矩阵理论与方法

内容提要 CONTENTS

- □ 课程信息
- □ 课程介绍
- □ 矩阵理论与方法

第3章 矩阵分析及其应用

其他微分

函数对矩阵的导数

函数矩阵对矩阵的导数

其它微分概念

函数对矩阵的导数(包括向量)

定义:设$X=(\xi_{ij})_{m\times n}$,mn元函数

$$f(X)=f(\xi_{11},\xi_{12},\cdots,\xi_{1n},\cdots,\xi_{m\times n})$$

定义$f(X)$对矩阵X的导数为

$$\frac{df}{dX}=\left(\frac{\partial f}{\partial \xi_{ij}}\right)_{m\times n}=\begin{bmatrix}\frac{\partial f}{\partial \xi_{11}} & \cdots & \frac{\partial f}{\partial \xi_{1n}}\\ \vdots & & \vdots \\ \frac{\partial f}{\partial \xi_{m1}} & \cdots & \frac{\partial f}{\partial \xi_{mn}}\end{bmatrix}$$

例11: $\boldsymbol{x}=\begin{bmatrix}\xi_1\\ \vdots\\ \xi_n\end{bmatrix}$: $\boldsymbol{f}(\boldsymbol{x})=\boldsymbol{f}(\xi_1,\xi_2,\cdots,\xi_n)$

$$\frac{d\boldsymbol{f}}{d\boldsymbol{x}}=\begin{bmatrix}\dfrac{\partial f}{\partial \xi_1}\\ \vdots\\ \dfrac{\partial f}{\partial \xi_n}\end{bmatrix}$$

例11: $\boldsymbol{x}=\begin{bmatrix}\xi_1\\ \vdots\\ \xi_n\end{bmatrix}$: $\boldsymbol{f}(\boldsymbol{x})=\boldsymbol{f}(\xi_1,\xi_2,\cdots,\xi_n)$ $\quad \dfrac{\boldsymbol{df}}{\boldsymbol{dx}}=\begin{bmatrix}\dfrac{\partial f}{\partial \xi_1}\\ \vdots\\ \dfrac{\partial f}{\partial \xi_n}\end{bmatrix}$

例12: $\boldsymbol{A}=\left(\boldsymbol{a}_{ij}\right)_{m\times n}$, $\boldsymbol{x}=\begin{bmatrix}\xi_1\\ \vdots\\ \xi_n\end{bmatrix}$: $\boldsymbol{f}(\boldsymbol{x})=\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}$, 求 $\dfrac{\boldsymbol{df}}{\boldsymbol{dx}}$

例11: $x=\begin{bmatrix}\xi_1\\ \vdots\\ \xi_n\end{bmatrix}$: $f(x)=f(\xi_1,\xi_2,\cdots,\xi_n)$ $\quad \dfrac{df}{dx}=\begin{bmatrix}\dfrac{\partial f}{\partial \xi_1}\\ \vdots\\ \dfrac{\partial f}{\partial \xi_n}\end{bmatrix}$

例12: $A=(a_{ij})_{m\times n}$, $x=\begin{bmatrix}\xi_1\\ \vdots\\ \xi_n\end{bmatrix}$: $f(x)=x^TAx$, 求 $\dfrac{df}{dx}$

$$f(x)=\xi_1\sum_{j=1}^{n}a_{1j}\xi_j+\cdots+\xi_k\sum_{j=1}^{n}a_{kj}\xi_j+\cdots+\xi_n\sum_{j=1}^{n}a_{nj}\xi_j$$

$$\frac{\partial f}{\partial \xi_k}=\xi_1a_{1k}+\cdots+\xi_{k-1}a_{k-1,k}+\left(\sum_{j=1}^{n}a_{kj}\xi_j+\xi_ka_{kk}\right)$$

$$+\xi_{k+1}a_{k+1,k}+\cdots+\xi_na_{nk}\quad=\sum_{j=1}^{n}a_{kj}\xi_j+\sum_{i=1}^{n}a_{ik}\xi_i$$

例12: $A=\left(a_{ij}\right)_{m\times n}, x=\begin{bmatrix}\xi_1\\ \vdots\\ \xi_n\end{bmatrix}$: $f(x)=x^T Ax$, 求 $\dfrac{df}{dx}$

$$\therefore \frac{df}{dx}=\left(A+A^T\right)x$$

如果 $A=A^T$ ，有 $\dfrac{df}{dx}=2Ax$

矩阵对矩阵的导数

定义:设 $X=\left(\xi_{ij}\right)_{m\times n}, f_{kl}(X)=f_{kl}\left(\xi_{11},\xi_{12},\cdots,\xi_{1n},\cdots,\xi_{m\times n}\right)$

$$F=\begin{bmatrix} f_{11} & \cdots & f_{1s} \\ \vdots & & \vdots \\ f_{r1} & \cdots & f_{rs} \end{bmatrix}, \qquad \frac{\partial F}{\partial \xi_{ij}}=\begin{bmatrix} \frac{\partial f_{11}}{\partial \xi_{ij}} & \cdots & \frac{\partial f_{1s}}{\partial \xi_{ij}} \\ \vdots & & \vdots \\ \frac{\partial f_{r1}}{\partial \xi_{ij}} & \cdots & \frac{\partial f_{rs}}{\partial \xi_{ij}} \end{bmatrix},$$

