**5.7** Compute the derivatives $df/dx$ of the following functions by using the chain rule. Provide the dimensions of every single partial derivative. Describe your steps in detail.

a.

$$f(z) = log(1+z), \qquad z = \mathbf{x}^T\mathbf{x}, \qquad \mathbf{x} \in \mathbb{R}^D$$

**Solution.**

$$\frac{d(f(z))}{dx} = \frac{d(log(1+z))}{dx} = \frac{d(log(1+z))}{d(1+z)} * \frac{d(1+z)}{dz} * \frac{dz}{dx}$$

$$\implies \frac{d(f(z))}{dx} = \left(\frac{1}{(1+z)} * 1 * \left(\frac{d\mathbf{x}^T\mathbf{x}}{d\mathbf{x}}\right)\right)$$

$$\implies \frac{d(f(z))}{dx} = \left(\frac{1}{(1+\mathbf{x}^T\mathbf{x})} * (2\mathbf{x}^T)\right) = \frac{2\mathbf{x}^T}{(1+\mathbf{x}^T\mathbf{x})}$$

b.

$$f(\mathbf{z}) = sin(\mathbf{z}), \qquad \mathbf{z} = \mathbf{A}\mathbf{x} + \mathbf{b}, \qquad \mathbf{A} \in \mathbb{R}^{E\times D}, \mathbf{x} \in \mathbb{R}^D, \mathbf{b} \in \mathbb{R}^E$$

where $sin(\cdot)$ is applied to every element of $\mathbf{z}$.

**Solution.**

$$\frac{d(sin(\mathbf{z}))}{d\mathbf{x}} = \frac{d(sin(\mathbf{z}))}{d\mathbf{z}} * \frac{d\mathbf{z}}{d\mathbf{x}}$$

$$\implies \frac{d(sin(\mathbf{z}))}{d\mathbf{x}} = \left(\frac{d(sin(\mathbf{z}))}{d\mathbf{z}}\right) * \frac{d(\mathbf{A}\mathbf{x}+\mathbf{b})}{d\mathbf{x}} = \left(\frac{d(sin(\mathbf{z}))}{d\mathbf{z}}\right) * \left(\frac{d(\mathbf{A}\mathbf{x})}{d\mathbf{x}} + \frac{d\mathbf{b}}{d\mathbf{x}}\right)$$

$$= \left(\frac{d(sin(\mathbf{z}))}{d\mathbf{z}}\right) * \left(\frac{d(\mathbf{A}\mathbf{x})}{d\mathbf{x}} + 0\right) = \left(\frac{d(sin(\mathbf{z}))}{d\mathbf{z}}\right) * \left(\frac{d(\mathbf{A}\mathbf{x})}{d\mathbf{x}}\right)$$

$$\text{Now, } \mathbf{Ax} = \begin{bmatrix} \text{———} & \mathbf{a}_1 & \text{———} \\ \text{———} & \mathbf{a}_2 & \text{———} \\ \text{———} & \vdots & \text{———} \\ \text{———} & \mathbf{a}_E & \text{———} \end{bmatrix} \begin{bmatrix} | \\ \mathbf{x} \\ | \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1\mathbf{x} \\ \mathbf{a}_2\mathbf{x} \\ \vdots \\ \mathbf{a}_E\mathbf{x} \end{bmatrix}$$

Applying 5.105, we get,

$$\frac{d(\mathbf{Ax})}{d\mathbf{x}} = \begin{bmatrix} \text{———} & \mathbf{a}_1 & \text{———} \\ \text{———} & \mathbf{a}_2 & \text{———} \\ \text{———} & \vdots & \text{———} \\ \text{———} & \mathbf{a}_E & \text{———} \end{bmatrix}$$

$$\text{Now, } \sin(\mathbf{z}) = \begin{bmatrix} sin(z_1) \\ sin(z_2) \\ \vdots \\ sin(z_E) \end{bmatrix}$$

$$\frac{d(sin(\mathbf{z}))}{d\mathbf{z}} = \begin{bmatrix} \frac{sin(z_1)}{dz_1} & \frac{sin(z_1)}{dz_2} & \dots & \frac{sin(z_1)}{dz_E} \\ \frac{sin(z_2)}{dz_1} & \frac{sin(z_2)}{dz_2} & \dots & \frac{sin(z_2)}{dz_E} \\ \vdots & & & \\ \frac{sin(z_E)}{dz_1} & \frac{sin(z_E)}{dz_2} & \dots & \frac{sin(z_E)}{dz_E} \end{bmatrix} = \begin{bmatrix} cos(z_1) & 0 & \dots & 0 \\ 0 & cos(z_2) & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & cos(z_E) \end{bmatrix}$$

$$
= \begin{bmatrix} cos(\mathbf{a}_1\mathbf{x} + b_1) & 0 & \dots & 0 \\ 0 & cos(\mathbf{a}_2\mathbf{x} + b_2) & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & cos(\mathbf{a}_E\mathbf{x} + b_E) \end{bmatrix}
$$

$$
\implies \left(\frac{d(sin(\mathbf{z}))}{d\mathbf{z}}\right) * \left(\frac{d(\mathbf{A}\mathbf{x})}{d\mathbf{x}}\right) = \begin{bmatrix} cos(\mathbf{a}_1\mathbf{x} + b_1) & 0 & \dots & 0 \\ 0 & cos(\mathbf{a}_2\mathbf{x} + b_2) & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & cos(\mathbf{a}_E\mathbf{x} + b_E) \end{bmatrix} \begin{bmatrix} \rule{2em}{0.4pt} & \mathbf{a}_1 & \rule{2em}{0.4pt} \\ \rule{2em}{0.4pt} & \mathbf{a}_2 & \rule{2em}{0.4pt} \\ \rule{2em}{0.4pt} & \vdots & \rule{2em}{0.4pt} \\ \rule{2em}{0.4pt} & \mathbf{a}_E & \rule{2em}{0.4pt} \end{bmatrix}
$$

$$
\implies \left(\frac{d(sin(\mathbf{z}))}{d\mathbf{z}}\right) * \left(\frac{d(\mathbf{A}\mathbf{x})}{d\mathbf{x}}\right) = \begin{bmatrix} \rule{2em}{0.4pt} & cos(\mathbf{a}_1\mathbf{x} + b_1)\mathbf{a}_1 & \rule{2em}{0.4pt} \\ \rule{2em}{0.4pt} & cos(\mathbf{a}_2\mathbf{x} + b_2)\mathbf{a}_2 & \rule{2em}{0.4pt} \\ \rule{2em}{0.4pt} & \vdots & \rule{2em}{0.4pt} \\ \rule{2em}{0.4pt} & cos(\mathbf{a}_E\mathbf{x} + b_E)\mathbf{a}_E & \rule{2em}{0.4pt} \end{bmatrix}
$$