2

Introduction to Riemannian Geometry

This chapter is intended to help those with little previous exposure to differential geometry by providing a rather informal summary of background for our purposes in the sequel and pointers for those who wish to pursue more geometrical features of the spaces of probability density functions that are our focus in the sequel. In fact, readers who are comfortable with doing calculations of curves and their arc length on surfaces in $\mathbb{R}^3$ could omit this chapter at a first reading.

A topological space is the least structure that can support arguments concerning continuity and limits; our first experiences of such analytic properties is usually with the spaces $\mathbb{R}$ and $\mathbb{R}^n$. A manifold is the least structure that can support arguments concerning differentiability and tangents–that is, calculus. Our prototype manifold is the set of points we call Euclidean n-space $\mathbb{E}^n$ which is based on the real number n-space $\mathbb{R}^n$ and carries the Pythagorean distance structure. Our common experience is that a 2-dimensional Euclidean space can be embedded in $\mathbb{E}^3$, (or $\mathbb{R}^3$) as can curves and surfaces. Riemannian geometry generalizes the Euclidean geometry of surfaces to higher dimensions by handling the intrinsic properties like distances, angles and curvature independently of any environing simpler space.

We need rather little geometry of Riemannian manifolds in order to provide background for the concepts of information geometry. Dodson and Poston [70] give an introductory treatment with many examples, Spivak [194, 195] provides a six-volume treatise on Riemannian geometry while Gray [99] gave very detailed descriptions and computer algebraic procedures using *Mathematica* [215] for calculating and graphically representing most named curves and surfaces in Euclidean $\mathbb{E}^3$ and code for numerical solution of geodesic equations. Our Riemannian spaces actually will appear as subspaces of $\mathbb{R}^n$ so global properties will not be of particular significance and then the formulae and Gray's procedures easily generalize to more variables.

K. Arwini, C.T.J. Dodson, *Information Geometry.*
Lecture Notes in Mathematics 1953,

2.0.2 Manifolds

A smooth n-manifold M is a (Hausdorff) topological space together with a collection of smooth maps (the charts)

$$\{\phi_\alpha : U_\alpha \longrightarrow \mathbb{R}^n \mid \alpha \in A\}$$

from open subsets U_α of M, which satisfy:

i) $\{U_\alpha \mid \alpha \in A\}$ is an open cover of M;
ii) each ϕ_α is a homeomorphism onto its image;
iii) whenever $U_\alpha \cap U_\beta \neq \emptyset$, then the maps between subsets of $\mathbb{R}^n$

$$\phi_\alpha \circ \phi_\beta^{-1} : \phi_\beta(U_\alpha \cap U_\beta) \longrightarrow \phi_\alpha(U_\alpha \cap U_\beta)\,,$$

$$\phi_\beta \circ \phi_\alpha^{-1} : \phi_\alpha(U_\alpha \cap U_\beta) \longrightarrow \phi_\beta(U_\alpha \cap U_\beta)\,,$$

have continuous derivatives of all orders (are C^∞ or smooth).

We call $\{(U_\alpha, \phi_\alpha) \mid \alpha \in A\}$ an atlas of charts for M; the properties of M are not significantly changed by adding more charts. The simplest example is the n-manifold $\mathbb{R}^n$ with atlas consisting of one chart, the identity map.

Intuitively, an n-manifold consists of open subsets of $\mathbb{R}^n$, the $\phi_\alpha(U_\alpha)$, pasted together in a smooth fashion according to the directions given by the $\phi_\alpha \circ \phi_\beta^{-1}$. For example, the unit circle $\mathbb{S}^1$ with its usual structure can be presented as a 1-manifold by pasting together two open intervals, each like $(-\pi, \pi)$. Similarly, the unit 2-sphere $\mathbb{S}^2$ has an atlas consisting of two charts

