Equivalent definitions of the tangent space

Alex Taylor
The University of Illinois at Urbana-Champaign

Abstract

We prove the equivalence of several common definitions of the tangent space to a smooth manifold, and then show that this equivalence is "natural" in the sense that the differentials are related via factorization through the corresponding isomorphism between any two tangent spaces.

Contents

1 Preliminaries

We denote smooth charts on an n-dimensional smooth manifold by $\varphi : U \subseteq M \to \mathbb{R}^n$. Given $p \in M$, without any loss of generality we will assume that any chart taken around p is centered at p, meaning that $\varphi(p) = 0$.

Given a smooth map $f : M \to \mathbb{R}$ we will call the map $\widetilde{f} = f \circ \varphi^{-1} : \varphi(U) \subseteq \mathbb{R}^n \to \mathbb{R}$ a local ***coordinate representation*** for f with respect to the smooth chart φ. In particular, the standard Euclidean coordinate functions $\widetilde{x}^i : \mathbb{R}^n \to \mathbb{R}$ are local coordinate representations for the functions $x^i = \widetilde{x}^i \circ \varphi : U \subseteq M \to \mathbb{R}$ with respect to some smooth chart φ. Often, by an intentional (but harmless) abuse of notation, we will denote both points in M and points in $\mathbb{R}^n$ under the image of a smooth chart φ by the same symbol x.

Throughout this note we will utilize Taylor's theorem for smooth functions $f : \mathbb{R}^n \to \mathbb{R}$ with an integral form for the remainder. In order to state the theorem precisely we fix the following standard notation: for any m-tuple $(i_1, i_2, \ldots, i_m)$ of indices $1 \leq i_j \leq n$, let $m = I$ denote the number of indices and

$$\partial_I = \frac{\partial}{\partial x_{i_1} \cdots \partial x_{i_m}}$$
$$(x-p)^I = (x_{i_1} - p_{i_1}) \cdots (x_{i_m} - p_{i_m})$$

for any fixed $p = (p_1, \ldots, p_n) \in \mathbb{R}^n$. The the ***kth order Taylor polynomial*** of f centered at p is the function

$$T_k(x) = f(p) + \sum_{m=1}^{k} \frac{1}{m!} \sum_{|I|=m} \partial_I f(p)(x-p)^I.$$

Theorem 1 (Taylor's theorem)**.** *Let $U \subseteq \mathbb{R}^n$ be an open set with a convex subset $W \subseteq U$. Fix some $p \in U$ and suppose that $f \in C^{k+1}(U)$ for some $k \geq 0$. Then for any $x \in W$ we have*

$$f(x) = T_k(x) + R_k(x)$$

where R_k is the kth order remainder term given by

$$R_k(x) = \frac{1}{k!} \sum_{|I|=k+1} (x-p)^I \int_0^1 (1-t)^k \partial_I f(p + t(x-p))\ dt$$

In particular for $k = 1$, the first order Taylor polynomial of f centered at p is the linear approximation

$$T_1(x) = f(p) + \nabla f(p)^T x$$

so in this context Taylor's theorem says that

$$f(x) = f(p) + \nabla f(p)^T x + \sum_{i,j=1}^{n} \left[(x_i - p_i)(x_j - p_j) \int_0^1 (1-t) \frac{\partial^2 f}{\partial x_i \partial x_j}(p + t(x-p))\ dt \right].$$

2 Tangent vectors in Euclidean space

Geometrically speaking, vectors in $\mathbb{R}^n$ are visualized as arrows attached to points, and this really means that we are thinking about a vector attached to a point $p \in \mathbb{R}^n$ as living in a copy of $\mathbb{R}^n$ with its origin translated to p. More precisely, this is the space

$$p \times \mathbb{R}^n = \{(p, v) : v \in \mathbb{R}^n = \mathbb{R}^n_p.$$

Thus, for instance, a tangent vector to the sphere $S^n \subset \mathbb{R}^n$ at a point $p \in S^n$ lives in the space $\mathbb{R}^n_p$. Notice that $\mathbb{R}^n_p$ really is a vector space with respect to the operations

$$(p, v_1) + (p, v_2) = (p, v_1 + v_2)$$
$$c(p, v) = (p, cv)$$

for every $v_1, v_2 \in \mathbb{R}^n$ and $c \in \mathbb{R}$. We will call this the space of ***geometric tangent vectors*** to $\mathbb{R}^n$ at p. It's naturally isomorphic to $\mathbb{R}^n$ via the map $v \mapsto (p, v)$, and in particular it has a standard basis $\{(p, e_1), (p, e_2), \ldots, (p, e_n)\}$.

