A COORDINATE-FREE VIEW OF THE DERIVATIVE OF A MAP BETWEEN EUCLIDEAN SPACES

SLOBODAN N. SIMIĆ

As a prelude to a definition of the derivative of a smooth map between smooth manifolds, let us reexamine the definition of the derivative of a smooth map between Euclidean spaces,

$$f : \mathbb{R}^n \to \mathbb{R}^m.$$

In multivariable calculus, $Df(a)$ is usually thought of simply as a matrix of partial derivatives,

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \vdots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix},$$

where $f = (f_1, \ldots, f_m)$. However, this identification of $Df(a)$ with its Jacobian matrix depends on the linear structure of $\mathbb{R}^n$ and the choice of the standard basis in particular, so we need to look more a more coordinate-free view of the derivative.

We can do this by reexamining the domain and codomain of $Df(a)$ and understanding how $Df(a)$ acts on the elements of the domain. First, $Df(a)$ is a linear transformation $\mathbb{R}^n \to \mathbb{R}^m$, so it takes vectors $v \in \mathbb{R}^n$ and produces vectors $Df(a)v \in \mathbb{R}^m$. However, here we are again using the special features of Euclidean spaces, namely, the natural identification of the tangent spaces at any point with the space itself. But since $Df(a)$ is the derivative of f **at** a, it makes sense to think of it as a linear transformation

$$Df(a) : T_a\mathbb{R}^n \to T_b\mathbb{R}^m,$$

where $b = f(a)$. The advantage of this viewpoint is that we have two equivalent coordinate-free characterizations (or definitions) of the tangent space to $\mathbb{R}^n$ at a point. In the first one, $T_a\mathbb{R}^n = \mathscr{C}(a)/\sim$, where $\mathscr{C}(a)$ is the space of smooth curves $t \mapsto \gamma(t)$ defined on a neighborhood of $0 \in \mathbb{R}$ such that $\gamma(0) = a$. Two such curves are identified if their velocity vectors at zero are the same:

$$\gamma_1 \sim \gamma_2 \quad \Leftrightarrow \quad \dot{\gamma}_1(0) = \dot{\gamma}_2(0).$$

The second definition of $T_a\mathbb{R}^n$ is as the space of derivation of the algebra $C^\infty(a)$ of smooth functions defined in a neighborhood of a, with two such functions identified if they coincide on the intersection of their domains.

Accordingly, we have these two coordinate-free views of $Df(a)$.

The derivative as a map of velocity vectors

Let us regard an arbitrary $X \in T_a\mathbb{R}^n$ as a velocity vector of a curve passing through a: $X = \dot{\gamma}(0)$, for some $\gamma \in \mathscr{C}(a)$. Since $\tilde{\gamma} = f \circ \gamma \in \mathscr{C}(b)$, the most natural way to define $Y = Df(a)X$ is by

$$Y = \dot{\tilde{\gamma}}(0).$$

Date: April 2, 2009.

Let us now think of Y as a derivation and let $u \in C^\infty(b)$ be arbitrary. What is Yu? Answer: the directional derivative of u in the direction of Y, where Y is regarded as the velocity vector of $\tilde{\gamma}$ at zero:

$$\begin{aligned} Yu &= \left.\frac{d}{dt}\right|_0 (u \circ \tilde{\gamma})(t) \\ &= \left.\frac{d}{dt}\right|_0 (u \circ f \circ \gamma)(t) \\ &= \left.\frac{d}{dt}\right|_0 [(u \circ f) \circ \gamma](t) \\ &= X(u \circ f). \end{aligned}$$

