Elementary differential geometry

Zhengchao Wan

Peking University

May 9, 2016

Information geometry

Information geometry is a branch of mathematics that applies the techniques of differential geometry to the field of probability theory. This is done by taking probability distributions for a statistical model as the points of a Riemannian manifold, forming a statistical manifold. The Fisher information metric provides the Riemannian metric.

Information geometry reached maturity through the work of Shun'ichi Amari and other Japanese mathematicians in the 1980s. Amari and Nagaoka's book, **Methods of Information Geometry**, is cited by most works of the relatively young field due to its broad coverage of significant developments attained using the methods of information geometry up to the year 2000.

Applications

Information geometry can be applied where parametrized distributions play a role. Here an incomplete list:

- statistical inference
- time series and linear systems
- quantum systems
- neural networks
- machine learning
- statistical mechanics
- biology
- statistics
- mathematical finance

Overview

- Differentiable manifolds
- Tangent vectors and tangent spaces
- Vector fields and tensor fields
- Connections
- Flatness
- Submanifolds
- Riemannian connection

Manifolds

Definition (Topological manifolds)

A manifold M of dimension m, or m-manifold, is a topological space with the following properties:

- *M is Hausdorff,*
- *M is locally Euclidean of dimension m, and*
- *M has a countable basis of open sets.*

(U, φ) is called a coordinate neighborhood of M, where U is an open set of M and φ is a homeomorphism of U to an open subset of R^m.

Differentiable manifolds

Definition (C^∞-compatible)

Two coordinate neighborhoods (U, φ) and (V, ψ) are called C^∞-compatible if $U \cap V$ nonempty implies that $\varphi \circ \psi^{-1}$ and $\psi \circ \varphi^{-1}$ are diffeomorphisms of the open subsets $\varphi(U \cap V)$ and $\psi(U \cap V)$ of R^m.

Definition (Differentiable structure)

A differentiable or C^∞ (or smooth) structure on a manifold M is a family $\mathcal{U} = (U_\alpha, \varphi_\alpha)$ of coordinate neighborhoods such that:

- *the U_α cover M.*
- *$(U_\alpha, \varphi_\alpha)$ and (U_β, φ_β) are C^∞-compatible for any α, β.*
- *any coordinate neighborhood (V, ψ) compatible with every $(U_\alpha, \varphi_\alpha) \in \mathcal{U}$ is itself in $\mathcal{U}$.*

Differentiable manifolds

Theorem

Let M be a Hausdorff space with a countable basis of open sets. If $V = \{V_\beta, \psi_\beta\}$ is a covering of M by C^∞-compatible charts, then there is a unique C^∞ structure on M containing these charts.

Theorem

For any two or finite points on a smooth connected manifold M, there exists a chart $(U, \varphi) \in \mathcal{U}$ containing them.

Examples

- Suppose M be an open set of R^m. Let $U = M$, and $\varphi : U \to R^m$ be an embedding. (U, φ) becomes a C^∞ cover on M, which makes up for a differentiable structure on M.
- Suppose $f : R^{n+1} \to R$ is C^∞. If the gradient $\text{grad}f = (\frac{\partial f}{\partial x^1}, \cdots, \frac{\partial f}{\partial x^{n+1}})$ never vanishes on a level set $M_C = \{p \in R^{n+1}; f(p) = c\}$, then M_c is an n-dimensional smooth manifold.
- Suppose M is an m-dimensional manifold with d-structure $(U_\alpha, \varphi_\alpha)$. Then an open set U of M is also an m-dimensional manifold with d-structure (V_α, ψ_α), where $V_\alpha = U \cap U_\alpha, \psi_\alpha = \varphi_\alpha|_{V_\alpha}$.

Smooth functions and mappings

Definition (Smooth functions)

Let $f : M \to R$ be a function on a smooth manifold M, then within a given coordinate neighborhood (U, φ), $\bar{f} = f \circ \varphi^{-1}$ is a function from $\varphi(U)$ to R. f is called C^∞ at $p \in U$ iff $\tilde{f}$ is C^∞ at $\varphi(p)$.

Definition (Smooth mappings)

Suppose M and N are two smooth manifolds, and $f : M \to N$ is a mapping, $p \in M$. If there exist two coordinate neighborhoods (U, φ) and (V, ψ), where $p \in U$, $f(p) \in V$ and $f(U) \subset V$. If

$$\tilde{f} = \psi \circ f \circ \varphi^{-1} : \varphi(U) \leftarrow \psi(V) \tag{1}$$

is a C^∞ mapping, then f is called C^∞ at p.

Examples

- Suppose $(U, \varphi; x^i)$ is a chart of M, then U is also a manifold. It's easy to verify that $x^i : U \to R$ are smooth functions.
- Let I be a closed interval of R, $\gamma : I \to M$ a mapping. If there exists an open interval (a, b) and smooth mapping $\tilde{\gamma} : (a, b) \to M$, s.t. $I \subset (a, b)$, and $\tilde{\gamma}|_I = \gamma$, we call γ a smooth curve on M. I could also be open or half open.

Tangent vectors

Definition (Tangent vectors)

A tangent vector v at point $p \in M$ is a mapping from C_p^∞ to R that satisfies:

- $\forall f, g \in C_p^\infty, \forall \lambda \in R, v(f + \lambda g) = v(f) + \lambda v(g)$;
- $\forall f, g \in C_p^\infty, v(fg) = v(f)g(p) + f(p)v(g)$.

