LECTURE 5: THE RIEMANNIAN CONNECTION

1. Linear connections on tensor fields

Now let M be a smooth manifold, and ∇ a linear connection on (vector fields of) M. We will extend ∇ to a linear connection on all tensor fields. This is very easy for $(0,0)$-tensor fields (= functions), since we already have a nice one,

$$\nabla : \Gamma(TM) \times C^\infty(M) \to C^\infty(M), \quad (X, f) \mapsto \nabla_X f := Xf = df(X),$$

which obviously satisfies the two conditions in the definition of linear connections.

According to the remark in the previous lecture, a linear connection on (r,s)-tensor fields is a bi-linear map

$$\nabla : \Gamma(TM) \times \Gamma(\otimes^{r,s} TM) \to \Gamma(\otimes^{r,s} TM), \quad (X, T) \mapsto \nabla_X T,$$

that satisfies

(1) $\nabla_{fX} T = f \nabla_X T$,
(2) $\nabla_X(fT) = f\nabla_X T + (Xf)T$.

Again there are too much choices of linear connections in general, and most of them are not interesting. However, if we impose extra assumptions that all these connections ∇ on (r,s)-tensor fields are related in the following *natural* way:

(3) ∇ coincide with the given connections on $\Gamma(TM)$ and $C^\infty(M)$,
(4) $\nabla_X(T_1 \otimes T_2) = (\nabla_X T_1) \otimes T_2 + T_1 \otimes \nabla_X T_2$,
(5) $C(\nabla_X T) = \nabla_X C(T)$, where

$$C : \Gamma(\otimes^{r,s} TM) \to \Gamma(\otimes^{r-1,s-1} TM)$$

is the contraction map that pairs the first vector with the first covector.

Then one can prove

Theorem 1.1. *Given any linear connection ∇ (on vector fields), there is a unique linear connection on all tensor fields that satisfies conditions (1)-(5) above.*

Sketch of proof. First we use the conditions (3)-(5) to derive the formula of ∇ on 1-forms. Let $\omega \in \Omega^1(M) = \Gamma(T^*M)$ be any 1-form, then by (3) and (5) we must have

$$X(\omega(Y)) = \nabla_X(\omega(Y)) = \nabla_X(C(\omega \otimes Y)) = C(\nabla_X(\omega \otimes Y)).$$

Now use (4), we get

$$C(\nabla_X(\omega \otimes Y)) = C(\nabla_X \omega \otimes Y + \omega \otimes \nabla_X Y) = (\nabla_X \omega)(Y) + \omega(\nabla_X Y).$$

So we conclude

$$(\nabla_X \omega)(Y) = X(\omega(Y)) - \omega(\nabla_X Y), \tag{1.1}$$

Second we can use (4) iteratively to show that for any (r,s)-tensor field T,

$$\begin{aligned}(\nabla_X T)(\omega_1, \cdots, \omega_r, Y_1, \cdots, Y_s) =& X(T(\omega_1, \cdots, \omega_r, Y_1, \cdots, Y_s)) \\ &- \sum_i T(\omega_1, \cdots, \nabla_X \omega_i, \cdots, \omega_r, Y_1, \cdots, Y_s) \\ &- \sum_j T(\omega_1, \cdots, \omega_r, Y_1, \cdots, \nabla_X Y_j, \cdots, Y_s).\end{aligned} \tag{1.2}$$

This can be done by induction. In particular, this shows the uniqueness.

To prove the existence, one just need to check that the connections defined by equations (1.1) and (1.2) satisfies all conditions (1)-(5). $\square$

In particular, since a Riemannian metric g is a $(0,2)$-tensor field on M, we get

$$(\nabla_X g)(Y, Z) = X\langle Y, Z\rangle - \langle \nabla_X Y, Z\rangle - \langle Y, \nabla_X Z\rangle.$$

Definition 1.2. A tensor field T is called *parallel* if $\nabla_X T = 0$ for all $X \in \Gamma(TM)$.

