Manifold Optimization Chapter 5: Second-Order Geometry

Xu, Yuan (121090658)

May 20, 2025

1 Differentiating Vector Fields on Manifolds: Connections

Notion of derivative for vector fields on manifolds is called a *connection*, traditionally denoted by ∇ ("nabla"). Given a tangent vector $u \in \mathrm{T}_x \mathcal{M}$ and a vector field V, $\nabla_u V$ is the derivative of V at x along u. Formally, we should write $\nabla_{(x,u)} V$ where the base point x is typically clear from context.
Note that we do not need a Riemannian metric yet.

Definition 1.1. A *connection* on a manifold M is an operator

$$\nabla\colon \mathrm{T}\,\mathcal{M} \times \mathfrak{X}(\mathcal{M}) \to \mathrm{T}\,\mathcal{M} : (u, V) \mapsto \nabla_u V$$

where:

- $\mathrm{T}\,\mathcal{M}$ is the tangent vector space
- $\mathfrak{X}(\mathcal{M})$ denotes smooth vector fields on $\mathcal{M}$

This operator must satisfy four properties for all $u, w \in \mathrm{T}\,\mathcal{M}$, $U, V, W \in \mathfrak{X}(\mathcal{M})$, $a, b \in \mathbb{R}$, and $f \in C^\infty(\mathcal{M})$:

0. *Smoothness*: $(\nabla_U V)(x) \triangleq \nabla_{U(x)} V$ defines a smooth vector field $\nabla_U V$;
1. *Linearity in u*: $\nabla_{au+bw} V = a\nabla_u V + b\nabla_w V$;
2. *Linearity in V*: $\nabla_u(aV + bW) = a\nabla_u V + b\nabla_u W$;
3. *Leibniz rule*: $\nabla_u(fV) = \mathrm{D}\, f(x)[u] \cdot V(x) + f(x)\nabla_u V$.

The field $\nabla_U V$ is called the *covariant derivative* of V along U with respect to ∇.

Theorem 1.2. *Let $\mathcal{M}$ be an embedded submanifold of a Euclidean space $\mathcal{E}$. The operator ∇ defined by*

$$\nabla_u V = \mathrm{Proj}_x \left(\mathrm{D}\, \bar{V}(x)[u]\right)$$

is a connection on $\mathcal{M}$.

Proof. Let $\mathcal{M}$ be an embedded submanifold of a Euclidean space $\mathcal{E}$ by $\bar{\nabla}$. Then $\nabla_u V = \mathrm{Proj}_x \left(\bar{\nabla}_u \bar{V}\right)$. ☐

Proposition 1.3. *Let $\mathcal{M}$ be a maniflod with arbitrary connection ∇. Given a smooth vector field $V \in \mathfrak{X}(\mathcal{M})$ and a point $x \in \mathcal{M}$, if $V(x) = 0$ then $\nabla_u V = \mathrm{D}\, V(x)[u]$ for all $u \in \mathrm{T}_x \mathcal{M}$. In particular, $\mathrm{D}\, V(x)[u]$ is tangent at x.*

2 Riemannian Connections

Definition 2.1. For $U, V \in \mathfrak{X}(\mathcal{M})$ and $f \in \mathfrak{F}(\mathcal{U})$ with $\mathcal{U}$ open in $\mathcal{M}$, define:

- $Uf \in \mathfrak{F}(\mathcal{U})$ such that $(Uf)(x) = \mathrm{D}\, f((x)[U(x)]$;
- $[U, V] : \mathfrak{F}(\mathcal{U}) \to \mathfrak{F}(\mathcal{U})$ such that $[U, V]f = U(Vf) - V(Uf)$;
- $\langle U, V\rangle \in \mathfrak{F}(\mathcal{M})$ such that $\langle U, V\rangle(x) = \langle U(x), V(x)\rangle_x$.

The notation Uf captures the action of a smooth vector field U on a smooth function f through derivation, transforming f into another smooth function. The commutator $[U, V]$ of such action is called the Lie bracket. Even in linear spaces $[U, V]f$ is nonzero in general. Notice that $Uf = \langle \mathrm{grad}\, f, U\rangle$ owing to the definitions of $Uf, \langle V, U\rangle$ and $\mathrm{grad}\, f$.

Theorem 2.2. *On a Riemannian manifold $\mathcal{M}$, there exists a unique connection ∇ which satisfies two additional properties for all $U, V, W \in \mathfrak{X}(\mathcal{M})$:*

1. *Symmetry:* $[U,V]f = (\nabla_U V - \nabla_V U)f$ *for all* $f \in \mathfrak{F}(\mathcal{M})$*;*

2. *Compatibility with the metric:* $U\langle V,W\rangle = \langle \nabla_U V, W\rangle + \langle V, \nabla_U W\rangle$.