Tangent vectors

Lemma

Suppose $(U; x^i)$ is a chart containing p in M, and denote $x_0^i = x^i(p)$. Then $\forall f \in C_p^\infty$, there exist m smooth functions $g_i \in C_p^\infty$, s.t.

$$g_i(p) = \frac{\partial f}{\partial x^i}(p), \quad 1 \leq i \leq m, \tag{2}$$

and for any q near p, we have

$$f(q) = f(p) + \sum_{i=1}^{m}(x^i(q) - x_0^i)g_i(q). \tag{3}$$

Tangent vectors

Proof.

$$\tilde{f} = f \circ \varphi^{-1} \in C^{\infty}_{\varphi(0)}$$

On some global neighborhood $\tilde{W}$ around $x_0 = \varphi(p)$, we have

$$\begin{aligned}\tilde{f}(x) - \tilde{f}(x_0) &= \int_0^1 \frac{d}{dt}\tilde{f}(x_0 + t(x - x_0))dt \\ &= \sum_{i=1}^{m}(x^i - x_0^i)\int_0^1 \frac{\partial \tilde{f}}{\partial x^i}(x_0 + t(x - x_0))dt.\end{aligned}$$

We let

$$\tilde{g}_i(x) = \int_0^1 \frac{\partial \tilde{f}}{\partial x^i}(x_0 + t(x - x_0))dt, \quad g_i = \tilde{g}_i \circ \varphi.$$

Tangent vectors in R^m

Let $M = R^m, x_0 \in R^m$. For a vector $v \in R^m$, we define $D_v : C^\infty_{x_0} \to R$ as: $\forall f \in C^\infty_{x_0}$

$$D_v f = \frac{df(x_0 + tv)}{dt}\Big|_{t=0}. \tag{4}$$

Then D_v is a tangent vector of M.
Conversely, if any mapping $\sigma : C^\infty_{x_0} \to R$ is a tangent vector, then there exists a unique vector $v \in R^m$, such that $D_v = \sigma$.

Tangent vectors in R^m

Proof.

Using the lemma before, we have

$$f(x) = f(x_0) + \sum_{i=1}^{m}(x^i - x_0^i)g_i(x), \tag{5}$$

where $g_i(x_0) = \frac{\partial f}{\partial x^i}(x_0)$. For any constant function λ, we have $\sigma(\lambda) = 0$. According to Equation (5), we have

$$\sigma(f) = \sum_{i=1}^{m} \frac{\partial f}{\partial x^i}(x_0) \cdot \sigma(x^i).$$

We let $v = (\sigma(x^1), \cdots, \sigma(x^m))$. □

Natural basis

Suppose p is contained in a coordinate neighborhood (U, φ) and φ can be written as a coordinate vector $\varphi(p) = (x^1(p), \cdots, x^n(p))$. Denote $\frac{\partial}{\partial x^i}|_p$ as an operator which maps $f \mapsto (\frac{\partial f}{\partial x^i})_p \triangleq (\frac{\partial \bar{f}}{\partial x^i} \circ \varphi)_p$. It can be verified that $\frac{\partial}{\partial x^i}|_p$ is a tangent vector.

Theorem

Denote T_pM as the set of all tangent vectors at p, then T_pM forms a vector space which is called the tangent space. Suppose (U, φ) is a neighborhood containing p, then

$$\left.\frac{\partial}{\partial x^i}\right|_p, \quad 1 \leq i \leq n, \tag{6}$$

form a basis of T_pM; Moreover, $\dim T_pM = n$.

Natural basis

Proof.

- For any $v \in T_pM$ and $f \in C_p^\infty$, first we have

$$f = f(p) + \sum_{i=1}^{m}(x^i - x_0^i)g_i, \quad g_i(p) = \frac{\partial f}{\partial x^i}(p),$$

where $x_0^i = x^i(p)$. Then

$$\begin{aligned} v(f) &= v(f(p) + \sum_{i=1}^{m}(x^i - x_0^i)g_i) \\ &= \sum_{i=1}^{m} v(x^i)\frac{\partial}{\partial x^i}\Big|_p(f). \end{aligned}$$

Natural basis

Proof.

- For any $a^1, \cdots, a^m \in R$, if $\sum_i a^i \frac{\partial}{\partial x^i}|_p = 0$, then for all $j, 1 \leq j \leq m$,

$$0 = \left(\sum_{i=1}^{m} a^i \frac{\partial}{\partial x^i} \Big|_p \right)(x^j) = \sum_{i=1}^{m} a^i \frac{\partial x^j}{\partial x^i}(p) = a^j.$$

Therefore, $\{\frac{\partial}{\partial x^i}|_p : 1 \leq i \leq m\}$ is linear independent.

□

Differential

Definition (Differential)

Let $\lambda : M \to N$ be a smooth mapping from a manifold M to another manifold N. Given a tangent vector $D \in T_p(M)$ of M, then it can be verified that the mapping $D' : C^\infty_{\lambda(p)} \to R$ defined by $D'(f) = D(f \circ \lambda)$ belongs to $T_{\lambda(p)}N$. Representing this correspondence as $D' = (d\lambda)_p(D)$, the linear mapping $(d\lambda)_p : T_pM \to T_{\lambda(p)}N$ is called the differential of λ at p.