Example. One can view the identity map

$$I = Id : \Gamma(TM) \to \Gamma(TM)$$

as a $(1,1)$-tensor via

$$I(\omega, Y) = \omega(Y).$$

Then it is parallel since according to (1.1),

$$(\nabla_X I)(\omega, Y) = X(\omega(Y)) - (\nabla_X \omega)(Y) - \omega(\nabla_X Y) = 0.$$

2. The Levi-Civita connection

Now let (M, g) be a Riemannian manifold, and ∇ a linear connection on M.

Definition 2.1. We say ∇ is *compatible* with g if the Riemannian metric g is parallel. In other words, ∇ is compatible with g if for all $X, Y, Z \in \Gamma(TM)$,

$$X(\langle Y, Z\rangle) = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle.$$

Definition 2.2. A connection ∇ is on (M, g) is called a *Levi-Civita connection* (also called a *Riemannian connection*) if it is torsion-free and is compatible with g.

For example, if we let $M = \mathbb{R}^n$ with the canonical Riemannian metric g_0, then the canonical linear connection (i.e. the one with all Christoffel symbols $\Gamma_{ij}^l = 0$) is a Levi-Civita connection. An nontrivial example is

Example. Let $M = S^n$ equipped with the round metric $g = g_{round}$, i.e. the induced metric from the canonical metric in $\mathbb{R}^{n+1}$. We denote by $\overline{\nabla}$ the canonical (Levi-Civita) connection in $\mathbb{R}^{n+1}$. For any $X, Y \in \Gamma(T^*S^n)$, one can extend X, Y to smooth vector fields $\bar{X}$ and $\bar{Y}$ on $\mathbb{R}^{n+1}$, at least near S^n. By localities we proved last time, the vector

$$\overline{\nabla}_{\bar{X}}\bar{Y}$$

at any point $p \in S^n$ depends only on the vector $\bar{X}(p) = X(p)$ and the vectors $\bar{X}(q) = X(q)$ for $q \in S^n$. In other words, it is indepedent of the choice of the extension we chose. So for simplicy we will write $\overline{\nabla}_X Y$ instead of $\overline{\nabla}_{\bar{X}}\bar{Y}$ for points on S^n. It is a vector that is not necessary tangent to S^n. We define $\nabla_X Y$ be the "orthogonal projection" of $\overline{\nabla}_X Y$ onto the tangent space of S^n, i.e.

$$\nabla_X Y := \overline{\nabla}_X Y - \langle \overline{\nabla}_X Y, \vec{n} \rangle \vec{n},$$

where $\vec{n}$ (=$\vec{x}$) is the unit out normal vector on S^n. I claim that it is a (=the) Levi-Civita connection of (M, g).

To prove this, first notice that ∇ is bilinear, and $\nabla_{fX}Y = f\nabla_X Y$. Also

$$\begin{aligned}\nabla_X(fY) &= \overline{\nabla}_X(fY) - \langle \overline{\nabla}_X(fY), \vec{n} \rangle \vec{n} \\ &= f\overline{\nabla}_X(Y) - f\langle \overline{\nabla}_X(Y), \vec{n} \rangle \vec{n} + (Xf)Y - \langle (Xf)Y, \vec{n} \rangle \vec{n} \\ &= (Xf)Y + f\nabla_X Y,\end{aligned}$$

where we used the fact that Y is a tangent vector field of S^n and thus $\langle (Xf)Y, \overline{n} \rangle = 0$. So ∇ is a linear connection on S^n.

