COMP760: Geometry and Generative Models

Week 5: Distributions on Manifolds

Notes by: Joey Bose and Prakash Panangaden

1

RIEMANNIAN MANIFOLDS

In the previous chapters we viewed Riemannian geometry from an intrinsic perspective. This required us to specify charts for every manifold and then in these charts we identified a local-coordinate system. In some sense this is a more aesthetically pleasing way to understand the geometry, but in machine learning we are chiefly interested in implementing these ideas. As a result, using the intrinsic geometry—while elegant mathematically—can pose a challenge from an numerical stability and optimization perspective.

An alternate but fully compatible perspective with the intrinsic perspective is taking an extrinsic view of the geometry. Specifically, we can view any d-dimensional Riemannian manifold $(\mathcal{M}, g)$ as being embedded in a higher dimensional Euclidean space $\mathbb{R}^m$, where $m > d$. The extrinsic perspective is always realizable as any Riemannian manifold can be isometrically embedded by the *Nash embedding theorem* [Gunther, 1991]. The main benefit from taking an extrinsic view is that the metric g coincides with pullback of the Euclidean metric via the inclusion map. In the following section we present a thorough treatment of Riemannian geometry from an extrinsic view, with much of the material taken from the appendix of Huang et al. [2022], after which we turn to distributions on Riemannian manifolds in the following chapter.

Notation Convention. There are a number of different notations used in differential geometry and all have their place. The most abstract level is with tensors (of which forms and vectors are special cases) and it is the best for establishing general properties. We used an index-free notation in our previous treatment. A coordinate-based, but still intrinsic, description that uses local charts which describe explicit coordinate systems for *patches* of the manifold. For computational purposes it is convenient to view manifolds as hypersurfaces embedded in $\mathbb{R}^m$ even though this obscures the geometric meaning; these are called extrinsic coordinates which we use for actual implementations.

We use capital letters to denote vectors, and tilded letters to denote vectors and variables defined on the local patch.

1.1 Smooth manifolds and tangent vectors

We recall some preliminaries of smooth manifolds. See Lee [2013] for a more detailed and comprehensive account.

A smooth d-manifold is a topological space $\mathcal{M}$ (assumed to be paracompact, Hausdorff and second countable) and a family of pairs $\{(U_i, \varphi_i)\}$, where the U_i are open sets that together cover all of $\mathcal{M}$ and each φ_i is a homeomorphism from U_i to an open set in $\mathbb{R}^d$; these pairs are called *charts*. They are required to satisfy a compatibility condition: if U_i and U_j have non-empty intersection, say V, then $\varphi_i \circ \varphi_j^{-1}|_V$ has to be an infinitely differentiable map from $\varphi_j(V) \subset \mathbb{R}^d$ to $\varphi_i(V) \subset \mathbb{R}^d$. The use of charts allows one to talk about differentiability of functions or vectors fields, by moving to $\mathbb{R}^d$ as needed. A *smooth function* f on $\mathcal{M}$ has type $\mathcal{M} \to \mathbb{R}$ and is such that for any chart (U, φ) the map $f \circ \varphi^{-1} : \mathbb{R}^d \to \mathbb{R}$ is smooth [1]. The set of smooth functions on $\mathcal{M}$ is denoted $C^\infty(\mathcal{M})$.

Let $\mathcal{M}$ be a smooth manifold, and fix a point x in $\mathcal{M}$. A **derivation** at x is a linear operator $D : C^\infty(\mathcal{M}) \to \mathbb{R}$ satisfying the product rule

$$D(fg) = f(x)D(g) + g(x)D(f) \tag{1.1}$$

for all $f, g \in C^\infty(\mathcal{M})$. The set of all derivations at x is a d-dimensional real vector space called the **tangent space** $\mathcal{T}_x\mathcal{M}$, and the elements of $\mathcal{T}_x\mathcal{M}$ are called the **tangent vectors** (or tangents) at x. For the Euclidean space $\mathcal{M} = \mathbb{R}^d$, we have that $\mathcal{T}_x\mathbb{R}^d = \text{span}\{\frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_d}\}$. We now see how to use the Euclidean derivations to induce the tangent space of arbitrary Riemannian manifolds.

