LECTURE 5: THE DIFFERENTIAL OF A SMOOTH MAP

1. The tangent space

¶ The idea behind the definition.

Now suppose M, N are smooth manifolds, and $f : M \to N$ smooth. As in the Euclidean case we would like to define its differential df_p at p to be a linear map between corresponding tangent spaces, which serves as a linearization of f near p. But the first question is: what is the tangent space of a smooth manifold at a point?

If M is a concrete manifold that sits in some $\mathbb{R}^N$ (which is always true as we will see in the future), then we may choose a coordinate chart (φ, U, V) near p, so that $\varphi^{-1} : V \to U$ is a diffeomorphism. If we denote the embedding of M into $\mathbb{R}^N$ to be $\iota : M \hookrightarrow \mathbb{R}^N$, then we get a smooth map $\iota \circ \varphi^{-1} : V \to \mathbb{R}^N$ between Euclidean open sets, and we may define the tangent space T_pM to be the image $d(\iota \circ \varphi^{-1})_{\varphi(p)}(\mathbb{R}^n)$ of the differential (which is a linear map)

$$d(\iota \circ \varphi^{-1})_{\varphi(p)} : \mathbb{R}^n \to \mathbb{R}^N.$$

Of course one has to check that the space T_pM defined by this way is independent of the choice of coordinate charts, AND, one need to identify different T_pM's obtained by this way when we use two different embeddings of M into Euclidean spaces.

In what follows we will define the tangent space T_pM using data on M itself only, since we don't know a priori whether M can be embedded into an Euclidean space, and Since M is an abstract manifold, we don't have a simple nice "geometric picture". To get some idea on how to define the tangent space at each point on a smooth manifold, we shall take a closer look at the Euclidean case. The idea is:

- $\forall \vec{v} \in \mathbb{R}^n$ at a given point a can be identified with a directional derivative at a,
- there is a pure algebraic characterization of the directional derivative, which can be generalized to manifolds.

That vector space is the abstract tangent space we are looking for!

¶ The Euclidean directional derivative: an algebraic characterization.

We recall: for any $\vec{v} \in \mathbb{R}^n$, the directional derivative of f at x in the direction v is

$$D^a_{\vec{v}} f := df_x(\vec{v}) = \lim_{h \to 0} \frac{f(x + h\vec{v}) - f(x)}{h} = \left.\frac{d}{dt}\right|_{t=0} f(a + t\vec{v}).$$

So $\vec{v}$ gives us an operator $D^a_{\vec{v}} : C^\infty(\mathbb{R}^n) \to \mathbb{R}$. In coordinates, if $\vec{v} = \langle v^1, \cdots, v^n \rangle^T$, then

$$D^a_{\vec{v}} f = \sum v^i \frac{\partial f}{\partial x^i},$$

in other words, as an operator on $C^\infty(\mathbb{R}^n)$,

$$D^a_{\vec v} = \sum_i v^i \frac{\partial}{\partial x^i}.$$

Of course we know that $D^a_{\vec v} : C^\infty(\mathbb{R}^n) \to \mathbb{R}$ is a very special operator: it is a linear

$$D^a_{\vec v}(\alpha f + \beta g) = \alpha D^a_{\vec v} f + \beta D^a_{\vec v} g, \quad \forall \alpha, \beta \in \mathbb{R}$$

and it satisfies the *Leibnitz law* at a:

$$D^a_{\vec v}(fg) = f(a) D^a_{\vec v} g + g(a) D^a_{\vec v} f.$$

Conversely,

Proposition 1.1. *If $D : C^\infty(\mathbb{R}^n) \to \mathbb{R}$ is linear and satisfies the Leibnitz law a, i.e.*

$$D(fg) = f(a)D(g) + g(a)D(f),$$

then $D = D^a_{\vec v}$ for some vector $\vec v$ at a.

