Lecture Notes: Vector Derivative

Yufei Tao
Department of Computer Science and Engineering
Chinese University of Hong Kong
taoyf@cse.cuhk.edu.hk

1 Scalar and Vector Functions

Recall that a *function* f takes an *input*, and yields an *output*. For example, in $f(t) = t^2 + 2t$, the input is t, whereas the output is the real value resulting from the calculation $t^2 + 2t$. We say that f is a *scalar function* if its output is a real value.

The output of a function can also be a vector. In this case, we refer to the function as a *vector function*. For instance, consider $\boldsymbol{f}(t) = [t^2, 2t, t^3 - t]$. Its input is t. For every fixed t, $\boldsymbol{f}(t)$ outputs a 3d vector $[t^2, 2t, t^3 - t]$. We will adopt the convention of using boldfaces to represent vector functions.

An input to a function may consist of multiple parameters. For example, $f(x, y) = x^2 + xy + y^3$ and $\boldsymbol{f}(x, y, z) = [xyz, y^3 z + y^2]$. If a scalar function f takes d real values as its input, we say that f is a *scalar field* in $\mathbb{R}^d$. Similarly, if a vector function $\boldsymbol{f}$ takes d real values as its input, we say that f is a *vector field* in $\mathbb{R}^d$. For example, the $f(x, y)$ and $\boldsymbol{f}(x, y, z)$ shown earlier are a scalar field in $\mathbb{R}^2$ and a vector field in $\mathbb{R}^3$, respectively.

2 Limits and Continuity of One-Variable Vector Functions

Consider first a scalar function $f(t)$ that takes a single real value t as its input. Recall that its *limit* at t_0 is defined as follows:

Definition 1. *Suppose that a scalar function $f(t)$ is defined around*[1] t_0 *(but not necessarily at t_0). We say that*

$$\lim_{t \to t_0} f(t) = v$$

if for any real $\delta > 0$, we can find a real value $\epsilon > 0$ such that $|f(t) - v| < \delta$ for all t satisfying $0 < |t - t_0| < \epsilon$.

Now consider a vector function $\boldsymbol{f}(t)$ that takes a single real value t as its input. Suppose that the output of $\boldsymbol{f}(t)$ is a d-dimensional vector. By definition, we can write the output vector in its component form $[x_1(t), x_2(t), ..., x_d(t)]$. Now we extend Definition 1 to vector functions:

Definition 2. *Suppose that $\boldsymbol{f}(t) = [x_1(t), x_2(t), ..., x_d(t)]$ is defined around t_0 (but not necessarily at t_0). We say that*

$$\lim_{t \to t_0} \boldsymbol{f}(t) = [v_1, v_2, ..., v_d]$$

if it holds for each $i \in [1, d]$ that $\lim_{t \to t_0} x_i(t) = v_i$.

[1] This means that there is a $\rho > 0$ such that $f(t)$ is defined for t satisfying $0 < |t - t_0| < \rho$.

For example, suppose that $\boldsymbol{f}(t) = [t^2, \sin(t)/t]$. Since $\lim_{t\to 0} t^2 = 0$ and $\lim_{t\to 0} \frac{\sin(t)}{t} = 1$, we know that $\lim_{t\to 0} \boldsymbol{f}(t) = [0, 1]$.

Definition 3. *Suppose that $\boldsymbol{f}(t) = [x_1(t), x_2(t), ..., x_d(t)]$ is defined around t_0 and at t_0. We say that $\boldsymbol{f}(t)$ is* **continuous** *at t_0 if $\lim_{t\to t_0} \boldsymbol{f}(t) = \boldsymbol{f}(t_0)$.*

For example, $\boldsymbol{f}(t) = [t^2, \sin(t)/t]$ is *not* continuous at 0 because the function is undefined at $t = 0$. On the other hand, $\boldsymbol{f}(t) = [t^2, \sqrt{t}+1]$ is continuous at $t = 0$. However, the following function is *not* continuous at $t = 0$:

$$\boldsymbol{f}(t) = \begin{cases} [t^2, \sqrt{t}+1] & \text{if } t \neq 0 \\ [0, 2] & \text{if } t = 0 \end{cases}$$

This is because $\lim_{t\to 0} \boldsymbol{f}(t) = [0, 1] \neq \boldsymbol{f}(0)$.

