What does the Cauchy Distribution "Look Like": Metric, Curvature and Geodesics on its Statistical Manifold

1 Abstract

A family of statistical distributions can be associated to a manifold using the Fisher-information metric. Given by the equation below, the Cauchy Distribution is an important distribution that arises in the study of forced resonance.

$$f(x;\gamma,x_0)=\frac{1}{\pi\gamma\left[1+\left(\frac{x-x_0}{\gamma}\right)^2\right]},\qquad \gamma\in\mathbb{R}^+,x_0\in\mathbb{R}$$

In this paper, we consider the statistical manifold defined by the Fisher-information metric on the family of Cauchy Distributions, and ask what the manifold 'looks like'. In order to give a geometric description of this statistical manifold, we work towards finding the Gaussian Curvature and the geodesics on the manifold.

2 Introduction

Information Geometry is a young field pioneered largely pioneered by Japanese mathematicians Amari and Nagaoka, and much of the work in the field stems from the seminal text on the subject [1]. Information Geometry is a branch of Differential Geometry that is concerned with manifolds whose points are probability distribution. It provides an insightful way to connect geometry and statistics.

This scope of manifolds can be naturally equipped with the Fisher-information metric. This metric arises from the Fisher-information matrix which is often used to compare probability distributions, and it naturally satisfies the conditions of a Riemannian metric. From this, curvatures may be calculated, and other properties of these manifolds may be explored.

In this paper, we explore the structure relating to the family of distributions known as the Cauchy distribution and given by the following equation:

$$f(x;\gamma,x_0)=\frac{1}{\pi\gamma\left[1+\left(\frac{x-x_0}{\gamma}\right)^2\right]},\qquad \gamma\in\mathbb{R}^+,x_0\in\mathbb{R} \tag{1}$$

This distribution arises in Brownian motion and as the ratio of two normally distributed variables.

An aside on notation. We will assume a probability distribution is a function $p : \chi \to \mathbb{R}$ for a set χ and $p(x) \geq 0$, $\int p(x)dx = 1$. If p depends on real parameters then we write $p(x;\xi)$ for $\xi \in \mathbb{R}^n$, and at times it may be convenient to suppress $x \in \chi$ and write $p(x;\xi) := p_\xi$.

3 Geometric Structure of the Cauchy Distribution

We begin by giving a rigorous definition for the main object of study within Information Geometry.

Definition 1. *Let M be a family of n-parametrized probability distributions, $p(-;\xi) : \chi \to \mathbb{R}$, such that*

$$M = \{p(x;\xi) : \xi = (\xi_1, \dots, \xi_n) \in X \subset \mathbb{R}^n\},$$

*and $\xi \to p(-;\xi)$ is injective. Then we call M a **statistical model**. If M further satisfies the properties of a differentiable manifold, then we may call it a **statistical manifold** (and if dim$(M) = n = 2$ then we call M a **statistical surface**).*

Now we may consider the family of Cauchy distributions $S = \{f(x;\gamma,x_0) : \gamma \in \mathbb{R}, x_0 \in \mathbb{R}^+\}$ with $f(x;\gamma,x_0)$ as in (1). Clearly (γ, x_0) is the upper half plane of $\mathbb{R}^2$ and is thus an open set of the plane. Furthermore, $(\gamma, x_0) \to f(-,\gamma,x_0)$ is injective and thus S is a statistical model. We note that $f(x;\gamma,x_0)$ and thus, if some fixed point x' has $f(x';\gamma,x_0)$ for some γ, x_0, it has a positive probability of happening for all γ, x_0. While seemingly innocuous, this condition is important as the support of f being constant makes the analysis much easier.

Now, we would like to show that S is a differentiable surface. First, we may endow S with some Hausdorff topology, say the quotient topology from $\mathbb{R}^2$ (it may later make sense to define the topology using the Fisher metric). Clearly, there is a bijective correspondence between $f(x;\gamma,x_0)$ and $\mathbb{R}^+ \times \mathbb{R} \underset{open}{\subset} \mathbb{R}^2$. Then we have the map $\varphi(f_{\gamma,x_0}) = (\gamma, x_0)$ which is bijective, continuous and smooth. The continuity of the map is trivial given that the map defines the topology on S. Then we need only check that $\varphi \in C^\infty$.

