第13卷第1期　　数学研究与评论　　Vol. 13 No. 1
1993年2月　　JOURNAL OF MATHEMATICAL RESEARCH AND EXPOSITION　　Feb. 1993

关于矩阵微分法的评注*

梁传广　　荆海英　　荆芷萍
（大连理工大学应用数学系）　（东北工学院数学系）　（辽宁省电教馆）

摘　要

本文简单评论了现有的矩阵微分法的定义及结论，在此基础上给出了一种新的矩阵微分定义，并导出了在新定义下的矩阵微分运算法则及计算公式.

§1 引　言

近年来，矩阵微分法在应用数学的各个领域里正日益显示出它的重要性，特别是在大系统的优化理论、灵敏度分析及 Robust 控制等方面，它已成为不可缺少的重要工具.

目前关于矩阵微分的定义有三种. 在每种定义下都已推出了若干求导法则及计算公式. 但是，我们发现在现有矩阵微分定义下有些求导法则和计算公式过于复杂，既不便于计算，又不便于记忆. 从矩阵只不过是一个"表格"这个观点上看问题，我们感到矩阵的微分定义应视具体情况，使其成为便于使用的工具而定. 由此，本文给出了一种新的矩阵微分定义，并在此基础上导出了若干求导法则及计算公式. 可以证明，本文关于数量函数相对于向量的求导定义与现有的定义相同，而关于矩阵函数相对于矩阵的求导公式与现有的公式在初等变换下是等价的.

§2.　记号、定义及其它准备工作

本文用$[a_{ij}]$表示以a_{ij}为它的位于第i行第j列的元素的矩阵. 设$A=[a_{ij}]\in R^{n\times m}$, $B=[b_{ij}]\in R^{m\times r}$.

先给出现有的三种矩阵微分定义.

定义 1[1]　$\dfrac{dA}{dB}\triangleq[\dfrac{dA}{db_{ij}}]$，这里$\dfrac{dA}{db_{ij}}\triangleq\begin{bmatrix}\dfrac{da_{11}}{db_{ij}} & \dfrac{da_{12}}{db_{ij}} & \cdots & \dfrac{da_{1m}}{db_{ij}}\\ \cdots & & & \cdots\\ \dfrac{da_{n1}}{db_{ij}} & \dfrac{da_{n2}}{db_{ij}} & \cdots & \dfrac{da_{nm}}{db_{ij}}\end{bmatrix}$.

定义 2[2]　$\dfrac{dA}{dB}\triangleq[\dfrac{da_{ij}}{dB}]$，这里$\dfrac{da_{ij}}{dB}\triangleq\begin{bmatrix}\dfrac{da_{ij}}{db_{11}} & \dfrac{da_{ij}}{db_{12}} & \cdots & \dfrac{da_{ij}}{db_{1r}}\\ \cdots & & & \cdots\\ \dfrac{da_{ij}}{db_{m1}} & \dfrac{da_{ij}}{db_{m2}} & \cdots & \dfrac{da_{ij}}{db_{mr}}\end{bmatrix}$.

* 1991年10月12日收到.

定义 3[3] $\frac{dA}{dB} \triangleq \begin{bmatrix} \frac{d\vec{A}_r}{db_{11}} \\ \frac{d\vec{A}_r}{db_{12}} \\ \cdots \\ \frac{d\vec{A}_r}{db_{mr}} \end{bmatrix}$，这里$\frac{d\vec{A}_r}{db_{ij}} \triangleq [\frac{da_{11}}{db_{ij}} \cdots \frac{da_{1m}}{db_{ij}} \frac{da_{21}}{db_{ij}} \cdots \frac{da_{nm}}{db_{ij}}]$.

显然，定义 1 与定义 2 差别不大，因而运算法则也很相似，故这里仅以定义 1 为例来讨论. 在第一种定义下，$\frac{dA}{dA}$，$\frac{dA^T}{dA}$，$\frac{dA}{dA^T}$ 都不是单位阵，因而所导出的类似于$\frac{dA^TA}{dA} = \frac{dA^T}{dA}\mathrm{diag}\{A, \cdots, A\} + \mathrm{diag}\{A^T, \cdots, A^T\}\frac{dA}{dA}$这样的公式形式上简单，但实际上并不好用. 同样，在第三种定义下，虽然$\frac{dA}{dA}$是单位阵，但$\frac{dA^T}{dA}$，$\frac{dA}{dA^T}$都不是，故对于这样定义下的求导公式：$\frac{dA^TA}{dA} = \frac{dA^T}{dA}(I \otimes A) + (A^T \otimes I)$ 也同样不便使用. 那么，能否"改造"一下原有的矩阵导数定义，使$\frac{dA^T}{dA}$，$\frac{dA}{dA^T}$，$\frac{dA}{dA}$这些矩阵是单位阵，或者具有简单的表达式?这正是本文的意图之一.