$$\{(U_N, \phi_N),\ (U_S, \phi_S)\}$$

where U_N consists of $\mathbb{S}^2$ with the north pole removed, U_S consists of $\mathbb{S}^2$ with the south pole removed, and the chart maps are stereographic projections. Thus, if $\mathbb{S}^2$ is the unit sphere in $\mathbb{R}^3$ centered at the origin then:

$$\phi_N : \mathbb{S}^2 \setminus \{n.p.\} \longrightarrow \mathbb{R}^2 : (x, y, z) \longmapsto \frac{1}{1+z}(x, y)$$

$$\phi_S : \mathbb{S}^2 \setminus \{s.p.\} \longrightarrow \mathbb{R}^2 : (x, y, z) \longmapsto \frac{1}{1-z}(x, y)\,.$$

Similar chart maps work also for the higher dimensional spheres.

2.0.3 Tangent Spaces

From elementary analysis we know that the derivative of a function is a linear approximation to that function, at the chosen point. Thus, we need vector spaces to define linearity and these are automatically present in the form of the vector space $\mathbb{R}^n$ at each point of Euclidean point space $\mathbb{E}^n$. At each point x of a manifold M we construct a vector space T_xM, called the tangent space to M at x. For this we employ equivalence classes $[T_{\phi_\alpha(x)}\mathbb{R}^n]$ of tangent spaces to the images of x, $\phi_\alpha(x)$, under chart maps defined at x. That is, we borrow the vector space structure from $\mathbb{R}^n$ via each chart (U_α, ϕ_α) with $x \in U_\alpha$, then identify the borrowed copies. The result, for $x \in \mathbb{S}^2$ embedded in $\mathbb{R}^3$, is simply a vector space isomorphic to the tangent plane to $\mathbb{S}^2$ at x. This works here because $\mathbb{S}^2$ embeds isometrically into $\mathbb{R}^3$, but not all 2-manifolds embed in $\mathbb{R}^3$, some need more dimensions; the Klein bottle is an example [70]. Actually, the formal construction is independent of M being embedded in this way; however, the Whitney Embedding Theorem [211] says that an embedding of an n-manifold is always possible in $\mathbb{R}^{2n+1}$.

Once we have the tangent space T_xM for each $x \in M$ we can present it in coordinates, via a choice of chart, as a copy of $\mathbb{R}^n$. The derivatives of the change of chart maps, like

$$\frac{\partial}{\partial x^i_\beta}(\phi_\alpha \circ \phi_\beta^{-1})\ (x^1_\beta, x^2_\beta, \cdots, x^n_\beta)\,,$$

provide linear transformations among the representations of T_xM. Next, we say that a map between manifolds

$$f : M \longrightarrow N$$

is differentiable at $x \in M$, if for some charts (U, ϕ) on M and (V, ψ) on N with $x \in U,\ f(x) \in V$, the map

$$\psi \circ f|_U \circ \phi^{-1} : \phi(U) \longrightarrow \psi(V)$$

is differentiable as a map between subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$, if M is an n-manifold and N is an m-manifold. This property turns out to be independent of the choices of charts, so we get a linear map

$$T_xf : T_xM \longrightarrow T_{f(x)}N\,.$$

Moreover, if we make a choice of charts then T_xf appears in matrix form as the set of partial derivatives of $\psi \circ f \circ \phi^{-1}$. The notation T_xf for the derivative of f at x is precise, but in many texts it may be found abbreviated to Df, f_*, f' or Tf, with or without reference to the point of application. When f is a curve in M, that is, a map from some interval

$$f : [0, 1] \longrightarrow M : t \mapsto f(t)\,,$$

then T_tf is sometimes denoted by $\dot{f}_t$. This is the tangent map to f at t and the result of its application to the standard unit vector to $\mathbb{R}$ at t, $\dot{f}_t(\hat{1})$, is the tangent vector to f at t . It is quite common for this tangent vector also to be abbreviated to $\dot{f}_t$.