Proposition 2. *Let S be the family of Cauchy distributions and let $\varphi : S \to \mathbb{R}^2$ be defined as $\varphi(p_{m,T}) = (m, T)$, then φ is a smooth map and S is a statistical surface.*

Now we move on to defining a Riemannian metric on our manifold S. Since our points in S are probability distributions, then it is natural to define our metric using the Fisher Information Matrix, which is widely used in the field of statistics to compare the content of distributions.

Definition 3. *fisher Let $G(\xi) := [g_{ij}(\xi)]$ be an $n \times n$ matrix where $g_{ij}(\xi)$ is defined below. Then we call $G(\xi)$ the **Fisher Information Matrix**.*

$$g_{ij}(\xi) := \int \partial_i \log f(x;\xi) \partial_j \log f(x;\xi) dx \tag{2}$$

We may also write the expected value of a function g with respect to a probability distribution f_ξ as $E_\xi[g] := \int g(x)f(x;\xi)dx$. Then (2) can be rewritten as

$$g_{ij}(\xi) = E_\xi[\partial_i \log f(x;\xi)\partial_j \log f(x;\xi)] \tag{3}$$

Now one may consider this matrix in order to define our Riemannian metric, which is a function $g = \langle -.- \rangle : T_pS \times T_pS \to \mathbb{R}^+$ that is symmetric. We make the following claim.

Proposition 4. *Let $g_{ij} = \langle \partial_i, \partial_j \rangle$, then this determines a unique Riemannian metric*

$$G = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2\gamma^2} \end{pmatrix} \tag{4}$$

on the family of Cauchy distributions, S.

Proof. This proof will be heavily computational and let $\xi = (x_0, \gamma)$.

$$\begin{aligned}
\frac{\partial}{\partial x_0} \log p(x;\xi) &= \frac{\partial}{\partial x_0} \log \frac{1}{\pi\gamma\left[1 + (\frac{x-x_0}{\gamma})^2\right]} \\
&= -\frac{\partial}{\partial x_0} \log \pi\gamma\left[1 + (\frac{x-x_0}{\gamma})^2\right] \\
&= \frac{1}{\pi\gamma\left[1 + (\frac{x-x_0}{\gamma})^2\right]} 2\pi\gamma \frac{x-x_0}{\gamma} \\
&= \frac{2\pi(x-x_0)}{\pi\gamma\left[1 + (\frac{x-x_0}{\gamma})^2\right]} = 2\pi(x-x_0)p(x;\xi)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial}{\partial \gamma} \log p(x;\xi) &= -\frac{1}{\pi\gamma\left[1 + (\frac{x-x_0}{\gamma})^2\right]} \left[\pi\left(1 + (\frac{x-x_0}{\gamma})^2\right) - 2\pi\gamma(x-x_0)^2\frac{1}{\gamma^2}\right] \\
&= -p(x;\xi)\left[\pi + -\pi\gamma\frac{(x-x_0)^2}{\gamma^2}\right] \\
&= -\pi\left(1 - (\frac{x-x_0}{\gamma})^2\right)p(x;\xi)
\end{aligned}$$

$$\begin{aligned}
g_{11}(\xi) &= \int \left[\frac{\partial}{\partial x_0} \log p(x;\xi)\right]\left[\frac{\partial}{\partial x_0} \log p(x;\xi)\right] p(x;\xi)dx \\
g_{11} &= \int 4\pi^2(x-x_0)^2 p(x;\xi)^3 dx
\end{aligned}$$

$$= \int_{-\infty}^{\infty} \frac{4(x-x_0)^2}{\pi\gamma^3\left[1+\left(\frac{x-x_0}{\gamma}\right)^2\right]^3} dx$$

$$= \frac{4}{\pi\gamma^3}\int_{-\infty}^{\infty} \frac{x^2}{\left(1+\frac{x^2}{\gamma^2}^3\right)} dx$$

$$= \frac{4}{\pi\gamma^3}\frac{\pi\gamma^3}{8} = \frac{1}{2}$$

$$g_{12}(\xi) = g_{21}(\xi) = \int -2\pi^2(x-x_0)\left(1-(\frac{x-x_0}{\gamma})^2\right)p(x;\xi)^3 dx = 0$$

$$g_{22}(\xi) = \frac{1}{2\gamma^2}$$

Then we have

$$G = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2\gamma^2} \end{pmatrix}$$