除此之外，考虑到在实际事物并不都是仅受到一种参数的影响，而常常是受到一组参数的制约，因此，以相对于向量的导数作为定义矩阵导数的基础是具有实际意义的. 有鉴于此，我们给出如下矩阵微分定义.

定义 4　设 $A = \begin{bmatrix} a_1^T \\ a_2^T \\ \cdots \\ a_n^T \end{bmatrix}$，$B = [b_1, b_2, \cdots, b_r]$，其中 a_i^T 是行向量，$i = 1, 2, \cdots, n$，b_j 是列向量，$j = 1, 2, \cdots, r$.

定义　$\frac{dA}{dB} \triangleq [\frac{dA}{db_1}, \frac{dA}{db_2}, \cdots, \frac{dA}{db_r}]$，这里$\frac{dA}{db_j} \triangleq \begin{bmatrix} \frac{da_1^T}{db_j} \\ \frac{da_2^T}{db_j} \\ \cdots \\ \frac{da_n^T}{db_j} \end{bmatrix}$，$j = 1, 2, \cdots, r$. 其中$\frac{da_i^T}{db_j}$的定义与文[2]一致.

显然，上述定义中的$\frac{dA}{dB}$是 $nm \times mr$ 维矩阵. 上述定义的特点是以相对于向量的求导作为基础的，这样既便于记忆，又可以直接利用向量的求导法则进行计算，从而使计算简化.

为了以下讨论方便，我们在这里还要给出矩阵的拉直与矩阵的 kronecker 积的定义.

定义 5　设 $A \in R^{n \times m}$，称 $\vec{A}_r = [a_{11} \cdots a_{1m} a_{21} \cdots a_{2m} \cdots a_{n1} \cdots a_{nm}]$ 为矩阵 A 的按行拉直. 称 $\vec{A}_c = [a_{11} \cdots a_{n1} a_{12} \cdots a_{n2} \cdots a_{1m} \cdots a_{nm}]^T$ 为矩阵 A 的按列拉直.

定义 6[1]　设 $A = [a_{ij}] \in R^{n \times m}$，$B = [b_{ij}] \in R^{p \times q}$，称如下的分块矩阵：

$$A \otimes B \triangleq \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1m}B \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1}B & a_{n2}B & \cdots & a_{nm}B \end{bmatrix} \in R^{np \times mq}$$

是 A 与 B 的 kronecker 积.

§3 新定义下的求导公式及运算法则

使用定义 4 及向量求导法则，我们可直接推出如下计算公式.

定理 1 只要下面的运算可行，则有：

(1) 设 $A \in R^{n\times m}$，则 $\frac{dA^T}{dA} = I_{mn}$，$\frac{dA}{dA^T} = I_{mn}$，$\frac{dA}{dA} = [\vec{I}_n, 0, \vec{I}_n, \cdots, 0, \vec{I}_n]$ 这里 $\vec{I}_n$ 表示 n 阶单位矩阵的拉直(按列)，0 表示 O 矩阵.

(2) $\frac{dA^TB}{dA} = I_m \otimes B$，$A \in R^{n\times m}$，$B \in R^{n\times r}$.

(3) $\frac{dAB^T}{dB} = A \otimes I_r$，$A \in R^{n\times m}$，$B \in R^{r\times m}$.

(4) $\frac{dA^TA}{dA} = [\vec{A}_c\ 0\ \vec{A}_c\ 0 \cdots 0\ \vec{A}_c] + I_m \otimes A$，其中 0 表示 O 矩阵.

(5) $\frac{dA^TBA}{dA} = (I_m \otimes B^T)[\vec{A}_c\ 0\ \vec{A}_c \cdots 0\ \vec{A}_c] + I_m \otimes BA$，其中 0 表示 O 矩阵，$B \in R^{n\times n}$.

证明 (1) 式由定义可直接证出.

