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	x	$x, \mathbf{x}, X$
dependent variable	y or 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

$$\underbrace{\nabla_{\substack{\mathbf{x}\\ \uparrow \\ n\times 1}} y}_{n\times 1} = \begin{bmatrix} \frac{\partial y}{\partial x_1} \\ \vdots \\ \frac{\partial y}{\partial x_n} \end{bmatrix} \tag{1}$$

$$\underbrace{J_{\substack{\mathbf{x}\\ \uparrow \\ n\times 1}} \overset{\substack{m\times 1\\ \downarrow}}{\mathbf{y}}}_{m\times n} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_m}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \frac{\partial \mathbf{y}}{\partial x_1} & \cdots & \frac{\partial \mathbf{y}}{\partial x_n} \end{bmatrix} = \begin{bmatrix} (\nabla_{\mathbf{x}} y_1)^T \\ \vdots \\ (\nabla_{\mathbf{x}} y_m)^T \end{bmatrix} \tag{2}$$

$$\underbrace{\nabla^2_{\substack{\mathbf{x}\\ \uparrow \\ n\times 1}} y}_{n\times n} = \begin{bmatrix} \frac{\partial^2 y}{\partial {x_1}^2} & \cdots & \frac{\partial^2 y}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 y}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 y}{\partial {x_n}^2} \end{bmatrix} = \begin{bmatrix} \nabla_{\mathbf{x}} \frac{\partial \mathbf{y}}{\partial x_1} & \cdots & \nabla_{\mathbf{x}} \frac{\partial \mathbf{y}}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \left(\frac{\partial}{\partial x_1} \nabla_{\mathbf{x}} y\right)^T \\ \vdots \\ \left(\frac{\partial}{\partial x_n} \nabla_{\mathbf{x}} y\right)^T \end{bmatrix} \tag{3}$$

$$\underbrace{\nabla_{\substack{X\\ \uparrow \\ m\times n}} y}_{m\times n} = \begin{bmatrix} \frac{\partial y}{\partial x_{11}} & \cdots & \frac{\partial y}{\partial x_{1n}} \\ \vdots & \ddots & \vdots \\ \frac{\partial y}{\partial x_{m1}} & \cdots & \frac{\partial^2 y}{\partial {x_{mn}}^2} \end{bmatrix} \tag{4}$$

2 Matrix-Vector Derivatives

$$\underbrace{\nabla_{\mathbf{x}} \overset{\overset{1\times n}{\downarrow}}{\mathbf{a}^T} \underset{\underset{n\times 1}{\uparrow}}{\mathbf{x}}}_{n\times 1} = \underset{\underset{n\times 1}{\uparrow}}{\mathbf{a}} \tag{5}$$

$$\underbrace{\nabla_{\mathbf{x}} \underset{\underset{1\times n}{\uparrow}}{\mathbf{x}^T} \overset{\overset{n\times n}{\downarrow}}{A} \underset{\underset{n\times 1}{\uparrow}}{\mathbf{x}}}_{n\times 1} = \underbrace{(A + A^T)}_{n\times n} \underset{\underset{n\times 1}{\uparrow}}{\mathbf{x}} \tag{6}$$

$$\underbrace{\nabla^2_{\mathbf{x}} \underset{\underset{1\times n}{\uparrow}}{\mathbf{x}^T} \overset{\overset{n\times n}{\downarrow}}{A} \underset{\underset{n\times 1}{\uparrow}}{\mathbf{x}}}_{n\times 1} = \underbrace{A + A^T}_{n\times n} \tag{7}$$

$$\underbrace{J_{\mathbf{x}} \overset{\overset{m\times n}{\downarrow}}{A} \underset{\underset{n\times 1}{\uparrow}}{\mathbf{x}}}_{m\times n} = \underset{\underset{m\times n}{\uparrow}}{A} \tag{8}$$

$$\underbrace{\nabla_X \underset{\underset{1\times m}{\uparrow}}{\mathbf{a}^T} \overset{\overset{m\times n}{\downarrow}}{X} \underset{\underset{n\times 1}{\uparrow}}{\mathbf{b}}}_{m\times n} = \underset{\underset{m\times 1}{\uparrow}}{\mathbf{a}} \overset{\overset{1\times n}{\downarrow}}{\mathbf{b}^T} \tag{9}$$