Let $\mathcal{N}$ be another smooth manifold. For any tangent $V \in \mathcal{T}_x\mathcal{M}$ and smooth map $\varphi : \mathcal{M} \to \mathcal{N}$, the **differential** $d\varphi_x : \mathcal{T}_x\mathcal{M} \to \mathcal{T}_{\varphi(x)}\mathcal{N}$ is defined as the pushforward of V acting on $f \in C^\infty(N)$:

$$d\varphi_x(v)(f) = V(f \circ \varphi). \tag{1.2}$$

Note that, if φ is a diffeomorphism, $d\varphi_x$ is an isomorphism between $\mathcal{T}_x\mathcal{M}$ and $\mathcal{T}_{\varphi(x)}\mathcal{N}$, and the inverse map satisfies $(d\varphi_x)^{-1} = d(\varphi^{-1})_{\varphi(x)}$. Furthermore, differentials follow the chain rule, i.e the differential of a composite is the composite of the differentials.

Let $\tilde{x} = (\tilde{x}_1, \cdots, \tilde{x}_d) = \varphi(x)$ be a local coordinate. Since $d\varphi_x : \mathcal{T}_x\mathcal{M} \to \mathcal{T}_{\varphi(x)}\mathbb{R}^d$ is an isomorphism, we can characterize $\mathcal{T}_x\mathcal{M}$ via inversion. We

[1] Strictly speaking this map has to be restricted to $\varphi(U)$ but we will assume that the appropriate restrictions are always intended rather than cluttering up the notation with restrictions all the time.

define the basis vector $\tilde{E}_i$ of $\mathcal{T}_x\mathcal{M}$ by

$$\tilde{E}_i = (d\varphi_x)^{-1}\left(\frac{\partial}{\partial \tilde{x}_i}\right) = (d\varphi^{-1})_{\varphi(x)}\left(\frac{\partial}{\partial \tilde{x}_i}\right), \tag{1.3}$$

which means

$$\tilde{E}_i(f) = \frac{\partial}{\partial \tilde{x}_i} f(\varphi^{-1}(\tilde{x})). \tag{1.4}$$

The tangent space $\mathcal{T}_x\mathcal{M}$ of $\mathcal{M}$ at x is spanned by $\left\{\tilde{E}_1, \cdots, \tilde{E}_d\right\}$. This means any tangent vector V can be represented by $\sum_{i=1}^d \tilde{v}_i \tilde{E}_i$ for some coordinate-dependent coefficients $\tilde{v}_i$.

A manifold $\mathcal{M}$ is said to be *embedded* in $\mathbb{R}^m$ if there is an inclusion map $\iota : \mathcal{M} \to \mathbb{R}^m$ such that $\mathcal{M}$ is homeomorphic to $\iota(\mathcal{M})$ and the differential at every point is injective. Every smooth manifold can be embedded in some $\mathbb{R}^m$ with $m > d$ for some suitably chosen m.

When $\mathcal{M}$ is embedded in $\mathbb{R}^m$, we can view $\mathcal{T}_x\mathcal{M}$ as a linear subspace of $\mathcal{T}_x\mathbb{R}^m$; note that this map has trivial kernel. Let $\iota : \mathcal{M} \to \mathbb{R}^m$ denote the inclusion map, i.e. $\iota(x) = x \in \mathbb{R}^m$ for $x \in \mathcal{M}$. Then

$$\tilde{E}_i = (d\varphi^{-1})_{\varphi(x)}\left(\frac{\partial}{\partial \tilde{x}_i}\right) = (d\iota^{-1})_{\iota(x)}(d\iota \circ \varphi^{-1})_{\varphi(x)}\left(\frac{\partial}{\partial \tilde{x}_i}\right) = \sum_{j=1}^m \frac{\partial \varphi_j^{-1}}{\partial \tilde{x}_i}\frac{\partial}{\partial x_j}. \tag{1.5}$$