This connection is torsion free because (we use $[X, Y] \perp \vec{n}$ here!)

$$\begin{aligned}\nabla_X Y - \nabla_Y X &= \overline{\nabla}_X Y - \overline{\nabla}_Y X - \langle \overline{\nabla}_X Y - \overline{\nabla}_Y X, \vec{n} \rangle \vec{n} \\ &= [X, Y] - \langle [X, Y], \vec{n} \rangle \vec{n} \\ &= [X, Y].\end{aligned}$$

Finally this connection is compatible with the metric g, since

$$X\langle Y, Z \rangle = \langle \overline{\nabla}_X Y, Z \rangle + \langle Y, \overline{\nabla}_X Z \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle,$$

where we used the fact that the difference between $\nabla_X Y$ and $\overline{\nabla}_X Y$ is a vector in the normal direction, and thus is perpendicular to Z.

Remark. By the same argument, one can prove that if (X, g) is a Riemannian manifold, with a Levi-Civita connection ∇^M, and if $(N, \iota^* g)$ is a Riemannian submanifold of (M, g), then the "orthogonal projection" of ∇^M onto TN,

$$\nabla^N_X Y := (\nabla^M_{\bar{X}} \bar{Y})^T,$$

defines a Levi-Civita connection on $(N, \iota^* g)$.

Remark. Since any Riemannian manifold can be embedded to the standard Euclidian space isometrically, the arguments in the previous remark immediately implies that on any Riemannian manifold, there exists Levi-Civita connection!

Our main theorem is to prove

Theorem 2.3 (The fundamental theorem of Riemannian geometry)**.** *On any Riemannian manifold (M, g), there is a unique Levi-Civita connection.*

Remark. Roughly speaking, smooth manifolds are the underlying space, and in geometry we are interested in various extra geometric structures defined on manifold. Given any smooth manifold, one has

- infinitely many different distance function (all compatible with the underlying topology),
- infinitely many different measures,
- infinitely many differnt Riemannian structures,
- infinitely many linear connections etc.

However, for the first five lectures in this course, we proved that if you fix a Riemannian metric, then you will get

- a canonical distance function (the Riemannian distance)
- a canonical measures (the Riemannian measure),
- a canonical linear connections (the Levi-Civita connection).

First proof (coordinate free). Assume the Levi-Civita connection exists. Then

$$\begin{aligned}\langle \nabla_X Y, Z\rangle =& X(\langle Y, Z\rangle) - \langle Y, \nabla_X Z\rangle \\ =& X(\langle Y, Z\rangle) - \langle Y, \nabla_Z X\rangle - \langle Y, [X, Z]\rangle \\ =& X(\langle Y, Z\rangle) - Z(\langle Y, X\rangle) + \langle \nabla_Z Y, X\rangle - \langle Y, [X, Z]\rangle \\ =& X(\langle Y, Z\rangle) - Z(\langle Y, X\rangle) + \langle \nabla_Y Z, X\rangle + \langle [Z, Y], X\rangle - \langle Y, [X, Z]\rangle \\ =& X(\langle Y, Z\rangle) - Z(\langle Y, X\rangle) + Y(\langle Z, X\rangle) - \langle Z, \nabla_Y X\rangle \\ & + \langle [Z, Y], X\rangle - \langle Y, [X, Z]\rangle \\ =& X(\langle Y, Z\rangle) - Z(\langle Y, X\rangle) + Y(\langle Z, X\rangle) - \langle Z, \nabla_X Y\rangle - \langle Z, [Y, X]\rangle \\ & + \langle [Z, Y], X\rangle - \langle Y, [X, Z]\rangle\end{aligned}$$

It follows that $\nabla_X Y$ must be the vector satisfying

$$\begin{aligned}2\langle \nabla_X Y, Z\rangle =& X(\langle Y, Z\rangle) - Z(\langle Y, X\rangle) + Y(\langle Z, X\rangle) \\ & - \langle Z, [Y, X]\rangle + \langle [Z, Y], X\rangle - \langle Y, [X, Z]\rangle.\end{aligned} \tag{2.1}$$

The right hand side is determined by the metric. So the uniqueness is proved. [The last formula is called the *Koszul* formula.]