定义 $$\frac{dF}{dX}=\begin{bmatrix} \frac{\partial F}{\partial \xi_{11}} & \cdots & \frac{\partial F}{\partial \xi_{1n}} \\ \vdots & & \vdots \\ \frac{\partial F}{\partial \xi_{m1}} & \cdots & \frac{\partial F}{\partial \xi_{mn}} \end{bmatrix}$$

例15: $\boldsymbol{x}=\begin{bmatrix}\xi_1\\ \vdots\\ \xi_n\end{bmatrix}$, $\boldsymbol{F}(\boldsymbol{x})=\left[f_1(\boldsymbol{x}),f_2(\boldsymbol{x}),\cdots,f_l(\boldsymbol{x})\right]$

$$\frac{d\boldsymbol{F}}{d\boldsymbol{x}}=$$

例15: $\boldsymbol{x}=\begin{bmatrix}\xi_1\\ \vdots\\ \xi_n\end{bmatrix}$, $\boldsymbol{F}(\boldsymbol{x})=\left[f_1(\boldsymbol{x}),f_2(\boldsymbol{x}),\cdots,f_l(\boldsymbol{x})\right]$

$$\frac{d\boldsymbol{F}}{d\boldsymbol{x}}=\begin{bmatrix}\dfrac{\partial f_1}{\partial \xi_1} & \cdots & \dfrac{\partial f_l}{\partial \xi_1}\\ \vdots & & \vdots\\ \dfrac{\partial f_1}{\partial \xi_n} & \cdots & \dfrac{\partial f_l}{\partial \xi_n}\end{bmatrix}$$

例15: $x = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix}$, $F(x) = \left[f_1(x), f_2(x), \cdots, f_l(x) \right]$

$$\frac{dF}{dx} = \begin{bmatrix} \dfrac{\partial f_1}{\partial \xi_1} & \cdots & \dfrac{\partial f_l}{\partial \xi_1} \\ \vdots & & \vdots \\ \dfrac{\partial f_1}{\partial \xi_n} & \cdots & \dfrac{\partial f_l}{\partial \xi_n} \end{bmatrix}$$

例16: $A = (a_{ij})_{n \times n}$, $x = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix}$ $Ax = \begin{bmatrix} \sum_{j=1}^{n} a_{1j}\xi_j \\ \vdots \\ \sum_{j=1}^{n} a_{nj}\xi_j \end{bmatrix}$

$$\frac{d(Ax)}{dx^T} =$$

例15: $x=\begin{bmatrix}\xi_1\\ \vdots\\ \xi_n\end{bmatrix}$, $F(x)=\left[f_1(x),f_2(x),\cdots,f_l(x)\right]$

$$\frac{dF}{dx}=\begin{bmatrix}\frac{\partial f_1}{\partial \xi_1} & \cdots & \frac{\partial f_l}{\partial \xi_1}\\ \vdots & & \vdots\\ \frac{\partial f_1}{\partial \xi_n} & \cdots & \frac{\partial f_l}{\partial \xi_n}\end{bmatrix}$$

例16: $A=\left(a_{ij}\right)_{n\times n}$, $x=\begin{bmatrix}\xi_1\\ \vdots\\ \xi_n\end{bmatrix}$ $Ax=\begin{bmatrix}\sum_{j=1}^{n} a_{1j}\xi_j\\ \vdots\\ \sum_{j=1}^{n} a_{nj}\xi_j\end{bmatrix}$

$$\frac{d(Ax)}{dx^T}=\begin{bmatrix}a_{11} & \cdots & a_{1n}\\ \vdots & & \vdots\\ a_{n1} & \cdots & a_{nn}\end{bmatrix}=A$$

Types of Matrix Derivatives

Types	Scalar	Vector	Matrix
Scalar	$\frac{\partial y}{\partial x}$	$\frac{\partial \mathbf{y}}{\partial x}$	$\frac{\partial \mathbf{Y}}{\partial x}$
Vector	$\frac{\partial y}{\partial \mathbf{x}}$	$\frac{\partial \mathbf{y}}{\partial \mathbf{x}}$	
Matrix	$\frac{\partial y}{\partial \mathbf{X}}$		

Vector-by-scalar [edit]

The derivative of a vector $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$, by a scalar x is written (in numerator layout notation) as

$$\frac{\partial \mathbf{y}}{\partial x} = \begin{bmatrix} \frac{\partial y_1}{\partial x} \\ \frac{\partial y_2}{\partial x} \\ \vdots \\ \frac{\partial y_m}{\partial x} \end{bmatrix}.$$

Scalar-by-vector [edit]

The derivative of a scalar y by a vector $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, is written (in denominator layout notation) as $\frac{\partial y}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial y}{\partial x_1} \\ \frac{\partial y}{\partial x_2} \\ \vdots \\ \frac{\partial y}{\partial x_n} \end{bmatrix}$.

