Matrix Algebra Marathon

J2200071 Ryuto Saito

May 5, 2024

1 Derivatives of Function of Vectors

1.1 Exercise

Exercise 2.12.2.

If we define a function $f\colon \mathbb{R}^2 \to \mathbb{R}$ as $f(\boldsymbol{x}) = 3x_1 - x_2$, show that the derivative is given by

$$\nabla_{\boldsymbol{x}} f(\boldsymbol{x}) = [3, -1]^{\mathrm{T}}. \tag{1}$$

Proof. Let $\boldsymbol{x} \in \mathbb{R}^2$.

$$\nabla_{\boldsymbol{x}} f(\boldsymbol{x}) = \begin{bmatrix} \nabla_{x_1} f(\boldsymbol{x}) \\ \nabla_{x_2} f(\boldsymbol{x}) \end{bmatrix} = \begin{bmatrix} \nabla_{x_1}(3x_1 - x_2) \\ \nabla_{x_2}(3x_1 - x_2) \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \end{bmatrix} = [3, -1]^{\mathrm{T}}$$

□

Exercise 2.12.4.

Let a function $f\colon \mathbb{R}^n \to \mathbb{R}$ be defined as

$$f(\boldsymbol{x}) \coloneqq \sum_{i=1}^{n} \sin(x_i). \tag{2}$$

Show that the derivative is given by

$$\nabla_{\boldsymbol{x}} f(\boldsymbol{x}) = [\cos(x_1), \ldots, \cos(x_n)]^{\mathrm{T}}. \tag{3}$$

Proof. Let $\boldsymbol{x} \in \mathbb{R}^n$.

$$
\begin{aligned}
\nabla_{\boldsymbol{x}} f(\boldsymbol{x}) &= \begin{bmatrix} \nabla_{x_1} f(\boldsymbol{x}) \\ \vdots \\ \nabla_{x_n} f(\boldsymbol{x}) \end{bmatrix} \\
&= \begin{bmatrix} \nabla_{x_1} (\sum_{i=1}^{n} \sin(x_i)) \\ \vdots \\ \nabla_{x_n} (\sum_{i=1}^{n} \sin(x_i)) \end{bmatrix} \\
&= \begin{bmatrix} \nabla_{x_1} \sin(x_1) \\ \vdots \\ \nabla_{x_n} \sin(x_n) \end{bmatrix} \\
&= \begin{bmatrix} \cos(x_1) \\ \vdots \\ \cos(x_n) \end{bmatrix} \\
&= [\cos(x_1), \ldots, \cos(x_n)]^{\mathrm{T}}
\end{aligned}
$$

□

Exercise 2.12.7.

Generate a sample and verify $\nabla_{\boldsymbol{x}} \langle \boldsymbol{a}, \boldsymbol{x} \rangle = \boldsymbol{a}$ in Exercise 2.12.6.

Proof. Let $\boldsymbol{a} = [1, 2]^{\mathrm{T}}$ and $\boldsymbol{x} \in \mathbb{R}^2$.

$$
\begin{aligned}
\nabla_{\boldsymbol{x}} \langle \boldsymbol{a}, \boldsymbol{x} \rangle &= \begin{bmatrix} \nabla_{x_1} \langle \boldsymbol{a}, \boldsymbol{x} \rangle \\ \nabla_{x_2} \langle \boldsymbol{a}, \boldsymbol{x} \rangle \end{bmatrix} \\
&= \begin{bmatrix} \nabla_{x_1} (x_1 + 2x_2) \\ \nabla_{x_2} (x_1 + 2x_2) \end{bmatrix} \\
&= \begin{bmatrix} 1 \\ 2 \end{bmatrix} \\
&= [1, 2]^{\mathrm{T}} \\
&= \boldsymbol{a}
\end{aligned}
$$

We obtain that $\nabla_{\boldsymbol{x}} \langle \boldsymbol{a}, \boldsymbol{x} \rangle = \boldsymbol{a}$ for this sample. □

Exercise 2.12.9.