设 $A = [a_1, a_2, \cdots, a_m]$，则 $A^T = \begin{bmatrix} a_1^T \\ a_2^T \\ \cdots \\ a_m^T \end{bmatrix}$，故：

$$\frac{dA^TB}{dA} = [\frac{dA^TB}{da_1}, \frac{dA^TB}{da_2}, \cdots, \frac{dA^TB}{da_m}] = \begin{bmatrix} \frac{da_1^TB}{da_1} & \cdots & \frac{da_1^TB}{da_m} \\ \cdots & & \cdots \\ \frac{da_m^TB}{da_1} & \cdots & \frac{da_m^TB}{da_m} \end{bmatrix}$$

$$= \begin{bmatrix} B & & & 0 \\ & B & & \\ & & \ddots & \\ 0 & & & B \end{bmatrix} = I_m \otimes B.$$

设 $B = [b_1, b_2, \cdots, b_m]$，则 $B^T = \begin{bmatrix} b_1^T \\ \cdots \\ b_m^T \end{bmatrix}$，故有：$\frac{dAB^T}{dB} = \begin{bmatrix} \frac{d\hat{a}_1^TB^T}{db_1} & \cdots & \frac{d\hat{a}_1^TB^T}{db_m} \\ \cdots & & \cdots \\ \frac{d\hat{a}_n^TB^T}{db_1} & \cdots & \frac{d\hat{a}_n^TB^T}{db_m} \end{bmatrix}$，这里

$A = \begin{bmatrix} \hat{a}_1^T \\ \hat{a}_2^T \\ \cdots \\ \hat{a}_n^T \end{bmatrix}$，由于 $\frac{d\hat{a}_i^TB^T}{db_j} = \frac{d(a_{i1}b_1^T + \cdots + a_{im}b_m^T)}{db_j} = a_{ij}I$，则(3) 式成立.

设 $A=[a_1,a_2,\cdots,a_m]$，则 $A^TA=\begin{bmatrix} a_1^Ta_1 & a_1^Ta_2 & \cdots & a_1^Ta_m \\ a_2^Ta_1 & a_2^Ta_2 & \cdots & a_2^Ta_m \\ \cdots & & & \cdots \\ a_m^Ta_1 & a_m^Ta_2 & \cdots & a_m^Ta_m \end{bmatrix}$，由于

$$\frac{dA^TA}{da_j}=\begin{bmatrix} 0 & \cdots & 0 & a_1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & a_2 & 0 & \cdots & 0 \\ & \cdots & & \cdots & & \cdots & \\ a_1 & \cdots & a_{j-1} & 2a_j & a_{j+1} & \cdots & a_m \\ 0 & \cdots & 0 & a_{j+1} & 0 & \cdots & 0 \\ & \cdots & & \cdots & & \cdots & \\ 0 & \cdots & 0 & a_m & 0 & \cdots & 0 \end{bmatrix}$$

$$\frac{dA^TA}{dA}=\begin{bmatrix} 2a_1 & a_2 & \cdots & a_m & 0 & a_1 & \cdots & 0 & \cdots & 0 & \cdots & a_1 \\ a_2 & 0 & & & a_1 & 2a_2 & \cdots & a_m & \cdots & 0 & \cdots & a_2 \\ & & & 0 & a_3 & 0 & & & & & & \\ & & & \cdots & & & \cdots & & & & & \\ a_m & 0 & & & 0 & a_m & & & 0 & a_1 & \cdots & 2a_m \end{bmatrix}$$

$$=[\vec{A}_c\,0\,\vec{A}_c\cdots 0\,\vec{A}_c]+I_m\otimes A$$

使用与上述相同的处理手段，可以证出：

$$\frac{dA^TBA}{dA}=\begin{bmatrix} (B+B^T)a_1 & Ba_2 & \cdots & Ba_m & 0 & 0 & \cdots & B^Ta_1 \\ B^Ta_2 & & Ba_1 & (B+B^T)a_2 & Ba_3 & \cdots\ Ba_m & \cdots & B^Ta_2 \\ & \cdots & 0 & B^Ta_3 & & 0 & & \\ & & & \cdots & & \cdots & & \\ B^Ta_m & & & B^Ta_m & \cdots\ Ba_1 & \cdots & Ba_{m-1} & (B+B^T)a_m \end{bmatrix}$$

$$=(I_m\otimes B^T)_c[\vec{A}_c\,0\,\vec{A}_c\cdots 0\,\vec{A}_c]+[I_m\otimes BA]$$

在上述证明中，我们使用了文[2]的运算法则：$\frac{da^Tb}{dx}=\frac{da^T}{dx}b+\frac{db^T}{dx}a$，这里 a,b,x 均为列向量. 证毕.

以下我们给出新定义下的矩阵微分运算法则.