We recall the following basic notion from calculus. For any $(p, v) \in \mathbb{R}^n_p$, define the ***directional derivative*** at p in the direction of v as the function $D_v\big|_p : C^\infty(\mathbb{R}^n) \to \mathbb{R}$ given by

$$\begin{aligned} D_v\big|_p(f) &= (D_v f)(p) \\ &= \lim_{t\to 0} \frac{f(p+tv) - f(p)}{t}. \end{aligned}$$

This operation is linear over $\mathbb{R}$ and satisfies the product rule:

$$(D_v fg)(p) = g(p)(D_v f)(p) + f(p)(D_v g)(p).$$

Moreover, the directional derivative can be expressed in coordinates as follows: given $v = \sum_{i=1}^n v_i e_i$, let $g(t) = p + tv$, then we have

$$\begin{aligned} (D_v f)(p) &= \lim_{t\to 0} \frac{f(p+tv) - f(p)}{t} \\ &= \frac{d}{dt}\bigg|_{t=0} f(g(t)) \\ &= \big[\nabla f(g(t)) g'(t)\big]\,\big|_{t=0} \qquad \text{(by the chain rule)} \\ &= \sum_{i=1}^n v_i \frac{\partial f}{\partial x_i}(p) \end{aligned}$$

and in particular, $(D_{e_i} f)(p) = (\partial f/\partial x_i)(p)$, so the directional derivatives along the standard basis vectors are simply the partial derivatives:

$$D_{e_i}\big|_p = \frac{\partial}{\partial x_i}\bigg|_p$$

for any $(p, e_i) \in \mathbb{R}^n_p$. Motivated by the directional derivative operators, we make the following definition. Given a fixed point $p \in \mathbb{R}^n$, $\omega : C^\infty(\mathbb{R}^n) \to \mathbb{R}$ is called a ***derivation*** of $C^\infty(\mathbb{R}^n)$ at p if:

(i) ω is linear over $\mathbb{R}$, and

(ii) ω satisfies the product rule

$$\omega(fg) = f(p)\omega(g) + g(p)\omega(g).$$

We let $T_p\mathbb{R}^n$ denote the set of all derivations of $C^\infty(\mathbb{R}^n)$ at p, called the ***tangent space*** to $\mathbb{R}^n$ at p. This is clearly a vector space with respect to pointwise function addition and scalar multiplication:

$$\begin{aligned} (\omega_1 + \omega_2)(f) &= \omega_1(f) + \omega_2(f) \\ (c\omega)(f) &= c(\omega f) \end{aligned}$$

for every $f \in C^\infty(\mathbb{R}^n)$ and $c \in \mathbb{R}$. With this terminology, our preceding discussion about directional derivatives implies that the set of directional derivative operators (at p) form a subspace of $T_p\mathbb{R}^n$. Moreover, we have a linear map

$$\mathbb{R}^n_p \to T_p\mathbb{R}^n : (p, v) \mapsto D_v\big|_p,$$

and in fact it's not hard to see that this map is injective. Indeed, suppose we have a vector $v = \sum_{i=1}^n v_i e_i$ such that $D_v\big|_p$ is the zero derivation in $T_p\mathbb{R}^n$. Then for any coordinate function $x^j : \mathbb{R}^n \to \mathbb{R}$ we use the coordinate expression for the directional derivative to obtain

$$0 = (D_v x^j)(p) = \sum_{i=1}^{n} v_i \frac{\partial x^j}{\partial x_i}(p) = \sum_{i=1}^{n} v_i \delta_{ij} = v_j$$

and since this holds for each $1 \leq j \leq n$, we conclude that $v = 0$. To summarize the discussion thus far, we've constructed an isomorphism of $\mathbb{R}^n_p$ with the directional derivative operators, which are themselves a subspace of $T_p\mathbb{R}^n$,

$$\mathbb{R}^n_p \simeq \{\text{directional derivatives at } p\} \subseteq T_p\mathbb{R}^n.$$

We aim to prove next that the map $(p, v) \mapsto D_v\big|_p$ is actually surjective, i.e. that every derivation is actually a directional derivative (in some direction). This will establish a three-way equivalence between geometric tangent vectors at p, directional derivatives at p, and derivations at p. But first we will prove some useful basic properties of derivations.

Fact 1 (Properties of derivations)**.** *Let $p \in \mathbb{R}^n$, $\omega \in T_p\mathbb{R}^n$, and $f, g \in \mathrm{C}^\infty(\mathbb{R}^n)$.*

(a) If f is constant then $\omega(f) = 0$.

(b) If $f(p) = g(p) = 0$ then $\omega(fg) = 0$.

(c) Derivations are locally determined: if f and g agree on a neighborhood of p then $\omega(f) = \omega(g)$.

Proof. (a) By linearity it suffices to check that $\omega(f_1) = 0$ for $f_1(x) = 1$. By the product rule we have

$$\omega(f_1) = \omega(f_1 \cdot f_1) = 2f_1(p)\omega(f_1) = 2\omega(f_1)$$

hence $\omega(f_1) = 0$.

(b) By the product rule again we have

$$\omega(fg) = f(p)\omega(g) + g(p)\omega(f) = 0 + 0 = 0$$

(c) Later we prove this in the more general context of smooth functions on a smooth manifold M and derivations of $\mathrm{C}^\infty(M)$, see Lemma 2.

■

Theorem 2. *Let $p \in \mathbb{R}^n$. The map $(p, v) \mapsto D_v\big|_p$ is a linear isomorphism $\mathbb{R}^n \to T_p\mathbb{R}^n$.*

Proof. We already know that the map in question is linear and injective, so it suffices to show that it is surjective. Let $\omega \in T_p\mathbb{R}^n$ be any derivation of $\mathrm{C}^\infty(\mathbb{R}^n)$ at p – we need to show that $\omega = D_v\big|_p$ for some vector $v \in \mathbb{R}^n$. Let $f \in \mathrm{C}^\infty(\mathbb{R}^n)$, then by Taylor's theorem we know that

$$f(x) = f(p) + \nabla f(p)^T(x - p) + \text{Remainder}$$

where we have an explicit integral formula for the remainder,

$$\text{Remainder} = \sum_{i,j=1}^{n} \left[(x_i - p_i)(x_j - p_j) \int_0^1 (1 - t) \frac{\partial^2 f}{\partial x_i \partial x_j}(p + t(x - p)) \, dt \right].$$