To prove the existence, one only need to check that the $\nabla_X Y$ defined by the above formula satisfies all conditions of Levi-Civita connections. □

Second proof (local coordinate). Again we first prove uniqueness. Let ∇ be a Levi-Civita connection. Pick a coordinate neighborhood and let Γ_{ij}^k be the functions

$\nabla_{\partial_i}\partial_j = \Gamma_{ij}^k\partial_k$. Then it is enough to prove that the Γ_{ij}^k's are determined by the metric g. First we note that by torsion free property,

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

Second we calculate

$$\begin{aligned}\partial_i g_{jk} = \partial_i(g(\partial_j, \partial_k)) &= g(\nabla_{\partial_i}\partial_j, \partial_k) + g(\partial_j, \nabla_{\partial_i}\partial_k) \\ &= g(\Gamma_{ij}^l\partial_l, \partial_k) + g(\partial_j, \Gamma_{ik}^l\partial_l) = \Gamma_{ij}^l g_{lk} + \Gamma_{ik}^l g_{jl}.\end{aligned}$$

Similarly one can prove

$$\partial_j g_{ki} = \Gamma_{jk}^l g_{li} + \Gamma_{ji}^l g_{kl} \quad \text{and} \quad \partial_k g_{ij} = \Gamma_{ki}^l g_{lj} + \Gamma_{kj}^l g_{il}.$$

So we get

$$\partial_j g_{ki} + \partial_i g_{jk} - \partial_k g_{ij} = 2 g_{lk}\Gamma_{ij}^l.$$

It follows

$$2\Gamma_{ij}^l = g^{lk}(\partial_j g_{ki} + \partial_i g_{jk} - \partial_k g_{ij}). \tag{2.2}$$

This proves the uniqueness.

For the existence, we can define locally (for $X = X^i\partial_i$ and $Y = Y^j\partial_j$)

$$\nabla_X Y = X^i\partial_i Y^j\partial_j + X^i Y^j \Gamma_{ij}^l\partial_l,$$

where Γ_{ij}^l is the function given by (2.2). By tedious compotations one can check that this give a Levi-Civita connection whose Christophel symbols are exactly the Γ_{ij}^l's. [In particular we immediately see that the connection is torsion free.] $\square$

The local expression (2.2) for Γ_{ij}^l in terms of g_{ij}'s is very useful in computations. For example, we have

Proposition 2.4. *Let ∇ be the Levi-Civita connection on (M, g). Then*

$$\Gamma_{ji}^j = \frac{1}{\sqrt{G}}\partial_i\sqrt{G}.$$

[Note: for the left hand side we are using the Einstein summation convection!]

Proof. We first use the formula (2.2) to get

$$2\Gamma_{ji}^j = g^{jk}(\partial_i g_{kj} + \partial_j g_{ik} - \partial_k g_{ij}) = g^{jk}\partial_i g_{kj} = \mathrm{Tr}((g^{rs})\partial_i(g_{kj})).$$

We need

Lemma 2.5. *Let $A = A(t)$ be a family of nonsingular matrices that depends smoothly on t, then*

$$\operatorname{Tr}(A^{-1}\frac{d}{dt}A) = \frac{d}{dt}\ln\det A. \tag{2.3}$$

Proof. By the standard perturbation trick, it is enough to prove the theorem for diagonalizable matrices. We write $A = P^{-1}DP$, where D is the diagonal matrix whose entries are the eigenvalues of A. Then the left hand side becomes [We use A' to represent $\frac{dA}{dt}$]

$$\begin{aligned}\operatorname{Tr}(A^{-1}A') &= \operatorname{Tr}\left(P^{-1}D^{-1}P[(P^{-1})'DP + P^{-1}D'P + P^{-1}DP']\right)\\ &= \operatorname{Tr}\left(P(P^{-1})' + D^{-1}D' + P^{-1}P'\right)\\ &= \operatorname{Tr}\left(D^{-1}D'\right),\end{aligned}$$

where in the last step we used the fact

$$P(P^{-1})' + P^{-1}P' = (P^{-1}P)' = 0.$$

For the right hand side we have

$$(\ln\det A)' = (\ln\det D)'.$$

So the problem is converted to prove (2.3) for the diagonal matrix D (whose entries depends on t), which is trivially true after straightforward computations. □