In a natural way we can provide a topology and differential structure for the set of all tangent vectors in all tangent spaces to an n-manifold M:

$$TM = \bigcup_{x \in M} T_x M \,;$$

details are given in [70]. So, it actually turns out that TM is a $2n$-manifold, called the tangent bundle to M. For example, if $M = \mathbb{R}^n$ then $TM = \mathbb{R}^n \times \mathbb{R}^n$. Similarly, if $M = \mathbb{S}^1$ with the usual structure then TM is topologically (and as a manifold) equivalent to the infinite cylinder $\mathbb{S}^1 \times \mathbb{R}$. The technical term for an n-manifold M that has a trivial product tangent bundle $TM \cong M \times \mathbb{R}^n$ is parallelizable and this property is discussed in the cited texts.

On the other hand, this simple situation is quite rare and it is rather a deep result that for spheres

$$T\mathbb{S}^n \text{ is equivalent to } \mathbb{S}^n \times \mathbb{R}^n \text{ only for } n = 1, 3, 7\,.$$

For other spheres, their tangent bundles consist of twisted products of copies of $\mathbb{R}^n$ over $\mathbb{S}^n$. In particular, $T\mathbb{S}^2$ is such a twisted product of $\mathbb{S}^2$ with one copy of $\mathbb{R}^2$ at each point. An intuitive picture of a 2-manifold that is a twisted product of $\mathbb{R}^1$ (or an interval from it) over $\mathbb{S}^1$ is a Möbius strip, which we know does not embed into $\mathbb{R}^2$ but does embed into $\mathbb{R}^3$.

A map $f : M \to N$ between manifolds is just called differentiable if it is differentiable at every point of M, and a diffeomorphism if it is differentiable with a differentiable inverse; in the latter case M and N are said to be diffeomorphic manifolds. Diffeomorphism implies homeomorphism, but not conversely. For example, the sphere $\mathbb{S}^2$ is diffeomorphic to an ellipsoid, but only homeomorphic to the surface of a cube because the latter is not a smooth manifold: it has corners and sharp edges so no well-defined tangent space structure. We note one generalisation however, sometimes we want a smooth manifold to have a boundary. For example a circular disc obviously cannot have its edge points homeomorphic to open sets in $\mathbb{R}^2$; so we relax our definition for charts to allow the chart maps to patch together open subsets like $\{(x, y) \in \mathbb{R}^2 | 0 < x \leq 1, 0 < y, < 1\}$ to deal with edge points. This is easily generalized to higher dimensions.

2.0.4 Tensors and Forms

For finite-dimensional real vector spaces it is easily shown that the set of all real-valued linear maps on the space is itself a real vector space, the dual space and similarly multilinear real-valued maps form real vector spaces; multilinear real-valued maps are called tensors. Elementary linear algebra introduces the notion of a real vector space X and its dual space X^* of real-valued linear functions on X; on manifolds we combine these types of spaces in a smooth way using tensor and exterior products to obtain the necessary composite bundle structures that can support the range of multilinear operations needed for geometry. Exterior differentiation, is the fundamental operation in the calculus on manifolds and it recovers all of vector calculus in $\mathbb{R}^3$ and extends it to arbitrary dimensional manifolds.

An m-form is a purely antisymmetric, real-valued, multilinear function on an argument of m tangent vectors, defined smoothly over the manifold. The space of m-forms becomes a vector bundle $\Lambda^m M$ over M with coordinate charts induced from those on M. A 0-form is a real valued function on the manifold. Thus, the space $\Lambda^0 M$ of 0-forms on M consists of sections of the trivial bundle $M \times \mathbb{R}$. The space $\Lambda^1 M$ of 1-forms on M consists of sections of the cotangent bundle T^*M, and $\Lambda^k M$ consists of sections of the antisymmetrized tensor product of k copies of T^*M. Locally, a 1-form has the local coordinates of an n-vector, a 2-form has the local coordinates of an antisymmetric $n \times n$ matrix. A k-form on an n-manifold has $\binom{n}{k}$ independent local coordinates. It follows that the only k-forms for $k > n$ are the zero k-forms. We summarize some definitions.