This means we can rewrite a tangent vector using the ambient space's basis

$$\sum_{i=1}^d \tilde{v}_i \tilde{E}_i = \sum_{i=1}^d \sum_{j=1}^m \tilde{v}_i \frac{\partial \varphi_j^{-1}}{\partial \tilde{x}_i}\frac{\partial}{\partial x_j} = \sum_{j=1}^m \bar{v}_j \frac{\partial}{\partial x_j} \tag{1.6}$$

where $\bar{v}_j = \sum_{i=1}^d \tilde{v}_i \frac{\partial \varphi_j^{-1}}{\partial \tilde{x}_i}$ is the coefficient corresponding to the j'th ambient space coordinate. What exactly is φ_j^{-1}? Note that $\iota \circ (\varphi^{-1})$ is a map from $\mathbb{R}^d$ to $\mathbb{R}^m$ and it takes $\varphi(x)$ to $\iota(x)$. It is this that we are writing as φ_j^{-1}.

In matrix-vector form, we can write $\bar{v} = \frac{\mathrm{d}\varphi^{-1}}{\mathrm{d}\tilde{x}}\tilde{v}$, where $\bar{v}$ is a vector that represents the m-dimensional coefficients in the ambient space. This also means $\bar{v}$ lies in the linear subspace spanned by the column vectors of the Jacobian $\frac{\partial \varphi^{-1}}{\partial \tilde{x}_i}$. This linear subspace is isomorphic to $\mathcal{T}_x\mathcal{M}$, which itself is a subspace of $\mathcal{T}_x\mathbb{R}^m$. We refer to this linear subspace as the **tangential linear subspace**. Intuitively, this means a particle traveling at speed $\bar{v}$

and position x can only move tangentially on the surface. Therefore it is restricted to move on the manifold.

A vector field V is a continuous map that assigns a tangent vector to each point on the manifold; that is $V(x) \in \mathcal{T}_x\mathcal{M}$. We abuse the notation a bit and use capital letters to denote both vector fields and vectors. It should be clear in the context whether it is meant to be a function of points on the manifold or not. Such a vector field can also map a smooth function to a function, via the assignment $x \in \mathcal{M} \mapsto V(x)(f) \in \mathbb{R}$. If it maps smooth functions to smooth functions we say that the vector field is smooth. The space of smooth vector fields on $\mathcal{M}$ is denoted by by $\mathfrak{X}(\mathcal{M})$.

1.2 Riemannian metric

A Riemannian manifold $(\mathcal{M}, g)$ is a d-dimensional smooth manifold $\mathcal{M}$ equipped with an inner product $g_x : \mathcal{T}_x\mathcal{M} \times \mathcal{T}_x\mathcal{M} \to \mathbb{R}$ on the tangent space of each $x \in \mathcal{M}$ [Lee, 2018]. g_x is called the metric tensor at x.

A **metric tensor field** is an assignment of a metric tensor to each point x of $\mathcal{M}$; we denote it by g. The metric tensor field g is said to be **smooth** if for any smooth vector fields u and v, $g(U, V)(x) = g_x(U(x), V(x))$ is a smooth function of x. When it is clear from the context, we suppress the subscript for simplicity. Since g is an inner product, we also write $g(u, v) = \langle u, v\rangle_g$.

The Euclidean metric $\bar{g}$ for $\mathbb{R}^m$ is defined as the Euclidean inner product, characterized by the delta function

$$\left\langle \frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j} \right\rangle = \delta_{ij}, \tag{1.7}$$

which is equal to 1 if $i = j$; otherwise it is equal to 0. This means for any $U, V \in T_x\mathcal{M}$,

$$\langle U, V\rangle_{\bar{g}} = \left\langle \sum_{i=1}^m \bar{u}_i \frac{\partial}{\partial x_i}, \sum_{j=1}^m \bar{v}_j \frac{\partial}{\partial x_j} \right\rangle_{\bar{g}} = \sum_{i=1}^m \bar{u}_i \bar{v}_i = \bar{u}^\top \bar{v}. \tag{1.8}$$

Generally, given a set of basis vectors, such as $\tilde{E}_i$, the metric tensor can be represented in a matrix form, via

$$g_{ij} := \langle \tilde{E}_i, \tilde{E}_j\rangle_g \tag{1.9}$$

This allows us to write the metric using the patch coordinates

$$\langle U, V\rangle_g = \sum_{i,j} \tilde{u}_i \tilde{v}_j \langle \tilde{E}_i, \tilde{E}_j\rangle_g = \sum_{i,j} \tilde{u}_i \tilde{v}_j g_{ij} = \tilde{u}^\top G \tilde{v} \tag{1.10}$$

where G is a matrix whose i'th row and j'th column corresponds to g_{ij}.