Notice that each term in the remainder is a product of two functions of x which vanish at $x = p$, thus by Fact 1 the derivation ω annihilates the remainder part. It also annihilates the constant term $f(p)$ and thus $\omega(f)$ is completely determined by the first order part,

$$\begin{aligned}\omega(f) &= \omega(f(p)) + \omega\left(\nabla f(p)^T (x-p)\right) \\ &= 0 + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(p)\omega(x_i - p_i) \\ &= \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(p)\omega(x^i).\end{aligned}$$

Hence we can choose $v = \sum_{i=1}^n \omega(x^i) e_i$ so that

$$\omega(f) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(p) v_i = (D_v f)(p),$$

which is to say, $\omega = D_v\big|_p$. ■

As a result of this theorem we have proven the aforementioned three-way equivalence between geometric tangent vectors at p, directional derivatives at p, and derivations at p,

$$\mathbb{R}^n_p \simeq \{\text{directional derivatives at } p\} = T_p\mathbb{R}^n.$$

In particular, by taking the standard basis $\{(p, e_1), (p, e_2), \ldots, (p, e_n)\}$ for $\mathbb{R}^n_p$ and applying the isomorphism of Theorem 2, we obtain a basis $\{D_{e_1}\big|_p, D_{e_2}\big|_p, \ldots, D_{e_n}\big|_p\}$ for $T_p\mathbb{R}^n$. As we previously observed, these directional derivatives along standard basis vectors are just partial derivative operators, so we have the following corollary.

Corollary 1. *For any $p \in \mathbb{R}^n$, the tangent space $T_p\mathbb{R}^n$ admits a basis of partial derivative operators,*

$$\left\{ \frac{\partial}{\partial x_1}\bigg|_p, \frac{\partial}{\partial x_2}\bigg|_p, \ldots, \frac{\partial}{\partial x_n}\bigg|_p \right\}$$

which means in particular that it has dimension n.

3 Tangent vectors as derivations of smooth functions

Motivated by the definition of the tangent space at a point in $\mathbb{R}^n$ as derivations of $C^\infty(\mathbb{R}^n)$, we will define the tangent spaces to a smooth manifold M as derivations of $C^\infty(M)$. Explicitly, given a fixed point $p \in M$, a ***derivation*** of $C^\infty(M)$ at p is a linear map $\omega : C^\infty(M) \to \mathbb{R}$ such that

(i) ω is linear over $\mathbb{R}$, and

(ii) ω satisfies the product rule

$$\omega(fg) = g(p)\omega(f) + f(p)\omega(g)$$

for every $f, g \in C^\infty(M)$.

Just as in Section 2, the set of all derivations of $C^\infty(M)$ at p form a vector space called the ***tangent space*** to M at p, denoted by T_pM. Thus, tangent vectors at p are precisely the derivations at p. Following Lee, we will take this definition of the tangent space as the primary one, and then construct the other conceptions of tangent space in terms of this one. In each section we will describe how the differentials and related constructions can be defined independently of this T_pM definition, in the event that one prefers to take another definition as the starting point.

The analogue of Fact 1 regarding derivations of $C^\infty(\mathbb{R}^n)$ holds in exactly the same way in this more general context of derivations of $C^\infty(M)$.

Fact 2 (Properties of tangent vectors). *Let M be a smooth manifold, $p \in M$, $\omega \in T_pM$, and $f, g \in C^\infty(M)$.*

(a) If f is constant then $\omega(f) = 0$.

(b) If $f(p) = g(p) = 0$ then $\omega(fg) = 0$.

(c) Tangent vectors are locally determined: if f and g agree on a neighborhood of p then $\omega(f) = \omega(g)$.

Proof. The proofs of (a) and (b) follow exactly the same pattern as in Fact 1, just by applying the product rule. As for (c), suppose f and g agree on a neighborhood of p and set $h = f - g$, so that h is a smooth function vanishing on a neighborhood of p. Take a smooth bump function $\eta \in C^\infty(M)$ that is identically equal to 1 on the support of h and has $\operatorname{supp}\eta = M \setminus \{p\}$. Then the product ηh is identically equal to h because $\eta = 1$ whenever h is nonzero, and since $\eta(p) = h(p) = 0$ we conclude from part (b) that

$$\omega(f - g) = \omega(h) = \omega(\eta h) = 0$$

for any tangent vector $\omega \in T_pM$. Hence $\omega(f) = \omega(g)$ by linearity. ∎

Given a smooth map $F : M \to N$ and $p \in M$, define the differential of F at p as the linear map $dF_p : T_pM \to T_{F(p)}N$ given by

$$dF_p(\omega)(g) = \omega(g \circ F)$$

where $\omega \in T_pM$ and $g \in C^\infty(N)$. In particular, the differential of a smooth function $f : M \to \mathbb{R}$ at p is the linear map $df_p : T_pM \to \mathbb{R}$ given by

$$df_p(\omega) = \omega(f)$$

for any $\omega \in T_pM$ (here the number $\omega(f)$ on the right-hand side is being surreptitiously identified with the linear map $\mathbb{R} \to \mathbb{R}$ that acts via multiplication by $\omega(f)$). What does this linear map look like in local coordinates? Take any smooth chart $\varphi : U \subseteq M \to \mathbb{R}^n$ around p and write $\omega = \sum_{i=1}^n \omega_i \left(\partial/\partial x_i\big|_p\right)$ with respect to this chart. Then

$$df_p(\omega) = \omega(f) = \sum_{i=1}^n \omega_i \frac{\partial f}{\partial x_i}(p) = \nabla f(p)^T(\omega_1, \omega_2, \ldots, \omega_n)$$

so in local coordinates the differential df_p is represented by the gradient of f. In general the differential dF_p is represented by the Jacobian matrix of F.