There are three fundamental operations on finite-dimensional vector spaces (in addition to taking duals): direct sum $\oplus$, tensor product $\otimes$, and exterior product $\wedge$ on a space with itself. Let F, G be two vector spaces, of dimensions n, m respectively. Take any bases $\{b_1, \cdots, b_n\}$ for F, $\{c_1, \cdots, c_m\}$ for G, then we can obtain bases

$$\{b_1, \cdots, b_n,\ c_1, \cdots, c_m\} \quad \text{for} \quad F \oplus G\,,$$
$$\{b_i \otimes c_j \ \mid\ i = 1, \cdots, n;\ j = 1, \cdots, m\} \quad \text{for} \quad F \otimes G\,,$$
$$\{b_i \wedge b_j \ =\ b_i \otimes b_j - b_j \otimes b_i \ \mid\ i = 1, \cdots, n;\ i < j\} \quad \text{for} \quad F \wedge F\,.$$

So, $F \oplus G$ is essentially the disjoint union of F and G with their zero vectors identified. In a formal sense (*cf.* Dodson and Poston [70], p. 104), $F \otimes G$ can be viewed as the vector space $L(F^*, G)$ of linear maps from the dual space $F^* = L(F, \mathbb{R})$ to G. Recall also the natural equivalence $(F^*)^* \cong F$. By taking the antisymmetric part of $F \otimes F$ we obtain $F \wedge F$. We deduce immediately:

$$\begin{aligned}
\dim F \oplus G &= \dim F + \dim G\,,\\
\dim F \otimes G &= \dim F \cdot \dim G\,,\\
\dim F \wedge F &= \frac{1}{2}\,\dim F(\dim F - 1)\,.
\end{aligned}$$

Observe that only for $\dim F = 3$ can we have $\dim F = \dim(F \wedge F)$. Actually, this is the reason for the existence of the vector cross product $\times$ on $\mathbb{R}^3$ only, giving the uniquely important isomorphism

$$\mathbb{R}^3 \wedge \mathbb{R}^3 \longrightarrow \mathbb{R}^3 : x \wedge y \longmapsto x \times y$$

and its consequences for geometry and vector calculus on $\mathbb{R}^3$.

Each of the operations $\oplus, \otimes$ and $\wedge$ induces corresponding operations on linear maps between spaces. Indeed, the operations are thoroughly universal and categorical, so they should and do behave well in linear algebraic contexts. Briefly, suppose that we have linear maps $f, h \in L(F, J)$ $g \in (G, K)$ then the induced linear maps in $L(F \oplus G,\ J \oplus K)$, $L(F \otimes G,\ J \otimes K)$ and $L(F \wedge F,$ $J \wedge J)$ are

$$\begin{aligned} f \oplus g : x \oplus y &\longmapsto f(x) \oplus g(y)\,, \\ f \otimes g : x \otimes y &\longmapsto f(x) \otimes g(y)\,, \\ f \wedge h : x \wedge y &\longmapsto f(x) \wedge h(y)\,. \end{aligned}$$

Local coordinates about a point in M induce bases for the tangent vector spaces and their spaces. The construction of the tangent spaces, directly from the choice of the differentiable structure for the manifold, induces a definite role for tangent vectors. An element $v \in T_x M$ turns out to be a derivation on smooth real functions defined near $x \in M$. In a chart about x, v is expressible as a linear combination of the partial derivations with respect to the chart coordinates $x^1, x^2, \ldots, x^n$ as

$$v \;=\; v^1 \partial_1 + v^2 \partial_2 + \cdots + v^n \partial_n$$

with $\partial_i = \frac{\partial}{\partial x^i}$, for some $v^i \in \mathbb{R}$.