Example

In fact, if we replace the above N with R, then $T_{\lambda(p)}R = R$, and thus $(d\lambda)_p \in T_p^*M$, where T_p^*M is the dual vector space of T_pM. Now, we let the coordinate function $x^i = \lambda : M \to R$. Then

$$(dx^i)_p \frac{\partial}{\partial x^j} = \left(\frac{\partial}{\partial x^j}\right)'_p,$$

where for any $f \in C^\infty_{x^i(p)}$

$$\left(\frac{\partial}{\partial x^j}\right)'_p(f) = f'(x^i(p))\left(\frac{\partial x^i}{\partial x^j}\right)\Big|_p = f'\delta_{ij},$$

therefore we have $\langle dx^i, \frac{\partial}{\partial x^j}\rangle_p = dx^i(\frac{\partial}{\partial x^j})|_p = \delta_{ij}$.

Example

dx^i form a basis of T_p^*M. Generally,for any $\alpha \in T_p^*M$, we have

$$\alpha = \sum_{i=1}^{m} \alpha_i dx^i|_p = \sum_{i=1}^{m} \langle \frac{\partial}{\partial x^i}\Big|_p, \alpha \rangle dx^i|_p.$$

Especially, for any $f \in C_p^\infty$,

$$df|_p = \sum_{i=1}^{m} \frac{\partial f}{\partial x^i}(p) dx^i|_p.$$

Vector fields

Denote $TM = \cup_{p \in M} T_pM$.

A vector field of M is a mapping $X : M \mapsto TM$, such that $\forall p \in M, X(p) \in T_pM$.

For example, in a given coordinate neighborhood (U, φ, x^i), $\frac{\partial}{\partial x^i}$ is a vector field on U.

Smooth vector fields

Definition (Smooth vector fields)

Let $X : M \to TM$ be a vector field on M. X is called a smooth vector field if $\forall p \in M$, there exists a coordinate neighborhood (U, φ, x^i) containing p and

$$X|_U = \sum_{i=1}^{n} X^i \frac{\partial}{\partial x^i}, \tag{7}$$

such that $X^i, i = 1, \cdots, n$ are smooth functions on U.

Obviously, the coordinate vector fields $\left.\frac{\partial}{\partial x^i}\right|_p$ are smooth on U. Denote $\mathcal{T}(M)$ as the set of all smooth vector fields on M.

Tensors

A (r,s)-tensor τ at $p \in M$ is refered to a multilinear map

$$\tau : \underbrace{T_p^*M \times \cdots \times T_p^*M}_{r \quad copies} \times \underbrace{T_pM \times \cdots \times T_pM}_{s \quad copies} \to R \tag{8}$$

Define $T_s^r(p)$ as the set of all (r,s)-tensors at $p \in M$:

$$T_s^r(p) = \mathcal{L}(\underbrace{T_p^*M, \cdots, T_p^*M}_{r \quad copies}, \underbrace{T_pM, \cdots, T_pM}_{s \quad copies}; R), \tag{9}$$

thus $T_s^r(p)$ is a vector space of dimension m^{r+s}.

Tensors

Under a coordinate neighborhood $(U; x^i)$, $T^r_s(p)$ has a natural basis:

$$\frac{\partial}{\partial x^{i_1}}\Big|_p \otimes \cdots \otimes \frac{\partial}{\partial x^{i_r}}\Big|_p \otimes \cdots \otimes dx^{j_1}|_p \otimes \cdots \otimes dx^{j_s}|_p, \tag{10}$$

$$1 \le i_1, \cdots, i_r, j_1, \cdots, j_s \le m. \tag{11}$$

Here if $(df)|_p \in T^*_p(M)$ and $v \in T_p(M)$, then

$$\frac{\partial}{\partial x^i}\Big|_p \otimes dx^j|_p((df)|_p, v) = \langle \frac{\partial}{\partial x^i}, df\rangle_p \langle v, dx^j\rangle_p. \tag{12}$$

Now we rewrite $T^r_x(p)$ as

$$T^r_s(p) = \underbrace{T_pM \otimes \cdots \otimes T_pM}_{r \quad copies} \otimes \underbrace{T^*_p(M) \otimes \cdots \otimes T^*_pM}_{s \quad copies}. \tag{13}$$

Tensor fields

Denote $T^r_s(M) = \cup_{p\in M} T^r_s(p)$.

A tensor field of M is a mapping $\tau : M \mapsto T^r_s(M)$, such that $\forall p \in M, \tau(p) \in T^r_s(p)$.

Smooth tensor fields

Definition (Smooth tensor fields)

Let $\tau : M \to T^r_s M$ be a (r,s)-tensor field on M. τ is called a smooth tensor field if $\forall p \in M$, there exists a coordinate neighborhood (U, φ, x^i) containing p and

$$\tau|_U = \tau^{i_1\cdots i_r}_{j_1\cdots j_s} \frac{\partial}{\partial x^{i_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{i_r}} \otimes \cdots \otimes dx^{j_1} \otimes \cdots \otimes dx^{j_s}, \quad (14)$$

such that $\tau^{i_1\cdots i_r}_{j_1\cdots j_s}$ are smooth functions on U.

Denote $\mathcal{T}^r_s(M)$ as the set of all smooth tensor fields on M.