4 Tangent vectors as derivations of germs

Since the action of tangent vectors on smooth functions is determined locally, i.e. $\omega(f) = \omega(g)$ whenever f and g agree near $p \in M$, we may as well consider a tangent vector $\omega \in T_pM$ as acting on an equivalence class of functions that agree near p. This idea leads directly to the definition of tangent vectors as derivations of the space of germs on a smooth manifold. More precisely, fixing a point $p \in M$ we consider the set

$$S_p = \{\text{smooth } f : U \subseteq M \to \mathbb{R} \text{ defined} \\ \text{on an open neighborhood of } p\}$$

and define an equivalence relation on S_p by $f \sim g$ if and only if f and g agree on a neighborhood of p. Then we call the equivalence class $[f]_p$ of smooth functions agreeing with f near p the ***germ*** of f at p, and the set of equivalence classes is

$$S_p/{\sim} = C_p^\infty(M) = \text{space of germs of smooth functions at } p$$

Note that $C_p^\infty(M)$ is indeed a vector space (over $\mathbb{R}$) with respect to the operations

$$c[f]_p = [cf]_p$$
$$[f]_p + [g]_p = [f + g]_p$$

and moreover it's an algebra over $\mathbb{R}$ with respect to the product $[f]_p \cdot [g]_p = [fg]_p$. Just as before, a derivation of $C_p^\infty(M)$ is a linear map $\omega : C_p^\infty(M) \to \mathbb{R}$ satisfying the product rule

$$\omega[fg]_p = g(p)\omega[f]_p + f(p)\omega[g]_p$$

for every $f, g \in S_p$. Then the tangent space to M at p is the space of derivations of $C_p^\infty(M)$, which we denote by D_pM to distinguish it from the other definitions of the tangent space. Defining the tangent space as D_pM has the benefit of making the local nature of derivations more explicit, because by definition the derivations act only on equivalence classes of functions that agree near p. On the downside, working with equivalence classes of functions is slightly less concrete. Our main goal in this section, realized by the following theorem, is to prove that T_pM and D_pM are isomorphic.

Theorem 3. *The map $\phi : D_pM \to T_pM$ defined by $\phi(\omega)(f) = \omega[f]_p$ is a linear isomorphism.*

- For any $\omega \in D_pM$, $\phi(\omega)$ is a derivation of $\mathrm{C}^\infty(M)$ because linearity and the product rule both follow directly from the analogous properties of ω.
- ϕ is linear. If $\omega_1, \omega_2 \in D_pM$ and $c \in \mathbb{R}$, then for any $f \in \mathrm{C}^\infty(M)$ we have

$$\begin{aligned} \phi(c\omega_1 + \omega_2)(f) &= (c\omega_1 + \omega_2)[f]_p \\ &= c\omega_1[f]_p + \omega_2[f]_p \\ &= c\phi(\omega_1)(f) + \phi(\omega_2)(f) \end{aligned}$$

- ϕ is injective. If $\omega \in D_pM$ is a derivation such that $\phi(\omega) = 0 \in T_pM$ then by definition we have $\omega[f]_p = 0$ for every f defined in a neighborhood of p, hence $\omega = 0$ and $\ker \phi = 0$.

- ϕ is surjective. We need to show that every tangent vector $\nu \in T_pM$ can be represented by some derivation of germs $\omega \in D_pM$. Given $\nu \in T_pM$, simply define $\omega \in D_pM$ by $\omega[f]_p = \nu(f)$. Then ω is a well-defined derivation of germs because $[f]_p = [g]_p$ means that f and g agree in some neighborhood of p, which implies that $\nu(f) = \nu(g)$ since by Fact 2 the tangent vector ν is determined locally. Therefore $\phi(\omega) = \nu$ and we conclude that ϕ is surjective.

  There is one additional point that should be clarified. In the equation $\omega[f]_p = \nu(f)$, the function f on the left-hand side is a priori only defined in a neighborhood of p, whereas the tangent vector on the right-hand side acts on functions defined on all of M. This is easy to fix: simply extend the function to all of M by multiplying with a suitable bump function, and then apply ν.

Using this definition of the tangent space, it's straightforward to define the differential of a smooth map between manifolds in essentially the same way as before, just replacing functions with germs. Explicitly, given a smooth map $F : M \to N$ and $p \in M$, define the differential of F at p as the linear map $\widetilde{d}F_p : D_pM \to D_{F(p)}N$ given by

$$\widetilde{d}F_p(\omega)([g]_{F(p)}) = \omega[g \circ F]_p$$

where $\omega \in D_pM$ and $[g]_{F(p)} \in C^\infty_{F(p)}(N)$. We use the symbol $\widetilde{d}$ here to distinguish this differential from the differential we defined in Section 3 with respect to the tangent space T_pM. In Section 7 we explain the relationship between the differentials d and $\widetilde{d}$.