Using the components of the metric tensor, we can define the **dual basis** $\tilde{E}^i = \sum_j g^{ij}\tilde{E}_j$, where g^{ij} stands for the (i,j)'th entry of the inverse matrix G^{-1}. $(\tilde{E}^1, \cdots, \tilde{E}^d)$ is called the dual basis for $(\tilde{E}_1, \cdots, \tilde{E}_d)$ since they form a bi-orthogonal system, meaning

$$\langle \tilde{E}^i, \tilde{E}_j\rangle_g = \left\langle \sum_k g^{ik}\tilde{E}_k, \tilde{E}_j \right\rangle_g = \sum_k g^{ik}\langle \tilde{E}_k, \tilde{E}_j\rangle_g = \sum_k g^{ik} g_{kj} = (G^{-1}G)_{ij} = \delta_{ij}. \tag{1.11}$$

If $\mathcal{M}$ is a submanifold, e.g. if it is embedded in an ambient space, it automatically inherits the ambient manifold's metric. Suppose $\mathcal{M} \subset \mathbb{R}^m$, where $m > d$ is the dimensionality of the ambient space. Then $g = \iota^*\bar{g}$ is a metric induced by the inclusion map, defined by

$$g_x(u, v) = \bar{g}(d\iota_x(u), d\iota_x(v)).$$

Unwinding the definitions, we have

$$g_{ij} = \left\langle d\iota_x(\tilde{E}_i), d\iota_x(\tilde{E}_j)\right\rangle_{\bar{g}} = \left\langle \sum_{k=1}^m \frac{\partial \varphi_k^{-1}}{\partial \tilde{x}_i}\frac{\partial}{\partial x_k}, \sum_{k'=1}^m \frac{\partial \varphi_{k'}^{-1}}{\partial \tilde{x}_j}\frac{\partial}{\partial x_{k'}} \right\rangle_{\bar{g}} = \sum_{k=1}^m \frac{\partial \varphi_k^{-1}}{\partial \tilde{x}_i}\frac{\partial \varphi_k^{-1}}{\partial \tilde{x}_j}. \tag{1.12}$$

That is, if $\psi = \varphi^{-1}$ is the inverse map of φ, we can write $G = \frac{\mathrm{d}\psi}{\mathrm{d}\tilde{x}}^\top \frac{\mathrm{d}\psi}{\mathrm{d}\tilde{x}}$, which can be equivalently deduced from equating (1.8) and (1.10).

1.2.1 Riemannian gradient and divergence

Riemannian gradient. Another crucial structure closely related to the metric is the **Riemannian gradient**. The definition of Riemannian gradient $\nabla_g : f \in C^\infty(\mathcal{M}) \mapsto \nabla_g f \in \mathfrak{X}(\mathcal{M})$ is motivated by the directional derivative in Euclidean space, satisfying

$$\langle \nabla_g f, V\rangle_g = V(f) \tag{1.13}$$

for any $V \in \mathfrak{X}(\mathcal{M})$.

To obtain an explicit formula for the Riemannian gradient, we expand both sides of (1.13):

$$\langle \nabla_g f, V\rangle_g = \sum_{i,j=1}^d \tilde{u}_i \tilde{v}_j g_{ij} \tag{1.14}$$

where we let $\tilde{u}_i$ and $\tilde{v}_j$ denote the coefficients of the gradient and V respectively. And,

$$V(f) = \sum_{j=1}^{d} \tilde{v}_j \frac{\partial}{\partial \tilde{x}_j} f \circ \varphi^{-1}. \tag{1.15}$$

Since v is arbitrary, this means for all j

$$\sum_{i=1}^{d} \tilde{u}_i g_{ij} = \frac{\partial}{\partial \tilde{x}_j} f \circ \varphi^{-1} \quad \Longrightarrow \quad \tilde{u}_i = \sum_{j=1}^{d} g^{ij} \frac{\partial}{\partial \tilde{x}_j} f \circ \varphi^{-1}. \tag{1.16}$$

Riemannian divergence. We can define the Riemannian divergence using the patch coordinates which we later show has a coordinate-free form and can be computed in the ambient space if the manifold is embedded. The following theorem extends the Stokes theorem to Riemannian manifolds.