3 Derivatives of Vector Functions

Recall that derivatives of scalar functions are defined as follows:

Definition 4. *Suppose that scalar function $f(t)$ is defined around t_0 and at t_0. If the following limit exists:*

$$\lim_{\Delta t\to 0} \frac{f(t_0 + \Delta t) - f(t_0)}{\Delta t}$$

then we say that

- *$f(t)$ is* **differentiable** *at t_0.*
- *the above limit, denoted as $f'(t_0)$, is the* **derivative** *of $f(t)$ at $t = t_0$.*

We now extend the definition to vectors:

Definition 5. *Suppose that vector function $\boldsymbol{f}(t)$ is defined around t_0 and at t_0. If the following limit exists:*

$$\lim_{\Delta t\to 0} \frac{\boldsymbol{f}(t_0 + \Delta t) - \boldsymbol{f}(t_0)}{\Delta t}$$

then we say that

- *$\boldsymbol{f}(t)$ is* **differentiable** *at t_0.*
- *the above limit, denoted as $\boldsymbol{f}'(t_0)$, is the* **derivative** *of $\boldsymbol{f}(t)$ at $t = t_0$.*

The next important lemma provides another view of the above definition through components:

Lemma 1. *Suppose that $\boldsymbol{f}(t) = [x_1(t), x_2(t), ..., x_d(t)]$ is differentiable at t_0 such that $\boldsymbol{f}'(t_0) = [y_1(t_0), y_2(t_0), ..., y_d(t_0)]$. Then, $y_i(t_0) = x_i'(t_0)$ for each $i \in [1, d]$.*

Proof. By definition of vector subtraction:

$$\boldsymbol{f}(t_0 + \Delta t) - \boldsymbol{f}(t_0) = [x_1(t_0 + \Delta t) - x_1(t_0), x_2(t_0 + \Delta t) - x_2(t_0), ..., x_d(t_0 + \Delta t) - x_d(t_0)].$$

Since

$$\lim_{\Delta t \to 0} \frac{\boldsymbol{f}(t_0 + \Delta t) - \boldsymbol{f}(t_0)}{\Delta t} = [y_1(t_0), y_2(t_0), ..., y_d(t_0)] \tag{1}$$

we know

$$\begin{aligned}
& \lim_{\Delta t \to 0} \frac{\boldsymbol{f}(t_0 + \Delta t) - \boldsymbol{f}(t_0)}{\Delta t} \\
= & \lim_{\Delta t \to 0} \frac{[x_1(t_0 + \Delta t) - x_1(t_0), x_2(t_0 + \Delta t) - x_2(t_0), ..., x_d(t_0 + \Delta t) - x_d(t_0)]}{\Delta t} \\
\text{(scalar multiplication)} = & \lim_{\Delta t \to 0} \left[\frac{x_1(t_0 + \Delta t) - x_1(t_0)}{\Delta t}, \frac{x_2(t_0 + \Delta t) - x_2(t_0)}{\Delta t}, ..., \frac{x_1(t_0 + \Delta t) - x_1(t_0)}{\Delta t}\right] \\
\text{(from (1))} = & [y_1(t_0), y_2(t_0), ..., y_d(t_0)].
\end{aligned}$$

It thus follows from Definition 2 that, for each $i \in [1, d]$:

$$\lim_{\Delta t \to 0} \frac{x_i(t_0 + \Delta t) - x_i(t_0)}{\Delta t} = y_i(t_0).$$

The left hand side of the above is precisely $x_i'(t_0)$ by Definition 4. We thus complete the proof. □

The above lemma provides a convenient and intuitive way to compute the derivative of a vector function. For example, consider $\boldsymbol{f}(t) = [\sin^2 t, \cos^2 t]$. Then we immediately know $\boldsymbol{f}'(t_0) = [2\sin(t_0)\cos(t_0), -2\sin(t_0)\cos(t_0)]$. Note that we will often replace t_0 with t in $\boldsymbol{f}(t_0)$ (after all, t_0 is nothing but a variable name). For instance, in this example, $\boldsymbol{f}'(t) = [2\sin(t)\cos(t), -2\sin(t)\cos(t)]$.