Vector-by-vector [edit]

Each of the previous two cases can be considered as an application of the derivative of a vector with respect to a vector, using a vector of
involving vectors in a corresponding way.

The derivative of a vector function (a vector whose components are functions) $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$, with respect to an input vector, $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$,

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_m}{\partial x_1} & \frac{\partial y_m}{\partial x_2} & \cdots & \frac{\partial y_m}{\partial x_n} \end{bmatrix}.$$

Matrix-by-scalar [edit]

The derivative of a matrix function **Y** by a scalar x is known as the **tangent matrix** and is given (in numerator layout notation) by

$$\frac{\partial \mathbf{Y}}{\partial x} = \begin{bmatrix} \frac{\partial y_{11}}{\partial x} & \frac{\partial y_{12}}{\partial x} & \cdots & \frac{\partial y_{1n}}{\partial x} \\ \frac{\partial y_{21}}{\partial x} & \frac{\partial y_{22}}{\partial x} & \cdots & \frac{\partial y_{2n}}{\partial x} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_{m1}}{\partial x} & \frac{\partial y_{m2}}{\partial x} & \cdots & \frac{\partial y_{mn}}{\partial x} \end{bmatrix}.$$

Scalar-by-matrix [edit]

The derivative of a scalar y function of a $p \times q$ matrix **X** of independent variables, with respect to the matrix **X**, is given (in numerator layout notation) by

$$\frac{\partial y}{\partial \mathbf{X}} = \begin{bmatrix} \frac{\partial y}{\partial x_{11}} & \frac{\partial y}{\partial x_{21}} & \cdots & \frac{\partial y}{\partial x_{p1}} \\ \frac{\partial y}{\partial x_{12}} & \frac{\partial y}{\partial x_{22}} & \cdots & \frac{\partial y}{\partial x_{p2}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y}{\partial x_{1q}} & \frac{\partial y}{\partial x_{2q}} & \cdots & \frac{\partial y}{\partial x_{pq}} \end{bmatrix}.$$

Result of differentiating various kinds of aggregates with other kinds of aggregates

	Scalar y		Vector y (size m)		Matrix Y (size $m\times n$)	
	Notation	Type	Notation	Type	Notation	Type
Scalar x	$\frac{\partial y}{\partial x}$	scalar	$\frac{\partial \mathbf{y}}{\partial x}$	(numerator layout) size-m column vector (denominator layout) size-m row vector	$\frac{\partial \mathbf{Y}}{\partial x}$	(numerator layout) $m\times n$ matrix
Vector x (size n)	$\frac{\partial y}{\partial \mathbf{x}}$	(numerator layout) size-n row vector (denominator layout) size-n column vector	$\frac{\partial \mathbf{y}}{\partial \mathbf{x}}$	(numerator layout) $m\times n$ matrix (denominator layout) $n\times m$ matrix	$\frac{\partial \mathbf{Y}}{\partial \mathbf{x}}$	
Matrix X (size $p\times q$)	$\frac{\partial y}{\partial \mathbf{X}}$	(numerator layout) $q\times p$ matrix (denominator layout) $p\times q$ matrix	$\frac{\partial \mathbf{y}}{\partial \mathbf{X}}$		$\frac{\partial \mathbf{Y}}{\partial \mathbf{X}}$	

分子布局

Numerator-layout notation [edit]

Using numerator-layout notation, we have:[1]

$$\frac{\partial y}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial y}{\partial x_1} & \frac{\partial y}{\partial x_2} & \cdots & \frac{\partial y}{\partial x_n} \end{bmatrix}.$$

$$\frac{\partial \mathbf{y}}{\partial x} = \begin{bmatrix} \frac{\partial y_1}{\partial x} \\ \frac{\partial y_2}{\partial x} \\ \vdots \\ \frac{\partial y_m}{\partial x} \end{bmatrix}.$$

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_m}{\partial x_1} & \frac{\partial y_m}{\partial x_2} & \cdots & \frac{\partial y_m}{\partial x_n} \end{bmatrix}.$$

$$\frac{\partial y}{\partial \mathbf{X}} = \begin{bmatrix} \frac{\partial y}{\partial x_{11}} & \frac{\partial y}{\partial x_{21}} & \cdots & \frac{\partial y}{\partial x_{p1}} \\ \frac{\partial y}{\partial x_{12}} & \frac{\partial y}{\partial x_{22}} & \cdots & \frac{\partial y}{\partial x_{p2}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y}{\partial x_{1q}} & \frac{\partial y}{\partial x_{2q}} & \cdots & \frac{\partial y}{\partial x_{pq}} \end{bmatrix}.$$