5 Tangent vectors as equivalence classes of smooth curves

Given a smooth curve $\gamma : I \to M$, recall that the tangent vector to γ at time $t \in I$ is the derivation $\gamma'(t) : \mathrm{C}^\infty(M) \to \mathbb{R}$ defined by $\gamma'(t)(f) = (f \circ \gamma)'(t)$.

Here's the idea: let M be a smooth manifold and fix some $p \in M$. A tangent vector $v \in T_pM$ should be uniquely determined by all the curves on M passing through p with velocity vector v at p. We can construct the tangent space in this way as follows. Consider the set of smooth curves on M which pass through p at time $t = 0$,

$$\Gamma_p = \{\text{smooth } \gamma : I \to M : \gamma(0) = p\}$$

and define an equivalence relation $\sim$ on Γ_p by

$$\gamma_1 \sim \gamma_2 \iff (f \circ \gamma_1)'(0) = (f \circ \gamma_2)'(0)$$

for every smooth function $f : M \to \mathbb{R}$ defined in an open neighborhood of p. Note that $f \circ \gamma$ is a function between subsets of $\mathbb{R}$, so we know what its derivative is without any reference to the differential of a smooth map between manifolds (hence the definition of V_pM is independent of T_pM). Let $V_pM = \Gamma_p/\sim$ denote the set of equivalence classes of smooth curves with respect to this relation. We will show that V_pM is precisely the tangent space to M at p, but first of all we need to check that it's even a vector space. Unlike the previous two definitions involving derivations, the vector space structure in this case is not obvious.

A natural property that we want the sum $[\alpha] + [\beta]$ to have is that the velocity of the resulting curve be the sum of the velocities $\alpha'(0) + \beta'(0)$, and similarly for scalar multiplication. So that's exactly how we'll define the operations in V_pM: given $\alpha, \beta \in \Gamma_p$ and $c \in \mathbb{R}$, define

- $[\alpha] + [\beta] = [\gamma]$ where $\gamma \in \Gamma_p$ is any smooth curve with $\gamma'(0) = \alpha'(0) + \beta'(0)$.
- $c[\alpha] = [\delta]$ where $\delta \in \Gamma_p$ is any smooth curve with $\delta'(0) = c\alpha'(0)$.

Note that curves γ and δ with these properties can always be constructed by the existence and uniqueness theorem for ordinary differential equations. These operations are also well-defined. Suppose $\gamma_1, \gamma_2 \in \Gamma_p$ satisfy

$$\gamma_1'(0) = \alpha'(0) + \beta'(0) = \gamma_2'(0)$$

then $[\gamma_1] = [\gamma_2]$ because by the chain rule we have

$$(f \circ \gamma_1)'(0) = \gamma_1'(0)f = \gamma_2'(0)f = (f \circ \gamma_2)'(0).$$

for any $f \in C^\infty(M)$. Similarly, if $\delta_1, \delta_2 \in \Gamma_p$ satisfy

$$\delta_1'(0) = c\alpha'(0) = \delta_2'(0)$$

then $[\delta_1] = [\delta_2]$ because

$$(f \circ \delta_1)'(0) = \delta_1'(0)f = \delta_2'(0)f = (f \circ \delta_2)'(0)$$

for any $f \in C^\infty(M)$. Furthermore, it's easy to check that these operations satisfy the vector space axioms, and therefore turn V_pM into a vector space over $\mathbb{R}$. If the space V_pM is going to capture our intuition for the tangent space as velocity vectors of smooth curves, we need to know that every tangent vector at p is tangent to some curve γ at p and that it's uniquely associated with the class $[\gamma] \in V_pM$. This is affirmed by the following theorem:

Theorem 4. *The map $\psi : V_pM \to T_pM$ given by $\psi[\gamma] = \gamma'(0)$ is a linear isomorphism.*

Proof. Let $\alpha, \beta \in \Gamma_p$.

- ψ is well-defined. If $[\alpha] = [\beta]$ then for any $f \in C^\infty(M)$ we have

  $$\alpha'(0)f = (f \circ \alpha)'(0) = (f \circ \beta)'(0) = \beta'(0)f$$

  so $\psi[\alpha] = \alpha'(0) = \beta'(0) = \psi[\beta]$.
- ψ is linear. For $c \in \mathbb{R}$, let $\gamma \in \Gamma_p$ be any smooth curve with $\gamma'(0) = c\alpha'(0) + \beta'(0)$. Then we have

  $$\psi(c[\alpha] + [\beta]) = \psi[\gamma] = \gamma'(0) = c\alpha'(0) + \beta'(0) = c\psi[\alpha] + \psi[\beta].$$

- ψ is injective. This is the same calculation as the one we made earlier using the chain rule. We must have $[\alpha] = [\beta]$ whenever $\alpha'(0) = \beta'(0)$ because the differential df_p of any smooth function agrees on these tangent vectors.
- ψ is surjective. This is the only nontrivial aspect of the proof: we must show that every tangent vector at p is the initial velocity of some smooth curve passing through p. Let $v \in T_pM$, and choose a smooth chart φ :

$U \subseteq M \to \mathbb{R}^n$ around p with $\varphi(p) = 0$. The differential is an isomorphism $d\varphi_p : T_pM \to \mathbb{R}^n_0$. With respect to these coordinates we can write

$$v = \sum_{i=1}^{n} v_i \frac{\partial}{\partial x_i}\bigg|_p = \sum_{i=1}^{n} v_i d\varphi_p^{-1} \left(D_{e_i}\big|_0\right).$$