Vector derivatives obey some rules that are reminiscent of the corresponding rules on scalar functions:

1. $(\boldsymbol{f}(t) + \boldsymbol{g}(t))' = \boldsymbol{f}'(t) + \boldsymbol{g}'(t)$.
2. $(\boldsymbol{f}(t) \cdot \boldsymbol{g}(t))' = \boldsymbol{f}'(t) \cdot \boldsymbol{g}(t) + \boldsymbol{f}(t) \cdot \boldsymbol{g}'(t)$.
3. Suppose that the outputs of $\boldsymbol{f}(t)$ and $\boldsymbol{g}(t)$ are 3d vectors. Then, $(\boldsymbol{f}(t) \times \boldsymbol{g}(t))' = \boldsymbol{f}'(t) \times \boldsymbol{g}(t) + \boldsymbol{f}(t) \times \boldsymbol{g}'(t)$.

Next, we will prove Rules 1 and 2 in full. The proof for Rule 3 is very tedious but not difficult; we will outline its main ideas.

Proof of Rule 1. Let $\boldsymbol{f}(t) = [x_1(t), ..., x_d(t)]$ and $\boldsymbol{g}(t) = [y_1(t), ..., y_d(t)]$. From Lemma 1, we know that $\boldsymbol{f}'(t) = [x_1'(t), ..., x_d'(t)]$ and $\boldsymbol{g}'(t) = [y_1'(t), ..., y_d'(t)]$. We have:

$$\begin{aligned}
(\boldsymbol{f}(t) + \boldsymbol{g}(t))' &= [x_1(t) + y_1(t), ..., x_d(t) + y_d(t)]' \\
\text{(by Lemma 1)} &= \left[(x_1(t) + y_1(t))', ..., (x_d(t) + y_d(t))'\right] \\
&= \left[x_1'(t) + y_1'(t), ..., x_d'(t) + y_d'(t)\right] \\
&= \boldsymbol{f}'(t) + \boldsymbol{g}'(t).
\end{aligned}$$

□

Proof of Rule 2. Let $\boldsymbol{f}(t) = [x_1(t), ..., x_d(t)]$ and $\boldsymbol{g}(t) = [y_1(t), ..., y_d(t)]$. From Lemma 1, we know that $\boldsymbol{f}'(t) = [x_1'(t), ..., x_d'(t)]$ and $\boldsymbol{g}'(t) = [y_1'(t), ..., y_d'(t)]$. We have:

$$\begin{aligned}
(\boldsymbol{f}(t) \cdot \boldsymbol{g}(t))' &= \left(\sum_{i=1}^{d} x_i(t) \cdot y_i(t)\right)' \\
&= \sum_{i=1}^{d} \left(x_i'(t) \cdot y_i(t) + x_i(t) \cdot y_i'(t)\right) \\
&= \sum_{i=1}^{d} x_i'(t) \cdot y_i(t) + \sum_{i=1}^{d} x_i(t) \cdot y_i'(t) \\
&= \boldsymbol{f}'(t) \cdot \boldsymbol{g}(t) + \boldsymbol{f}(t) \cdot \boldsymbol{g}'(t).
\end{aligned}$$

□

Proof of Rule 3 (Sketch). Let $\boldsymbol{f}(t) = [x_1(t), x_2(t), x_3(t)]$ and $\boldsymbol{g}(t) = [y_1(t), y_2(t), y_3(t)]$. The key to the proof is to write out both sides of Rule 2 in their component forms. For the left hand side, we know:

$$\begin{aligned}
& (\boldsymbol{f}(t) \times \boldsymbol{g}(t))' \\
= \; & [x_2(t)y_3(t) - x_3(t)y_2(t), x_3(t)y_1(t) - x_1(t)y_3(t), x_1(t)y_2(t) - x_2(t)y_1(t)]' \\
= \; & \left[(x_2(t)y_3(t))' - (x_3(t)y_2(t))', (x_3(t)y_1(t))' - (x_1(t)y_3(t))', (x_1(t)y_2(t))' - (x_2(t)y_1(t))'\right].
\end{aligned}$$

You want to unfold the right hand side $\boldsymbol{f}'(t) \times \boldsymbol{g}(t) + \boldsymbol{f}(t) \times \boldsymbol{g}'(t)$ into similar forms. Then, you will see that both sides are equivalent. □