

Information Geometry

G H Lathrom

October 2, 2021

1 Background

Information Geometry is a area which combines Differential Geometry, Probability Theory and Information Theory into one discipline. Because of this, it is somewhat difficult to get a grasp on all of the concepts which go into Information Theory. By molding these three disciplines together we also see connections between Physics and Machine Learning also with implications for Artificial Intelligence. Let's just see where all of these things take us, while supplementing necessary background along the way. One of the books of interest will be [AD08] along with a collection of papers available here. We begin with the connection between Information Geometry and Mechanics using [?] as the guide.

$$\Gamma_{k}^{ij} = A \cdots$$

2 Parallel Transport

Suppose we are given a manifold M with a coordinate chart

$$x : U \subseteq \mathbb{R}^n \rightarrow M.$$

The tangent space at a point $p \in M$ is a vector spanned by the vectors

$$\partial_i = \left. \frac{\partial x}{\partial u^i} \right|_p$$

or the column space of the Jacobian matrix, J , evaluated at the point p , denoted $T_p(M)$.

We have n of the vectors ∂_i which may live in some unknown dimensional vector space in which the manifold is embedded. In order to simplify things we define a metric tensor $g = J^t J$ which gives a linear transformation

$$g : T_p(M) \times T_p(M) \rightarrow \mathbb{R}.$$

This means that our metric tensor can actually be viewed as an element

$$g = g_{ij} dx^i \otimes dx^j \in T_p^*(M) \otimes T_p^*(M).$$

For tangent vectors $v^k \partial_k$ and $w^l \partial_l$ we have

$$\begin{aligned}
 g(v, w) &= g_{ij} dx^i \otimes dx^j (v^k \partial_k, w^l \partial_l) \\
 &= g_{ij} v^k w^l dx^i (\partial_k) dx^j (\partial_l) \\
 &= g_{ij} v^k w^l \delta_k^i \delta_l^j \\
 &= g_{ij} v^i w^j = \langle v | w \rangle_g.
 \end{aligned}$$

Because each tangent space is attached at a point to the manifold, there is no natural way to move a vector from one tangent space to another. The mechanism which does this is found in a natural way and is called a **parallel transport**.

If we are in the space \mathbb{R}^m then we simply move the tangent vector from the point p to the point q . The problem with this is that the tangent space, has moved as we have gone from one point to another. Therefore, what we need in order to move one tangent space to another is a projection onto the tangent space at q . This will give us a linear map $T_p(M) \rightarrow T_q(M)$.

Definition 1. A linear operator $P : V \rightarrow V$ is a **projection** if $P^2 = P$.

If M were an n -sphere, then for any point $p = (p^1, p^2, \dots, p^n)$ then the point diametrically opposed to p , say $q = (-p^1, -p^2, \dots, -p^n)$ will have parallel tangent spaces.

If a tangent vector happens to be in both tangent spaces, then we would want the image under the projection to be the identity map.

References

- [AD08] K. Arwini and C. Dodson. *Information Geometry*. Lecture Notes in Mathematics 1953. Springer, 2008.