$$\int_D g(x_1, x_2, \cdots, x_n) dx_1 dx_2 \cdots dx_n \qquad D \subset R^n$$

$$= \int_T g(f^{-1}(y)) \left| J(x \to y) \right| dy$$

式中$T = \{y \mid y = f(x), x \in D\}$, $\left| J(x \to y) \right| = \left| \frac{\partial x'}{\partial y} \right|_+$

f是一一变换，其中$|A|_+$表示A的行列式的绝对值.

定义1.4.5 设$X \in R^{m\times n}$是矩阵变量，

$Y = F(X) \in R^{m\times n}$是一一变换，

$F(X)$可微，则变换$Y = F(X)$的

雅可比行列式定义为

$$J(Y \to X) = J(Y : X) = \left|\frac{\partial Y}{\partial X}\right|_{+}$$

（一）、当矩阵变量为一般矩阵的情况

性质1

若$Y = AXB$，且A为m阶非奇异方阵，

B为n阶非奇异方阵，则有

$$J(Y \to X) = \left|(\det(A))^n (\det(B))^m\right| = |A|_+^n |B|_+^m$$

特别$y = Ax, J(y \to x) = |A|_+$

证明：因为 $vec(Y) = (B' \otimes A)vec(X)$

所以

$$J(Y \to X) = \left|\frac{\partial Y}{\partial X}\right|_+ = \left|\frac{\partial [vec(Y)]'}{\partial vec(X)}\right|_+$$

$$= \left|B \otimes A'\right|_+ = \left|B\right|_+^m \left|A'\right|_+^n = \left|A\right|_+^n \left|B\right|_+^m$$

性质2

若X为n阶可逆方阵，作变换 $Y = X^{-1}$,

则有

$$J(Y \to X) = |X|^{-2n}$$

证明：由矩阵对矩阵的微商中的性质4知，

$$\frac{\partial X^{-1}}{\partial X} = -[X^{-1} \otimes (X^{-1})']$$

所以有

$$J(Y \to X) = \left|\frac{\partial Y}{\partial X}\right|_{+} = \left|-\left(X^{-1} \otimes (X^{-1})'\right)\right|_{+}$$

$$= \left|X^{-1}\right|_{+}^{n} \left|(X^{-1})'\right|_{+}^{n} = \left|X\right|^{-2n}$$

(二)、当矩阵变量为三角矩阵的情况

性质3

设G为给定的n阶非奇异下三角矩阵，

X为n阶下三角矩阵变量，令 $Y = GX$,

则有$J(Y \to X) = \left|\prod_{i=1}^{n} g_{ii}^{i}\right|$, 其中$g_{11}, g_{22}, \cdots,$

g_{nn}是G的主对角元.

$$Y = GX$$

$$J(Y \to X) = \left| \prod_{i=1}^{n} g_{ii}^{i} \right|$$

性质4

条件同性质3，令变换$Y = XG$，则有

$$\mathrm{J}(Y \to X) = \left|\prod_{i=1}^{n} g_{ii}^{n-i+1}\right|$$

性质5

设X为下三角矩阵变量，令变换

$Y = XX'$，则有

$$J(Y \to X) = 2^n \left| \prod_{i=1}^{n} x_{ii}^{n-i+1} \right|$$

注意：若要求X的对角线元素非负，则

$Y = XX'$是一一变换.

类似对**X**为上三角矩阵时，**G**为非奇异上三角矩阵，有类似性质**3**、性质**4**和性质**5**的结论。

(三)、当矩阵变量为对称矩阵的情况

性质6

设$X' = \underset{n\times n}{X}$，$G$是非奇异上三角矩阵，

令$Y = G'XG$，则有

$$J(Y \to X) = \left|G\right|_{+}^{n+1}$$

性质7

设$X' = X$，G是非奇异下三角矩阵，

令$Y = G'XG$，则有

$$J(Y \to X) = \left|G\right|_{+}^{n+1}$$

性质8

设$X' = X$，P是非奇异矩阵，令

$Y = P'XP$，则有

$$\mathrm{J}(Y \to X) = |P|_+^{n+1}$$

作业

1、证明AB和BA有相同的非零特征根。

2、设P为p阶非奇异矩阵，U为$p\times q$矩阵，V为$q\times p$矩阵，而p阶方阵$Q=P+UV$和q阶方阵$I_q+VP^{-1}U$也非奇异，则

$$Q^{-1}=\left(P+UV\right)^{-1}$$

$$=P^{-1}-P^{-1}U\left(I_q+VP^{-1}U\right)^{-1}VP^{-1}$$

特别 $\left(P+xy'\right)^{-1}=P^{-1}-\left(1+y'P^{-1}x\right)^{-1}P^{-1}xy'P^{-1}$

3、若A>0(或≥0)，则存在B>0(或≥0)，使得 $A = B^2$ ，称B为A的平方根矩阵，记为 $B = A^{\frac{1}{2}}$.

4、证明加号逆的性质**5**，即

$$(A^T A)^+ = A^+ (A^+)^T$$

5、若$\boldsymbol{A}$**>0**，将$\boldsymbol{A}$剖分为

$$\boldsymbol{A}=\begin{pmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12}\\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22}\end{pmatrix}.$$

其中A_{11}为方阵，则

$$\boldsymbol{A}_{11}>0,\ \boldsymbol{A}_{22}>0,\ \boldsymbol{A}_{11\cdot 2}>0,\ \boldsymbol{A}_{22\cdot 1}>0.$$

6、证明：

$$\frac{\partial\{X^{-1}\}}{\partial t}=-X^{-1}\frac{\partial\{X\}}{\partial t}X^{-1}$$