[**Divergence theorem**] For any compactly supported $f \in \mathfrak{X}(\mathcal{M})$, $\int_{\mathcal{M}} \nabla_g \cdot f \, \mathrm{d}\mu_g = 0$.

Proof. Let $\{(\Psi_i, U_i)\}$ be a partition of unity. By compactness, we can choose a finite subcover over the support of f, so the index set of i is finite.

$$\int_{\mathcal{M}} \nabla_g \cdot f \, \mathrm{d}\mu_g = \int_{\mathcal{M}} \nabla_g \cdot \left(\sum_i \Psi_i f \right) \mathrm{d}\mu_g \tag{1.17}$$

$$= \sum_i \int_{U_i} \nabla_g \cdot (\Psi_i f) \, \mathrm{d}\mu_g \tag{1.18}$$

$$= \sum_i \int_{\varphi_i(U_i)} \nabla \cdot (|G|^{\frac{1}{2}} \Psi_i f) \circ \varphi^{-1} \, \mathrm{d}\tilde{x}. \tag{1.19}$$

All of the finitely many summands equal 0 by an application of Stokes' theorem in $\mathbb{R}^d$ [Rudin et al., 1976, Theorem 10.33]. This is because the support of $\Psi_i \circ \varphi_i^{-1}$ is contained in $\varphi_i(U_i)$; therefore at the boundary of $\varphi_i(U_i)$, $\Psi_i \circ \varphi_i^{-1}$ is equal to 0.

□

The Riemannian divergence satisfies the following product rule.

[**Product rule**] Assume $V \in \mathfrak{X}(\mathcal{M})$ and $f \in C^\infty(\mathcal{M})$. Then

$$\nabla_g \cdot (fV) = V(f) + f\nabla_g \cdot V. \tag{1.20}$$

Proof. The product rule of the Affine connection (see Appendix 1.3),

$$\nabla_g \cdot (fV) = \sum_{j=1}^{d} \langle \nabla_{\tilde{E}_j}(fV), \tilde{E}^j \rangle_g \tag{1.21}$$

$$= \sum_{j=1}^{d} \langle f\nabla_{\tilde{E}_j} V + \tilde{E}_j(f)V, \tilde{E}^j \rangle_g \tag{1.22}$$

$$= f\sum_{j=1}^{d} \langle \nabla_{\tilde{E}_j} V, \tilde{E}^j \rangle_g + \sum_{j=1}^{d} \tilde{E}_j(f) \left\langle \sum_{j'=1}^{d} \tilde{v}_{j'} \tilde{E}_{j'}, \tilde{E}^j \right\rangle_g \tag{1.23}$$

$$= f\nabla_g \cdot V + \sum_{j,j'=1}^{d} \tilde{E}_j(f)\tilde{v}_{j'} \langle \tilde{E}_{j'}, \tilde{E}^j \rangle_g \tag{1.24}$$

$$= f\nabla_g \cdot V + \sum_{j,j'=1}^{d} \tilde{E}_j(f)\tilde{v}_{j'}\delta_{jj'} \tag{1.25}$$

$$= f\nabla_g \cdot V + \sum_{j}^{d} \tilde{E}_j(f)\tilde{v}_j = f\nabla_g \cdot V + V(f). \tag{1.26}$$

□

[**Expanding Riemannian gradient**] Let V denote the tangential projection matrix. Then for any $f \in C^\infty(\mathcal{M})$

$$\sum_{k=1}^{d} V_k(f)V_k = \nabla_g f. \tag{1.27}$$

1.3 Covariant derivative

An **affine connection** allows us to compare values of a vector field at nearby points. It is a differential operator denoted by $\nabla : \mathfrak{X}(\mathcal{M}) \times \mathfrak{X}(\mathcal{M}) \to$

$\mathfrak{X}(\mathcal{M})$ and written as $U, V \mapsto \nabla_U V$ for $U, V \in \mathfrak{X}(\mathcal{M})$, satisfying the following defining properties:

1. Linearity in U: $\nabla_{fU_1+gU_2} V = f\nabla_{U_1} V + g\nabla_{U_2} V$ for $f, g \in C^\infty(M)$ and $U_1, U_2, V \in \mathfrak{X}(M)$.
2. Linearity in V: $\nabla_U(aV_1 + bV_2) = a\nabla_U V_1 + b\nabla_U V_2$ for $a, b \in \mathbb{R}$ and $U, V_1, V_2 \in \mathfrak{X}(M)$.
3. Product rule: $\nabla_U(fV) = f\nabla_U V + U(f)V$ for $f \in C^\infty(M)$ and $U, V \in \mathfrak{X}(M)$.

$\nabla_U V$ is called the **covariant derivative** of V in the U-direction.

If $U, V \in \mathfrak{X}(\mathbb{R}^m)$, the Euclidean connection $\overline{\nabla}$ is defined as

$$\overline{\nabla}_U V = \sum_{i=1}^{m} \sum_{j=1}^{m} \bar{u}_j \frac{\partial \bar{v}_i}{\partial x_j} \frac{\partial}{\partial x_i}. \tag{1.28}$$

It can be verified that the Euclidean connection is indeed an affine connection.

We can express a connection internally in terms of a coordinate system $\tilde{E}_i$. For any pair of indices i and j, we define the connection coefficients of ∇, denoted by Γ, as d^3 smooth functions satisfying

$$\nabla_{\tilde{E}_i} \tilde{E}_j = \sum_{k=1}^{d} \Gamma_{ij}^k \tilde{E}_k. \tag{1.29}$$

Then for any $U, V \in \mathfrak{X}(\mathcal{M})$, we have

$$\begin{aligned}
\nabla_U V &= \nabla_U \sum_{j=1}^{d} \tilde{v}_j \tilde{E}_j &(1.30)\\
&= \sum_{j=1}^{d} \tilde{v}_j \nabla_U \tilde{E}_j + U(\tilde{v}_j)\tilde{E}_j &(1.31)\\
&= \sum_{i,j=1}^{d} \tilde{u}_i \tilde{v}_j \nabla_{\tilde{E}_i} \tilde{E}_j + \sum_{j=1}^{d} U(\tilde{v}_j)\tilde{E}_j &(1.32)\\
&= \sum_{i,j,k=1}^{d} \tilde{u}_i \tilde{v}_j \Gamma_{ij}^k \tilde{E}_k + \sum_{j=1}^{d} U(\tilde{v}_j)\tilde{E}_j. &(1.33)
\end{aligned}$$

Now given a metric tensor, we say that ∇ is a **Levi-Civita connection of** g if it is

1. Compatible with g: $U(g(V,W)) = g(\nabla_U V, W) + g(V, \nabla_U W)$.
2. Symmetric: $\nabla_u v - \nabla_V U = [U,V]$, where $[U,V] := \sum_{i=1}^{d} U(V_i)\tilde{E}_i - V(U_i)\tilde{E}_i$ is the Lie bracket.

The first condition looks messy but it essentially says that the Levi-Civita connection leaves the metric invariant. It is equivalent to saying that the covariant derivative of g in any direction is zero.

> [**Fundamental Theorem of Riemannian Geometry**] Let $(\mathcal{M}, g)$ be a Riemannian manifold. There exists a unique Levi-Civita connection of g.

See Lee [2018, Theorem 5.10] for proof. The connection coefficients of the Levi-Civita connection are called the **Christoffel symbols** of g. They are symmetric in the lower indices, i.e.$\Gamma_{ij}^k = \Gamma_{ji}^k$. A by-product of the proof of the fundamental theorem is the following identity, which will turn out to be useful in deriving the identity for the Riemannian divergence:

$$\frac{\partial}{\partial \tilde{x}_j} g_{ki} = \sum_{l=1}^{d} \Gamma_{jk}^l g_{li} + \Gamma_{ji}^l g_{lk}. \tag{1.34}$$