分母布局

Denominator-layout notation [edit]

Using denominator-layout notation, we have:[3]

$$\frac{\partial y}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial y}{\partial x_1} \\ \frac{\partial y}{\partial x_2} \\ \vdots \\ \frac{\partial y}{\partial x_n} \end{bmatrix}.$$

$$\frac{\partial \mathbf{y}}{\partial x} = \begin{bmatrix} \frac{\partial y_1}{\partial x} & \frac{\partial y_2}{\partial x} & \cdots & \frac{\partial y_m}{\partial x} \end{bmatrix}.$$

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_2}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_1} \\ \frac{\partial y_1}{\partial x_2} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_m}{\partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_1}{\partial x_n} & \frac{\partial y_2}{\partial x_n} & \cdots & \frac{\partial y_m}{\partial x_n} \end{bmatrix}.$$

$$\frac{\partial y}{\partial \mathbf{X}} = \begin{bmatrix} \frac{\partial y}{\partial x_{11}} & \frac{\partial y}{\partial x_{12}} & \cdots & \frac{\partial y}{\partial x_{1q}} \\ \frac{\partial y}{\partial x_{21}} & \frac{\partial y}{\partial x_{22}} & \cdots & \frac{\partial y}{\partial x_{2q}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y}{\partial x_{p1}} & \frac{\partial y}{\partial x_{p2}} & \cdots & \frac{\partial y}{\partial x_{pq}} \end{bmatrix}.$$

Identities: vector-by-vector $\frac{\partial \mathbf{y}}{\partial \mathbf{x}}$

Condition	Expression	Numerator layout, i.e. by y and x^T	Denominator layout, i.e. by y^T and x
a is not a function of **x**	$\frac{\partial \mathbf{a}}{\partial \mathbf{x}} =$	$\mathbf{0}$	
	$\frac{\partial \mathbf{x}}{\partial \mathbf{x}} =$	$\mathbf{I}$	
A is not a function of **x**	$\frac{\partial \mathbf{A}\mathbf{x}}{\partial \mathbf{x}} =$	$\mathbf{A}$	$\mathbf{A}^\top$
A is not a function of **x**	$\frac{\partial \mathbf{x}^\top \mathbf{A}}{\partial \mathbf{x}} =$	$\mathbf{A}^\top$	$\mathbf{A}$
a is not a function of **x**, **u** = **u**(**x**)	$\frac{\partial a\mathbf{u}}{\partial \mathbf{x}} =$	$a\frac{\partial \mathbf{u}}{\partial \mathbf{x}}$	
a = a(**x**), **u** = **u**(**x**)	$\frac{\partial a\mathbf{u}}{\partial \mathbf{x}} =$	$a\frac{\partial \mathbf{u}}{\partial \mathbf{x}} + \mathbf{u}\frac{\partial a}{\partial \mathbf{x}}$	$a\frac{\partial \mathbf{u}}{\partial \mathbf{x}} + \frac{\partial a}{\partial \mathbf{x}}\mathbf{u}^\top$
A is not a function of **x**, **u** = **u**(**x**)	$\frac{\partial \mathbf{A}\mathbf{u}}{\partial \mathbf{x}} =$	$\mathbf{A}\frac{\partial \mathbf{u}}{\partial \mathbf{x}}$	$\frac{\partial \mathbf{u}}{\partial \mathbf{x}}\mathbf{A}^\top$
u = **u**(**x**), **v** = **v**(**x**)	$\frac{\partial (\mathbf{u}+\mathbf{v})}{\partial \mathbf{x}} =$	$\frac{\partial \mathbf{u}}{\partial \mathbf{x}} + \frac{\partial \mathbf{v}}{\partial \mathbf{x}}$	
u = **u**(**x**)	$\frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{x}} =$	$\frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{u}}\frac{\partial \mathbf{u}}{\partial \mathbf{x}}$	$\frac{\partial \mathbf{u}}{\partial \mathbf{x}}\frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{u}}$
u = **u**(**x**)	$\frac{\partial \mathbf{f}(\mathbf{g}(\mathbf{u}))}{\partial \mathbf{x}} =$	$\frac{\partial \mathbf{f}(\mathbf{g})}{\partial \mathbf{g}}\frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{u}}\frac{\partial \mathbf{u}}{\partial \mathbf{x}}$	$\frac{\partial \mathbf{u}}{\partial \mathbf{x}}\frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{u}}\frac{\partial \mathbf{f}(\mathbf{g})}{\partial \mathbf{g}}$

Identities: scalar-by-vector $\frac{\partial y}{\partial \mathbf{x}} = \nabla_{\mathbf{x}} y$