Connections

Definition (Connections)

Suppose M is a smooth manifold with dimension m. A connection is a map

$$\nabla : \mathcal{T}(M) \times \mathcal{T}(M) \to \mathcal{T}(M), \tag{15}$$

written $(X, Y) \mapsto \nabla_X Y$, satisfying the following properties:

1. *$\nabla_X Y$ is linear over $C^\infty(M)$ in X:* $\nabla_{fX_1+gX_2} Y = f\nabla_{X_1} Y + g\nabla_{X_2} Y$ *for $f, g \in C^\infty(M)$;*
2. *$\nabla_X Y$ is linear over R in Y:* $\nabla_X(aY_1 + bY_2) = a\nabla_X Y_1 + b\nabla_X Y_2$ *for $a, b \in R$;*
3. *∇ satisfies the following product rule:* $\nabla_X(fY) = f\nabla_X Y + (Xf)Y$ *for $f \in C^\infty(M)$.*

Example on R^m

Define the **Euclidean connection** by

$$\overline{\nabla}_X(Y^j\partial_j) = (XY^j)\partial_j. \tag{16}$$

Hence, $\overline{\nabla}_X Y$ is just the vector field whose components are the ordinary directional derivatives of the components of Y in the direction X.

Christoffel symbols

Let $\{E_i\}$ be a local frame for TM on an open subset $U \subset M$. We usually work with a coordinate frame $E_i = \partial_i$. Expand $\nabla_{E_i} E_j$ in terms of this same frame:

$$\nabla_{E_i} E_j = \Gamma_{ij}^k E_k. \tag{17}$$

This defines n^3 functions Γ_{ij}^k on U, called the **Christoffel symbols** of ∇ with respect to this frame.

Let $X, Y \in \mathcal{T}(U)$ be expressed in terms of a local frame by $X = X^i E_i$, $Y = Y^j E_j$. Then we have

$$\nabla_X Y = (XY^k + X^i Y^j \Gamma_{ij}^k) E_k. \tag{18}$$

Vector fields along curves

Let $\gamma : I \to M$ be a curve, where I is an open interval. At any time $t \in I$, the **velocity** $\dot{\gamma}(t)$ of γ is invariantly defined as $(d\gamma)_t(d/dt)$. It acts on functions by

$$\dot{\gamma}(t)f = \frac{d}{dt}(f \circ \gamma)(t). \tag{19}$$

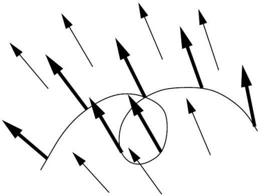
If we write the coordinate representation of γ as $\gamma(t) = (\gamma^1(t), \cdots, \gamma^n(t))$, then

$$\dot{\gamma}(t) = \dot{\gamma}^i(t)\partial_i. \tag{20}$$

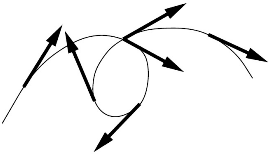
A vector field along a curve $\gamma : I \to M$ is a smooth map $V : I \to TM$ such that $V(t) \in T_{\gamma(t)}M$ for every $t \in I$. We let $\mathcal{T}(\gamma)$ denote the space of vector fields along γ.

Extendible vector field

Suppose $\gamma : I \to M$ is a curve, and $\tilde{V} \in \mathcal{T}(M)$ is a vector field on M. For each $t \in I$, let $V(t) = \tilde{V}_{\gamma(t)}$. A vector field V along γ is said to be **extendible** if it can be extended to a vector field $\tilde{V}$ on a neighborhood of the image of γ.



(a) Extendible vector field.



(b) Nonextendible vector field.

Covariant derivative along curves

Theorem

Let ∇ be a linear connection on M. For each curve $\gamma : I \to M$, ∇ determines a unique operator

$$D_t : \mathcal{T}(\gamma) \to \mathcal{T}(\gamma) \tag{21}$$

satisfying the following properties:

1. *Linearity over R: $D_t(aV + bV) = aD_tV + bD_tW$ for $a, b \in R$.*
2. *Product rule: $D_t(fV) = \dot{f}V + fD_tV$ for $f \in C^\infty(I)$.*
3. *If V is extendible, then for any extension $\tilde{V}$ of V,*

$$D_tV(t) = \nabla_{\dot{\gamma}(t)}\tilde{V}. \tag{22}$$

For any $V \in \mathcal{T}(\gamma)$, D_tV is called the covariant derivative of V along γ.

Parallel translation

A vector field V along a curve γ is said to be **parallel along** γ with respect to ∇ if $D_t V \equiv 0$.

Lemma (Parallel Translation)

Given a curve $\gamma : I \to M$, $t_0 \in I$, and a vector $V_0 \in T_{\gamma(t_0)}M$, there exists a unique parallel vector field V along γ such that $V(t_0) = V_0$.

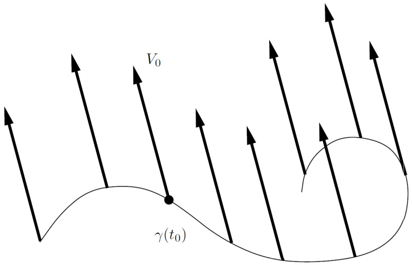


Figure: Parallel translate of V_0 along γ.

