1. Basic algebra of vector fields

Let V be a finite dimensional vector space over $\mathbb{R}$. Recall that $V^* = \{L : V \to \mathbb{R}\}$ is defined to be the set of all linear maps to $\mathbb{R}$. V^* is isomorphic to V, but there is no canonical isomorphism.

An *inner product* on V is a map $v, w \mapsto \langle v, w\rangle \in \mathbb{R}$ which is *bilinear, symmetric*, and *positive definite*. An inner product always exists because a basis always exists. An inner product defines an isomorphism $V \to V^*$ by $v \mapsto \langle v, \cdot\rangle$. Notice that, with respect to some choice of basis on V, the inner product can be defined by a positive definite symmetric matrix $\langle v, w\rangle = v^T A w$.

We would like to study these structures on manifolds. First notice that one can choose an inner product $\langle \cdot, \cdot\rangle_p$ on each vector space TM_p. In a local chart $M \supseteq U \cong \mathbb{R}^n$, $\langle \cdot, \cdot\rangle_p$ looks like a matrix of functions (of p) which is positive definite and symmetric. If these are smooth functions of p, we say that $\langle \cdot, \cdot\rangle$ is smooth. A smooth inner product on TM is also called a *Riemannian metric*, or just *metric* for short.

Every manifold has a metric. To see this, notice that the restriction of a metric to a submanifold is a metric on that submanifold. Since $\mathbb{R}^N$ has a metric and every manifold embeds in $\mathbb{R}^N$ this completes the claim. Without appealing to the Whitney embedding theorem, we can also show that an abstract manifold has a metric by using a partition of unity, noting that the set of metrics is a convex subset in the space of all bilinear functions.

Define $T^*M_p = (TM_p)^*$, and $T^*M = \cup_{p \in M} T^*M_p$. Then T^*M is defined to be a set which is a union of vector spaces, but we can also define it to be a smooth manifold as follows. If $\langle \cdot, \cdot\rangle$ is a metric on M, then it induces a bijection $TM \to T^*M$ (which is a linear isomorphism on each $TM_p \to T^*M_p$). We define a smooth structure on TM by defining this map to be a diffeomorphism. One easily checks that any two choices of (smooth!) metric define diffeomorphic smooth structures on T^*M. We call T^*M the *cotangent bundle* of M, and smooth sections of this bundle $\sigma : M \to T^*M$ (satisfying $\sigma(x) \in T^*M_x$) are called either *covector fields* or 1-*forms*. We denote the space of all 1-forms by $\Omega^1(M)$.

Suppose we are given a smooth function $f : M \to \mathbb{R}$. The derivative is a map $df : TM \to T\mathbb{R}$. Notice that $T\mathbb{R}_x$ is *canonically* isomorphic to $\mathbb{R}$ (here it is important that the codomain of f is $\mathbb{R}$, not simply a 1-manifold diffeomorphic to $\mathbb{R}$). Thus the map $v \mapsto df_p(v) \in T\mathbb{R}_x \cong \mathbb{R}$ is a linear function $TM_p \to \mathbb{R}$, and we can therefore think of df as a covector field.

If V is a vector field, then we can evaluate $df(V)$ which is a smooth function on M. We also write $df(V) = V(f)$ if we wish to think of V as an operator on smooth functions f, rather than evaluation of df on vector fields V. In this way, we can think of V as a map $V : C^\infty(M) \to C^\infty(M)$ (here $C^\infty(M)$ is the set of all smooth functions from M to $\mathbb{R}$). This is simply the familiar directional derivative from vector calculus. Notice that $V(f)(p)$ only depends on the vector $V_p \in TM_p$, but it depends on f in a *neighborhood* of p. Phrased differently, this means that $V(f)$ is *tensorial* in V but not in f.

Notice that the function $V : C^\infty(M) \to C^\infty(M)$ is linear and also Leibnitzian, that is, it satisfies the equality $V(fg) = fV(g) + gV(f)$. In fact, these properties characterize vector fields.