Condition	Expression	Numerator layout, i.e. by $\mathbf{x}^{\mathsf{T}}$; result is row vector	Denominator layout, i.e. by $\mathbf{x}$; result is column vector
a is not a function of $\mathbf{x}$	$\frac{\partial a}{\partial \mathbf{x}} =$	$\mathbf{0}^{\top}$ [4]	$\mathbf{0}$ [4]
a is not a function of $\mathbf{x}$, $u = u(\mathbf{x})$	$\frac{\partial au}{\partial \mathbf{x}} =$	$a\frac{\partial u}{\partial \mathbf{x}}$	
$u = u(\mathbf{x})$, $v = v(\mathbf{x})$	$\frac{\partial (u+v)}{\partial \mathbf{x}} =$	$\frac{\partial u}{\partial \mathbf{x}} + \frac{\partial v}{\partial \mathbf{x}}$	
$u = u(\mathbf{x})$, $v = v(\mathbf{x})$	$\frac{\partial uv}{\partial \mathbf{x}} =$	$u\frac{\partial v}{\partial \mathbf{x}} + v\frac{\partial u}{\partial \mathbf{x}}$	
$u = u(\mathbf{x})$	$\frac{\partial g(u)}{\partial \mathbf{x}} =$	$\frac{\partial g(u)}{\partial u}\frac{\partial u}{\partial \mathbf{x}}$	
$u = u(\mathbf{x})$	$\frac{\partial f(g(u))}{\partial \mathbf{x}} =$	$\frac{\partial f(g)}{\partial g}\frac{\partial g(u)}{\partial u}\frac{\partial u}{\partial \mathbf{x}}$	
$\mathbf{u} = \mathbf{u}(\mathbf{x})$, $\mathbf{v} = \mathbf{v}(\mathbf{x})$	$\frac{\partial (\mathbf{u} \cdot \mathbf{v})}{\partial \mathbf{x}} = \frac{\partial \mathbf{u}^{\top}\mathbf{v}}{\partial \mathbf{x}} =$	$\mathbf{u}^{\top}\frac{\partial \mathbf{v}}{\partial \mathbf{x}} + \mathbf{v}^{\top}\frac{\partial \mathbf{u}}{\partial \mathbf{x}}$ • assumes numerator layout of $\frac{\partial \mathbf{u}}{\partial \mathbf{x}}, \frac{\partial \mathbf{v}}{\partial \mathbf{x}}$	$\frac{\partial \mathbf{u}}{\partial \mathbf{x}}\mathbf{v} + \frac{\partial \mathbf{v}}{\partial \mathbf{x}}\mathbf{u}$ • assumes denominator layout of $\frac{\partial \mathbf{u}}{\partial \mathbf{x}}, \frac{\partial \mathbf{v}}{\partial \mathbf{x}}$
$\mathbf{u} = \mathbf{u}(\mathbf{x})$, $\mathbf{v} = \mathbf{v}(\mathbf{x})$, $\mathbf{A}$ is not a function of $\mathbf{x}$	$\frac{\partial (\mathbf{u} \cdot \mathbf{A}\mathbf{v})}{\partial \mathbf{x}} = \frac{\partial \mathbf{u}^{\top}\mathbf{A}\mathbf{v}}{\partial \mathbf{x}} =$	$\mathbf{u}^{\top}\mathbf{A}\frac{\partial \mathbf{v}}{\partial \mathbf{x}} + \mathbf{v}^{\top}\mathbf{A}^{\top}\frac{\partial \mathbf{u}}{\partial \mathbf{x}}$ • assumes numerator layout of $\frac{\partial \mathbf{u}}{\partial \mathbf{x}}, \frac{\partial \mathbf{v}}{\partial \mathbf{x}}$	$\frac{\partial \mathbf{u}}{\partial \mathbf{x}}\mathbf{A}\mathbf{v} + \frac{\partial \mathbf{v}}{\partial \mathbf{x}}\mathbf{A}^{\top}\mathbf{u}$ • assumes denominator layout of $\frac{\partial \mathbf{u}}{\partial \mathbf{x}}, \frac{\partial \mathbf{v}}{\partial \mathbf{x}}$
	$\frac{\partial^2 f}{\partial \mathbf{x} \partial \mathbf{x}^T} =$		$\mathbf{H}$, the Hessian matrix[5]