This is often abbreviated to $v = v^i \partial_i$, where summation is to be understood over repeated upper and lower indices, the summation convention of Einstein. The dual base to $\{\partial_i\}$ is written $\{dx^i\}$ and defined by

$$dx^j(\partial_i) = \delta_i^j = \begin{cases} 1 & \text{if } i = j\,, \\ 0 & \text{if } i \neq j\,. \end{cases}$$

So a 1-form $\alpha \in T_x^* M$ is locally expressible as

$$\alpha = \alpha_1 dx^1 + \alpha_2 dx^2 + \cdots + \alpha_n dx^n = \alpha_i dx^i$$

for some $\alpha_i \in \mathbb{R}$, but a 2-form γ as

$$\gamma = \sum_{i<j} \gamma_{ij} dx^i \wedge dx^j$$

for some $\gamma_{ij} \in \mathbb{R}$. The common summation convention here is $\gamma = \gamma_{[ij]} dx^i \wedge dx^j$. A symmetric 2-tensor would use (ij).

Since the ∂_i and dx^i are well-defined in some chart (U, ϕ) about x, they serve also as basis vectors [70] at other points in U. Hence, they act as basis fields for the restrictions of sections of $TM \to M$ and $T^*M \to M$ to U, generating thereby local basis fields for sections of all tensor product bundles $T^k_m M \to M$ and exterior product bundles of forms $\Lambda^k M \to M$, restricted to U. The spaces of bases or frames form a structure called the frame bundle over a manifold, details of its geometry may be found in Cordero, Dodson and deLeon [43].

Given two vector fields u, v on M their commutator or Lie bracket is the new vector field $[u, v]$ defined as a derivation on real functions f by

$$[u, v](f) = u(v(f)) - v(u(f)) \,.$$

Locally in coordinates using basis fields, for $u = u^i \partial_i$ and $v = v^j \partial_j$,

$$[u, v] = (u^i \partial_i v^j - v^i \partial_i u^j)\partial_j \,.$$

The exterior derivative is a linear map on k-forms satisfying

(i) $d : \Lambda^k M \to \Lambda^{k+1} M$ (d has degree $+1$);
(ii) $df = \operatorname{grad} f$ if $f \in \Lambda^0 M$ (locally, $df = \partial_i f \, dx^i$);
(iii) if $\alpha \in \Lambda^a M$ and $\beta \in \Lambda^* M$, then

$$d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^a \alpha \wedge d\beta \,;$$

(iv) $d^2 = 0$.

This d is unique in satisfying these properties.

2.0.5 Riemannian Metric

We recall the importance of inner products on vector spaces—these allow the definition of lengths or norms of vectors and angles between vectors. The corresponding entity for the tangent vectors to an n-manifold M is a smooth choice of inner products over its family of vector spaces $\{T_x M \mid x \in M\}$. Such a smooth choice is called a Riemannian metric on M. Formally, a Riemannian metric g on n-manifold M is a smooth family of maps

$$g|_x : T_x M \times T_x M \to \mathbb{R}, \; x \in M$$

that is bilinear, symmetric and positive definite on each tangent space. Then we call the pair (M, g) a Riemannian n-manifold. Locally, at each $x \in M$, each $g|_x$ appears in coordinates as a symmetric $n \times n$ matrix $[g_{ij}]$ that is positive definite, so it has positive determinant. For each $v \in T_x M$, the norm of v is defined to be $||v|| = \sqrt{g(v, v)}$.

We can measure the angle θ between any two vectors u, v in the same tangent space by means of

$$\cos\theta = \frac{g(u, v)}{\sqrt{g(u, u)\, g_(v, v)}}.$$

For a smooth curve in (M, g)

$$c : [0, 1] \longrightarrow M : t \longmapsto c(t)$$

with tangent vector field

$$\dot{c} : [0, 1] \longrightarrow TM : t \longmapsto \dot{c}(t)$$

the arc length is the integral of the norm of its tangent vector along the curve:

$$L_c(t) = \int_0^1 \sqrt{g_{c(t)}(\dot{c}(t), \dot{c}(t))}\, dt\,.$$

The arc length element ds along a curve can be expressed in terms of coordinates (x^i) by

$$ds^2 = \sum_{i,j} g_{ij}\, dx^i\, dx^j \tag{2.1}$$

which is commonly abbreviated to

$$ds^2 = g_{ij}\, dx^i\, dx^j \tag{2.2}$$

using the convention to sum over repeated indices.