Proposition 1.1. *Let $F : C^\infty(M) \to C^\infty(M)$ be any function which is linear, Leibnitzian, and continuous. Then F is induced by a unique vector field V.*

Proof: First, notice that if F and G are linear and Leibnitzian, and $\lambda \in C^\infty(M)$, then $F+G$ and λF are also linear and Leibnitzian. Therefore we can work locally: if every function $F: C^\infty(\mathbb{R}^n) \to C^\infty(\mathbb{R}^n)$ is induced by a vector field on $\mathbb{R}^n$, then for an arbitrary manifold M we can use a partition of unity to decompose F into a function which is zero outside of one chart.

Let $f_i : \mathbb{R}^n \to \mathbb{R}$ be defined by $F(x_i)$ (where $(x_1, \ldots, x_n)$ are coordinates on $\mathbb{R}^n$), and let $V = f_1 V_1 + \ldots + f_n V_n$, where $V_1, \ldots, V_n$ are the standard basis vectors on $\mathbb{R}^n$. By definition F and V are the same map on the functions x_i, and since they are both linear and Leibnitzian it follows that they are equal on all polynomials. But every function can be approximated arbitrarily closely by polynomials (in the C^∞ topology on compact subsets), therefore since F and V are both continuous functions on $C^\infty(\mathbb{R}^n)$ it follows that they are equal everywhere.

For uniqueness, suppose that V and W are two vector fields inducing the same operation $C^\infty(M) \to C^\infty(M)$. Then the vector field $V - W$ induces the zero operation. Suppose that $V_p - W_p \neq 0$ at some point $p \in M$. Then we can choose coordinates so that $V_p - W_p = V_1$ are the origin, but then the function x_1 (cut off to be zero outside of this chart) cannot be sent to zero. $\square$

If we choose a metric on M, the diffeomorphism $TM \to T^*M$ shows that every smooth vector field corresponds to a unique 1-form, and vice versa. Nevertheless, you should not think of vector fields and 1-forms as being essentially the same thing! Since this correspondence depends on a choice of metric, and a diffeomorphism will only very rarely preserve a metric, the properties of vector fields up to diffeomorphism are quite different than the properties of 1-forms.

For example, a vector field generates an isotopy by taking the flow of the vector field, but nothing like this is true for 1-forms. In particular, this implies that a non-vanishing vector field is locally standard: any non-vanishing vector field can be given as $\frac{\partial}{\partial x_1}$ for some choice of local coordinate system (prove this!). On the other hand, non-vanishing covector fields do have local invariants! Some non-vanishing covector fields can be locally described as the differential of a function, and some cannot. If it can, we say that it is *closed* or sometimes *conservative*.

To make things quite clear, on $\mathbb{R}^2$ the 1-form dx is not locally diffeomorphic to the 1-form $(y^2+1)dx$, because the latter cannot be written as df for any function $f \in C^\infty(\mathbb{R}^2)$, even locally (why?). However, using the standard metric, these 1-forms are respectively dual to the vector fields $\frac{\partial}{\partial x}$ and $(y^2+1)\frac{\partial}{\partial x}$, and the diffeomorphism $(x, y) \mapsto ((y^2+1)x, y)$ maps the former to the latter (globally even).

If a metric is chosen on a manifold M, we define the *gradient* of $f : M \to \mathbb{R}$ to be the vector field ∇f which is dual to df, that is, $\langle \nabla f, v\rangle = df(v)$ for all vectors v.

Lie bracket and Hessian

Definition 1.2. Let X and Y be vector fields on a manifold M, and let $f : M \to \mathbb{R}$ be a smooth function. We define $[X, Y](f) = X(Y(f)) - Y(X(f))$. This defines $[X, Y]$ as a linear Leibnitzian operator on $C^\infty(M)$, therefore $[X, Y]$ is a vector field on M, called the *Lie bracket* of X and Y.

