

Convention:

variable type	notation	examples
scalar	regular lowercase	$x, f, n$
vector	bold lowercase	$\mathbf{y}, \mathbf{f}, \mathbf{b}$
matrix	uppercase	$X, Y, A$
independent variable	$\mathbf{x}$	$x, \mathbf{x}, X$
dependent variable	$y$ or $\mathbf{z}$	$y, \mathbf{z}, Y$
function	$f$	$f(\mathbf{x}), \mathbf{f}(y)$
constant	$a, b$ , or $c$	$A, \mathbf{b}, c$
dimension	$n, m$ , or $k$	$n, m, k$

## 1 Definitions

$$\underbrace{\nabla_{\mathbf{x}}}_{n \times 1} y = \begin{bmatrix} \frac{\partial y}{\partial x_1} \\ \vdots \\ \frac{\partial y}{\partial x_n} \end{bmatrix} \quad (1)$$

$$\underbrace{J_{\mathbf{x}}}_{m \times n} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_m}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \frac{\partial \mathbf{y}}{\partial x_1} & \cdots & \frac{\partial \mathbf{y}}{\partial x_n} \end{bmatrix} = \begin{bmatrix} (\nabla_{\mathbf{x}} y_1)^T \\ \vdots \\ (\nabla_{\mathbf{x}} y_m)^T \end{bmatrix} \quad (2)$$

$$\underbrace{\nabla_{\mathbf{x}}^2}_{n \times n} y = \begin{bmatrix} \frac{\partial^2 y}{\partial x_1^2} & \cdots & \frac{\partial^2 y}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 y}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 y}{\partial x_n^2} \end{bmatrix} = \begin{bmatrix} \nabla_{\mathbf{x}} \frac{\partial y}{\partial x_1} & \cdots & \nabla_{\mathbf{x}} \frac{\partial y}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \left( \frac{\partial}{\partial x_1} \nabla_{\mathbf{x}} y \right)^T \\ \vdots \\ \left( \frac{\partial}{\partial x_n} \nabla_{\mathbf{x}} y \right)^T \end{bmatrix} \quad (3)$$

$$\underbrace{\nabla_X}_{m \times n} y = \begin{bmatrix} \frac{\partial y}{\partial X_{11}} & \cdots & \frac{\partial y}{\partial X_{1n}} \\ \vdots & \ddots & \vdots \\ \frac{\partial y}{\partial X_{m1}} & \cdots & \frac{\partial y}{\partial X_{mn}^2} \end{bmatrix} \quad (4)$$

## 2 Matrix-Vector Derivatives

$$\underbrace{\nabla_{\mathbf{x}} \mathbf{a}^T}_{n \times 1} \underbrace{\mathbf{x}}_{n \times 1} = \underbrace{\mathbf{a}}_{n \times 1} \quad (5)$$

$$\underbrace{\nabla_{\mathbf{x}} \mathbf{x}^T}_{n \times 1} \underbrace{A}_{n \times n} \underbrace{\mathbf{x}}_{n \times 1} = \underbrace{(A + A^T)}_{n \times n} \underbrace{\mathbf{x}}_{n \times 1} \quad (6)$$

$$\underbrace{\nabla_{\mathbf{x}}^2 \mathbf{x}^T}_{n \times 1} \underbrace{A}_{n \times n} \underbrace{\mathbf{x}}_{n \times 1} = \underbrace{A + A^T}_{n \times n} \quad (7)$$

$$\underbrace{J_{\mathbf{x}}}_{m \times n} \underbrace{A}_{m \times n} \underbrace{\mathbf{x}}_{n \times 1} = \underbrace{A}_{m \times n} \quad (8)$$

$$\underbrace{\nabla_X \mathbf{a}^T}_{m \times n} \underbrace{X}_{m \times n} \underbrace{\mathbf{b}}_{n \times 1} = \underbrace{\mathbf{a}}_{m \times 1} \underbrace{\mathbf{b}^T}_{1 \times n} \quad (9)$$

$$\underbrace{\nabla_X \mathbf{a}^T}_{m \times n} \underbrace{X^T}_{n \times m} \underbrace{\mathbf{b}}_{m \times 1} = \underbrace{\mathbf{b}}_{m \times 1} \underbrace{\mathbf{a}^T}_{1 \times n} \quad (10)$$