a is not a function of **x**	$\frac{\partial(\mathbf{a}\cdot\mathbf{x})}{\partial\mathbf{x}} = \frac{\partial(\mathbf{x}\cdot\mathbf{a})}{\partial\mathbf{x}} =$ $\frac{\partial\mathbf{a}^\top\mathbf{x}}{\partial\mathbf{x}} = \frac{\partial\mathbf{x}^\top\mathbf{a}}{\partial\mathbf{x}} =$	$\mathbf{a}^\top$	$\mathbf{a}$
A is not a function of **x** **b** is not a function of **x**	$\frac{\partial\mathbf{b}^\top\mathbf{A}\mathbf{x}}{\partial\mathbf{x}} =$	$\mathbf{b}^\top\mathbf{A}$	$\mathbf{A}^\top\mathbf{b}$
A is not a function of **x**	$\frac{\partial\mathbf{x}^\top\mathbf{A}\mathbf{x}}{\partial\mathbf{x}} =$	$\mathbf{x}^\top(\mathbf{A}+\mathbf{A}^\top)$	$(\mathbf{A}+\mathbf{A}^\top)\mathbf{x}$
A is not a function of **x** **A** is symmetric	$\frac{\partial\mathbf{x}^\top\mathbf{A}\mathbf{x}}{\partial\mathbf{x}} =$	$2\mathbf{x}^\top\mathbf{A}$	$2\mathbf{A}\mathbf{x}$
A is not a function of **x**	$\frac{\partial^2\mathbf{x}^\top\mathbf{A}\mathbf{x}}{\partial\mathbf{x}^2} =$	$\mathbf{A}+\mathbf{A}^\top$	
A is not a function of **x** **A** is symmetric	$\frac{\partial^2\mathbf{x}^\top\mathbf{A}\mathbf{x}}{\partial\mathbf{x}^2} =$	$2\mathbf{A}$	
	$\frac{\partial(\mathbf{x}\cdot\mathbf{x})}{\partial\mathbf{x}} = \frac{\partial\mathbf{x}^\top\mathbf{x}}{\partial\mathbf{x}} =$	$2\mathbf{x}^\top$	$2\mathbf{x}$
a is not a function of **x**, **u** = **u**(**x**)	$\frac{\partial(\mathbf{a}\cdot\mathbf{u})}{\partial\mathbf{x}} = \frac{\partial\mathbf{a}^\top\mathbf{u}}{\partial\mathbf{x}} =$	$\mathbf{a}^\top\frac{\partial\mathbf{u}}{\partial\mathbf{x}}$ • assumes numerator layout of $\frac{\partial\mathbf{u}}{\partial\mathbf{x}}$	$\frac{\partial\mathbf{u}}{\partial\mathbf{x}}\mathbf{a}$ • assumes denominator layout of $\frac{\partial\mathbf{u}}{\partial\mathbf{x}}$
a, **b** are not functions of **x**	$\frac{\partial\,\mathbf{a}^\top\mathbf{x}\mathbf{x}^\top\mathbf{b}}{\partial\,\mathbf{x}} =$	$\mathbf{x}^\top(\mathbf{a}\mathbf{b}^\top+\mathbf{b}\mathbf{a}^\top)$	$(\mathbf{a}\mathbf{b}^\top+\mathbf{b}\mathbf{a}^\top)\mathbf{x}$
A, **b**, **C**, **D**, **e** are not functions of **x**	$\frac{\partial\,(\mathbf{A}\mathbf{x}+\mathbf{b})^\top\mathbf{C}(\mathbf{D}\mathbf{x}+\mathbf{e})}{\partial\,\mathbf{x}} =$	$(\mathbf{D}\mathbf{x}+\mathbf{e})^\top\mathbf{C}^\top\mathbf{A}+(\mathbf{A}\mathbf{x}+\mathbf{b})^\top\mathbf{C}\mathbf{D}$	$\mathbf{D}^\top\mathbf{C}^\top(\mathbf{A}\mathbf{x}+\mathbf{b})+\mathbf{A}^\top\mathbf{C}(\mathbf{D}\mathbf{x}+\mathbf{e})$
a is not a function of **x**	$\frac{\partial\,\lVert\mathbf{x}-\mathbf{a}\rVert}{\partial\,\mathbf{x}} =$	$\frac{(\mathbf{x}-\mathbf{a})^\top}{\lVert\mathbf{x}-\mathbf{a}\rVert}$	$\frac{\mathbf{x}-\mathbf{a}}{\lVert\mathbf{x}-\mathbf{a}\rVert}$

