Partial Derivatives of Vector-valued Functions

Vector-valued functions. Consider a general vector-valued function $f : \mathbb{R}^m \to \mathbb{R}^n$:

$$f(x) = \begin{pmatrix} f^1(x) \\ f^2(x) \\ \vdots \\ f^m(x) \end{pmatrix} \text{ where } x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

and $f^i : \mathbb{R}^n \to \mathbb{R}$ for all $i = 1, ..., m$. These real-valued functions $f^1, ..., f^m$ are called *components* of the vector-valued function f.

As an example, you could think of the constraint function $G(x)$.

Partial derivatives of f. As stated in the textbook, we adopt the convention that when the argument of a function is a column vector, the vector of partial derivatives is a row vector, and vice versa.[1]

1. **Differentiate the component function f^i with respect to x.**

$$f^i_x(x) = \begin{pmatrix} f^i_1(x) & ... & f^i_n(x) \end{pmatrix}$$

   The subscript j denotes the partial derivative with respect to x_j for $j = 1, ..., n$.

2. **Differentiate the vector-valued function f with respect to x_j.**

$$f_j(x) = \begin{pmatrix} f^1_j(x) \\ f^2_j(x) \\ \vdots \\ f^m_j(x) \end{pmatrix}$$

3. **Differentiate the vector-valued function f with respect to x.**

$$f_x(x) = \begin{pmatrix} f^1_x(x) \\ f^2_x(x) \\ \vdots \\ f^m_x(x) \end{pmatrix} = \begin{bmatrix} f^1_1(x) & ... & f^1_n(x) \\ f^2_1(x) & ... & f^2_n(x) \\ \vdots & \ddots & \vdots \\ f^m_1(x) & ... & f^m_n(x) \end{bmatrix}.$$

[1]Other books may adopt differnt conventions.

Examples.

Example 1. *Consider the following function $f : \mathbb{R}^2 \to \mathbb{R}^4$:*

$$f(x) = \begin{pmatrix} f^1(x) \\ f^2(x) \\ f^3(x) \\ f^4(x) \end{pmatrix} = \begin{pmatrix} x_1^2 + x_2^3 \\ x_1 + x_2 \\ x_1^2 - x_2^2 \\ 2x_1x_2 \end{pmatrix}, \text{ where } x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

We could calculate $f_x^1(x)$, $f_1(x)$ and $f_x(x)$.

(i) f^1 is a real-valued function and x is a vector, so $f_x^1(x)$ belongs to case 1.

$$f_x^1(x) = \begin{pmatrix} f_1^1(x) & f_2^1(x) \end{pmatrix} = \begin{pmatrix} 2x_1 & 3x_2^2 \end{pmatrix}$$

(ii) f is a vector-valued function and x_1 is a scalar, so $f_1(x)$ belongs to case 2.

$$f_1(x) = \begin{pmatrix} f_1^1(x) \\ f_1^2(x) \\ f_1^3(x) \\ f_1^4(x) \end{pmatrix} = \begin{pmatrix} 2x_1 \\ 1 \\ 2x_1 \\ 2x_2 \end{pmatrix}$$

(iii) f is a vector-valued function and x is a vector, so $f_x(x)$ belongs to case 3.

$$f_x(x) = \begin{bmatrix} f_1^1(x) & f_2^1(x) \\ f_1^2(x) & f_2^2(x) \\ f_1^3(x) & f_2^3(x) \\ f_1^4(x) & f_2^4(x) \end{bmatrix} = \begin{bmatrix} 2x_1 & 3x_2^2 \\ 1 & 1 \\ 2x_1 & -2x_2 \\ 2x_2 & 2x_1 \end{bmatrix}.$$

Example 2. *Consider the following function $f : \mathbb{R}^n \to \mathbb{R}$:*

$$f(x) = -wx, \quad \text{where } w = \begin{pmatrix} w_1 & \dots & w_n \end{pmatrix} \text{ and } x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

We want to calculate f_x and f_{xw}.

Solution. *f is a real-valued function and x is a vector, so $f_x(x)$ belongs to case 1. Then we have*

$$f_x(x) = \begin{pmatrix} f_1(x) & \dots & f_n(x) \end{pmatrix} = \begin{pmatrix} -w_1 & \dots & -w_n \end{pmatrix}$$

$f_x(x)$ is a vector-valued function and w is a vector, so $f_{xw}(x)$ belongs to case 3. Here, the function f_x is a row vector and w is also a row vector. So by the convention we adopt, the rows and columns in case 3 need to be switched.

$$f_{xw}(x) = \begin{pmatrix} \frac{\partial(-w_1)}{\partial w} & \dots & \frac{\partial(-w_n)}{\partial w} \end{pmatrix} = \begin{bmatrix} \frac{\partial(-w_1)}{\partial w_1} & \dots & \frac{\partial(-w_n)}{\partial w_1} \\ \frac{\partial(-w_1)}{\partial w_2} & \dots & \frac{\partial(-w_n)}{\partial w_2} \\ \vdots & \ddots & \vdots \\ \frac{\partial(-w_1)}{\partial w_n} & \dots & \frac{\partial(-w_n)}{\partial w_n} \end{bmatrix} = \begin{bmatrix} -1 & 0 & \dots & 0 \\ 0 & -1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & -1 \end{bmatrix} = -I.$$

Remark. *$F_{xw}(x^*, w) = -I$ in Example 8.4 Part I follows similarly.*

Remark. *In some books, they adopt the convention that vectors are viewed as column matrices. So the vector-valued function is written as $f : \mathbb{R}^m \to \mathbb{R}^n$:*

$$f(x) = \begin{pmatrix} f^1(x) & f^2(x) & \dots & f^m(x) \end{pmatrix} \text{ where } x = \begin{pmatrix} x_1 & x_2 & \dots & x_n \end{pmatrix}.$$