LECTURE 4: THE LINEAR CONNECTION

1. Linear connections

Let M be any smooth manifold (so no Riemannian structure is assumed). As usual we will use denote by $\Gamma(TM)$ the set of all smooth sections of TM, i.e. the set of all smooth vector fields on M. For any $X, Y \in \Gamma(TM)$ we would like to study the "directional derivative" of Y along the direction of X.

To illustrate, let's start with the Euclidian case. Let $X = X^i\partial_i$ and $Y = Y^j\partial_j$ be two smooth vector fields on $\mathbb{R}^m$. Then at a point x the directional derivative of Y along X is the limit

$$D_X Y = \lim_{t\to 0} \frac{Y(x+tX) - Y(x)}{t}, \tag{1}$$

where both x and X are viewed as vectors in $\mathbb{R}^m$. A simple computation yields

$$D_X Y = X^i\partial_i(Y^j)\partial_j.$$

It is easy to see that for any smooth function f, one has

$$D_{fX}Y = fD_XY \quad \text{and} \quad D_X(fY) = fD_XY + (Xf)Y. \tag{2}$$

Note that these two properties are both natural: X represents the direction that we want to take "derivative", so one should have pointwise linearity; Y represents the vector field to be differentiated, so one should have a Leibniz law. [Note: The Lie derivative is NOT a directional derivative since it does not satisfy the first equation.]

On a smooth manifold M, to define the "directional derivative" of a vector field Y along X, our first candidate is the Lie derivative

$$\mathcal{L}_X Y = [X, Y].$$

Unfortunately it does not satisfy the first equation in (2). So we can't use it as a directional derivative. A second candidate is to use formula (1). Here the first problem we will face is that $x+tX$ is no longer a point on the manifold, so $Y(x+tX)$ makes no sense at all. This problem is easy to solve, for example, one can replace $x + tX$ by a curve $\gamma(t)$ on M with $\gamma(0) = x$, so that the tangent vector of γ at $\gamma(t)$ is the vector $X(\gamma(t))$. Then we consider the limit

$$\lim_{t\to 0} \frac{Y(\gamma(t)) - Y(x)}{t}.$$

However, even if one did this, still the expression above does not make sense, since the vectors $Y(\gamma(t))$ and $Y(x)$ belong to different vector spaces, and thus one cann't take the difference. [So, to define a directional derivative, one need a new structure on M to identify the tangent spaces at different points. We will explore the meaning of this sentence in more detail later.]

Another possible way to solve this problem is to embed M into a large Euclidean space $\mathbb{R}^N$. For any $X, Y \in \Gamma(TM)$, one can also extend them to vector fields $\widetilde{X}, \widetilde{Y}$ on $\mathbb{R}^N$. Then one define $D_X Y$ to be some kind of "projection" of $D_{\widetilde{X}}\widetilde{Y}$ onto the tangent space of M. It turns out that for all reasonable definitions as this, such a "directional derivative" satisfies (2). [Of course one has to explain what do we mean by "projection". And also the non-uniqueness of embedding tells us the non-uniqueness of "directional derivatives".]

In conclusion,to differential a vector field Y along a vector field X, one need a new structure on M. And the equations (2) are the guiding rule for this new structure.

Definition 1.1. A *linear connection* ∇ on M is a bilinear map

$$\nabla : \Gamma(TM) \times \Gamma(TM) \to \Gamma(TM), \quad (X, Y) \mapsto \nabla_X Y$$

such that for any $X, Y \in \Gamma(TM)$ and any $f \in C^\infty(M)$,

- $\nabla_{fX} Y = f\nabla_X Y$,
- $\nabla_X(fY) = f\nabla_X Y + (Xf)Y.$

The vector field $\nabla_X Y$ is called the *covariant derivative* of Y along X (with respect to the connection ∇).

Example. Let $M = \mathbb{R}^m$. Then the usual directional derivative

$$\nabla_X Y = \nabla_{X^i\partial_i}(Y^j\partial_j) = X^i\partial_i(Y^j)\partial_j.$$

give rise to a linear connection. More generally, for any choice of m^3 smooth functions γ_{ij}^k, one can check that

$$\nabla_X Y = X^i\partial_i(Y^j)\partial_j + X^iY^j\gamma_{ij}^k\partial_k.$$

defines a linear connection on $\mathbb{R}^m$.

In fact, if we let $(U, x^1, \cdots, x^m)$ be a coordinate chart, then since $\nabla_{\partial_i}\partial_j$ is a smooth vector field on U (the meaning of this "locally defined connection" will be explained in a while, see locality I below), there exists functions Γ_{ij}^k on U such that

$$\nabla_{\partial_i}\partial_j = \Gamma_{ij}^k\partial_k.$$

Definition 1.2. We will call the functions Γ_{ij}^k the *Christoffel symbols* of ∇ with respect to the given chart.

Obviously the functions γ_{ij}^k's are exactly the Christoffel symbols. In particular for the canonical linear connection on $\mathbb{R}^m$, the Christophel symbols are all zero. I will leave it as an exercise to write down the coordinate change formula for Chirstophel symbols.