Therefore,

$$(Df(a)X)u = X(u \circ f).$$

The derivative as a map of derivations

Let us now think of $X \in T_a\mathbb{R}^n$ as a derivation of the algebra $C^\infty(a)$ and compute $Y = Df(a)X \in T_b\mathbb{R}^m$ as a derivation of the algebra $C^\infty(b)$. Since $Df(a)$ is linear, it is enough to look at its action on the elements of a basis of $T_a\mathbb{R}^n$, e.g., the natural basis

$$\mathscr{X} = \left\{ \frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n} \right\},$$

where all the "partials" are taken at a. Recall that $\frac{\partial}{\partial x_i} \equiv E_i$, the i^{th} element of the standard basis for $\mathbb{R}^n$, thought of as a derivation of $C^\infty(a)$. Similarly, we have the natural basis

$$\mathscr{Y} = \left\{ \frac{\partial}{\partial y_1}, \dots, \frac{\partial}{\partial y_m} \right\}$$

of $T_b\mathbb{R}^m$, where the "partials" are taken at b. Again, $\frac{\partial}{\partial y_i} \equiv F_i$, the i^{th} element of the standard basis for $\mathbb{R}^m$, thought of as a derivation of $T_b\mathbb{R}^m$. Then:

$$\begin{aligned} Df(a)\frac{\partial}{\partial x_i} &\equiv Df(a)E_i \\ &= i^{th}\text{column of } \; Df(a) \\ &= \sum_{j=1}^{n} \frac{\partial f_j}{\partial x_i}(a) F_j \\ &\equiv \sum_{j=1}^{n} \frac{\partial f_j}{\partial x_i}(a) \frac{\partial}{\partial y_j}. \end{aligned}$$

For any $u \in C^\infty(b)$, we obtain

$$\begin{aligned}\left(Df(a)\frac{\partial}{\partial x_i}\right)u &= \sum_{j=1}^n \frac{\partial f_j}{\partial x_i}(a)\frac{\partial u}{\partial y_j}(b) \\ &= \sum_{j=1}^n \frac{\partial u}{\partial y_j}(b)\frac{\partial f_j}{\partial x_i}(a) \\ &= \frac{\partial}{\partial x_i}(u \circ f),\end{aligned}$$

by the chain rule. Therefore,

$$\left(Df(a)\frac{\partial}{\partial x_i}\right)u = \frac{\partial}{\partial x_i}(u \circ f),$$

for every $u \in C^\infty(b)$. By linearity of $Df(a)$, this extends to all $X \in T_a\mathbb{R}^n$, i.e.,

$$(Df(a)X)u = X(u \circ f). \tag{$*$}$$

Therefore, in both cases we have the coordinate-free formula $(*)$. This is how we will define the derivative of a smooth map between manifolds. In Boothby's notation

$$Df(a) = f_*.$$

I prefer a more functorial notation $T_a f$, since to each smooth $f : M \to N$ we assign its derivative or tangent map $Tf : TM \to TN$.

The point of this story is that the notion of the derivative of a map between manifolds is a natural extension of the notion of the derivative between Euclidean spaces. Boothby doesn't really point this out explicitly, but this observation makes several things easier to prove. For example, once the property

$$(h \circ g)_* = h_* \, g_*$$

is (easily) established for every $g : M \to N$ and $h : N \to L$, it is almost immediate that the matrix of the derivative of $f : M \to N$ at $p \in M$ relative to the coordinate frames of T_pM and $T_{f(p)}N$ is exactly the Jacobian matrix of the corresponding expression of f in local coordinates. Namely, let (U, φ) and (V, ψ) be coordinate neighborhoods of p and $q = f(p)$, respectively and let

$$\hat{f} = \psi \circ f \circ \varphi^{-1} : \varphi(U) \to \psi(V).$$

Since φ^{-1} and ψ^{-1} are diffeomorphisms, their derivatives are linear isomorphisms, so they map $\mathscr{X}$ and $\mathscr{Y}$ above (with $a = \varphi(p)$ and $b = \psi(q)$) to bases $(\varphi^{-1})_*(\mathscr{X}), (\psi^{-1})_*(\mathscr{Y})$ of T_pM and T_qN respectively. Since

$$\hat{f}_* = \psi_* f_* \varphi_*^{-1}$$

the matrix of $T_pf = f_* : T_pM \to T_qN$ relative to these two bases is the same as the matrix of $D\hat{f}(a)$ relative to the bases $\mathscr{X}, \mathscr{Y}$, i.e., the Jacobian matrix of $D\hat{f}(a)$.

Department of Mathematics, San José State University, San José, CA 95192-0103
E-mail address: simic@math.sjsu.edu