This connection is called the Levi-Civita or Riemannian connection.

Theorem 2.3. *The Riemannian connection on a Euclidean space* $\mathcal{E}$ *with any Euclidean metric* $\langle\cdot,\cdot\rangle$ *is* $\nabla_u V = \mathrm{D}\,V(x)[u]$*: the canonical Euclidean connection.*

Theorem 2.4. *Let* $\mathcal{M}$ *be an embedded submanifold of a Euclidean space* $\mathcal{E}$*. The connection* ∇ *defined by 1.2 is symmetric on* $\mathcal{M}$*.*

Theorem 2.5. *Let* M *be a Riemannian submanifold of a Euclidean space. The connection* ∇ *defined by 1.2 is the Riemannian connection on* $\mathcal{M}$*.*

Proposition 2.6. *Let* U, V *be two smooth vector fields on a manifold* $\mathcal{M}$*. There exists a unique smooth vector field* W *on* $\mathcal{M}$ *such that* $[U,V]f = Wf$ *for all* $f \in \mathfrak{F}(\mathcal{M})$*. Therefore, we identify* $[U,V]$ *with that smooth vector field. Explicitly, if* ∇ *is any symmetric connection, then* $[U,V] = \nabla_U V - \nabla_V U$*.*

3 Riemannian Hessians

Definition 3.1. Let $\mathcal{M}$ be a Riemannian manifold with its Riemannian con nection ∇. The Riemannian Hessian of $f \in \mathfrak{F}(\mathcal{M})$ at $x \in \mathcal{M}$ is the linear map $\operatorname{Hess} f(x) : \mathrm{T}_x \mathcal{M} \to \mathrm{T}_x \mathcal{M}$ defined as follows:

$$\operatorname{Hess} f(x)[u] = \nabla_u \operatorname{grad} f.$$

Equivalently, $\operatorname{Hess} f$ maps $\mathfrak{X}(\mathcal{M})$ to $\mathfrak{X}(\mathcal{M})$ as $\operatorname{Hess} f[U] = \nabla_U \operatorname{grad} f$.

Proposition 3.2. *The Riemannian Hessian is self-adjoint with respect to the Riemannian metric. That is, for all* $x \in \mathcal{M}$ *and* $u, v \in \mathrm{T}_x \mathcal{M}$*,* $\langle \operatorname{Hess} f(x)[u], v\rangle_x = \langle u, \operatorname{Hess} f(x)[v]\rangle_x$*.*

Corollary 3.3. *Let* $\mathcal{M}$ *be a Riemannian submanifold of a Euclidean space. Consider a smooth function* $f : \mathcal{M} \to \mathbb{R}$*. Let* $\bar{G}$ *be a smooth extension of* $\operatorname{grad} f$ *— that is,* $\bar{G}$ *is any smooth vector field defined on a neighborhood of* $\mathcal{M}$ *in the embedding space such that* $\bar{G}(x) = \operatorname{grad} f(x)$ *for all* $x \in \mathcal{M}$*. Then,* $\operatorname{Hess} f(x)[u] = \operatorname{Proj}_x\left(\mathrm{D}\,\bar{G}(x)[u]\right)$*.*

4 Connections as Pointwise Derivatives*

Definition 4.1. A connection on a manifold $\mathcal{M}$ is an operator

$$\nabla\colon \mathfrak{X}(\mathcal{M}) \times \mathfrak{X}(\mathcal{M}) \to \mathfrak{X}(\mathcal{M}) : (U,V) \mapsto \nabla_U V$$

which has three properties for all $U, V, W \in \mathfrak{X}(\mathcal{M})$, $f, g \in \mathfrak{F}(\mathcal{M})$ and $a, b \in \mathbb{R}$:

1. $\mathfrak{F}(\mathcal{M})$-linearity in U: $\nabla_{fU+gW} V = f\nabla_U V + g\nabla_W V$;

2. $\mathbb{R}$-linearity in V: $\nabla_U(aV + bW) = a\nabla_U V + b\nabla_U W$; and

3. Leibniz rule: $\nabla_U(fV) = (Uf)V + f\nabla_U V$.

The field $\nabla_U V$ is the covariant derivative of V along U with respect to ∇.