Identities: vector-by-scalar $\frac{\partial \mathbf{y}}{\partial x}$

Condition	Expression	Numerator layout, i.e. by y, result is column vector	Denominator layout, i.e. by $\mathbf{y}^T$, result is row vector
a is not a function of *x*	$\frac{\partial \mathbf{a}}{\partial x} =$	$\mathbf{0}$[4]	
a is not a function of *x*, **u** = **u**(*x*)	$\frac{\partial a\mathbf{u}}{\partial x} =$	$a\frac{\partial \mathbf{u}}{\partial x}$	
A is not a function of *x*, **u** = **u**(*x*)	$\frac{\partial \mathbf{A}\mathbf{u}}{\partial x} =$	$\mathbf{A}\frac{\partial \mathbf{u}}{\partial x}$	$\frac{\partial \mathbf{u}}{\partial x}\mathbf{A}^\top$
u = **u**(*x*)	$\frac{\partial \mathbf{u}^\top}{\partial x} =$	$\left(\frac{\partial \mathbf{u}}{\partial x}\right)^\top$	
u = **u**(*x*), **v** = **v**(*x*)	$\frac{\partial (\mathbf{u}+\mathbf{v})}{\partial x} =$	$\frac{\partial \mathbf{u}}{\partial x}+\frac{\partial \mathbf{v}}{\partial x}$	
u = **u**(*x*), **v** = **v**(*x*)	$\frac{\partial (\mathbf{u}^\top \times \mathbf{v})}{\partial x} =$	$\left(\frac{\partial \mathbf{u}}{\partial x}\right)^\top \times \mathbf{v} + \mathbf{u}^\top \times \frac{\partial \mathbf{v}}{\partial x}$	$\frac{\partial \mathbf{u}}{\partial x} \times \mathbf{v} + \mathbf{u}^\top \times \left(\frac{\partial \mathbf{v}}{\partial x}\right)^\top$
u = **u**(*x*)	$\frac{\partial \mathbf{g}(\mathbf{u})}{\partial x} =$	$\frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{u}}\frac{\partial \mathbf{u}}{\partial x}$	$\frac{\partial \mathbf{u}}{\partial x}\frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{u}}$
		Assumes consistent matrix layout; see below.	
u = **u**(*x*)	$\frac{\partial \mathbf{f}(\mathbf{g}(\mathbf{u}))}{\partial x} =$	$\frac{\partial \mathbf{f}(\mathbf{g})}{\partial \mathbf{g}}\frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{u}}\frac{\partial \mathbf{u}}{\partial x}$	$\frac{\partial \mathbf{u}}{\partial x}\frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{u}}\frac{\partial \mathbf{f}(\mathbf{g})}{\partial \mathbf{g}}$
		Assumes consistent matrix layout; see below.	
U = **U**(*x*), **v** = **v**(*x*)	$\frac{\partial (\mathbf{U} \times \mathbf{v})}{\partial x} =$	$\frac{\partial \mathbf{U}}{\partial x} \times \mathbf{v} + \mathbf{U} \times \frac{\partial \mathbf{v}}{\partial x}$	$\mathbf{v}^\top \times \left(\frac{\partial \mathbf{U}}{\partial x}\right) + \frac{\partial \mathbf{v}}{\partial x} \times \mathbf{U}^\top$

Identities: scalar-by-matrix $\frac{\partial y}{\partial \mathbf{X}}$

<table>
<tr><th>Condition</th><th>Expression</th><th>Numerator layout, i.e. by $\mathbf{X}^\mathsf{T}$</th><th>Denominator layout, i.e. by $\mathbf{X}$</th></tr>
<tr><td>a is not a function of $\mathbf{X}$</td><td>$\frac{\partial a}{\partial \mathbf{X}} =$</td><td>$\mathbf{0}^\top$ [6]</td><td>$\mathbf{0}$ [6]</td></tr>
<tr><td>a is not a function of $\mathbf{X}$, $u = u(\mathbf{X})$</td><td>$\frac{\partial au}{\partial \mathbf{X}} =$</td><td colspan="2">$a\frac{\partial u}{\partial \mathbf{X}}$</td></tr>
<tr><td>$u = u(\mathbf{X})$, $v = v(\mathbf{X})$</td><td>$\frac{\partial (u+v)}{\partial \mathbf{X}} =$</td><td colspan="2">$\frac{\partial u}{\partial \mathbf{X}} + \frac{\partial v}{\partial \mathbf{X}}$</td></tr>
<tr><td>$u = u(\mathbf{X})$, $v = v(\mathbf{X})$</td><td>$\frac{\partial uv}{\partial \mathbf{X}} =$</td><td colspan="2">$u\frac{\partial v}{\partial \mathbf{X}} + v\frac{\partial u}{\partial \mathbf{X}}$</td></tr>
<tr><td>$u = u(\mathbf{X})$</td><td>$\frac{\partial g(u)}{\partial \mathbf{X}} =$</td><td colspan="2">$\frac{\partial g(u)}{\partial u}\frac{\partial u}{\partial \mathbf{X}}$</td></tr>
<tr><td>$u = u(\mathbf{X})$</td><td>$\frac{\partial f(g(u))}{\partial \mathbf{X}} =$</td><td colspan="2">$\frac{\partial f(g)}{\partial g}\frac{\partial g(u)}{\partial u}\frac{\partial u}{\partial \mathbf{X}}$</td></tr>
<tr><td rowspan="2">$\mathbf{U} = \mathbf{U}(\mathbf{X})$</td><td rowspan="2">[5] $\frac{\partial g(\mathbf{U})}{\partial X_{ij}} =$</td><td>$\operatorname{tr}\left(\frac{\partial g(\mathbf{U})}{\partial \mathbf{U}}\frac{\partial \mathbf{U}}{\partial X_{ij}}\right)$</td><td>$\operatorname{tr}\left(\left(\frac{\partial g(\mathbf{U})}{\partial \mathbf{U}}\right)^\top\frac{\partial \mathbf{U}}{\partial X_{ij}}\right)$</td></tr>
<tr><td colspan="2">Both forms assume numerator layout for $\frac{\partial \mathbf{U}}{\partial X_{ij}}$,
i.e. mixed layout if denominator layout for $\mathbf{X}$ is being used.</td></tr>
<tr><td>a and b are not functions of $\mathbf{X}$</td><td>$\frac{\partial a^\top \mathbf{X} b}{\partial \mathbf{X}} =$</td><td>$ba^\top$</td><td>$ab^\top$</td></tr>
<tr><td>a and b are not functions of $\mathbf{X}$</td><td>$\frac{\partial a^\top \mathbf{X}^\top b}{\partial \mathbf{X}} =$</td><td>$ab^\top$</td><td>$ba^\top$</td></tr>
<tr><td>a, b and $\mathbf{C}$ are not functions of $\mathbf{X}$</td><td>$\frac{\partial (\mathbf{X}a+b)^\top \mathbf{C}(\mathbf{X}a+b)}{\partial \mathbf{X}} =$</td><td>$((\mathbf{C}+\mathbf{C}^\top)(\mathbf{X}a+b)a^\top)^\top$</td><td>$(\mathbf{C}+\mathbf{C}^\top)(\mathbf{X}a+b)a^\top$</td></tr>
<tr><td>a, b and $\mathbf{C}$ are not functions of $\mathbf{X}$</td><td>$\frac{\partial (\mathbf{X}a)^\top \mathbf{C}(\mathbf{X}b)}{\partial \mathbf{X}} =$</td><td>$(\mathbf{C}\mathbf{X}ba^\top + \mathbf{C}^\top\mathbf{X}ab^\top)^\top$</td><td>$\mathbf{C}\mathbf{X}ba^\top + \mathbf{C}^\top\mathbf{X}ab^\top$</td></tr>
</table>