Proof. For any $f \in C^\infty(\mathbb{R}^n)$, we have

$$f(x) = f(a) + \int_0^1 \frac{d}{dt} f(a + t(x-a))dt = f(a) + \sum_{i=1}^n (x^i - a^i) h_i(x),$$

where

$$h_i(x) = \int_0^1 \frac{\partial f}{\partial x^i}(a + t(x-a))dt.$$

Note that by the Leibnitz law, when we apply D to the constant function 1, we must have $D(1) = 0$ since

$$D(1) = D(1 \cdot 1) = 2D(1).$$

By linearity, $D(c) = 0$ for any constant c. So

$$D(f) = 0 + \sum_{i=0}^n D(x^i) h_i(a) + \sum_{i=0}^n (a^i - a^i) D(h_i) = \sum_{i=0}^n D(x^i) \frac{\partial f}{\partial x^i}(a).$$

It follows that as an operator on $C^\infty(\mathbb{R}^n)$,

$$D = \sum_{i=1}^n D(x^i) \left. \frac{\partial}{\partial x^i} \right|_{x=a}.$$

In other words, if we let $\vec v = \langle D(x^1), \cdots, D(x^n) \rangle$, then $D = D^a_{\vec v}$. □

This motivates the following "algebraic" definition of derivatives:

Definition 1.2. Any linear operator $D^a : C^\infty(\mathbb{R}^n) \to \mathbb{R}$ satisfying the Leibnitz law at a is called a *derivative* at a.

It is obvious that the set $\mathcal{D}$ of all derivatives at a is a vector space. Now consider the correspondence

$$\vec v \rightsquigarrow D^a_{\vec v}.$$

We have

- It is a linear map from the vector space $\mathbb{R}^n$ (of all tangent vectors at a) to the vector space $\mathcal{D}$ of all derivatives at a:

$$D^a_{\alpha\vec{v}+\beta\vec{w}} = \alpha D^a_{\vec{v}} + \beta D^a_{\vec{w}}.$$

- It is injective: if $\vec{v}_1 \neq \vec{v}_2$, then $D^a_{\vec{v}_1} \neq D^a_{\vec{v}_2}$ (try to prove it).
- It is surjective: This follows from Proposition 1.1.

So the space of tangent vectors $\vec{v}$ at a is linearly isomorphic to the space of derivatives at a, i.e. we can identify the set (vector space) of all tangent vectors at a with the set (vector space) of all derivatives at a!

¶ Tangent vector on manifolds.

Now back to the study of manifolds. Although we don't have the geometric vectors in the abstract setting, we do have the space $C^\infty(M)$ of all smooth functions. So as in the Euclidean case we may define derivatives at a point and simply call them tangent vectors at that point:

Definition 1.3. Let M be an n-dimensional smooth manifold. A *tangent vector* at a point $p \in M$ is a $\mathbb{R}$-linear map $X_p : C^\infty(M) \to \mathbb{R}$ satisfying the Leibnitz law

$$X_p(fg) = f(p)X_p(g) + X_p(f)g(p) \tag{1}$$

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

It is easy to see that the set of all tangent vectors of M at p is a linear space. We will denote this set by T_pM, and call it the *tangent space* T_pM to M at p..

As argued above, we can easily prove: if $f \equiv c$ is a constant function, then $X_p(f) = 0$. More generally,

Lemma 1.4. *If $f = c$ in a neighborhood of p, then $X_p(f) = 0$.*

Proof. Let φ be a smooth function on M that equals 1 near p, and equals 0 at points where $f \neq c$. (How to construct such an f?) Then

$$(f - c)\varphi \equiv 0.$$

So

$$0 = X_p((f - c)\varphi) = (f(p) - c)X_p(\varphi) + X_p(f)\varphi(p) = X_p(f).$$ □

As a consequence, we see that if $f = g$ in a neighborhood of p[1], then $X_p(f) = X_p(g)$. In other words, $X_p(f)$ is determined by the values of f near p. So one can replace $C^\infty(M)$ in Definition 1.3 by $C^\infty(U)$, where U is any open set that contains p. In other words, as linear spaces,

Proposition 1.5. *If M is a smooth manifold, $p \in U \subset M$, where U is open, then*

$$T_pM \simeq T_pU.$$

[1]There is a notion "germ": we say f and g define the same germ at p if $f = g$ in a neighborhood of p. It is easy to see that "germ at p" gives an equivalence relation on $C^\infty(M)$ (or more generally on $C^\infty(M, N)$). When studying local properties, it is always enough to work on germs.

2. The Differentials

¶ The differential of smooth maps between smooth manifolds.