Let $\boldsymbol{x}, \boldsymbol{b} \in \mathbb{R}^n$. Show that

$$
\nabla_{\boldsymbol{x}} \|\boldsymbol{x} + \boldsymbol{b}\|^2 = 2(\boldsymbol{x} + \boldsymbol{b}).
$$

Proof. Let $\boldsymbol{x}, \boldsymbol{b} \in \mathbb{R}^n$.

$$
\begin{aligned}
\nabla_{\boldsymbol{x}} \|\boldsymbol{x} + \boldsymbol{b}\|^2 &= \begin{bmatrix} \nabla_{x_1} \|\boldsymbol{x} + \boldsymbol{b}\|^2 \\ \vdots \\ \nabla_{x_n} \|\boldsymbol{x} + \boldsymbol{b}\|^2 \end{bmatrix} \\
&= \begin{bmatrix} \nabla_{x_1} \left(\sum_{i=1}^n (x_i + b_i)^2 \right) \\ \vdots \\ \nabla_{x_n} \left(\sum_{i=1}^n (x_i + b_i)^2 \right) \end{bmatrix} \\
&= \begin{bmatrix} \nabla_{x_1} (x_1 + b_1)^2 \\ \vdots \\ \nabla_{x_n} (x_n + b_n)^2 \end{bmatrix} \\
&= \begin{bmatrix} \nabla_{x_1} (x_1^2 + b_1^2 + 2x_1 b_1) \\ \vdots \\ \nabla_{x_n} (x_n^2 + b_n^2 + 2x_n b_n) \end{bmatrix} \\
&= \begin{bmatrix} 2(x_1 + b_1) \\ \vdots \\ 2(x_n + b_n) \end{bmatrix} \\
&= 2 \left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} \right) \\
&= 2(\boldsymbol{x} + \boldsymbol{b})
\end{aligned}
$$

□

Exercise 2.12.11.

Let $\boldsymbol{x}, \boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$. Show that

$$\nabla_{\boldsymbol{x}} \langle \boldsymbol{a}, \boldsymbol{x} + \boldsymbol{b} \rangle = \boldsymbol{a}.$$

Proof. Let $\boldsymbol{x}, \boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$.

$$\begin{aligned}
\nabla_{\boldsymbol{x}} \langle \boldsymbol{a}, \boldsymbol{x} + \boldsymbol{b} \rangle &= \begin{bmatrix} \nabla_{x_1} \langle \boldsymbol{a}, \boldsymbol{x} + \boldsymbol{b} \rangle \\ \vdots \\ \nabla_{x_n} \langle \boldsymbol{a}, \boldsymbol{x} + \boldsymbol{b} \rangle \end{bmatrix} \\
&= \begin{bmatrix} \nabla_{x_1} \sum_{i=1}^n a_i(x_i + b_i) \\ \vdots \\ \nabla_{x_n} \sum_{i=1}^n a_i(x_i + b_i) \end{bmatrix} \\
&= \begin{bmatrix} \nabla_{x_1} \sum_{i=1}^n a_i(x_i + b_i) \\ \vdots \\ \nabla_{x_n} \sum_{i=1}^n a_i(x_i + b_i) \end{bmatrix} \\
&= \begin{bmatrix} \nabla_{x_1} a_1(x_1 + b_1) \\ \vdots \\ \nabla_{x_n} a_n(x_n + b_n) \end{bmatrix} \\
&= \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \\
&= \boldsymbol{a}
\end{aligned}$$

□

Exercise 2.12.13.