Proposition 4.2. *For any connection* ∇ *and smooth vector fields* U, V *on a manifold* $\mathcal{M}$*, the vector field* $\nabla_U V$ *at* x *depends on* U *only through* $U(x)$*.*

Lemma 4.3. *Given any real numbers* $0 < r_1 < r_2$ *and any point* x *in a Euclidean space* $\mathcal{E}$ *with norm* $\|\cdot\|$*, there exists a smooth function* $b : \mathcal{E} \to \mathbb{R}$ *such that*

- $b(y) = 1$ *if* $\| y - x \| \le r_1$*;*
- $b(y) = 0$ *if* $\| y - x \| \ge r_2$*; and*
- $b(y) \in (0,1)$ *if* $\| y - x \| \in (r_1, r_2)$*.*

Using bump functions, we can show that $(\nabla_U V)(x)$ depends on U and V only through their values in a neighborhood around x. This is the object of the two following lemmas.

Lemma 4.4. *Let V_1, V_2 be smooth vector fields on a manifold M equipped with a connection ∇. If $V_1 \mid_{\mathcal{U}} = V_2 \mid_{\mathcal{U}}$ on some open set $\mathcal{U}$ of $\mathcal{M}$, then $(\nabla_U V_1) \mid_{\mathcal{U}} = (\nabla_U V_2) \mid_{\mathcal{U}}$ for all $U \in \mathfrak{X}(\mathcal{M})$.*

Lemma 4.5. *Let U_1, U_2 be smooth vector fields on a manifold $\mathcal{M}$ equipped with a connection ∇. If $U_1 \mid_{\mathcal{U}} = U_2 \mid_{\mathcal{U}}$ on some open set $\mathcal{U}$ of $\mathcal{M}$, then $(\nabla_{U_1} V) \mid_{\mathcal{U}} = (\nabla_{U_2} V) \mid_{\mathcal{U}}$ for all $V \in \mathfrak{X}(\mathcal{M})$.*

Lemma 4.6. *Let U be a neighborhood of a point x on a manifold $\mathcal{M}$. Given a smooth function $f \in \mathfrak{F}(\mathcal{U})$, there exists a smooth function $g \in \mathfrak{F}(\mathcal{M})$ and a neighborhood $\mathcal{U}' \subseteq \mathcal{U}$ of x such that $g \mid \mathcal{U}' = f \mid \mathcal{U}'$.*

Lemma 4.7. *Let U be a neighborhood of a point x on a manifold $\mathcal{M}$. Given a smooth vector field $U \in \mathfrak{X}(\mathcal{U})$, there exists a smooth vector field $V \in \mathfrak{X}(\mathcal{M})$ and a neighborhood $\mathcal{U}' \subseteq \mathcal{U}$ of x such that $V \mid_{\mathcal{U}'} = U \mid_{\mathcal{U}'}$.*

Lemma 4.8. *Let U, V be two smooth vector fields on a manifold $\mathcal{M}$ equipped with a connection ∇. Further let $\mathcal{U}$ be a neighborhood of $x \in \mathcal{M}$ such that $U|_{\sqcap} = g_1 W_1 + \cdots + g_n W_n$ for some $g_1, \ldots, g_n \in \mathfrak{F}(\mathcal{U})$ and $W_1, \ldots, W_n \in \mathfrak{X}(\upsilon)$. Then,*

$$(\nabla_U V)(x) = g_1(x)(\nabla_{W_1} V)(x) + \cdots + g_n(x)(\nabla_{W_n} V)(x),$$

where each vector $(\nabla_{W_i} V)(x)$ is understood to mean $(\nabla_{\widetilde{W_i}} V)(x)$ with $\widetilde{W_i}$ any smooth extension of W_i to $\mathcal{M}$ around x.

5 Differentiating Vector Fields on Curves

Definition 5.1. Let $c : I \to \mathcal{M}$ be a smooth curve on $\mathcal{M}$ defined on an open interval I. A map $Z : I \to \mathrm{T}\,\mathcal{M}$ is a vector field on c if $Z(t)$ is in $\mathrm{T}_{c(t)}\,\mathcal{M}$ for all $t \in I$. Moreover, Z is a smooth vector field on c if it is also smooth as a map from I to $\mathrm{T}\,\mathcal{M}$. The set of smooth vector fields on c is denoted by $X(c)$.