Parallel translation

If $\gamma : I \to M$ is a curve and $t_0, t_1 \in I$, parallel translation defines an operator

$$P_{t_0 t_1} : T_{\gamma(t_0)}M \to T_{\gamma(t_1)}M \tag{23}$$

by setting $P_{t_0 t_1} V_0 = V(t_1)$, where V is the parallel translate of V_0 along γ. This is a linear isomorphism between $T_{\gamma(t_0)}M$ and $T_{\gamma(t_1)}M$.

Lemma

Let ∇ be a linear connection on M, then we have

$$D_t V(t_0) = \lim_{t \to t_0} \frac{P^{-1}_{t_0 t} V(t) - V(t_0)}{t - t_0}. \tag{24}$$

In R^m, we have exactly $D_t V(t_0) = lim_{t \to t_0} \frac{V(t) - V(t_0)}{t - t_0}$.

Parallel translation

Proof.

Choose coordinates near $\gamma(t_0)$, and write $V(t) = V^j(t)\partial_j$ near t_0. Then by the properties of D_t, since ∂_j si extendible

$$\begin{aligned} D_t V(t_0) &= \dot{V}^j(t_0)\partial_j + \dot{V}^j(t_0)\nabla_{\dot{\gamma}(t_0)}\partial_j \\ &= (\dot{V}^k(t_0) + V^j(t_0)\dot{\gamma}^i(t_0)\Gamma^k_{ij}(\gamma(t_0)))\partial_k. \end{aligned}$$

Let $W(t_0) = P^{-1}_{t_0 t_1} V(t)$. Then

$$W(t_0) = W^k(t_0)\partial_k = (V^k(t) + (t_0 - t)\dot{V}^k(t) + o(t_0 - t))\partial_k$$

Since $W(t_0)$ is the parallel translation of $V(t)$,

$$\dot{V}^k(t) + V^j(t)\dot{\gamma}^i(t)\Gamma^k_{ij}(\gamma(t)) = 0,$$

Parallel translation

Proof.

therefore

$$W(t_0) = (V^k(t) + (t - t_0)V^j(t)\dot{\gamma}^i(t)\Gamma^k_{ij}(\gamma(t)) + o(t - t_0))\partial_k$$

$$lim_{t \to t_0} \frac{P^{-1}_{t_0 t} V(t) - V(t_0)}{t - t_0} = \lim_{t \to t_0} \frac{W(t_0) - V(t_0)}{t - t_0}$$
$$= \lim_{t \to t_0} (\frac{V^k(t) - V^k(t_0)}{t - t_0} + V^j(t)\dot{\gamma}^i(t)\Gamma^k_{ij}(\gamma(t)))\partial_k$$
$$= (\dot{V}^k(t_0) + V^j(t_0)\dot{\gamma}^i(t_0)\Gamma^k_{ij}(\gamma(t_0)))\partial_k = D_t V(t_0).$$

□

Parallel

Definition (Parallel)

Let $X \in \mathcal{T}(M)$ be a vector field on M. If for any curve γ on M, $X_\gamma :\mapsto X_{\gamma(t)}$ is parallel along γ (with respect to ∇), we say that X is parallel on M (with respect to ∇).

A necessary and sufficient condition for an $X = X^i \partial_i$ to be parallel is that $\nabla_Y X = 0$ for all $Y \in \mathcal{T}(M)$, or equivalently that

$$\partial_i X^k + X^j \Gamma_{ij}^k = 0. \tag{25}$$

Curvature and torsion

Let ∇ be a connection on M. Then for vector fields $X, Y, Z \in \mathcal{T}(M)$, if we define

$$R(X,Y)Z \triangleq \nabla_X(\nabla_Y Z) - \nabla_Y(\nabla_X Z) - \nabla_{[X,Y]}Z \tag{26}$$

$$T(X,Y) \triangleq \nabla_X Y - \nabla_Y X - [X,Y], \tag{27}$$

then these are also vector fields ($\in \mathcal{T}(M)$). Here, letting $X = X^i \partial_i$ and $Y = Y^i \partial_i$, we have defined $[X, Y]$ to be the vector field

$$[X,Y] = XY - YX = (X^j \partial_j Y^i - Y^j \partial_j X^i)\partial_i \tag{28}$$

(this does not depend on the choice of coordinate system).

Curvature and torsion

Now it can be proved that $R \in \mathcal{T}_3^1(M)$ and $T \in \mathcal{T}_2^1(M)$ and they are called curvature tensor field and torsion tensor field respectively. The component expressions of R and T with respect to coordinate system $(U; x^i)$ are given by

$$R(\partial_i, \partial_j)\partial_k = R^l_{ijk}\partial_l \quad \text{and} \quad T(\partial_i, \partial_j) = T^k_{ij}\partial_k, \tag{29}$$

and these may be computed in the following way:

$$R^l_{ijk} = \partial_i \Gamma^l_{jk} - \partial_j \Gamma^l_{ik} + \Gamma^l_{ih}\Gamma^h_{jk} - \Gamma^l_{jh}\Gamma^h_{ik} \quad \text{and} \tag{30}$$

$$T^k_{ij} = \Gamma^k_{ij} - \Gamma^k_{ji}. \tag{31}$$

Flat

Definition (Flat)

A connection ∇ of a smooth manifold M is called flat if for any point $p \in M$, there exist a coordinate neighborhood $(U; x^i)$, such that ∂_i are all parallel on U or equivalently, $\nabla_{\partial_i}\partial_j \equiv 0$ or $\{\Gamma^k_{ij}\}$ vanish on U. Such coordinate system is called an affine coordinate system.