Note that $[X, Y]$ is Leibnitzian by simply calculation:

$$\begin{aligned}[X,Y](fg) &= X(Y(fg)) - Y(X(fg)) \\ &= X(fY(g) + gY(f)) - Y(fX(g) + gX(f)) \\ &= fX(Y(g)) + X(f)Y(g) + gX(Y(f)) + X(g)Y(f)\end{aligned}$$

$$-fY(X(g)) - Y(f)X(g) - gY(X(f)) - Y(g)X(f))$$
$$= fX(Y(g)) + gX(Y(f)) - fY(X(g)) - gY(X(f))$$
$$= f[X,Y](g) + g[X,Y](f).$$

The Lie bracket should be thought of as an obstruction for a vector field to come from a coordinate system. If $M \supseteq U \cong \mathbb{R}^n$ is a chart, the coordinates of the diffeomorphism can be thought of as n functions $x_i : U \to \mathbb{R}$, $i = 1, \dots, n$. This gives a smooth basis of covectors dx_i, which are simply the composition of the differential of the diffeomorphism $U \to \mathbb{R}^n$ composed with the differential with the coordinate projections. This as defines basis vectors $\frac{\partial}{\partial x_i}$ as the dual basis to dx_i (which are the standard basis vectors in $\mathbb{R}^n$, pushed forward to U by the inverse of the diffeomorphism). Then $\left[\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}\right] = 0$ for any i and j, because partial derivatives commute on $\mathbb{R}^n$.

Much more interesting is the converse of this statement, which is called the Frobenius integrability theorem: any basis of vector fields which have zero Lie brackets can be realized as a local coordinate system. The key point in the proof is, assuming $[X,Y] = 0$, then $\varphi_t^X \circ \varphi_s^Y = \varphi_s^Y \circ \varphi_t^X$ for all $t, s \in \mathbb{R}$ (where φ denotes the flow of a vector field). Try proving both the lemma and the integrability theorem on your own!

The Lie bracket is also Leibnitzian with respect to scalar multiplication:

$$[X, fY] = X(f)Y + f[X,Y].$$

Then, as an example, let $X = \overrightarrow{j}$ and $Y = \overrightarrow{i} + y\overrightarrow{k}$ be vector fields on $\mathbb{R}^3$. Then $[X,Y] = \overrightarrow{k}$. If f and g are two smooth functions, then $[fX, gY] = fX(g)Y + g[fX,Y] = fX(g)Y + gY(f)X + fg[X,Y]$. Since X, Y and $[X,Y]$ are linear independent vectors at every point, we see that $[fX, gY] \neq 0$ for any choice of f and g. If follows that there is no surface in $\mathbb{R}^3$ which is tangent to the span of X and Y, even locally. Indeed, any surface can be described by $F(x,y,z) = 0$ If (u,v) are coordinates on the surface, then (u,v,F) are coordinates on $\mathbb{R}^3$. So $\left[\frac{\partial}{\partial u}, \frac{\partial}{\partial v}\right] = 0$, but is the surface is tangent to the span of X and Y then $\frac{\partial}{\partial u} = f_1X + g_1Y$ and $\frac{\partial}{\partial v} = f_2X + g_2Y$ for some smooth functions, which is a contradiction.

(This uses the language of Lie bracket, but the same proof can be expressed in much simpler notation. The vector $\overrightarrow{k}$ is never in the span of X and Y, so any surface which is tangent to X and Y can be expressed as a graph $\{z = F(x,y)\}$. The fact that this surface is tangent to X says that $\frac{\partial F}{\partial y} = 0$, and it is tangent to Y if and only if $\frac{\partial F}{\partial x} = y$. But then $0 = \frac{\partial^2 F}{\partial x \partial y} = \frac{\partial^2 F}{\partial y \partial x} = 1$.)

Proposition/Definition 1.3. Let $f : M \to \mathbb{R}$ be a smooth function, and let $p \in M$ be a critical point. Given $v, w \in TM_p$, we define $H_p(v,w) \in \mathbb{R}$ by extending v and w to vector fields $\widetilde{v}$ and $\widetilde{w}$ in a neighborhood of p, and then defining $H_p(v,w) = \widetilde{v}(\widetilde{w}(f))(p)$. Then H_p is a well defined, bilinear, symmetric function on TM_p, which we call the *Hessian* of f at p.