By the existence and uniqueness theorem for ordinary differential equations we can find a smooth curve $\alpha = (\alpha_1, \ldots, \alpha_n) : I \to \mathbb{R}^n$ which solves the system of equations

$$\begin{cases} \alpha_1'(0) & = v_1 \\ \alpha_2'(0) & = v_2 \\ & \vdots \\ \alpha_n'(0) & = v_n \end{cases}$$

with initial condition $\alpha(0) = 0$ (valid within some small open neighborhoods of $0 \in I$ and $0 \in \mathbb{R}^n$). Then (shrinking U if necessary) we get a smooth curve on M, $\varphi^{-1} \circ \alpha : I \to U \subseteq M$, which satisfies $(\varphi^{-1} \circ \alpha)(0) = p$ and

$$\begin{aligned} (\varphi^{-1} \circ \alpha)'(0) &= d\varphi_p^{-1}(\alpha'(0)) \\ &= d\varphi_p^{-1}(\alpha_1'(0), \ldots, \alpha_n'(0)) \\ &= d\varphi_p^{-1}(v_1, \ldots, v_n) \\ &= \sum_{i=1}^{n} v_i d\varphi_p^{-1} \left(D_{e_i}\big|_0\right) \\ &= \sum_{i=1}^{n} v_i \frac{\partial}{\partial x_i}\bigg|_p \\ &= v \end{aligned}$$

hence $\gamma = \varphi^{-1} \circ \alpha$ is a smooth curve in a neighborhood of p whose initial velocity is v; i.e. $\psi[\gamma] = \gamma'(0) = v$.

■

Thus, if we want to define the tangent space to M at p as the space V_pM, we can say that the tangent vector to a smooth curve $\gamma \in \Gamma_p$ is the equivalence class $[\gamma]$ – then by definition the tangent space at p is the space of all of these tangent vectors. Moreover, it's easy to define the differential of a smooth map $F : M \to N$ using this definition of the tangent space. Given $p \in M$, simply define $\widetilde{d}F_p : V_pM \to V_{F(p)}N$ by setting

$$\widetilde{d}F_p[\gamma] = [F \circ \gamma]$$

for any $[\gamma] \in V_pM$. The benefit of defining the tangent space as V_pM is that the geometric character of the tangent vectors is more obvious, but on the downside it involves more work in verifying the vector space structure of V_pM.

6 Tangent vectors as the dual of cotangent vectors

Usually authors tend to define the tangent space T_pM first and then the cotangent space T_p^*M as the dual of the tangent space, i.e. defining covectors as linear functionals on T_pM. We can also do things the other way around: first define the cotangent space T_p^*M and then define tangent vectors as linear functionals on T_p^*M. Interestingly, this approach seems to reveal the "linear approximation" aspect of the tangent space more clearly compared to the other approaches.

Take a smooth manifold M and fix a point $p \in M$. Define the following subspaces of $C^\infty(M)$:

$$\begin{aligned} I_p &= \{f \in C^\infty(M) : f(p) = 0\} \\ I_p^2 &= \operatorname{span}\{fg : f, g \in I_p\} \end{aligned}$$

and note that $I_p^2 \subseteq I_p \subseteq C^\infty(M)$. As the following fact states, the space I_p^2 can be characterized as the space of functions whose first-order Taylor polynomial vanishes.

Fact 3. *$f \in I_p^2$ if and only if the first-order Taylor polynomial of the coordinate representation of f in any smooth chart around p vanishes. In other words I_p^2 is the space of smooth functions that "vanish at first-order".*

Proof. First we'll prove that any $f \in I_p^2$ vanishes at first-order. We can write $f = \sum c_i(g_i h_i)$ for some $g_i, h_i \in I_p$. Take any smooth chart $\varphi : U \subseteq M \to \mathbb{R}^n$ centered at p, and let $T_1(f)$ denote the Taylor polynomial (centered at $\varphi(p) = 0$) of the coordinate representation of f with respect to this chart. By linearity of the Taylor polynomial it suffices to check that $T_1(gh) = 0$ for any pair $g, h \in I_p$. By definition of T_1, for any $x \in \mathbb{R}^n$ we have

$$\begin{aligned} T_1(gh)(x) &= g(\varphi^{-1}(0))h(\varphi^{-1}(0)) + \nabla(gh \circ \varphi^{-1})(0)^T x \\ &= g(p)h(p) + \left[h(p)\nabla(g \circ \varphi^{-1})(0)^T + g(p)\nabla(h \circ \varphi^{-1})(0)\right]^T x \\ &= 0 \end{aligned}$$

since $g(p) = h(p) = 0$, and hence $T_1(gh) = 0$ as desired.