Theorem 5.2. *Let $c \colon I \to M$ be a smooth curve on a manifold equipped with a connection ∇. There exists a unique operator $\frac{D}{dt} \colon \mathfrak{X}(c) \to \mathfrak{X}(c)$ which satisfies the following properties for all $Y, Z \in \mathfrak{X}(c)$, $U \in \mathfrak{X}(\mathcal{M})$, $g \in \mathfrak{F}(I)$, and $a, b \in \mathbb{R}$:*

1. *$\mathbb{R}$-linearity: $\frac{D}{dt}(aY + bZ) = a\frac{D}{dt}Y + b\frac{D}{dt}Z$;*
2. *Leibniz rule: $\frac{D}{dt}(gZ) = \frac{dg}{dt}Z + g\frac{D}{dt}Z$;*
3. *Chain rule: $\left(\frac{D}{dt}(U \circ c)\right)(t) = \nabla_{c'(t)} U$ for all $t \in I$;*
4. *Product rule: If M is a Riemannian manifold and ∇ is compatible with its metric (e.g., the Levi-Civita connection), then additionally:*
$$\frac{d}{dt}\langle Y, Z\rangle = \left\langle \frac{D}{dt}Y, Z \right\rangle + \left\langle Y, \frac{D}{dt}Z \right\rangle$$
where the inner product $\langle \cdot, \cdot \rangle$ along c is defined by $\langle Y, Z\rangle(t) = \langle Y(t), Z(t)\rangle_{c(t)}$.

We call $\frac{D}{dt}$ the induced covariant derivative (induced by ∇).

Proposition 5.3. *Let $\mathcal{M}$ be an embedded submanifold of a Euclidean space $\mathcal{E}$ with connection ∇ as in 1.2. The operator $\frac{D}{dt}$ defined by*

$$\frac{D}{dt}Z(t) = \mathrm{Proj}_{c(t)}\left(\frac{d}{dt}Z(t)\right)$$

is the induced covariant derivative, that is, it satisfies properties 1–3 in Theorem 5.2. If $\mathcal{M}$ is a Riemannian submanifold of $\mathcal{E}$, then $\frac{D}{dt}$ also satisfies property 4 in that same theorem.

6 Acceleration and Geodesics

Definition 6.1. Let $c : I \to \mathcal{M}$ be a smooth curve. Its velocity is the vector field $c' \in \mathfrak{X}(c)$. The acceleration of c is the smooth vector field $c'' \in \mathfrak{X}(c)$ defined by $c'' = \frac{D}{dt}c'$. We also call c'' the intrinsic acceleration of c.

Definition 6.2. On a Riemannian manifold $\mathcal{M}$, a geodesic is a smooth curve $c : I \to \mathcal{M}$ such that $c''(t) = 0$ for all $t \in I$, where I is an open interval of $\mathbb{R}$.

7 A Second-order Taylor Expansion on Curves

Lemma 7.1. *Let $c(t)$ be a geodesic connecting $x = c(0)$ to $y = c(1)$, and assume $\mathrm{Hess}\, f(c(t)) \succeq \mu I$ for some $\mu \in \mathbb{R}$ and all $t \in [0, 1]$. Then, $f(y) \geq f(x) + \langle \mathrm{grad}\, f(x), v\rangle_x + \frac{\mu}{2}\|v\|_x^2$.*

8 Second-order Retractions

Definition 8.1. A second-order retraction R on a Riemannian manifold $\mathcal{M}$ is a retraction such that, for all $x \in \mathcal{M}$ and all $v \in \mathrm{T}_x \mathcal{M}$, the curve $c(t) = R_x(tv)$ has zero acceleration at $t = 0$, that is, $c''(0) = 0$.

Proposition 8.2. *Consider a Riemannian manifold M equipped with any retraction R, and a smooth function $f : M \to \mathbb{R}$. If x is a critical point of f (that is, if* $\mathrm{grad} f(x) = 0$*), then*

$$f(R_x(s)) = f(x) + \frac{1}{2}\langle \mathrm{Hess} f(x)[s], s\rangle_x + O(\|s\|_x^3).$$

Also, if R is a second-order retraction, then for all points $x \in M$ we have

$$f(R_x(s)) = f(x) + \langle \mathrm{grad} f(x), s\rangle_x + \frac{1}{2}\langle \mathrm{Hess} f(x)[s], s\rangle_x + O(\|s\|_x^3).$$

Proposition 8.3. *If the retraction is second order or if* $\mathrm{grad}\, f(x) = 0$*, then*

$$\mathrm{Hess}\, f(x) = \mathrm{Hess}(f \circ R_x)(0),$$

where the right-hand side is the Hessian of $f \circ R_x : \mathrm{T}_x \mathcal{M} \to \mathbb{R}$ at $0 \in \mathrm{T}_x \mathcal{M}$. The latter is a "classical" Hessian since $\mathrm{T}_x \mathcal{M}$ is a Euclidean space.

9 Riemannian Submanifolds*

10 Metric Projection Retractions*