Arc length is often difficult to evaluate analytically because it contains the square root of the sum of squares of derivatives. Accordingly, we sometimes use the 'energy' of the curve instead of length for comparison between nearby curves. Energy is given by integrating the square of the norm of $\dot{c}$

$$E_c(a,b) = \int_a^b ||\dot{c}(t)||^2\, dt. \tag{2.3}$$

A diffeomorphism f between Riemannian manifolds $(M, g),\ (N, h)$ is called an isometry if its derivative Tf preserves the norms of all tangent vectors: $g(v, v) = h(Tf(v), Tf(v))$. A situation of common interest is when a manifold can be isometrically embedded as a submanifold of some Euclidean $\mathbb{E}^m$ or of $\mathbb{R}^m$ with some specified metric. Note that if we have a Riemannian manifold (M, g) then an open subset X of M inherits a manifold structure using the restriction of chart maps and the metric g induces a subspace metric $g|_X$ so $(X, g|_X)$ becomes a Riemannian submanifold of (M, g). For example, the unit sphere $\mathbb{S}^2$ in $\mathbb{E}^3$ inherits the subspace metric from the Euclidean metric but of course $\mathbb{S}^2$ has spherical not Euclidean geometry. Evidently, the dimension of a submanifold will not exceed the dimension of its host manifold.

2.0.6 Connections

In order to compare tangent vectors at different points along a curve in a manifold M we need to have a procedure that transports tangent space vectors along the curve, so providing a way to 'connect up' unambiguously the tangent spaces passed through. A smooth assignation of tangent vectors along a curve is called a vector field along the curve; one such field is the actual field of tangents to the curve. A suitable connecting entity in the limiting case at a point defines a derivative of a vector field with respect to the tangent to the curve, and gives the result as another tangent vector at the same point. Now, every tangent vector $u \in T_xM$ can be realised as the tangent vector to a curve through x and therefore we finish up with a smooth family of bilinear maps $\nabla = \{\nabla|_x, x \in M\}$ with the property

$$\nabla|_x : T_x \times T_x \to T_x : (u, v) \mapsto \nabla_u v, \quad \text{defined over } x \in M. \tag{2.4}$$

In coordinates, we have a basis of $T_x M$ given by the derivations (∂_i) and so for some real components (u^i), (v^j), using the summation convention for repeated indices and (∂_i) as basis vector fields $u = u^i \partial_i$, $v = v^j \partial_j$ and then

$$\nabla_u v = (u^i \partial_i v^j + u^k v^m \Gamma^j_{km})\partial_j \tag{2.5}$$

for a smooth $n \times n \times n$ array of functions Γ^j_{km} called the Christoffel symbols. It turns out that ∇ provides a derivative for vector valued maps on the manifold, that is of vector field maps $v : M \to TM$, and returns the answer as another vector field; this derivation operator is called the covariant derivative. The smooth family of bilinear maps (2.4) is called a linear connection and there are many ways to formalise its definition [70]. The important theorem here is that for a given Riemannian manifold there is a unique linear connection that preserves the metric and has symmetric Christoffel symbols, this is the Levi-Civita or symmetric metric connection.

Now, we have seen above §2.0.3 that the derivative of a smooth map between manifolds $f : M \to N$ gives a corresponding map $Tf : TM \to TN$. Also, a vector field v on M, is a section $v : M \to TM$ of the tangent bundle projection $\pi : TM \to M$; this means that $\pi \circ v$ is the identity map on M. Therefore the derivative of the vector field will not be another vector field but a map $Tv : TM \to TTM$. This is why we need the connection, it provides a projection of a derivative Tv back onto the the tangent bundle; the covariant derivative of a vector field is precisely the projection of a derivative.

