Directional derivatives

Definition 1. *Let $\boldsymbol{f}: \mathbb{R}^N \mapsto \mathbb{R}^M$. Let $\boldsymbol{x}_0 \in \mathbb{R}$ and $\boldsymbol{v} \in \mathbb{R}^N$. Then we say the directional derivative $\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{v}}$ is defined as the limit*

$$\lim_{h \to 0, h \neq 0} \frac{\boldsymbol{f}(\boldsymbol{x}_0 + h\,\boldsymbol{v}) - \boldsymbol{f}(\boldsymbol{x}_0)}{h}. \tag{1}$$

Example 2. Let $f(x, y) = x\,y$. Let $\boldsymbol{v}_1 = (1, 1), \boldsymbol{v}_2 = (2, 2)$. Calculate $\frac{\partial f}{\partial \boldsymbol{v}_1}(1, 1)$, $\frac{\partial f}{\partial \boldsymbol{v}_2}(1, 1)$.

Solution. We have

$$f((1, 1) + h\,(1, 1)) - f(1, 1) = (1 + h)^2 - 1 = 2\,h + h^2. \tag{2}$$

Now it is clear that $\frac{\partial f}{\partial \boldsymbol{v}_1}(1, 1) = 2$. Similarly we have $\frac{\partial f}{\partial \boldsymbol{v}_2}(1, 1) = 4$.

Exercise 1. Explain that partial derivatives are special cases of directional derivatives.

Exercise 2. Assume $\boldsymbol{f}$ is differentiable at $\boldsymbol{x}_0$. Prove that its directional derivative exists for all $\boldsymbol{v} \in \mathbb{R}^N$. Find a formula for its directional derivative using the Jacobian matrix of $\boldsymbol{f}$.

Exercise 3. Let $\boldsymbol{f}: \mathbb{R}^N \mapsto \mathbb{R}^M$ be such that its directional derivative exists for all $\boldsymbol{v} \in \mathbb{R}^N$ at some point $\boldsymbol{x}_0 \in \mathbb{R}^N$. Can we conclude that $\boldsymbol{f}$ is continuous at $\boldsymbol{x}_0$? Justify your answer.

Proposition 3. *Let $\boldsymbol{f}: \mathbb{R}^N \mapsto \mathbb{R}^M$ be differentiable at $\boldsymbol{x}_0$. Let $\boldsymbol{v} \in \mathbb{R}^N$ be any vector. Then*

$$\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{v}} = (D\,\boldsymbol{f}(\boldsymbol{x}_0))(\boldsymbol{v}). \tag{3}$$

Remark 4. Note that the left hand side is a vector in $\mathbb{R}^M$, while the right hand side is a linear function $D\,\boldsymbol{f}(\boldsymbol{x}_0)$ acting on a vector $\boldsymbol{v} \in \mathbb{R}^N$, thus is also a vector in $\mathbb{R}^M$.

Remark 5. Clearly, if A is the representation of $D\,\boldsymbol{f}(\boldsymbol{x}_0)$, we have

$$\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{v}} = A\,\boldsymbol{v}. \tag{4}$$

This time the right hand side is matrix-vector multiplication.

Proof. Exercise. □

Exercise 4. Let $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_N \in \mathbb{R}^N$ be such that $\|\boldsymbol{v}_i\| = 1$ for all i, $\boldsymbol{v}_i \cdot \boldsymbol{v}_j = 0$ for all $i \neq j$. Let $u: \mathbb{R}^N \mapsto \mathbb{R}$ be differentiable. Prove

$$\left(\frac{\partial u}{\partial \boldsymbol{v}_1}\right)^2 + \cdots + \left(\frac{\partial u}{\partial \boldsymbol{v}_N}\right)^2 = \left(\frac{\partial u}{\partial x_1}\right)^2 + \cdots + \left(\frac{\partial u}{\partial x_N}\right)^2. \tag{5}$$

Question 6. *If directional derivative linear in the direction, then differentiable?*

Geometric meaning of the differential

Let $\boldsymbol{f}: \mathbb{R}^N \mapsto \mathbb{R}^M$ be differentiable at $\boldsymbol{x}_0 \in \mathbb{R}^N$. Then $D\boldsymbol{f}(\boldsymbol{x}_0): \mathbb{R}^N \mapsto \mathbb{R}^M$ is a linear function and has a matrix representation, called the Jacobian. We can view the Jacobian $\left(\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}\right)$ row by row or column by column.