Conversely, suppose that $f \in C^\infty(M)$ vanishes at first-order, we need to show that $f \in I_p^2$. By assumption, for any smooth chart $\varphi : U \subseteq M \to \mathbb{R}^n$ centered at p, the first-order Taylor polynomial of $\widetilde{f} = f \circ \varphi^{-1}$ (centered at 0) vanishes. Hence

$$0 = T_1(f)(x) = f(p) + \nabla \widetilde{f}(0)^T x$$

which implies that $f(p) = 0$ and $\nabla \widetilde{f}(0) = 0$. This latter condition on the gradient can be expanded to see that

$$\begin{aligned} 0 &= \nabla \widetilde{f}(0) \\ &= \sum_{i=1}^n \frac{\partial \widetilde{f}}{\partial x_i}(0) e_i \\ &= \sum_{i=1}^n df_p\left(d\varphi_0^{-1}\left(D_{e_i}\big|_0\right)\right) e_i \\ &= \sum_{i=1}^n df_p\left(\left.\frac{\partial}{\partial x_i}\right|_p\right) e_i \\ &= \sum_{i=1}^n \frac{\partial f}{\partial x_i}(p) e_i \end{aligned}$$

and hence each partial derivative of f vanishes at p, which means that $df_p = 0$. Now by Taylor's theorem, since all of the first order terms vanish, within this smooth chart we can write

$$\widetilde{f}(x) = \sum_{i,j=1}^{n} \left[x^i x^j \int_0^1 (1-t) \frac{\partial^2 \widetilde{f}}{\partial x_i \partial x_j}(tx)\, dt \right].$$

Let $\widetilde{g}_i = x^i$ and $\widetilde{h}_j = x^j$, and denote by c_{ij} the coefficient of $\widetilde{g}_i \widetilde{h}_j$. With this notation we've expressed $\widetilde{f} = f \circ \varphi^{-1}$ as

$$\widetilde{f}(x) = (f \circ \varphi^{-1})(x) = \sum_{i,j=1}^{n} c_{ij} \widetilde{g}_i(x) \widetilde{h}_j(x)$$

where each $\widetilde{g}_i$ and $\widetilde{h}_j$ vanish at $\varphi(p) = 0$. Therefore

$$\begin{aligned} f(x) &= \sum_{i,j=1}^{n} c_{ij} \widetilde{g}_i(\varphi(x)) \widetilde{h}_j(\varphi(x)) \\ &= \sum_{i,j=1}^{n} c_{ij} g_i(x) h_i(x) \end{aligned}$$

with $g_i(p) = \widetilde{g}_i(\varphi(p)) = \widetilde{g}_i(0) = 0$ and $h_j(p) = \widetilde{h}_j(\varphi(p)) = \widetilde{h}_j(0) = 0$, i.e. $g_i, h_j \in I_p$ for every $1 \le i, j \le n$ – hence $f \in I_p^2$ as desired. ∎

Consider the linear map $\theta : I_p \to T_p^* M$ given by $\theta(f) = df_p$. Due to Fact 3, we know that $f \in \ker \theta$ holds if and only if $f \in I_p$ and $df_p = 0$ both hold; i.e. if and only if f vanishes to first order. Hence $\ker \theta = I_p^2$ and in fact, we will show in the next theorem that θ is surjective, so that θ induces an isomorphism $I_p / I_p^2 \simeq T_p^* M$.

Theorem 5. *The linear map $\theta : I_p \to T_p^* M$ given by $\theta(f) = df_p$ is surjective, and $\ker \theta = I_p^2$. Therefore θ induces a linear isomorphism $I_p / I_p^2 \simeq T_p^* M$.*

Proof. The characterization of the kernel of θ follows directly from Fact 3, and linearity is obvious, so we just need to show that θ is surjective. In other words, we need to show that any linear functional $\xi : T_p M \to \mathbb{R}$ can be realized as the differential of a smooth function $f : M \to \mathbb{R}$, which means that $\xi(\omega) = df_p(\omega) = \omega(f)$ for any $\omega \in T_p M$. This involves a fairly standard argument using bump functions.

Choose any smooth chart $\varphi : U \subseteq M \to \mathbb{R}^n$ centered at p, with corresponding coordinate functions $x^i = \widetilde{x}^i \circ \varphi : U \subseteq M \to \mathbb{R}$. Let $\eta : U \subseteq M \to \mathbb{R}$ be a smooth bump function which is compactly supported in U and which satisfies $\eta = 1$ identically in some neighborhood of p. Then define $f : U \subseteq M \to \mathbb{R}$ by

$$f(x) = \eta(x) \xi \left(\sum_{i=1}^{n} x^i(x) \left. \frac{\partial}{\partial x_i} \right|_p \right) = \eta(x) \sum_{i=1}^{n} x^i(x) \xi \left(\left. \frac{\partial}{\partial x_i} \right|_p \right).$$

Then extend f to all of M by setting $f(x) = 0$ for every $x \in M \setminus \operatorname{supp} \eta$ (note that $f(x) = 0$ already holds for every $x \in U \setminus \operatorname{supp} \eta$ by definition, so this extension makes sense). We claim that $\xi = df_p$. For any $\omega \in T_p M$, write in local coordinates

$\omega = \sum_{i=1}^n \omega_i \left(\partial/\partial x_i\big|_p \right)$ for some $\omega_i \in \mathbb{R}$, then

$$\begin{aligned} df_p(\omega) &= \omega(f) \\ &= \omega\left(\eta \cdot \sum_{i=1}^n x^i \cdot \xi\left(\frac{\partial}{\partial x_i}\bigg|_p \right) \right) \\ &= \sum_{i=1}^n \xi\left(\frac{\partial}{\partial x_i}\bigg|_p \right) \omega(\eta \cdot x^i) \\ &= \sum_{i=1}^n \xi\left(\frac{\partial}{\partial x_i}\bigg|_p \right) \left[\eta(p)\omega(x^i) + x^i(p)\omega(\eta) \right] \text{ (by the product rule)} \\ &= \sum_{i=1}^n \xi\left(\frac{\partial}{\partial x_i}\bigg|_p \right) \omega(x^i) \\ &= \xi(\omega) \end{aligned}$$

where the second-to-last line follows from the fact that η is constant in a neighborhood of p (so ω annihilates it) and $\eta(p) = 1$. Thus we conclude that $df_p = \xi$ and θ is surjective. ■