Proof: First, note that $H_p(v,w)$ does not depend on $\widetilde{v}$, since as we noted above the derivative of a function with respect to a vector field is tensorial in the vector field. Therefore, since H_p is well-defined and linear in the first slot, we only need to show that it is symmetric. By definition, $H_p(v,w) - H_p(w,v) = [\widetilde{v}, \widetilde{w}](f)(p)$. But since $[\widetilde{v}, \widetilde{w}]$ is a vector field and p is a critical point, $[\widetilde{v}, \widetilde{w}](f)(p) = df_p([\widetilde{v}, \widetilde{w}]_p) = 0$. □

Note. In order to show that H_p is well defined, we had to use the fact that p is a critical point. This isn't laziness, in fact we cannot take the second derivative of a smooth function at a non-critical point in any canonical way, without appealing to extra structure (such as a metric).

Definition 1.4. Let $f : M \to \mathbb{R}$ and let $p \in M$ be a critical point. We say that p is *non-degenerate* if H_p has maximal rank. We say that the *index* of a critical point is equal to the maximal dimension of a subspace of TM_p so that H_p is negative definite. We denote this index by $\mathrm{ind}(p) \in \{0, \dots, n\}$.

Remark 1.5. Since H_p is a bilinear functional on TM_p, we can also think of it as a map $H_p : TM_p \to T^*M_p$. If we also have a metric on M then we have an isomorphism between T^*M_p and TM_p, and therefore we can think of H_p as a self-adjoint linear operator. In this case we can decompose H_p into eigenspaces, and the index of p is equal to the number of negative eigenvalues of H_p. But without a metric, neither the eigenspaces or even the eigenvalues are well-defined.

Critical points and topology

The main theme of Morse theory is that we can study topology by looking at critical points of functions on manifolds. The next proposition is the simplest form of this idea: an absence of critical points implies the absence of topology.

Proposition 1.6. *Let M be a compact manifold, and let $f : M \to [0, 1]$ be a function with no critical points, so that $f^{-1}(0) \amalg f^{-1}(1) = \partial M$. Then M is diffeomorphic to $f^{-1}(0) \times [0, 1]$ (in particular, $f^{-1}(1) \cong f^{-1}(0)$).*

Proof: Let $N = f^{-1}(0)$. Choose an arbitrary metric on M, and let $V = \frac{\nabla f}{||\nabla f||^2}$. (Thus by definition $V(f) = 1$ everywhere. Of course $||w||^2 = \langle w, w\rangle$ by definition, and $||\nabla f||^2 \neq 0$ anywhere since f has no critical points.) For $x \in M$, let $\varphi_t(x)$ denote the flow of V acting on the point x for time t.

We study when this flow is defined. If $\varphi_{t_0}(x) \in M \setminus \partial M$ then there is some ε so that $\varphi_t(x)$ is defined for all $t \in [t_0, t_0 + \varepsilon)$, and since M is compact it is also defined for $t = t_0 + \varepsilon$ as well. Thus we see that, for every point $x \in M$ $\varphi_t(x)$ is defined for all $t \in [0, t_x]$, where $\varphi_{t_x}(x) \in \partial M$. Of course, if V points into the manifold at $\varphi_{t_x}(x)$ then we can extend the time domain of definition as well.

Since V points inward at N, we can restrict the flow to N and define $\varphi_t : N \to M$ for some interval $t \in [0, \varepsilon]$ with $\varepsilon > 0$. But for any fixed $x \in N$, the function $t \mapsto (f \circ \varphi_t)(x)$ is a function $[0, \varepsilon] \to [0, 1]$ satisfying $(f \circ \varphi_0)(x) = 0$ and $\frac{d}{dt}(f \circ \varphi_0)(x) = 1$ (by definition of a flow). Therefore we can take $\varepsilon = 1$. We claim that $\Phi : N \times [0, 1] \to M$ defined by $\Phi(x, t) = \varphi_t(x)$ is a diffeomorphism.