It can be proved that ∇ is flat iff $R = T = 0$.

The Euclidean connection $\overline{\nabla}$ of R^m is flat.
It's easy to see that $R^l_{ijk} = -R^l_{jik}$ and $T^k_{ij} = -T^k_{ji}$. Hence when M is 1-dimensional, $R = 0$ and $T = 0$ necessarily hold, and therefore M is flat.

Affine transformation

Suppose $(\tilde{U}, \phi; \tilde{x}^i)$ is another coordinate neighborhood of M with Christoffel symbols $\tilde{\Gamma}_{ij}^k$. When $U \cap \tilde{U} \neq \emptyset$, we have the following equation on $U \cap \tilde{U}$:

$$\tilde{\Gamma}_{ij}^k = \Gamma_{pq}^r \frac{\partial x^p}{\partial \tilde{x}^i} \frac{\partial x^q}{\partial \tilde{x}^j} \frac{\partial \tilde{x}^k}{\partial x^r} + \frac{\partial^2 x^r}{\partial \tilde{x}^i \partial \tilde{x}^j} \frac{\partial \tilde{x}^k}{\partial x^r}. \quad (32)$$

If $(U, \varphi; x^i)$ is affine, then we have $\frac{\partial^2 x^r}{\partial \tilde{x}^i \partial \tilde{x}^j} \frac{\partial \tilde{x}^k}{\partial x^r}$. Hence $(\tilde{U}, \phi; \tilde{x}^i)$ is affine iff $\frac{\partial^2 x^r}{\partial \tilde{x}^i \partial \tilde{x}^j} = 0$. This is equivalent to the condition that there exist an $m \times m$ matrix A and an $m-$dimensional vector B such that

$$\varphi(p) = A\phi(p) + B \quad (\forall p \in U \cap \tilde{U}), \quad (33)$$

which is called an **affine transformation**.

Riemannian metric

Definition (Riemannian metric)

A Riemannian metric on a smooth manifold M is a 2-tensor field $g \in \mathcal{T}_2^0(M)$ that is symmetric (i.e., $g(X, Y) = g(Y, X)$) and positive definite (i.e., $g(X, Y) > 0$ if $X \neq 0$).

A Riemannian metric thus determines an inner product on each tangent space T_pM, which is typically written $\langle X, Y \rangle \triangleq g(X, Y)$ for $X, Y \in T_pM$.

A manifold M together with a given Riemannian metric g is called a Riemannian manifold (M, g).

Riemannian metric

We define the **length** or **norm** of any tangent vector $X \in T_pM$ to be $|X| \triangleq \langle X, X\rangle^{1/2}$ and the **angle** between two nonzero vectors $X, Y \in T_pM$ to be the unique $\theta \in [0, \pi]$ satisfying $cos\theta = \langle X, Y\rangle/(|X||Y|)$. We say that X and Y are **orthogonal** if their angle is $\pi/2$, or equivalently if $\langle X, Y\rangle = 0$. Vectors $E_1, \cdots, E_k$ are called **orthonormal** if thay are of length 1 and pairwise orthogonal, or equivalently if $\langle E_i, E_j\rangle = \delta_{ij}$.

Given a local frame $(\partial_i, \cdots, \partial_m)$ for TM, and $(dx^1, \cdots, dx^m)$ is its dual coframe, a Riemannian metric can be written locally as $g = g_{ij}dx^i \otimes dx^j$, or $g = g_{ij}dx^i dx^j$, if we denote $dx^i dx^j = \frac{1}{2}(dx^i \otimes dx^j + dx^j \otimes dx^i)$ and $g_{ij} = \langle \partial_i, \partial_j\rangle$.

Examples

One obvious example of a Riemannian manifold is R^n with its Euclidean metric $\bar{g}$, which is just the usual inner product on each tangent space $T_x R^n$ under the natural identification $T_x R^n = R^n$. In standard coordinates, this can be written in several ways;

$$\bar{g} = \sum_i dx^i dx^i = \sum_i (dx^i)^2 = \delta_{ij} dx^i dx^j. \tag{34}$$

The matrix of $\bar{g}$ in these coordinates is thus $\bar{g}_{ij} = \delta_{ij}$.

Examples

Proposition

Let M, N be two smooth manifolds, and $f : M \to N$ a smooth mapping. If φ is a smooth $(0, r)$-tensor field on N, then we have a smooth $(0, r)$-tensor field $f^*\varphi$ on M: for $p \in M, \forall v_1, \cdots, v_r \in T_pM$,

$$((f^*\varphi)(p))(v_1, \cdots, v_r) = (\varphi(p))((df)_p(v_1), \cdots, (df)_p(v_r)).$$

Particularly, if f is an immersion, and h is a Riemannian metric on N, then $g = f^*h$ is a Riemannian metric on M. g is called the **induced metric of** *h.*

Examples

Proof.