定理 2　若下面的运算可行，则有：

(1)　设 $A_i\in R^{n\times m}$，$C\in R^{p\times q}$，则 $\frac{d(\sum_{i=1}^{k}\xi_iA_i)}{dc}=\sum_{i=1}^{k}\xi_i\frac{dA_i}{dc}$，这里 ξ_i 是非零常数，$i=1,\cdots,k$.

(2)　设 $A\in R^{n\times m}$，$B\in R^{m\times r}$，$C\in{}^{p\times q}$，则有：

$$\frac{dAB}{dc}=\frac{dA}{dc}(I_q\otimes B)+(I_r\otimes\frac{d\vec{B}}{dc})H(I_{qr}\otimes\vec{A}_cn^T)$$

这里 H 是某一确定的初等列变换阵，特殊地，当 $A=a^T$ 是行向量，$B=b$ 是列向量时有：

$$\frac{da^Tb}{dc}=\frac{da^T}{dc}(I_q\otimes b)+\frac{db^T}{dc}(I_q\otimes a)$$

(3)　设 $f=f(Y(A))$ 是复合矩阵函数，其中 $A\in R^{n\times m}$，$Y=Y(A)\in R^{p\times r}$，$f=f(Y)$ 是数量函数. 则：

$$\frac{df}{dA}=\frac{d\vec{Y}_c^T}{dA}H(I_r\otimes\frac{df}{d\vec{Y}_c})$$

这里 II 是某一确定的初等行变换阵. 特殊地，当 $Y \in R^{m\times 1}$ 时，有 $\frac{df}{dA} = \frac{dY^T}{dA}(I \otimes \frac{df}{dY})$

证明　按定义 4，(1) 显然成立.

欲证(2)，设 $C = [C_1, C_2, \cdots, C_q]$，$a^T \in R^{1\times m}$，$b \in R^{m\times 1}$，则：

$$\begin{aligned}\frac{da^Tb}{dc} &= [\frac{da^Tb}{dc_1}, \frac{da^Tb}{dc_2}, \cdots, \frac{da^Tb}{dc_q}] \\ &= [\frac{da^T}{dc_1}b + \frac{db^Tb}{dc_1}a, \frac{da^T}{dc_2}b + \frac{db^T}{dc_2}a, \cdots, \frac{da^T}{dc_q}b + \frac{db^T}{dc_q}a] \\ &= [\frac{da^T}{dc_1}, \cdots, \frac{da^T}{dc_q}]\begin{bmatrix} b & & \\ & b & \\ & & \ddots \\ & & & b\end{bmatrix} + [\frac{db^T}{dc_1}, \cdots, \frac{db^T}{dc_q}]\begin{bmatrix} a & & \\ & a & \\ & & \ddots \\ & & & a\end{bmatrix} \\ &= \frac{da^T}{dc}(I_q \otimes b) + \frac{db^T}{dc}(I_q \otimes a).\end{aligned}$$

又设 $A = \begin{bmatrix} a_1^T \\ a_2^T \\ \cdots \\ a_n^T \end{bmatrix}$，$B = [b_1, b_2, \cdots, b_r]$，则：

$$\frac{dAB}{dc_i} = \begin{bmatrix} \frac{da_1{}^T}{dc_i}b_1 + \frac{db_1{}^T}{dc_i}a_1 & \frac{da_1^T}{dc_i}b_2 + \frac{db_1^T}{dc_i}a_1 & \cdots & \frac{da_1^T}{dc_i}b_r + \frac{db_r^T}{dc_i}a_1 \\ & \cdots & & \cdots \\ \frac{da_n^T}{dc_i}b_1 + \frac{db_1^T}{dc_i}a_n & \frac{da_n^T}{dc_i}b_2 + \frac{db_1^T}{dc_i}a_n & \cdots & \frac{da_n^T}{dc_i}b_r + \frac{db_r^T}{dc_i}a_n \end{bmatrix}.$$

$$= [\frac{dA}{dc_i}b_1, \frac{dA}{dc_i}b_2, \cdots, \frac{dA}{dc_i}b_r] + \left[\begin{bmatrix} \frac{db_1{}^T}{dc_i} & & & \\ & \frac{db_1{}^T}{dc_i} & & \\ & & \ddots & \\ & & & \frac{db_1{}^T}{dc_i}\end{bmatrix}\vec{A}_c^T, \right.$$