Finally we try to define the differential of a smooth map between smooth manifolds. Recall that the differential of a smooth map $f : U \to V$ between open sets in Euclidean spaces at $a \in U$ is a linear map $df_a : T_aU = \mathbb{R}^n_x \to T_{f(a)}V = \mathbb{R}^m_y$ whose matrix is the Jacobian matrix $(\frac{\partial f_i}{\partial x^j})$ of f at a. To transplant this conception to smooth maps between smooth manifolds, we need to take a closer look at the two interpretation of T_aU: We have seen that we can identify the (geometric) vector $\vec{v} \in \mathbb{R}^n$ at a with the (algebraic) derivative $D^a_{\vec{v}} = \sum v^i \frac{\partial}{\partial x^i}$. Note that geometrically,

$$df_a(\vec{v}) = \left(\frac{\partial f_i}{\partial x^j}\right)\vec{v} = \left\langle \sum_j \frac{\partial f_1}{\partial x^j}v^j, \cdots, \sum_j \frac{\partial f_n}{\partial x^j}v^j \right\rangle^T.$$

The vector in the right hand side is a vector in $\mathbb{R}^m_y$. When interpreted as a derivative on V at $f(a)$, it is a map that maps $g \in C^\infty(\mathbb{R}^m_y)$ to

$$\sum_i \sum_j v^j \frac{\partial f_i}{\partial x^j}\frac{\partial g}{\partial y^i} = \sum_j v^j \frac{\partial}{\partial x^j}(g \circ f) = D^a_{\vec{v}}(g \circ f).$$

In other words, the derivative that corresponds to the vector $df_a(\vec{v})$ is the derivative at $f(a)$ that maps $g \in C^\infty(\mathbb{R}^m)$ to $D^a_{\vec{v}}(g \circ f)$.

Motivated by these computations, we define

Definition 2.1. Let $f : M \to N$ be a smooth map. Then for each $p \in M$, the *differential* of f is the linear map $df_p : T_pM \to T_{f(p)}N$ defined by

$$df_p(X_p)(g) = X_p(g \circ f)$$

for all $X_p \in T_pM$ and $g \in C^\infty(N)$.

¶ Properties of the differential.

The chain rule still holds for composition of smooth maps between smooth manifolds:

Theorem 2.2 (Chain rule)**.** *Suppose $f : M \to N$ and $g : N \to P$ are smooth maps, then $d(g \circ f)_p = dg_{f(p)} \circ df_p$.*

Proof. For any $X_p \in T_pM$ and $h \in C^\infty(P)$,

$$d(g \circ f)_p(X_p)(h) = X_p(h \circ g \circ f) = df_p(X_p)(h \circ g) = dg_{f(p)}(df_p(X_p))(h).$$ □

Obviously the differential of the identity map is the identity map between tangent spaces. So again the differential d is a functor from the category of smooth manifolds (with morphisms smooth maps) to the category of vector spaces (with morphisms linear maps). By the standard functorality argument (c.f. the proof of Theorem 2.1 in Lecture 4), we get

Corollary 2.3. *If $f : M \to N$ is a diffeomorphism, then $df_p : T_pM \to T_{f(p)}N$ is a linear isomorphism.*

In particular, we have

Corollary 2.4. *If* $\dim M = n$*, then* T_pM *is an* n*-dimensional linear space.*

Proof. Let (φ, U, V) be a chart near p. Then $\varphi : U \to V$ is a diffeomorphism. It follows that $\dim T_pM = \dim T_pU = \dim T_{f(p)}V = n$. □

Note that in the proof we really showed: For any local chart (φ, U, V) around p, we have

$$T_pM = \mathrm{span}\{\partial_1, \cdots, \partial_n\},$$

where $\partial_i := d\varphi^{-1}(\frac{\partial}{\partial x^i})$. We will abuse the notation and think of x^i as a function on U (which really should be $x^i \circ \varphi$). In other words, we will denote $\varphi = (x^1, \cdots, x^n)$ with each x^k the k^{th} coordinate function on U, and denote the chart by $\{U; x^1, \cdots, x^n\}$. In such coordinate charts, one has the following explicit formula for ∂_i:

$$\partial_i : C^\infty(U) \to \mathbb{R}, \quad \partial_i(f) = \frac{\partial(f \circ \varphi^{-1})}{\partial x^i}(\varphi(p)).$$