$d\Phi$ is surjective: $d\Phi_{(x,t)}|_{TN} = (d\varphi_t)_x$ whose image is $\ker df$, and $df(d\Phi(\frac{\partial}{\partial t})) = 1$ by construction. Therefore Φ is a local diffeomorphism. Φ is also injective: if $\Phi(x, t) = \Phi(y, s)$ then $t = s$ since $f(\Phi(x, t)) = t$, and $\varphi_t : N \to M$ is injective since it is a flow. Since $N \times [0, 1]$ is compact it follows that Φ is a closed map. Also Φ is an open map, since it is a local diffeomorphism and $\Phi(\partial(N \times [0, 1]) \subseteq \partial M$. Together, we see that Φ is a diffeomorphism onto its image, which is both open and closed.

It therefore suffices to show that every connected component of M intersects N. By compactness every component must contain a point m which is a minimum of f. m cannot be a interior point or else it would be a critical point, but also $f(m)$

cannot be 1 or else f would be constant on this component. Therefore $f(m) = 0$. □

If we have a manifold M whose topology we would like to understand better, one thing we might do is just choose a random function $f : M \to \mathbb{R}$. In some sense the above proposition tells us that all of the interesting topology of M is contained in neighborhoods of the critical points. In general, the topology near a critical point can be arbitrarily complicated (for example, the topology near the critical points of a constant map is all topology). But in the non-degenerate case, we can understand the picture well.

Proposition 1.7. *Let $f : M \to \mathbb{R}$, and let $p \in M$ be a non-degenerate critical point. Then there is a neighborhood U of p an coordinates $\varphi : U \to \mathbb{R}^n$ so that $\varphi(p) = 0$ and $(f \circ \varphi)(x) = -x_1^2 - \ldots - x_k^2 + x_{k+1}^2 + \ldots + x_n^2 + f(p)$. (It follows that there are no critical points in a neighborhood of p, and $k = \operatorname{ind}(p)$).*

This is Lemma 2.2 in Milnor, so we don't repeat the proof here.

Homotopy Type

Definition 1.8. Let X and Y be topological spaces. We say that X and Y are *homotopy equivalent* if there are continuous maps $f : X \to Y$ and $g : Y \to X$ so that $f \circ g$ is homotopic to the identity maps on Y, and $g \circ f$ is homotopic to the identity map on X. We say that f and g are *homotopy equivalences*, and also that X and Y have the same *homotopy type*.

Homotopy equivalence is an equivalence relation of topological spaces, a strictly weaker equivalence relation than homeomorphism. For example, The ball B^n is homotopy equivalent to a point, and the cylinder $S^1 \times [0,1]$ is homotopy equivalent to S^1. Even restricting to smooth manifolds of a fixed dimension, homotopy equivalence does not guarantee homeomorphism: for example, $S^2 \setminus \{3 \text{ open balls}\}$ is homotopy equivalent to $T^2 \setminus \{1 \text{ open ball}\}$, since both are homotopy equivalent to the figure 8 curve. But they cannot be homeomorphic because the boundary of the first is disconnected.

Even if we restrict ourselves to closed manifolds of a fixed dimension, there are manifolds which are homotopy equivalent but not homeomorphic. Also, as an aside, there are smooth manifolds which are homeomorphic but not diffeomorphic.

Definition 1.9. Let Y be a topological space and $X \subseteq Y$ a subspace. We say that X is a *deformation retract* of Y if there is a homotopy of maps $f_t : Y \to Y$, so that $f_0 = \mathrm{id}$, $f_1(Y) \subseteq X$, and $f_t(x) = x$ for all $x \in X$ and $t \in [0,1]$.

In particular, any deformation retraction is a homotopy equivalence, whose homotopy inverse is the inclusion map. The main theorem of Morse theory describes the homotopy type of a non-degenerate critical point.

Theorem 1.10. *Let $f : M \to [0,1]$ be a smooth function with exactly one critical point $p \in M$, so that $\partial M = f^{-1}(0) \amalg f^{-1}(1)$, and $f(p) \in (0,1)$. Then for $k = \operatorname{ind}(p)$, there is a disk $D^k \subseteq M$ with $\partial D^k \subseteq f^{-1}(0)$, so that $f^{-1}(0) \cup D^k$ is a deformation retract of M.*