Remark. One can regard a linear connection as a map

$$\nabla : \Gamma(TM) \to \Gamma(TM) \otimes \Gamma(T^*M)$$

so that

$$\nabla Y(X) := \nabla_X Y.$$

[More precisely, this means $\nabla Y(\omega, X) := \omega(\nabla_X Y)$.]

Remark. More generally for any vector bundle E over M, one can define a linear connection to be a bilinear map

$$\nabla : \Gamma(TM) \times \Gamma(E) \to \Gamma(E), \quad (X, Y) \mapsto \nabla_X Y$$

that satisfies the two properties as above, i.e.

$$\nabla_{fX} s = f\nabla_X s, \quad \text{and} \quad \nabla_X(fs) = f\nabla_X s + (Xf)s.$$

As in the case of Riemannian metrics, one has

Theorem 1.3. *There exists (plenty of) (nontrivial) connections on M.*

Sketch of proof. Of course 0 is a trivial connection. A semi-trivial connection can be constructed as follows: fix a coordinate patch U, then define a *local* linear connection ∇^U on U using the directional derivatives on $\mathbb{R}^m$ above. This can be extend *trivially* to a "not-so-interesting" linear connection on M. To get more complicated "non-trivial" connections on M, one can use the partition of unity trick to glue these not-so-interesting linear connections together.

The most interesting connection on M is constructed as follows: On any Riemannian manifold (M, g) we will construct explicitly a unique linear connection called the *Levi-Civita connection*, which is torsion free and is compatible with the Riemannian structure. Since there exist plenty of Riemannian metrics on M, there must be plenty of linear connections on M. $\square$

Remark. The set of all linear connections on M is not a linear space. However, it form a convex set. Namely, if $\nabla^{(1)}, \cdots, \nabla^{(k)}$ are linear connections on M, and $f_1, \cdots, f_k$ smooth functions on M such that $f_1 + \cdots + f_k = 1$. Then the sum $f_i\nabla^{(i)}$ is also a linear connection on M. [Check this!][This fact is used in the partition of unity argument above!]

2. Localities

Now let ∇ be a linear connection on a smooth manifold M. We shall prove that $\nabla_X Y$ depends only on local information of X and Y.

Proposition 2.1 (Locality I)**.** *For any open subset $U \subset M$, if $X|_U = \tilde{X}|_U$ and $Y|_U = \tilde{Y}|_U$, then $\nabla_X Y|_U = \nabla_{\tilde{X}} \tilde{Y}|_U$.*

Proof. It is enough to prove that if either $X = 0$ in U or if $Y = 0$ in U, then $\nabla_X Y = 0$ in U. To show this, for any point $p \in U$, we choose an open set V such that $p \in V \subset U$. We then choose a function $f \in C_0^\infty(U)$ such that $f = 1$ on V. Then either $X = (1-f)X$ or $Y = (1-f)Y$. It follows in the first case,

$$\nabla_X Y(p) = (1 - f(p))\nabla_X Y(p) = 0,$$

or in the second case,

$$\nabla_X Y(p) = X(1-f)(p)Y(p) + (1-f(p))\nabla_X Y(p) = 0,$$

which proves the assertion. □

Proposition 2.2 (Locality Ⅱ)**.** *If $X(p) = \tilde{X}(p)$, then $\nabla_X Y(p) = \nabla_{\tilde{X}} Y(p)$.*

Proof. It is enough to show that if $X(p) = 0$, then $\nabla_X Y(p) = 0$. By locality I above, we only need to prove this for X supported in an coordinate neighborhood U, with $x(p) = 0$ the origin of V. Now $X = X^i \partial_i$ and $X^i(0) = 0$ for $1 \le i \le m$. So by Taylor's theorem, there exists functions X^i_k such that $X^i = x^k X^i_k$ on U. So

$$\nabla_X Y(p) = \nabla_{x^k X^i_k \partial_i} Y(p) = x^k(p) \nabla_{X^i_k \partial_i} Y(p) = 0.$$

The proposition follows. □

As a consequence, for any vector $v \in T_m M$ and any vector field $Y \in \Gamma(TM)$, one can define $\nabla_v Y(p)$, the "directional derivative" of Y at p along the direction v, to be the vector $\nabla_X Y(p)$, where X is any vector field such that $X(p) = v$.