$$\underbrace{\nabla_X \underset{\underset{1\times n}{\uparrow}}{\mathbf{a}^T} \overset{\overset{n\times m}{\downarrow}}{X^T} \underset{\underset{m\times 1}{\uparrow}}{\mathbf{b}}}_{m\times n} = \underset{\underset{m\times 1}{\uparrow}}{\mathbf{b}} \overset{\overset{1\times n}{\downarrow}}{\mathbf{a}^T} \tag{10}$$

$$\underbrace{\nabla_X \underset{\underset{1\times n}{\uparrow}}{\mathbf{a}^T} \overset{\overset{n\times m}{\downarrow}}{X} \underset{\underset{m\times m}{\uparrow}}{C} \overset{\overset{m\times n}{\downarrow}}{X} \underset{\underset{n\times 1}{\uparrow}}{\mathbf{b}}}_{m\times n} = \underset{\underset{m\times m}{\uparrow}}{C^T} \overset{\overset{m\times n}{\downarrow}}{X} \underbrace{\mathbf{a}\mathbf{b}^T}_{n\times n} + \underset{\underset{m\times m}{\uparrow}}{C} \overset{\overset{m\times n}{\downarrow}}{X} \underbrace{\mathbf{b}\mathbf{a}^T}_{n\times n} \tag{11}$$

$$\frac{\partial}{\partial x}|\underset{\substack{\uparrow\\ n\times n}}{Y}| = |Y|\,\mathrm{Tr}\left(\underset{\substack{\uparrow\\ n\times n}}{Y^{-1}}\underset{\substack{\uparrow\\ n\times n}}{\frac{\partial Y}{\partial x}}\right) \tag{12}$$

$$\frac{\partial}{\partial x}\underset{\substack{\uparrow\\ n\times n}}{Y^{-1}} = -\underset{\substack{\uparrow\\ n\times n}}{Y^{-1}}\underset{\substack{\uparrow\\ n\times n}}{\frac{\partial Y}{\partial x}}\underset{\substack{\uparrow\\ n\times n}}{Y^{-1}} \tag{13}$$

3 Product and Chain Rule

$$\underbrace{J_{\underset{\substack{\uparrow\\ n\times 1}}{\mathbf{x}}}\ \overset{\substack{m\times k\\ \downarrow}}{A}\ \underset{\substack{\uparrow\\ k\times 1}}{\mathbf{y}}}_{m\times n} = \underset{\substack{\uparrow\\ m\times k}}{A}\ \underbrace{J_{\mathbf{x}}\mathbf{y}}_{k\times n} \tag{14}$$

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

$$\underbrace{\nabla_{\underset{\substack{\uparrow\\ n\times 1}}{\mathbf{x}}}\ \overset{\substack{1\times m\\ \downarrow}}{\mathbf{y}^T}\ \underset{\substack{\uparrow\\ m\times 1}}{\mathbf{z}}}_{n\times 1} = \underbrace{(J_{\mathbf{x}}\mathbf{y})^T}_{n\times m}\ \underset{\substack{\uparrow\\ m\times 1}}{\mathbf{z}} + \underbrace{(J_{\mathbf{x}}\mathbf{z})^T}_{n\times m}\ \underset{\substack{\uparrow\\ m\times 1}}{\mathbf{y}} \tag{16}$$

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

$$\underbrace{\frac{\partial}{\partial x}\ \overset{\substack{m\times 1\\ \downarrow}}{\mathbf{f}}\ (\underset{\substack{\uparrow\\ k\times 1}}{\mathbf{y}})}_{m\times 1} = \underbrace{(J_{\mathbf{y}}\mathbf{f}(\mathbf{y}))}_{m\times k}\ \underset{\substack{\uparrow\\ k\times 1}}{\frac{\partial \mathbf{y}}{\partial x}} \tag{18}$$

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

$$\underbrace{J_{\underset{\underset{n\times 1}{\uparrow}}{\mathbf{x}}} \overset{\overset{m\times 1}{\downarrow}}{\mathbf{f}} (\underset{\underset{k\times 1}{\uparrow}}{\mathbf{y}})}_{m\times n} = \underbrace{(J_{\mathbf{y}}\mathbf{f}(\mathbf{y}))}_{m\times k}\underbrace{J_{\mathbf{x}}\mathbf{y}}_{k\times n} \tag{20}$$