Suppose $r = 2$. $\forall p \in M$, take its chart $(U; x^i)$ in M and $f(p)$'s chart $(V; y^\alpha)$ in N, s.t. $f(U) \subset V$.

$$\varphi|_V = \sum_{\alpha,\beta=1}^{n} \varphi_{\alpha,\beta} dy^\alpha \otimes dy^\beta, \quad f^\alpha = y^\alpha \circ f,$$

where $\varphi_{\alpha,\beta} = \varphi\left(\frac{\partial}{\partial y^\alpha}, \frac{\partial}{\partial y^\beta}\right)$. Then we have

$$(f^*\varphi)|_U = \sum_{i,i=1}^{m} (f^*\varphi)\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right) dx^i \otimes dx^j$$

$$= \sum_{i,j} \sum_{\alpha,\beta} \frac{\partial f^\alpha}{\partial x^i} \frac{\partial f^\beta}{\partial x^j} (\varphi_{\alpha,\beta} \circ f) dx^i \otimes dx^j$$

Examples

Hypersurfaces in R^{n+1}
Suppose $f : N \to R^{n+1}$ is an immersion from n-dimensional manifold N into R^{n+1}. We call (f, N) the immersed hypersurface of R^n. Let $h = \langle \cdot, \cdot \rangle$ be the standard metric on R^{n+1}, and $(x^1, \cdots, x^{n+1})$ the Cartesian coordinate, then $h = \sum_{\alpha=1}^{n+1}(dx^\alpha)^2$. Under a chart $(U; u^i)$ on N, let $x^\alpha = f^\alpha(u^1, \cdots, u^n), 1 \leq \alpha \leq n+1$. Then we have

$$g|U = \sum_{\alpha,i,j} \frac{\partial f^\alpha}{\partial u^i}\frac{\partial f^\alpha}{\partial u^j} du^i du^j$$

Riemannian connection

Definition (Riemannian connection)

If for all $X, Y, Z \in \mathcal{T}(M)$*,*

$$Z\langle X, Y\rangle = \langle \nabla_Z X, Y\rangle + \langle X, \nabla_Z Y\rangle, \tag{35}$$

then we say that ∇ *is a metric connection with respect to* g *or Riemannian connection.*

If a metric connection is also symmetric, we call it the **Riemannian connection**.

Riemannian connection

Consider a curve $\gamma : t \mapsto \gamma(t)$ on M and two vector fields X and Y along γ. We rewrite Equation (35) as

$$\frac{d}{dt}\langle X(t), Y(t)\rangle = \langle D_t X(t), Y(t)\rangle + \langle X(t), D_t Y(t)\rangle. \quad (36)$$

Now if X and Y are both parallel on γ, then the right hand side of the equation above is 0.

$$\langle P_\gamma(X), P_\gamma(Y)\rangle_q = \langle X, Y\rangle_p. \quad (37)$$

Riemannian connection

Using $\partial_i, \partial_j, \partial_k$ in Equation (35), we have

$$\partial_k \langle \partial_i, \partial_j \rangle = \langle \nabla_{\partial_k} \partial_i, \partial_j \rangle + \langle \partial_i, \nabla_{\partial_k} \partial_j \rangle,$$

which is equivalent to

$$\begin{aligned} \partial_k g_{ij} &= \Gamma^h_{ki} g_{hj} + \Gamma^h_{kj} g_{hi} \\ &= \Gamma_{ki,j} + \Gamma_{kj,i} \end{aligned}$$

where $\Gamma_{ki,j} \triangleq \langle \nabla_{\partial_k} \partial_i, \partial_j \rangle = \Gamma^h_{ki} g_{hj}$. For Riemannian connection, which requires $\Gamma_{ij,k} = \Gamma_{ji,k}$, we have

$$\Gamma_{ij,k} = \frac{1}{2}(\partial_i g_{jk} + \partial_j g_{ki} - \partial_k g_{ij}). \tag{38}$$

Flat Riemannian connection

Suppose a Riemannian connection ∇ is flat and there exists an affine coordinate system $[x^i]$. Since ∂_i is parallel on M, $\langle \partial_i, \partial_j \rangle$ is constant on M. If

$$\langle \partial_i, \partial_j \rangle = \delta_{ij}, \tag{39}$$

we call such coordiante system a **Euclidean coordinate system**.

In fact, the Riemmanian connection is flat iff there exists a Euclidean coordinate system.

Fundamental lemma

Theorem (Fundamental lemma of Riemannian geometry)

Let (M, g) be a Riemannian manifold. There exists a unique linear connection ∇ on M that is compatible with g and symmetric. (A linear connection ∇ is symmetric if its torsion vanishes identically.)

Submanifolds

Definition (Regular submanifold)

Let $F : M \to N$ be a smooth injective map between two smooth manifolds satisfying the following properties:

1. $(dF)_p : T_pM \to T_{F(p)}N$ *is injective for all* $p \in M$;
2. $F : M \to F(M)$ *is a homeomorphism with respect to the induced topology from* N *onto* $F(M)$.

Then we call F *as a regular embedding and* M *as a regular submanifold of* N.

We can consider M as a subset of N.