$$\left. \begin{bmatrix} \frac{db_2{}^T}{dc_i} & & & \\ & \frac{db_2{}^T}{dc_i} & & \\ & & \ddots & \\ & & & \frac{db_2{}^T}{dc_i}\end{bmatrix}\vec{A}_c^T, \cdots, \begin{bmatrix} \frac{db_1{}^T}{dc_i} & & & \\ & \frac{db_1{}^T}{dc_i} & & \\ & & \ddots & \\ & & & \frac{db_1{}^T}{dc_i}\end{bmatrix}\vec{A}_c^T \right]$$

$$= \frac{dA}{dc_i}B + \text{blockdiag}\{\frac{d\vec{B}_r^T}{dc_i}, \cdots, \frac{d\vec{B}_r^T}{dc_i}\}II_1(I_r \otimes \frac{d\vec{A}_c^T}{dc_i})$$

这里 II_1 是某一确定的初等列变换阵. 故

$$\frac{dAB}{dc} = [\frac{dAB}{dc_1}, \frac{dAB}{dc_2}, \cdots, \frac{dAB}{dc_q}]$$

$$= [\frac{dA}{dc_1}B, \frac{dA}{dc_2}B, \cdots, \frac{dA}{dc_q}B] + [\text{blockdiag}\{\frac{d\vec{B}_r^T}{dc_1}, \cdots, \frac{d\vec{B}_r^T}{dc_1}\}H_1,$$

$$\cdots, \text{blockdiag}\{\frac{d\vec{B}_r^T}{dc_q}, \cdots, \frac{d\vec{B}_r^T}{dc_q}\}H_1][I_q \otimes (I_r \otimes \vec{A}_c^T]$$

$$= \frac{dA}{dc}(I_q \otimes B) + \text{blockdiag}\{\frac{d\vec{B}_r^T}{dc}, \cdots, \frac{d\vec{B}_r^T}{dc}\}H_2(I_q \otimes H_1)[I_{qr} \otimes [I_{qr} \otimes \vec{A}_c^T]$$

$$= \frac{dA}{dc}(I_q \otimes B) + (I_q \otimes \frac{d\vec{B}_r^T)}{dc})H(I_{qr} \otimes \vec{A}_c^T)$$

由于 H_1 与 H_2 均为确定的初等列变换阵，故 $H = H_i(I_q \otimes H_1)$ 也是确定的初等列变换阵. 因此，(2) 式成立.

欲证(3)，设 $Y = [Y_1, Y_2, \cdots, Y_r] = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1r} \\ y_{21} & y_{22} & \cdots & y_{2r} \\ & \cdots & \cdots & \\ y_{p1} & y_{p2} & \cdots & y_{pr} \end{bmatrix}$，又设 $A = [a_1, a_2, \cdots, a_m] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ & \cdots & \cdots & \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}$，则：

$$\frac{df}{dA} = [\frac{df}{da_1} \frac{df}{da_2} \cdots \frac{df}{da_m}]$$

$$= \begin{bmatrix} \sum_{i=1}^{r} \frac{\partial Y_i^T}{\partial a_{11}} \frac{\partial f}{\partial Y_i} & \sum_{i=1}^{r} \frac{\partial Y_i^T}{\partial a_{12}} \frac{\partial f}{\partial Y_i} & \cdots & \sum_{i=1}^{r} \frac{\partial Y_i^T}{\partial a_{1m}} \frac{\partial f}{\partial Y_i} \\ \cdots & & & \cdots \\ \sum_{i=1}^{r} \frac{\partial Y_i^T}{\partial a_{n1}} \frac{\partial f}{\partial Y_i} & \sum_{i=1}^{r} \frac{\partial Y_i^T}{\partial a_{n2}} \frac{\partial f}{\partial Y_i} & \cdots & \sum_{i=1}^{r} \frac{\partial Y_i^T}{\partial a_{nm}} \frac{\partial f}{\partial Y_i} \end{bmatrix}$$

$$= \sum_{i=1}^{r} \begin{bmatrix} \frac{\partial Y_i^T}{\partial a_{11}} \frac{\partial f}{\partial Y_i} & \frac{\partial Y_i^T}{\partial a_{12}} \frac{\partial f}{\partial Y_i} & \cdots & \frac{\partial Y_i^T}{\partial a_{1m}} \frac{\partial f}{\partial Y_i} \\ \cdots & & & \cdots \\ \frac{\partial Y_i^T}{\partial a_{n1}} \frac{\partial f}{\partial Y_i} & \frac{\partial Y_i^T}{\partial a_{n2}} \frac{\partial f}{\partial Y_i} & \cdots & \frac{\partial Y_i^T}{\partial a_{nm}} \frac{\partial f}{\partial Y_i} \end{bmatrix}$$