Let $\boldsymbol{x}, \boldsymbol{a} \in \mathbb{R}^n$. Show that

$$\nabla_{\boldsymbol{x}} \langle \boldsymbol{a}, \boldsymbol{x} \rangle^2 = 2 \langle \boldsymbol{a}, \boldsymbol{x} \rangle \boldsymbol{a}.$$

Proof. Let $\boldsymbol{x}, \boldsymbol{a} \in \mathbb{R}^n$.

$$\begin{aligned}
\nabla_{\boldsymbol{x}} \langle \boldsymbol{a}, \boldsymbol{x} \rangle^2 &= \begin{bmatrix} \nabla_{x_1} \langle \boldsymbol{a}, \boldsymbol{x} \rangle^2 \\ \vdots \\ \nabla_{x_n} \langle \boldsymbol{a}, \boldsymbol{x} \rangle^2 \end{bmatrix} \\
&= \begin{bmatrix} \nabla_{x_1} (\sum_{i=1}^n a_i x_i)^2 \\ \vdots \\ \nabla_{x_n} (\sum_{i=1}^n a_i x_i)^2 \end{bmatrix} \\
&= \begin{bmatrix} 2(\sum_{i=1}^n a_i x_i) a_1 \\ \vdots \\ 2(\sum_{i=1}^n a_i x_i) a_n \end{bmatrix} \\
&= 2 \left(\sum_{i=1}^n a_i x_i \right) \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \\
&= 2 \langle \boldsymbol{a}, \boldsymbol{x} \rangle \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \\
&= 2 \langle \boldsymbol{a}, \boldsymbol{x} \rangle \boldsymbol{a}
\end{aligned}$$

□

2 Derivatives of Vector-Valued Functions

2.1 Exercise

Exercise 2.13.2.

Let $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$. Let a vector-valued function $\boldsymbol{f} : \mathbb{R} \to \mathbb{R}^n$ be defined as

$$\boldsymbol{f}(x) = [\sin(b_1 x), \sin(b_2 x), \ldots, \sin(b_n x)]^{\mathrm{T}}.$$

Show that

$$\nabla_x \langle \boldsymbol{a}, \boldsymbol{f}(x) \rangle = \sum_{i=1}^n a_i b_i \cos(b_i x).$$

Proof. Let $x \in \mathbb{R}$ and $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$.

$$
\begin{aligned}
\nabla_x \langle \boldsymbol{a}, \boldsymbol{f}(x) \rangle &= \nabla_x \left(\sum_{i=1}^{n} a_i \sin(b_i x) \right) \\
&= \sum_{i=1}^{n} a_i \cos(b_i x) b_i \\
&= \sum_{i=1}^{n} a_i b_i \cos(b_i x)
\end{aligned}
$$

□

Exercise 2.13.4.

Let $\boldsymbol{f} : \mathbb{R} \to \mathbb{R}^n$ and $\boldsymbol{g} : \mathbb{R} \to \mathbb{R}^n$. Show that

$$
\nabla_x \langle \boldsymbol{f}(x), \boldsymbol{g}(x) \rangle = \langle \nabla_x \boldsymbol{f}(x), \boldsymbol{g}(x) \rangle + \langle \boldsymbol{f}(x), \nabla_x \boldsymbol{g}(x) \rangle. \tag{4}
$$

Proof. Let $x \in \mathbb{R}$.

$$
\begin{aligned}
\nabla_x \langle \boldsymbol{f}(x), \boldsymbol{g}(x) \rangle &= \nabla_x \left(\sum_{i=1}^{n} f_i(x) g_i(x) \right) \\
&= \sum_{i=1}^{n} (\nabla_x f_i(x) g_i(x) + f_i(x) \nabla_x g_i(x)) \\
&= \sum_{i=1}^{n} (\nabla_x f_i(x)) g_i(x) + \sum_{i=1}^{n} f_i(x) (\nabla_x g_i(x)) \\
&= \langle \nabla_x \boldsymbol{f}(x), \boldsymbol{g}(x) \rangle + \langle \boldsymbol{f}(x), \nabla_x \boldsymbol{g}(x) \rangle
\end{aligned}
$$

□