Formally, a linear connection ∇ gives a smooth bundle splitting at each $u \in TTM$ of the space $T_u TM$ into a direct sum

$$T_u TM \cong H_u TM \oplus V_u TM$$

where $V_u TM = \ker(T\pi : T_u TM \to T_{\pi(u)} M)$. We call $H_u TM$ the horizontal subspace (of TTM) at $u \in TM$ and $V_u TM$ the vertical subspace at $u \in TM$. They comprise the horizontal and vertical subbundles, respectively, of TTM.

$$TTM = HTM \oplus VTM.$$

For our purposes, the important role of a connection is that it induces isomorphisms called horizontal lifts from tangent spaces on the base M to horizontal subspaces of the tangent spaces to TM:

$$\uparrow \; : T_{\pi(u)} M \longrightarrow H_u TM \subset T_u TM : v \longmapsto v^{\uparrow}.$$

Technically, a connection splits the exact sequence of vector bundles

$$0 \longrightarrow VTM \longrightarrow TTM \longrightarrow TM \longrightarrow 0$$

by providing a bundle morphism $TM \to TTM$ with image the bundle of horizontal subspaces.

Along any curve $c : [0,1) \to M$ in M we can construct through each $u_0 \in \pi^{-1}(c(0)) \subset TM$ a unique curve $c^{\uparrow} : [0,1) \longrightarrow TM$ with horizontal tangent vector and $\pi \circ c^{\uparrow} = c$, $c^{\uparrow}(0) = u_0$. The map

$$\tau_t : \pi^{-1}(c(0)) \longrightarrow \pi^{-1}(c(t)) : u_0 \longmapsto c^{\uparrow}(t)$$

defined by the curve is called parallel transport along c. Parallel transport is always a linear isomorphism. An associated parallel transport map satisfies $\widetilde{\tau}_t \circ v(c(t)) = v(c(t))$. The covariant derivative of v along c is defined to be the limit, if it exists

$$\lim_{h \to 0} \frac{1}{h} \big(\widetilde{\tau}_h^{-1} \circ v(c(t+h)) - v(c(t)) \big)$$

and is usually denoted by $\nabla_{\dot{c}(t)} v$. Using integral curves c, this extends easily to $\nabla_w v$ for any vector field w. Evidently, the operator ∇ is linear and a derivation:

$$\nabla_w (u+v) = \nabla_w u + \nabla_w v \qquad \text{and} \qquad \nabla_w (fv) = w(f)v + f \nabla_w v\,;$$

it measures the departure from parallelism. The local appearance of ∇ on basis fields (∂_i) about $x \in M$ is

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

where the Γ_{ij}^k are the Christoffel symbols defined earlier.

For a linear connection we define two important tensor fields in terms of their action on tangent vector fields: the torsion tensor field T is defined by

$$T(u,v) = \nabla_u v - \nabla_v u - [u,v]$$

and the curvature tensor field is the section of $T_3^1 M$ defined by

$$R(u,v)w = \nabla_u \nabla_v w - \nabla_v \nabla_u w - \nabla_{[u,v]} w\,.$$

The connection is called torsion-free or symmetric when $T = 0$ and flat when $R = 0$.

In local coordinates with respect to base fields (∂_i),

$$\begin{aligned} T(\partial_j, \partial_k) &= (\Gamma_{jk}^i - \Gamma_{kj}^i)\partial_i\,, \\ R(\partial_k, \partial_l)\partial_j &= (\partial_k \Gamma_{lj}^i - \partial_l \Gamma_{kj}^i + \Gamma_{lj}^h \Gamma_{kh}^i - \Gamma_{kj}^h \Gamma_{lh}^i)\partial_i\,. \end{aligned}$$