$$= \sum_{i=1}^{r} [\frac{\partial Y_i^T}{\partial a_1} \frac{\partial f}{\partial Y_i}, \frac{\partial Y_i^T}{\partial a_2} \frac{\partial f}{\partial Y_i}, \cdots, \frac{\partial Y_i^T}{\partial a_m} \frac{\partial f}{\partial Y_i}] = \sum_{i=1}^{r} \frac{\partial Y_i^T}{\partial A}[I_m \otimes \frac{\partial f}{\partial Y_i}]$$

$$= \frac{d\vec{Y}_r^T}{dA} \begin{bmatrix} I_m \otimes \frac{df}{dY_1} \\ I_m \otimes \frac{df}{dY_2} \\ \cdots \\ I_m \otimes \frac{df}{dY_r} \end{bmatrix} = \frac{d\vec{Y}_r^T}{dA} H \begin{bmatrix} \frac{df}{d\vec{Y}_c} & & & 0 \\ & \frac{df}{d\vec{Y}_c} & & \\ & & \ddots & \\ 0 & & & \frac{df}{d\vec{Y}_c} \end{bmatrix}$$

$$= \frac{d\vec{Y}_r^T}{dA} H(I_r \otimes \frac{df}{d\vec{Y}_c}),$$

这里H是某一确定的初等行变换阵. 特殊地，当$Y\in R^{l\times 1}$时，则利用文[2]的数量函数相对于向量的复合函数求导法则：$\frac{df}{dx}=\frac{dY^T}{dx}\frac{df}{dY}$，这里$x$与$Y$均为向量，我们可推出：

$$\frac{df}{dA}=[\frac{df}{da_1},\frac{df}{da_2},\cdots,\frac{df}{da_m}]=[\frac{dY^T}{da_1}\frac{df}{dY},\frac{dY^T}{da_2}\frac{df}{dY},\cdots,\frac{dY^T}{da_m}\frac{df}{dY}]$$

$$=\frac{dY^T}{dA}(I_m\otimes\frac{df}{dY}).$$

定理至此全部证完.

§4 结束语

本文给出了一种新的矩阵微分定义，该定义的优点之一是可导出公式：$\frac{dA}{dA^T}=\frac{dA^T}{dA}=I$，这可使许多常用的求导公式变得简单好用. 另外，由于定义中采用了以相对于向量的微分作为最小计算单位的原则，故很易从定义出发直接推出许多常用的矩阵求导公式，从而可避免套用复杂的求导公式.

诚然，本文并非要否定现有的三种矩阵微分定义，事实上，每种定义都有其自己的优越性. 由于相对于矩阵的求导运算比较复杂，故宜于具体情况具体处理.

我们将本文的结果用于广义大系统的灵敏度分析与Robust控制等问题上，得出的结果是令人满意的. 例如，我们常遇到求$\frac{dx^T(E+BK)^T(E+BK)x}{dK}$的问题，这里$x\in R^{n\times 1}$，$E\in R^{n\times m}$，$B\in R^{n\times m}$，$K\in R^{m\times n}$，使用本文的定理2的(2)，有：

$$\frac{dx^T(E+BK)^T(E+BK)x}{dK}=2\frac{dx^T(E+BK)^T}{dK}(I_n\otimes(E+BK))$$

$$=2\frac{dx^TK^T}{dK}[I_n\otimes B^T][I_n\otimes(E+BK)].$$

利用本文定理1的(3)，我们可推出原式等于$2(x^T\otimes I_m)(I_n\otimes B^T(E+BK))=B^T(E+BK)[x^T\otimes I_n]$.

参 考 文 献

[1] 须田信英，自动控制系统中的矩阵理论，科学出版社，1979年.

[2] 谢绪恺，现代控制理论基础，辽宁人民出版社，1980年.

[3] 倪国熙，常用的矩阵理论和方法，上海科技出版社，1984年.

[4] Bellman, R., *Introduction to matrix analysis*, 1970, Mcgrawhill.

A Note on Matrix Differential Calculus

Liang Chuanguang (Dalian Univ. of Tech.)

Jing Haiying *Jing Zhiping* (Northeast Univ. of tech.)

Abstract

We present a new definition of matrix differential and the corresponding operation formula of matrix differential calculation by reviewing the concurrent definition and related conclusions of matrix differential calcalus.