Exercise 5.1 :

$$f(x) \; = \; \log\left(x^4\right)\sin\left(x^3\right) \;\forall\, x \in \mathbb{R}$$

$$f'(x) \; = \; \frac{4x^3}{x^4}\sin\left(x^3\right) + 3x^2\log\left(x^4\right)\cos\left(x^3\right) \; = \; \frac{4}{x}\sin\left(x^3\right) + 3x^2\log\left(x^4\right)\cos\left(x^3\right)$$

$$= \; \frac{4}{x}\sin\left(x^3\right) + 12x^2\log(x)\cos\left(x^3\right)$$

Exercise 5.2 :

$$f(x) \; = \; \frac{1}{1+e^{-x}} \;\; \forall\, x \in \mathbb{R}$$

$$f'(x) \; = \; \frac{e^x}{(1+e^x)^2}$$

Exercise 5.3 :

$$f(x) \; = \; \exp\left(\frac{1}{2\sigma^2}(x-\mu)^2\right) \;\forall\, x \in \mathbb{R}$$

$$f'(x) \; = \; \frac{1}{\sigma^2}(x-\mu)\exp\left(\frac{1}{2\sigma^2}(x-\mu)^2\right) \; = \; \frac{1}{\sigma^2}(x-\mu)f(x)$$

Exercise 5.5 :

a) $f_1(x) \; = \; \sin(x_1)\cos(x_2) \;\forall\, x \in \mathbb{R}^2$

$$\frac{\partial f_1}{\partial x} \; = \; \begin{cases} \dfrac{\partial f_1}{\partial x_1} = \cos x_1 \cos x_2 \\[2ex] \dfrac{\partial f_1}{\partial x_2} = -\sin(x_1)\sin(x_2) \end{cases}$$

Therefore, $\frac{\partial f_1}{\partial x}$ has 2 dimensions and $f_1'(x) \; = \; \cos(x_1 + x_2)$

$f_2(x,y) \; = \; x^\top y \;\forall\, (x,y) \in (\mathbb{R}^n)^2$

$$y \; = \; \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ x_n \end{pmatrix}, x \; = \; \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, x^\top \; = \; \begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix}$$

Now, $x^\top y = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \sum_{i=1}^{n} x_i y_i$

$$\frac{\partial f_2}{\partial x} = \begin{cases} \frac{\partial x^\top y}{\partial x_1} = y_1 \\ \frac{\partial x^\top y}{\partial x_2} = y_2 \\ \vdots \\ \frac{\partial x^\top y}{\partial x_n} = y_n \end{cases}$$

Therefore, $\frac{\partial f_2}{\partial x}$ has n dimensions and $f_2'(x) = \sum_{i=1}^{n} y_i$

$$f_3(x) = xx^\top \ \forall\, x \in \mathbb{R}^n, x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, x^\top = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix}$$

$$xx^\top = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix} = \begin{pmatrix} x_1^2 & x_2 x_1 & \cdots & x_n x_1 \\ x_1 x_2 & x_2^2 & \cdots & x_n x_2 \\ \vdots & \vdots & & \vdots \\ x_1 x_n & x_2 x_n & \cdots & x_n^2 \end{pmatrix}$$

$$\frac{\partial f_3}{\partial x} = \begin{cases} \frac{\partial f_3}{\partial x_1} = \begin{pmatrix} \frac{\partial x_1^2}{\partial x_1} & \frac{\partial x_2 x_1}{\partial x_1} & \cdots & \frac{\partial x_n x_1}{\partial x_1} \\ \frac{\partial x_1 x_2}{\partial x_1} & \frac{\partial x_2^2}{\partial x_1} & \cdots & \frac{\partial x_n x_2}{\partial x_1} \\ \vdots & \vdots & & \vdots \\ \frac{\partial x_1 x_n}{\partial x_1} & \frac{\partial x_2 x_n}{\partial x_1} & \cdots & \frac{\partial x_n^2}{\partial x_1} \end{pmatrix} = \begin{pmatrix} 2x_1 & x_2 & \cdots & x_n \\ x_2 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ x_n & 0 & \cdots & 0 \end{pmatrix} \\ \frac{\partial f_3}{\partial x_2} = \begin{pmatrix} \frac{\partial x_1^2}{\partial x_2} & \frac{\partial x_2 x_1}{\partial x_2} & \cdots & \frac{\partial x_n x_1}{\partial x_2} \\ \frac{\partial x_1 x_2}{\partial x_2} & \frac{\partial x_2^2}{\partial x_2} & \cdots & \frac{\partial x_n x_2}{\partial x_2} \\ \vdots & \vdots & & \vdots \\ \frac{\partial x_n x_n}{\partial x_2} & \frac{\partial x_2 x_n}{\partial x_2} & \cdots & \frac{\partial x_n^2}{\partial x_2} \end{pmatrix} = \begin{pmatrix} 0 & x_1 & 0 & \cdots & 0 \\ x_1 & 2x_2 & x_3 & \cdots & x_n \\ 0 & x_3 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & x_n & 0 & 0 & 0 \end{pmatrix} \\ \vdots \\ \frac{\partial f_3}{\partial x_n} = \begin{pmatrix} \frac{\partial x_1^2}{\partial x_n} & \frac{\partial x_2 x_1}{\partial x_n} & \cdots & \frac{\partial x_n x_1}{\partial x_n} \\ \frac{\partial x_1 x_2}{\partial x_n} & \frac{\partial x_2^2}{\partial x_n} & \cdots & \frac{\partial x_n x_2}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial x_1 x_n}{\partial x_n} & \frac{\partial x_2 x_n}{\partial x_n} & \cdots & \frac{\partial x_n^2}{\partial x_n} \end{pmatrix} = \begin{pmatrix} 0 & 0 & \cdots & x_1 \\ 0 & 0 & \cdots & x_2 \\ \vdots & \vdots & & \vdots \\ x_1 & x_2 & \cdots & 2x_n \end{pmatrix} \end{cases}$$