An example of a Levi-Civita connection is the Euclidean connection of $(\mathbb{R}^d, \bar{g})$. It can be checked that $\overline{\nabla}$ is both symmetric and compatible with $\bar{g}$. Furthermore, for any d-submanifold $\mathcal{M}$ embedded in $\mathbb{R}^m$ for $m > d$, we can define a **tangential connection**

$$\nabla_U^\top V = P \overline{\nabla}_{\overline{U}} \overline{V} \tag{1.35}$$

for $U, V \in \mathfrak{X}(\mathcal{M})$, where $\overline{U}$ and $\overline{V}$ are any[2] smooth extensions of U and V to $\mathbb{R}^m$. P is the tangential projection defined as

$$(PV)(x) = \sum_{j=1}^{m} (P_x \bar{v})_j \frac{\partial}{\partial x_j} \tag{1.36}$$

for any $V \in \mathfrak{X}(\mathbb{R}^m)$. Recall that P_x is the **orthogonal projection** onto the tangent space spanned by $\frac{\partial \psi}{\partial \tilde{x}_i}$. *The tangential connection $\nabla^\top$ is the Levi-Civita connection on the embedded submanifold $\mathcal{M}$* [Lee, 2018, Proposition 5.12].

[2]The value of the tangential connection is independent of the extensions chosen, so $\nabla^\top$ is well-defined.

2

DISTRIBUTIONS ON RIEMANNIAN MANIFOLDS

We are interested in elevating probability theory that is typically employed in $\mathbb{R}^n$ to Riemannian manifolds. To do so, we will have to generalize many familiar notions and start from the ground up. We will see two complementary of achieving this in the following two sections.

2.1 Perspective 1: Distribution on Manifolds

We start by first defining a probability space $(\Omega, \mathcal{B}(\Omega), \Pr)$, where $\mathcal{B}(\Omega)$ is a Borel σ-algebra and Pr is a measure on $\mathcal{B}(\Omega)$ with the condition $\Pr(\Omega) = 1$. A random point on a Riemannian manifold $\mathcal{M}$ is a Borel measurable function from $\Omega \to \mathcal{M}$. This allows us to induce a probability measure on the manifold itself. To go further we need to first understand the notion of a volume form. Essentially, a volume form Vol is an n-order differential form on an n-dimensional *orientable* manifold. Intuitively, a volume form gives us a mechanism to define integrals on manifolds including a measure for functions that can be integrated using the Lebesgue integral.

Note: This is basically a rehash of Definition 1 in Pennec [2006].

Note: A manifold is orientable if it has a coordinate atlas all of whose transitions have positive Jacobian determinant.

Recall, that in normal Euclidean spaces an infinitesimal volume element is the volume of the parallelepipedon. In two dimensions the area of a parallelogram is found by,

$$\begin{aligned}
\text{Area} &= ||a_1||||a_2|| \sin\theta && (2.37)\\
&= ||a_1||||a_2||\sqrt{1-\cos^2\theta} && (2.38)\\
&= \sqrt{||a_1||_2^2||a_2||_2^2 - \langle a_1, a_2\rangle_2^2}, && (2.39)
\end{aligned}$$

where a_1 and a_2 are the lengths of the adjacent sides and θ is the angle between them. If we stack the vectors in a matrix $A = [a_1, a_2]$ then the area can be expressed as the determinant of the Grammian matrix:

$$\text{Area} = \sqrt{||a_1||_2^2||a_2||_2^2 - \langle a_1, a_2\rangle_2^2} = \sqrt{\det A^T A}. \tag{2.40}$$

The benefit of this construction is that it can be generalized to define the volume of a parallelepipedon in n-dimensions and while the matrix A is not square the Grammian $A^T A$ is guaranteed to be square.

For Riemannian manifolds, we will follow a similar strategy except we will induce this volume form on the tangent space using the Riemannian metric. Concretely, as first outlined in equation 1.9 given a choice of basis vectors for the tangent space at a point on the manifold $\tilde{E} = \{\tilde{E}_1, \ldots, \tilde{E}_n\}$, the Riemannian metric tensor has components:

$$g_{ij} := \langle \tilde{E}_i, \tilde{E}_j \rangle_g \tag{2.41}$$

Without loss of generality assume that the set of basis vectors are orthonormal—if not we can always apply a Gram-Schimdt process to get an orthonormal basis. Then the matrix representation of the Riemannian metric tensor is given by $G = \tilde{E}^T \tilde{E}$. As a result the induced infinitesimal volume form on the tangent space is $d\mathrm{Vol} = \sqrt{\det G} dx$. It is also customary to write this as $d\mathcal{V} = \sqrt{|G(x)|} d\mathcal{M}x$ to make explicit that this is a infinitesimal volume form on $\mathcal{T}_x\mathcal{M}$.