Submanifolds

Proposition

Suppose M is an regular n-dimensional submanifold of an m-dimensional manifold N. For any $p \in M$, there exist **slice coordinates** $(x^1, \cdots, x^m)$ *on a neighborhood $\tilde{\mathcal{U}}$ of $p \in N$ s.t. $\tilde{\mathcal{U}} \cap M$ is given by $\{x : x^{n+1} = \cdots = x^m = 0\}$, and $(x^1, \cdots, x^n)$ form local coordinates for M. At each $q \in \tilde{\mathcal{U}} \cap M$, $T_q M$ can be naturally identified as the subspace of $T_q N$ spanned by the vectors $(\partial_1, \cdots, \partial_n)$.*

Submanifolds

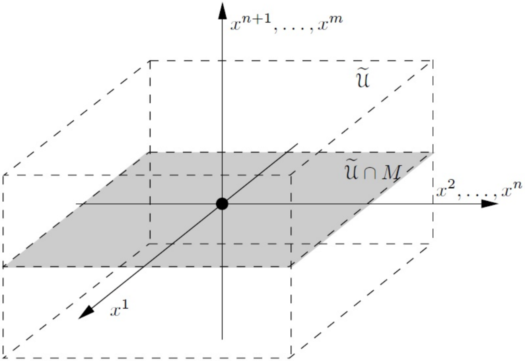


Figure: Slice coordinates.

Projection of connections

Let N be an n-dimensional manifold and M an m-dimensional submanifold of N. Suppose $T_pN = T_pM \oplus T_p^{\perp}M$, where T_pM is the tangent space of p on M, and $T_p^{\perp}M$ its orthogonal complement space with respect to g. Let $\pi^{\top} : TN|_M \to TM$ denote the orthogonal projection. Then we define $\nabla^{\top}$ on M as

$$(\nabla^{\top}_X Y)_p = \pi^{\top}((\nabla_X Y)_p). \tag{40}$$

It can be verified that $\nabla^{\top}$ is a connection on M, and we call such $\nabla^{\top}$ the **projection of ∇ onto M with respect to g.**

Autoparallel submanifolds

Let $(U, \varphi; x^\alpha)$ and $(V, \phi; y^i)$ be coordinate neighborhoods of M and N respectively, where $F(U) \subset V$.

Now letting $X, Y \in \mathcal{T}(M)$ and $X\big|_U = X^\alpha \partial_\alpha$ and $Y\big|_U = Y^\alpha \partial_\alpha$ be vector fields on M and ∇ the connection on N, generally, we don't have $\nabla_X Y \in \mathcal{T}(N)$ to be a vector field of M.

If, however, $\nabla_X Y \in \mathcal{T}(M)$ for $\forall X, Y \in \mathcal{T}(M)$, that is $\nabla_X Y = \nabla_X^\top Y$, then we say that M is **autoparallel** with respect to ∇ and ∇ can be viewed as the connection on M.

Autoparallel submanifolds

Using identities such as $\partial_\alpha = (\partial_\alpha y^i)\partial_i$, we have

$$\nabla_X Y = X^\alpha(\partial_\alpha Y^\beta)\partial_\beta + X^\alpha Y^\beta \nabla_{\partial_\alpha}\partial_\beta, \tag{41}$$

$$\nabla_{\partial_\alpha}\partial_\beta = \{(\partial_\alpha y^i)(\partial_\beta y^j)\Gamma_{ij}^k + \partial_\alpha\partial_\beta y^k\}\partial_k. \tag{42}$$

Thus, M is autoparallel iff $\nabla_{\partial_\alpha}\partial_\beta \in \mathcal{T}(M)$ holds for all α, β. This, in turn, is equivalent to there existing m^3 functions $\{\Gamma_{\alpha\beta}^\gamma\}(\in C^\infty(M))$ which satisfy

$$\nabla_{\partial_\alpha}\partial_\beta = \Gamma_{\alpha\beta}^\gamma \partial_\gamma. \tag{43}$$

Then we have

$$\Gamma_{\alpha\beta}^\gamma \partial_\gamma y^k = (\partial_\alpha y^i)(\partial_\beta y^j)\Gamma_{ij}^k + \partial_\alpha\partial_\beta y^k \tag{44}$$

Autoparallel flat submanifolds

Suppose N is flat with respect to ∇. Then it can be proved that autoparallel submanifolds are also flat. Hence without loss of generality we may assume that $[x^\alpha]$ and $[y^i]$ are affine coordinate systems in Equation (44).

Due to flatness, we have $\partial_\alpha \partial_\beta y^k = 0$, which is equivalent to there existing an $n \times m$ matrix A and an n-dimensional vector B that satisfies

$$\phi(p) = A\varphi(p) + B \quad (\forall p \in M) \tag{45}$$

A subspace of R^n which may be expressed as $\{Au + B | u \in R^m\}$ is called the affine subspace of R^n.

Autoparallel flat submanifolds

Theorem

If N is flat, then a necessary and sufficient condition for a submanifold M to be autoparallel is that M is expressed as an affine subspace of N with respect to an affine coordinate system.

Examples

Open subsets of N are autoparallel.

1-dimensional autoparallel submanifolds are called autoparallel curves or geodesics. For a curve $\gamma : t \mapsto \gamma(t)$, the condition in Equation (43) may be rewritten using Equation (22) as

$$D_t\dot{\gamma}(t) = \Gamma(t)\dot{\gamma}(t). \tag{46}$$

As noted before, connections on 1-dimensional manifolds are necessarily flat, thus $\Gamma(t) \equiv 0$ and $D_t\dot{\gamma}(t) = 0$, which can be expressed as

$$\ddot{\gamma}(t) + \dot{\gamma}^i(t)\dot{\gamma}^j(t)(\Gamma_{ij}^k)_{\gamma(t)} = 0. \tag{47}$$