Therefore, $\frac{df_3}{dx}$ has n dimensions and $f_3'(x) = \sum_{i=1}^{n} \frac{\partial f_3(x)}{\partial x_i}$

b) $J_{f_1}(x) = [\cos(x_1)\cos(x_2) \quad -\sin(x_1)\sin(x_2)] \; \forall\, x \in \mathbb{R}^{1\times 2}$

$$J_{f_{2(x,y)}} = \begin{bmatrix} \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{bmatrix} = \begin{bmatrix} y^\top & x^\top \end{bmatrix} \in \mathbb{R}^{1\times 2n}$$

$$J_{f_3}(x) = \begin{cases} \frac{\partial f_3(x)}{\partial x_1} \\ \frac{\partial f_3(x)}{\partial x_2} \\ \vdots \\ \frac{\partial f_3(x)}{\partial x_n} \end{cases} \forall\, x \in \mathbb{R}^{(n\times n)\times n}$$

Exercise 5.6 :

$f(t) = \sin\left(\log\left(t^\top t\right)\right) \; \forall\, t \in \mathbb{R}^D$

Taking $\sin(u)$, where $u = \log(t^T t)$:

$\frac{d}{du}\sin u = \cos u$

$\frac{d}{dt}(\log(t^T t)) = \frac{d}{dt}\left(\log\left(\sum_i t_i^2\right)\right) = \frac{1}{t^T t} \cdot 2t = \frac{2t}{t^T t}$

Therefore, $\frac{df}{dt} = \cos(\log(t^T t)) \cdot \frac{2t}{t^T t}$

$g(X) = \mathrm{tr}(AXB)$, $A \in \mathbb{R}^{D\times E}, X \in \mathbb{R}^{E\times F}, B \in \mathbb{R}^{F\times D}$

We know that, $\mathrm{tr}(T) = \sum_{i=1}^{D} T_{ii}$

therefore, $(AXB)_{pq} = \sum_{k=1}^{D}\sum_{i=1}^{E}\sum_{j=1}^{F} A_{pi}X_{ij}B_{jq}$

and $\mathrm{tr}(AXB) = \sum_{k=1}^{D} (AXB)_{kk} = \sum_{k=1}^{D}\left(\sum_{i=1}^{E}\sum_{j=1}^{F} A_{ki}X_{ij}B_{jk}\right)$

=> $\frac{\partial}{\partial X_{ij}}\mathrm{tr}(AXB) = \sum_k A_{ki}B_{jk} = (BA)_{ji}$

=> $\frac{\partial}{\partial X}\mathrm{tr}(AXB) = \underset{E\times D}{A^\top}\,\underset{D\times F}{B^\top}$

Exercise 5.7 :

a) We have $f(z) = \log(1+z)$ where $z = x^\top x$ and $x \in \mathbb{R}^D$

$$\frac{\mathrm{d}f}{\mathrm{d}x} = \frac{\partial f}{\partial z}\frac{\partial z}{\partial x}$$

$$\frac{\partial f}{\partial z} = \frac{1}{1+z} = \frac{1}{1+x^\top x}$$

$$\frac{\partial z}{\partial x} = 2x^\top$$

Plugging this in:

$$\frac{\mathrm{d}f}{\mathrm{d}x} = \frac{2x^\top}{1+x^\top x}$$

b) We have $z = Ax + b$ where $A \in \mathbb{R}^{E\times D},\ x \in \mathbb{R}^D,\ b \in \mathbb{R}^E$ and $f(z) = \sin(z)$

$$\frac{\mathrm{d}f}{\mathrm{d}x} = \frac{\partial f}{\partial z}\frac{\partial z}{\partial x}$$
$$\frac{\partial f}{\partial z} = \mathrm{diag}(\cos(z))$$
$$\frac{\partial z}{\partial x} = A$$

Plugging this in:

$$\frac{\mathrm{d}f}{\mathrm{d}x} = \mathrm{diag}(\cos(Ax+b))A$$

Exercise 5.9 :

Given $g(x,z,\nu) = \log(p(x,z)) - \log(g(z,\nu))$, with $z = t(\epsilon,\nu)$

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\nu} g(x,z,\nu) &= \frac{\mathrm{d}}{\mathrm{d}\nu}\left(\log p(x,z) - \log q(z,\nu)\right) = \frac{\mathrm{d}}{\mathrm{d}\nu}\log p(x,z) - \frac{\mathrm{d}}{\mathrm{d}\nu}\log q(z,\nu) \\
&= \frac{\partial}{\partial z}\log p(x,z)\frac{\partial t(\epsilon,\nu)}{\partial \nu} - \frac{\partial}{\partial z}\log q(z,\nu)\frac{\partial t(\epsilon,\nu)}{\partial \nu} - \frac{\partial}{\partial \nu}\log q(z,\nu) \\
&= \left(\frac{\partial}{\partial z}\log p(x,z) - \frac{\partial}{\partial z}\log q(z,\nu)\right)\frac{\partial t(\epsilon,\nu)}{\partial \nu} - \frac{\partial}{\partial \nu}\log q(z,\nu) \\
&= \left(\frac{1}{p(x,z)}\frac{\partial}{\partial z}p(x,z) - \frac{1}{q(z,\nu)}\frac{\partial}{\partial z}q(z,\nu)\right)\frac{\partial t(\epsilon,\nu)}{\partial \nu} - \frac{\partial}{\partial \nu}\log q(z,\nu)
\end{aligned}$$