What is the quotient I_p/I_p^2 anyways? If two cosets in this quotient space coincide, say $f + I_p^2 = g + I_p^2$, then their difference $f - g \in I_p^2$ vanishes to first order. Thus f and g agree at first order, and taking this quotient is the same as modding out by the relation "$f \sim g$ if and only if f and g agree at first order". In particular, this means that any function f is identified with its first-order linear approximation $T_1(f)$, and so the isomorphism θ realizes cotangent vectors as linear approximations of smooth functions.

Motivated by Theorem 5, we could conceivably define the cotangent space as $T_p^*M = I_p/I_p^2$ (independently of T_pM) and then define the tangent space as the dual of the cotangent space, $T_pM = (I_p/I_p^2)^*$. In this case we can first define the cotangent differential of a smooth map $F : M \to N$ at $p \in M$ as the linear map $\widetilde{d}F_p^* : T_{F(p)}^*N \to T_p^*M$ given by

$$\widetilde{d}F_p^*[g] = [g \circ F]$$

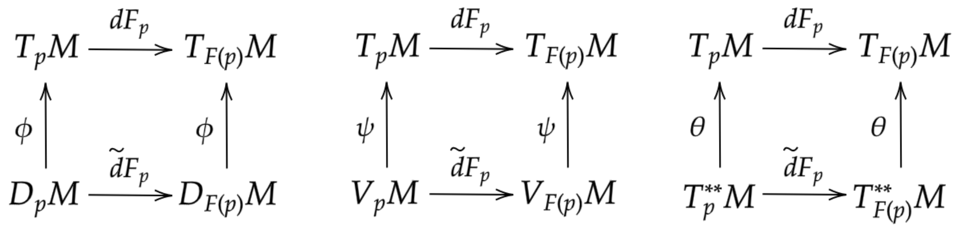
where $g \in I_{F(p)}(N)$ and $[g] = g + I_{F(p)}^2(N)$. Then the differential of F at p is the adjoint of $\widetilde{d}F_p^*$, i.e. the linear map $\widetilde{d}F_p : T_p^{**}M \to T_{F(p)}^{**}N$ given by

$$\widetilde{d}F_p(\omega)[g] = \omega\left(\widetilde{d}F_p^*[g] \right) = \omega[g \circ F]$$

where $\omega : T_p^*M \to \mathbb{R}$ is a linear functional on T_p^*M, $g \in I_{F(p)}(N)$ and $[g] = g + I_{F(p)}^2(N)$.

7 The relationship between differentials

In this section we will explain the relationship between the differential $dF_p : T_pM \to T_{F(p)}N$ (using the definition of tangent space described in Section 3) and the differentials $\widetilde{d}F_p$ (using the alternative definitions of tangent space). Actually, the relationship is very simple and is essentially the same in all three cases. The three situations we discussed above correspond to the three commutative diagrams:



$\widetilde{dF}_p$ acting on derivations of germs — $\widetilde{dF}_p$ acting on equivalence classes of smooth curves — $\widetilde{dF}_p$ acting on linear functionals

Because of this common pattern, all of these definitions are kind of "naturally equivalent" in the sense that the differentials factor through the corresponding isomorphism between any two tangent spaces.

Case 1: $\widetilde{dF}_p$ acting on derivations of germs. Recall that for any derivation of germs $\nu \in D_pM$ and for any smooth function $f \in C^\infty(M)$, we have $\phi(\nu)(f) = \nu[f]_p$. Thus for each $g \in C^\infty(N)$,

$$\begin{aligned} dF_p(\phi(\nu))(g) &= \phi(\nu)(g \circ F) \\ &= \nu([g \circ F]_p) \\ &= \widetilde{dF}_p(\nu)([g]_{F(p)}) \\ &= \phi(\widetilde{dF}_p(\nu))(g) \end{aligned}$$

so $dF_p \circ \phi = \phi \circ \widetilde{dF}_p$ as claimed.

Case 2: $\widetilde{dF}_p$ acting on equivalence classes of smooth curves. Recall that for any $[\gamma] \in V_pM$ we have $\psi[\gamma] = \gamma'(0)$, hence for any $g \in C^\infty(N)$ it follows that

$$\begin{aligned} dF_p(\psi[\gamma])(g) &= dF_p(\gamma'(0)(f) \\ &= \gamma'(0)(f \circ F) \\ &= (f \circ F \circ \gamma)'(0) \end{aligned}$$

and on the other hand,

$$\begin{aligned} \psi(\widetilde{dF}_p[\gamma])(f) &= \psi([F \circ \gamma])(f) \\ &= (F \circ \gamma)'(0)(f) \\ &= (f \circ F \circ \gamma)'(0) \end{aligned}$$

so $dF_p \circ \psi = \psi \circ \widetilde{dF}_p$ as claimed.

Case 3: $\widetilde{dF}_p$ acting on linear functionals. Tedious to write down, but similar to the above two cases.

8 References

In this note we mostly followed John Lee's *Intro to Smooth Manifolds* (pp. 50 - 75), filling in details to several exercises and problems along the way.