Identities: matrix-by-scalar $\frac{\partial \mathbf{Y}}{\partial x}$

Condition	Expression	Numerator layout, i.e. by Y
U = **U**(*x*)	$\frac{\partial a\mathbf{U}}{\partial x} =$	$a\frac{\partial \mathbf{U}}{\partial x}$
A, **B** are not functions of . **U** = **U**(*x*)	$\frac{\partial \mathbf{AUB}}{\partial x} =$	$\mathbf{A}\frac{\partial \mathbf{U}}{\partial x}\mathbf{B}$
U = **U**(*x*), **V** = **V**(*x*)	$\frac{\partial(\mathbf{U}+\mathbf{V})}{\partial x} =$	$\frac{\partial \mathbf{U}}{\partial x} + \frac{\partial \mathbf{V}}{\partial x}$
U = **U**(*x*), **V** = **V**(*x*)	$\frac{\partial(\mathbf{UV})}{\partial x} =$	$\mathbf{U}\frac{\partial \mathbf{V}}{\partial x} + \frac{\partial \mathbf{U}}{\partial x}\mathbf{V}$
U = **U**(*x*), **V** = **V**(*x*)	$\frac{\partial(\mathbf{U}\otimes\mathbf{V})}{\partial x} =$	$\mathbf{U}\otimes\frac{\partial \mathbf{V}}{\partial x} + \frac{\partial \mathbf{U}}{\partial x}\otimes\mathbf{V}$
U = **U**(*x*), **V** = **V**(*x*)	$\frac{\partial(\mathbf{U}\circ\mathbf{V})}{\partial x} =$	$\mathbf{U}\circ\frac{\partial \mathbf{V}}{\partial x} + \frac{\partial \mathbf{U}}{\partial x}\circ\mathbf{V}$
U = **U**(*x*)	$\frac{\partial \mathbf{U}^{-1}}{\partial x} =$	$-\mathbf{U}^{-1}\frac{\partial \mathbf{U}}{\partial x}\mathbf{U}^{-1}$
U = **U**(*x*,*y*)	$\frac{\partial^2 \mathbf{U}^{-1}}{\partial x \partial y} =$	$\mathbf{U}^{-1}\left(\frac{\partial \mathbf{U}}{\partial x}\mathbf{U}^{-1}\frac{\partial \mathbf{U}}{\partial y} - \frac{\partial^2 \mathbf{U}}{\partial x \partial y} + \frac{\partial \mathbf{U}}{\partial y}\mathbf{U}^{-1}\frac{\partial \mathbf{U}}{\partial x}\right)\mathbf{U}^{-1}$
any matrix function defined s its derivative, and **g**′(**X**) is	$\frac{\partial\, \mathbf{g}(x\mathbf{A})}{\partial x} =$	$\mathbf{A}\mathbf{g}'(x\mathbf{A}) = \mathbf{g}'(x\mathbf{A})\mathbf{A}$
A is not a function of *x*	$\frac{\partial e^{x\mathbf{A}}}{\partial x} =$	$\mathbf{A}e^{x\mathbf{A}} = e^{x\mathbf{A}}\mathbf{A}$

下课，谢谢大家！