$$\underbrace{\nabla_X \mathbf{a}^T}_{m \times n} \underbrace{X}_{n \times m} \underbrace{C}_{m \times m} \underbrace{X}_{m \times n} \underbrace{\mathbf{b}}_{n \times 1} = \underbrace{C^T}_{m \times m} \underbrace{X}_{m \times n} \underbrace{\mathbf{a} \mathbf{b}^T}_{n \times n} + \underbrace{C}_{m \times m} \underbrace{X}_{m \times n} \underbrace{\mathbf{b} \mathbf{a}^T}_{n \times n} \quad (11)$$

$$\frac{\partial}{\partial x} |Y| = |Y| \operatorname{Tr} \left( Y^{-1} \frac{\partial Y}{\partial x} \right) \quad (12)$$

$\begin{matrix} & \uparrow \\ n \times n & \end{matrix}$ 
 $\begin{matrix} & \uparrow \\ n \times n & \end{matrix}$ 
 $\begin{matrix} & \uparrow \\ n \times n & \end{matrix}$

$$\frac{\partial}{\partial x} Y^{-1} = -Y^{-1} \frac{\partial Y}{\partial x} Y^{-1} \quad (13)$$

$\begin{matrix} & \uparrow \\ n \times n & \end{matrix}$ 
 $\begin{matrix} & \uparrow \\ n \times n & \end{matrix}$ 
 $\begin{matrix} & \uparrow \\ n \times n & \end{matrix}$ 
 $\begin{matrix} & \uparrow \\ n \times n & \end{matrix}$

### 3 Product and Chain Rule

$$\underbrace{J_{\mathbf{x}}}_{n \times 1} \underbrace{A}_{m \times k} \underbrace{\mathbf{y}}_{k \times 1} = \underbrace{A}_{m \times k} \underbrace{J_{\mathbf{x}} \mathbf{y}}_{k \times n} \quad (14)$$

$$\underbrace{J_{\mathbf{x}}}_{n \times 1} \underbrace{z}_{m \times 1} \underbrace{\mathbf{y}}_{m \times 1} = \underbrace{\mathbf{y}}_{m \times 1} \underbrace{(\nabla_{\mathbf{x}} z)^T}_{1 \times n} + z \underbrace{J_{\mathbf{x}} \mathbf{y}}_{m \times n} \quad (15)$$

$$\underbrace{\nabla_{\mathbf{x}}}_{1 \times m} \underbrace{\mathbf{y}^T}_{n \times 1} \underbrace{\mathbf{z}}_{m \times 1} = \underbrace{(J_{\mathbf{x}} \mathbf{y})^T}_{n \times m} \underbrace{\mathbf{z}}_{m \times 1} + \underbrace{(J_{\mathbf{x}} \mathbf{z})^T}_{n \times m} \underbrace{\mathbf{y}}_{m \times 1} \quad (16)$$

$$\underbrace{\frac{\partial}{\partial x} f(\mathbf{y})}_{1 \times 1} = \underbrace{(\nabla_{\mathbf{y}} f(\mathbf{y}))^T}_{1 \times k} \underbrace{\frac{\partial \mathbf{y}}{\partial x}}_{k \times 1} \quad (17)$$

$$\underbrace{\frac{\partial}{\partial x} \mathbf{f}(\mathbf{y})}_{m \times 1} = \underbrace{(J_{\mathbf{y}} \mathbf{f}(\mathbf{y}))}_{m \times k} \underbrace{\frac{\partial \mathbf{y}}{\partial x}}_{k \times 1} \quad (18)$$

$$\underbrace{\nabla_{\mathbf{x}}}_{n \times 1} f(\mathbf{y}) = \underbrace{(J_{\mathbf{x}} \mathbf{y})^T}_{n \times m} \underbrace{\nabla_{\mathbf{y}} f(\mathbf{y})}_{m \times 1} \quad (19)$$

$$\underbrace{J_{\mathbf{x}}}_{m \times n} \underbrace{\mathbf{f}(\mathbf{y})}_{k \times 1} = \underbrace{(J_{\mathbf{y}} \mathbf{f}(\mathbf{y}))}_{m \times k} \underbrace{J_{\mathbf{x}} \mathbf{y}}_{k \times n} \quad (20)$$