Density in a chart. How can we relate the pdf of a point on $\mathcal{M}$ to that of a chart? Specifically, we $x \in \mathcal{M}$ be a random point on the manifold and let (U, φ) be a chart such that $\tilde{x} = \varphi(x)$. If we have a pdf p_x on $\mathcal{M}$ we can define a corresponding pdf with respect to the Lebesgue measure dx on $\mathbb{R}^n$, ρ_x. The two pdfs for a point y are related via:

$$\rho_x(y) = p_x(y)\sqrt{|G(y)|}, \tag{2.42}$$

where the density ρ_x is dependent on the choice of chart φ while p_x is intrinsic to the manifold.

2.2 Perspective 2: Distribution on Manifolds

An important use of the metric is to define a measure over measurable subsets of the manifold. Let (U, φ) be a chart and consider all functions smooth functions f supported in U. Then

$$f \mapsto \int_{\varphi(U)} (f\sqrt{|\det G|}) \circ \varphi^{-1} \, \mathrm{d}\tilde{x}$$

is a positive linear functional. Since $\mathcal{M}$ is Hausdorff and locally compact, by the *Riesz representation theorem* [Rudin, 1987, Theorem 2.14], there exists a unique Borel measure μ_g (over U) such that $\int_U f \, \mathrm{d}\mu_g$ is equal to the evaluation of the functional above. We can then apply a partition of unity [Lee, 2013, Theorem 2.23] to extend this construction of μ_g to be defined over the entire $\mathcal{M}$, which says that for any open cover $\{U_i\}$ of $\mathcal{M}$, there exists a set of continuous functions Φ_i satisfying the following properties:

1. $0 \leq \Phi_i(x)$ for all $x \in \mathcal{M}$.

2. $\mathrm{supp}\Phi_i \subseteq U_i$.
3. $\sum_i \Phi_i(x) = 1$ for all $x \in \mathcal{M}$.
4. Any $x \in \mathcal{M}$ has a neighborhood that intersects with only finitely many $\mathrm{supp}\Phi_i$.

By means of the partition, we can consider the following positive linear functional instead:

$$f \in C_c(\mathcal{M}) \mapsto \sum_i \int_{\varphi(U_i)} (\Psi_i f \sqrt{|\det G|}) \circ \varphi^{-1} \,\mathrm{d}\tilde{x}, \tag{2.43}$$

which is always well-defined since f is compactly supported in $\mathcal{M}$ (only finitely many summands are non-zero). $\sqrt{|\det G|}$ is called the **volume density**. We write $|G| = |\det G|$ for short. A probability density p over $\mathcal{M}$ can be thought of as a non-negative integrable function satisfying $\int_{\mathcal{M}} p \,\mathrm{d}\mu_g = 1$.

Bibliography

M. Gunther. Isometric embeddings of riemannian manifolds, kyoto, 1990. In *Proc. Intern. Congr. Math.*, pages 1137–1143. Math. Soc. Japan, 1991.

C.-W. Huang, M. Aghajohari, A. J. Bose, P. Panangaden, and A. Courville. Riemannian diffusion models. *arXiv preprint arXiv:2208.07949*, 2022.

J. M. Lee. *Introduction to Smooth Manifolds.* Springer, 2013.

J. M. Lee. *Introduction to Riemannian manifolds.* Springer, 2018.

X. Pennec. Intrinsic statistics on riemannian manifolds: Basic tools for geometric measurements. *Journal of Mathematical Imaging and Vision*, 25(1):127–154, 2006.

W. Rudin. *Real and Complex Analysis.* McGraw-Hill, 1987.

W. Rudin et al. *Principles of mathematical analysis*, volume 3. McGraw-hill New York, 1976.