On the other hand side, it is not too hard to construct vector fields $X, Y, \overline{Y}$ such that $Y(p) = \overline{Y}(p)$ but $\nabla_X Y(p) \neq \nabla_X \overline{Y}(p)$. However, we have

Proposition 2.3 (Locality Ⅲ)**.** *Let $\gamma : (-\varepsilon, \varepsilon) \to M$ be a smooth curve on M with*

$$\gamma(0) = p \qquad \text{and} \qquad \dot{\gamma}(0) = v.$$

Suppose $X, Y, \overline{Y}$ are vector fields on M such that $X(p) = v$ and

$$Y(\gamma(t)) = \overline{Y}(\gamma(t)), \qquad -\varepsilon < t < \varepsilon.$$

Then

$$\nabla_X Y(p) = \nabla_X \overline{Y}(p).$$

Proof. It suffices to prove that if $Y = 0$ along γ, then $\nabla_v Y(p) = 0$. Pick a local coordinate patch $(U, x^1, \cdots, x^m)$ near p such that $x(p) = 0$ and that the geometric curve γ has the defining equation $x^2 = \cdots = x^m = 0$ near p. Then $v = a\partial_1$ for some scalar a, and the condition "$Y = 0$ along γ" means $Y = Y^j \partial_j$ with $Y^j(x_1, 0, \cdots, 0) = 0$. In particular,

$$Y^j(p) = 0 \quad \text{and} \quad \partial_1 Y^j(p) = 0$$

for all j. It follows

$$\nabla_v Y(p) = \nabla_{a\partial_1} Y^j \partial_j(p) = a\left(\partial_1(Y^j)(p)\partial_j + Y^j(p)\nabla_{\partial_1}\partial_j\right) = 0.$$

□

3. Torsion and curvature tensors

Let ∇ be a linear connection on M. We define

$$T(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y],$$
$$R(X,Y)Z = -\nabla_X\nabla_Y Z + \nabla_Y\nabla_X Z + \nabla_{[X,Y]}Z.$$

This gives us two multilinear maps

$$T : \Gamma(TM)\times\Gamma(TM) \to \Gamma(TM), \quad (X,Y)\mapsto T(X,Y)$$
$$R : \Gamma(TM)\times\Gamma(TM)\times\Gamma(TM) \to \Gamma(TM), \quad (X,Y,Z)\mapsto R(X,Y)Z$$

Proposition 3.1. *T is a $(1,2)$-tensor, while R is a $(1,3)$-tensor.*

[Here when we say T is a $(1,2)$-tensor we really means $T(\omega,X,Y)=\omega(T(X,Y))$. Similar for R we means $R(\omega,X,Y,Z)=\omega(R(X,Y,Z))$.]

Proof. We need to prove

$$T(fX,Y) = T(X,fY) = fT(X,Y)$$

and

$$R(fX,Y)Z = R(X,fY)Z = R(X,Y)(fZ) = fR(X,Y)Z.$$

I will only check one of them:

$$\begin{aligned} R(fX,Y)Z &= -f\nabla_X\nabla_Y Z + \nabla_Y(f\nabla_X Z) + \nabla_{(fX)Y-Y(fX)}Z \\ &= f(-\nabla_X\nabla_Y Z + \nabla_Y\nabla_X Z + \nabla_{[X,Y]}Z) + (Yf)\nabla_X Z - (Yf)\nabla_X Z \\ &= fR(X,Y)Z. \end{aligned}$$

The others are similar and is left as exercise. ☐

Definition 3.2. We call T the the *torsion tensor*, and R the *curvature tensor* of ∇.

Locally if we write T as

$$T = T_{ij}^k \partial_k \otimes dx^i \otimes dx^j,$$

i.e. we let T_{ij}^k to be the functions such that

$$T(\partial_i,\partial_j) = T_{ij}^k\partial_k,$$

then from

$$T(\partial_i,\partial_j) = \nabla_{\partial_i}\partial_j - \nabla_{\partial_j}\partial_i - [\partial_i,\partial_j] = \Gamma_{ij}^k\partial_k - \Gamma_{ji}^k\partial_k$$

one gets

$$T_{ij}^k = \Gamma_{ij}^k - \Gamma_{ji}^k.$$

Definition 3.3. If $T=0$, we call ∇ *torsion free* (or *symmetric*) connection.

Obviously the above computation yields

Proposition 3.4. *∇ is a torsion free connection if and only if $\Gamma_{ij}^k = \Gamma_{ji}^k$ for all i,j.*

Remark. In particular, we see that the symmetry-condition "$\Gamma_{ij}^k = \Gamma_{ji}^k$ for all i, j" is independent of the choice of local coordinates.

Similarly for the curvature tensor R, if we can denote

$$R(\partial_i, \partial_j)\partial_k = R_{ijk}^l \partial_l,$$

i.e. we write the tensor R as

$$R = R_{ijk}^l \partial_l \otimes dx^i \otimes dx^j \otimes dx^k,$$

then the coefficients $=R_{ijk}^l$ is related to the Christoffel symbols by

$$R_{ijk}^l = -\Gamma_{jk}^s \Gamma_{is}^l + \Gamma_{ik}^s \Gamma_{js}^l - \partial_i \Gamma_{jk}^l + \partial_j \Gamma_{ik}^l,$$

which is a consequence of

$$R_{ijk}^l \partial_l = R(\partial_i, \partial_j)\partial_k = -\nabla_{\partial_i}(\Gamma_{jk}^s \partial_s) + \nabla_{\partial_j}(\Gamma_{ik}^s \partial_s).$$