Applying this to $A = (g_{ij})$, we get

$$2\Gamma^j_{ji} = \partial_i \ln\det(g_{kj}) = 2\partial_i \ln\sqrt{G} = 2\frac{1}{\sqrt{G}}\partial_i\sqrt{G},$$

and the conclusion follows. □

3. The Hessian

Now let (M, g) be a Riemannian manifold, and ∇ the Levi-Civita connection on M. For any vector field $X \in \Gamma(TM)$, we can define a linear map

$$\nabla X : \Gamma(TM) \to \Gamma(TM), \quad Y \mapsto \nabla_Y X.$$

According to locality 2, at each point p, ∇X is just a map from T_pM to T_pM. In particular, it makes sense to talk about the trace of ∇X at each p, which gives us a function on M.

Lemma 3.1. $\operatorname{div}(X) = \operatorname{Tr}(\nabla X)$.

Proof. Both sides are functions on M, so one only need to prove it at one point p. We pick a local coordinate system near p. Then

$$\nabla_{\partial_i} X = (\nabla_{\partial_i} X^j)\partial_j + X^j \nabla_{\partial_i}\partial_j = \partial_i(X^j)\partial_j + X^j\Gamma^k_{ij}\partial_k$$

implies

$$\mathrm{Tr}(\nabla X) = \partial_i(X^i) + X^i\Gamma^j_{ji} = \partial_i(X^i) + X^i\frac{1}{\sqrt{G}}\partial_i\sqrt{G} = \frac{1}{\sqrt{G}}\partial_i(X^i\sqrt{G}) = \mathrm{div}(X).$$

$\square$

Recall that $\Delta f = -\mathrm{div}\nabla f$. So by the proposition above, we get another formula for the Laplace-Beltrami operator:

$$\Delta f = -\mathrm{Tr}(\nabla^2 f).$$

Definition 3.2. For any $f \in C^\infty(M)$, we will call $\nabla^2 f = \nabla(\nabla f)$ the *Hessian* of f.

Remark. The ∇ in ∇f here represents the gradient, not the connection. The connection on functions is $\nabla_X f = Xf = df(X)$. In other words, the connection ∇ is $\nabla f = df$.

So $\nabla^2 f$ is a map

$$\nabla^2 f : \Gamma(TM) \to \Gamma(TM),$$

which can be identified with a $(1,1)$-tensor

$$\nabla^2 f(X, \omega) = \omega(\nabla_X \nabla f).$$

Using the metric g, one can also convert this $(1,1)$-tensor $\nabla^2 f$ into a $(0,2)$-tensor

$$\nabla^2 f(X,Y) = \nabla^2 f(X, \flat Y) = (\flat Y)(\nabla_X\nabla f) = \langle \nabla_X \nabla f, Y\rangle.$$

Proposition 3.3. *$\nabla^2 f$ is a symmetric $(0,2)$-tensor.*

Proof. By metric compatibility,

$$\nabla^2 f(X,Y) = \langle \nabla_X\nabla f, Y\rangle = \nabla_X(\langle \nabla f, Y\rangle) - \langle \nabla f, \nabla_X Y\rangle = X(Yf) - (\nabla_X Y)f.$$

On the other hand, by the torsion-free property,

$$X(Yf) - (\nabla_X Y)f = Y(Xf) - (\nabla_Y X)f.$$

So we conclude

$$\nabla^2 f(X,Y) = \nabla^2 f(Y,X).$$

$\square$

Remark. One could define the Hessian of f with respect to any linear connection (and without using the metric structure), by setting

$$\nabla^2 f(X,Y) := X(Yf) - (\nabla_X Y)f.$$

Then the proof above shows that the Hessian is symmetric if and only if the connection is torsion free. This gives another explaination of the torsion tensor.