1 Definitions

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

$$\underbrace{J_{\underset{n\times 1}{\underset{\uparrow}{\mathbf{x}}}} \overset{m\times 1}{\overset{\downarrow}{\mathbf{f}}}}_{m\times n} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \frac{\partial \mathbf{f}}{\partial x_1} & \cdots & \frac{\partial \mathbf{f}}{\partial x_n} \end{bmatrix} = \begin{bmatrix} (\nabla_{\mathbf{x}} f_1)^T \\ \vdots \\ (\nabla_{\mathbf{x}} f_m)^T \end{bmatrix} \tag{2}$$

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

2 Matrix-Vector Derivatives

$$\underbrace{\nabla_{\mathbf{x}} \overset{1\times n}{\overset{\downarrow}{\mathbf{a}^T}} \underset{n\times 1}{\underset{\uparrow}{\mathbf{x}}}}_{n\times 1} = \underset{n\times 1}{\underset{\uparrow}{\mathbf{a}}} \tag{4}$$

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

$$\underbrace{\nabla^2_{\mathbf{x}} \underset{\substack{\uparrow\\1\times n}}{\mathbf{x}^T} \; \overset{\substack{n\times n\\\downarrow}}{A} \; \underset{\substack{\uparrow\\n\times 1}}{\mathbf{x}}}_{n\times 1} = \underbrace{A + A^T}_{n\times n} \tag{6}$$

$$\underbrace{J_{\mathbf{x}} \; \overset{\substack{m\times n\\\downarrow}}{A} \; \underset{\substack{\uparrow\\n\times 1}}{\mathbf{x}}}_{m\times n} = \underset{\substack{\uparrow\\m\times n}}{A} \tag{7}$$

$$\underbrace{\nabla_X \, \underset{\substack{\uparrow\\1\times m}}{\mathbf{a}^T} \; \overset{\substack{m\times n\\\downarrow}}{X} \; \underset{\substack{\uparrow\\n\times 1}}{\mathbf{b}}}_{m\times n} = \underset{\substack{\uparrow\\m\times 1}}{\mathbf{a}} \; \overset{\substack{1\times n\\\downarrow}}{\mathbf{b}^T} \tag{8}$$

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

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

$$\frac{\partial}{\partial x} \underset{\substack{\uparrow\\n\times n}}{|Y|} = |Y| \operatorname{Tr}\left(\underset{\substack{\uparrow\\n\times n}}{Y^{-1}} \underset{\substack{\uparrow\\n\times n}}{\frac{\partial Y}{\partial x}} \right) \tag{11}$$

$$\frac{\partial}{\partial x} \underset{\substack{\uparrow\\n\times n}}{Y^{-1}} = - \underset{\substack{\uparrow\\n\times n}}{Y^{-1}} \underset{\substack{\uparrow\\n\times n}}{\frac{\partial Y}{\partial x}} \underset{\substack{\uparrow\\n\times n}}{Y^{-1}} \tag{12}$$

3 Product and Chain Rule

$$\underbrace{J_{\underset{\underset{n\times 1}{\uparrow}}{\mathbf{x}}} \overset{\overset{m\times k}{\downarrow}}{C} \underset{\underset{k\times 1}{\uparrow}}{\mathbf{y}}}_{m\times n} = \underset{\underset{m\times k}{\uparrow}}{C} \underbrace{J_{\mathbf{x}}\mathbf{y}}_{k\times n} \quad C \text{ is a constant} \tag{13}$$

$$\underbrace{J_{\underset{\underset{n\times 1}{\uparrow}}{\mathbf{x}}} z \underset{\underset{m\times 1}{\uparrow}}{\mathbf{y}}}_{m\times n} = \underset{\underset{m\times 1}{\uparrow}}{\mathbf{y}} \underbrace{(\nabla_{\mathbf{x}} z)^T}_{1\times n} + z \underbrace{J_{\mathbf{x}}\mathbf{y}}_{m\times n} \tag{14}$$

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

$$\underbrace{\frac{\partial}{\partial x} f(\underset{\underset{k\times 1}{\uparrow}}{\mathbf{y}})}_{1\times 1} = \underbrace{(\nabla_{\mathbf{y}} f(\mathbf{y}))^T}_{1\times k} \underset{\underset{k\times 1}{\uparrow}}{\frac{\partial \mathbf{y}}{\partial x}} \tag{16}$$

$$\underbrace{\frac{\partial}{\partial x} \overset{\overset{m\times 1}{\downarrow}}{\mathbf{f}} (\underset{\underset{k\times 1}{\uparrow}}{\mathbf{y}})}_{m\times 1} = \underbrace{(J_{\mathbf{y}}\mathbf{f}(\mathbf{y}))}_{m\times k} \underset{\underset{k\times 1}{\uparrow}}{\frac{\partial \mathbf{y}}{\partial x}} \tag{17}$$

$$\underbrace{\nabla_{\underset{\underset{n\times 1}{\uparrow}}{\mathbf{x}}} f(\underset{\underset{m\times 1}{\uparrow}}{\mathbf{y}})}_{n\times 1} = \underbrace{(J_{\mathbf{x}}\mathbf{y})^T}_{n\times m} \underbrace{\nabla_{\mathbf{y}} f(\mathbf{y})}_{m\times 1} \tag{18}$$

$$\underbrace{J_{\underset{\underset{n\times 1}{\uparrow}}{\mathbf{x}}} \overset{\overset{m\times 1}{\downarrow}}{\mathbf{f}} (\underset{\underset{k\times 1}{\uparrow}}{\mathbf{y}})}_{m\times n} = \underbrace{(J_{\mathbf{y}}\mathbf{f}(\mathbf{y}))}_{m\times k} \underbrace{J_{\mathbf{x}}\mathbf{y}}_{k\times n} \tag{19}$$