The connection form ω is an $\mathbb{R}^{n^2}$-valued linear function on vector fields and is expressible as a matrix valued 1-form with components

$$\omega_j^i = \Gamma_{jk}^i \, dx^k\,. \tag{2.6}$$

Hence

$$\begin{aligned} d\omega^i_j &= d(\Gamma^i_{jk}) \wedge dx^k \\ &= \partial_r \, \Gamma^i_{jk} \, dx^r \wedge dx^k \\ \omega^i_h \wedge \omega^h_j &= \Gamma^i_{hr} \, \Gamma^h_{jk} \, dx^r \wedge dx^k \end{aligned}$$

The curvature form Ω is an $\mathbb{R}^{n^2}$-valued antisymmetric bilinear function on pairs of vector fields and it has the local expression

$$\begin{aligned} \Omega^i_j &= \frac{1}{2} R^i_{jrk} \, dx^r \wedge dx^k \\ &= R^i_{jrk} \, dx^r \wedge dx^k \,. \end{aligned}$$

2.1 Autoparallel and Geodesic Curves

A curve $c : [0,1) \to M$ that has a parallel tangent vector field $\dot{c} = \dot{c}^j \partial_j$ satisfies:

$$\nabla_{\dot{c}(t)} \dot{c}(t) = 0 \tag{2.7}$$

which in coordinate components from (2.5) becomes

$$\ddot{c}^i + \Gamma^i_{jk} \, \dot{c}^j \, \dot{c}^k = 0 \quad \text{for each } i.$$

It is then called an autoparallel curve . In the case that the connection ∇ is the Levi-Civita connection of a Riemannian manifold (M, g), all the parallel transport maps are actually isometries and then the autoparallel curves c satisfying (2.7) are called geodesic curves (cf. [70] for more discussion of geodesic curves). Geodesic curves have extremal properties—between close enough points they provide uniquely shortest length curves. For example, in Euclidean $\mathbb{E}^3$ the geodesics are straight lines and so provide shortest distances between points; on the standard unit sphere $\mathbb{S}^2 \subset \mathbb{E}^3$ the geodesics are arcs of great circles and so between pairs of points the two arcs provide maximal and minimal geodesic distances.

2.2 Universal Connections and Curvature

A connection, §2.0.6 encodes geometrical choices, and through its curvature, underlying topological information. In some situations, both in geometry and in theoretical physics, it is necessary to consider a family of connections, for example with regard to stability of certain properties [36]. Also, it is common for statisticians to consider a number of linear connections on a given statistical manifold and so it can be important to be able to handle these connections as a geometrical family of some kind.

In general, the space of linear connections on a manifold is infinite dimensional, but Mangiarotti and Modugno [140, 152] introduced the idea of a system (or structure) of connections which gives a representation of the space of linear connections as a finite dimensional bundle. On this system there is a 'universal' connection and corresponding 'universal' curvature; then all linear connections and their curvatures are pullbacks of these universal objects.

A full account of the underlying geometry of jet bundles and their morphisms is beyond our present scope so we refer the interested reader to Mangiarotti and Modugno [140, 152]. Dodson and Modugno [69] provided a universal calculus for this context. An application of universal linear connections to a stability problem in spacetime geometry was given by Canarutto and Dodson [36] and further properties of the system of linear connections were given by Del Riego and Dodson [53]. An explicit set of geometrical examples with interesting topological properties was provided by Cordero, Dodson and Parker [44]. The first application to information geometry was given by Dodson [59] for the system of α-connections.

The technical details would take us too far from our present theme but our recent results on statistical manifolds are given in Arwini, Del Riego and Dodson [16]. There we describe the system of all linear connections on the manifold of exponential families, using the tangent bundle, §2.0.3, to give the system space. We provide formulae for the universal connections and curvatures and give an explicit example for the manifold of gamma distributions, §3.5. It seems likely that there could be significant developments from the results on universal connections for exponential families §3.2, for example in the context of group actions on random variables.

Springer

http://www.springer.com/978-3-540-69391-8

Information Geometry
Near Randomness and Near Independence
Arwini, K.; Dodson, C.T.J.
2008, X, 260 p. 103 illus., Softcover
ISBN: 978-3-540-69391-8