- Column-by-column.

$$\left(\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}\right) = \begin{pmatrix} \frac{\partial \boldsymbol{f}}{\partial x_1} & \cdots & \frac{\partial \boldsymbol{f}}{\partial x_N} \end{pmatrix} \tag{6}$$

  This point of view is more convenient when $N < M$. The basic understanding is that each vector $\frac{\partial \boldsymbol{f}}{\partial x_i}$ is a tangent vector to the image of $\boldsymbol{f}$, which is a surface in $\mathbb{R}^M$.

- Row-by-row.

$$\left(\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}\right) = \begin{pmatrix} (\text{grad } f_1)^T \\ \vdots \\ (\text{grad } f_M)^T \end{pmatrix} \tag{7}$$

  where the "gradient" is defined for any scalar function $f: \mathbb{R}^N \mapsto \mathbb{R}$ through

$$(\text{grad } f)(\boldsymbol{x}_0) := \begin{pmatrix} \frac{\partial f}{\partial x_1}(\boldsymbol{x}_0) \\ \vdots \\ \frac{\partial f}{\partial x_N}(\boldsymbol{x}_0) \end{pmatrix} \tag{8}$$

  The geometric meaning of grad f will be discussed later.

The case $N < M$

There are two special sub-cases where the geometric meaning is particularly clear: $N = 1$ and $N = M - 1$.

$N = 1$

In the case $N = 1$ we often denote the variable by t, that is

$$\boldsymbol{f}(t) = \begin{pmatrix} f_1(t) \\ \vdots \\ f_M(t) \end{pmatrix}. \tag{9}$$

It is easy to see then that the matrix representation of $D\boldsymbol{f}$ is $\begin{pmatrix} f_1'(t) \\ \vdots \\ f_M'(t) \end{pmatrix}$ which can be seen as a vector in $\mathbb{R}^M$.

Exercise 5. Prove the above claim: The matrix representation of $D\boldsymbol{f}$ is $\begin{pmatrix} f_1'(t) \\ \vdots \\ f_M'(t) \end{pmatrix}$

To understand the geometric meaning of this vector, we need to first understand the geometric meaning of $\boldsymbol{f}(t)$.

Definition 7. (Curve in $\mathbb{R}^M$) *A curve in $\mathbb{R}^M$ is the image of a continuous function $\boldsymbol{f}: \mathbb{R} \mapsto \mathbb{R}^M$. If $\boldsymbol{f}$ is furthermore one-to-one then it is called a simple curve.*

Example 8. The unit circle in $\mathbb{R}^2$ is a curve.

We notice that the image of

$$\boldsymbol{f}(t) := \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} \tag{10}$$

is exactly the unit circle.

Now from the definition

$$\boldsymbol{f}'(t_0) = \lim_{t \longrightarrow t_0} \frac{\boldsymbol{f}(t) - \boldsymbol{f}(t_0)}{t - t_0} \tag{11}$$

we see that the line:

$$\{\boldsymbol{f}(t_0) + s\,\boldsymbol{f}'(t_0) |\; s \in \mathbb{R}\} \tag{12}$$

should be the tangent line of the curve $\boldsymbol{f}(t)$. One can also write the equation for this line in coordinates:

$$\frac{x_1 - f_1(t_0)}{f_1'(t_0)} = \cdots = \frac{x_M - f_M(t_0)}{f_M'(t_0)}. \tag{13}$$

Exercise 6. How should we understand the above equation if some $f_i'(t_0) = 0$?

Example 9. Consider

$$\boldsymbol{f}(t) := \begin{pmatrix} R\cos t \\ R\sin t \\ t \end{pmatrix}. \tag{14}$$

Find the equation for its tangent.

Solution. We have

$$\boldsymbol{f}'(t) = \begin{pmatrix} -R\sin t \\ R\cos t \\ 1 \end{pmatrix} \tag{15}$$

so the equation is

$$\frac{x - R\cos t_0}{-R\sin t_0} = \frac{y - R\sin t_0}{R\cos t_0} = z - t_0. \tag{16}$$

Remark 10. Note that if we identify $\boldsymbol{f}(t)$ as a curve in $\mathbb{R}^M$, then the size of $\boldsymbol{f}'(t)$ does not matter, as it only represents details of parametrization; On the other hand the direction $\boldsymbol{f}'(t)/\|\boldsymbol{f}'(t)\|$ is very informative. Therefore in classical differential geometry, we often use the so-called "arc length" parametrization, that is do a change of variable $t \longrightarrow s$ where s is determined through

$$\frac{\mathrm{d}s}{\mathrm{d}t} = \|\boldsymbol{f}'(t)\|. \tag{17}$$

Exercise 7. Prove that after this change of variable, $\|\boldsymbol{f}'(s)\| = 1$.