Now it only remains to show that G is a Riemannian metric. First, $G^t = G$, so G is symmetric, and the eigenvalues of G are positive so its output is positive definite. Finally, $u^T Gv \in \mathbb{R}$ for $u, v \in TM$ and hence $G \in Sym^2(M)$. Hence G is a Riemannian metric. □

4 Guassian Curvature of the family of Cauchy Distributions

Now that we have calculated our Fisher metric we need to find the corresponding covariant derivative in order to calculate the curvature tensor. In particular, we would like to find the Levi-Civita connection which corresponds to our metric G on the family of Cauchy distributions. In order to do this it will first be necessary to calculate the corresponding Christoffel symbols for our connection using the following well-known equation from Riemannian geometry,

$$\Gamma_{ij}^m = \frac{1}{2}g^{ml}\left(\frac{\partial}{\partial x_j}g_{il} + \frac{\partial}{\partial x_i}g_{lj} - \frac{\partial}{\partial x_l}g_{ij}\right) \tag{5}$$

Then we omit the computation but present the following computed values.

$$G^{-1} = (g^{ij}) = \begin{pmatrix} 2 & 0 \\ 0 & 2\gamma^2 \end{pmatrix}.$$

$$\Gamma_{11}^1 = 0 \qquad \Gamma_{12}^1 = 0$$
$$\Gamma_{11}^2 = 0 \qquad \Gamma_{12}^2 = 0$$
$$\Gamma_{21}^1 = 0 \qquad \Gamma_{22}^1 = 0$$

$$\Gamma^2_{21} = 0 \qquad \Gamma^2_{22} = \frac{1}{\gamma}$$

Then the Levi-Civita connection can be written as

$$\nabla_{e_i} e_j = \sum_k \Gamma^k_{ij} e_k \tag{6}$$

Then we can calculate $\nabla_{e_1} e_1 = 0, \nabla_{e_2} e_1 = 0, \nabla_{e_1} e_2 = 0, \nabla_{e_2} e_2 = e_2$. Now we may use this information to easily calculate the Gaussian curvature

$$K = \frac{\langle (\nabla_{e_2} \nabla_{e_1} - \nabla_{e_1} \nabla_{e_2}) e_1, e_2 \rangle}{\det G} = 0 \tag{7}$$

So the Gaussian curvature of the family of Cauchy distributions under the Fisher metric is $K = 0$.

5 Geodesics on the Statistical Manifold of Cauchy Distribution

Now that we have defined our Levi-Civita connection corresponding to our metric, we can then calculate the geodesics on our statistical manifold. Geodesics are important as it will better help us to understand why the Gaussian Curvature $K = 0$. Due to the structure of our distributions, intuition shows us that our space is not flat. To formalize the intuition, we find and graph the geodesics to show this presently.

Due to the theorem for the existence and uniqueness of geodesics, we get the following systems of differential equations that will give our geodesics. Suppose the geodesic is $c(t) = (c_1(t), c_2(t))$.

$$\begin{cases} c_1''(t) + \Gamma^1_{ij} c_i'(t) c_j'(t) = 0 \\ c_2''(t) + \Gamma^2_{ij} c_i'(t) c_j'(t) = 0 \end{cases}$$

$$\Longrightarrow$$

$$\begin{cases} c_1''(t) = 0 \\ c_2''(t) + \frac{1}{c_2(t)} (c_2'(t))^2 = 0 \end{cases}$$

The solution to this system of differential equations is

$$\begin{cases} c_1(t) = k_1 t + k_2 \\ c_2(t) = k_3 \sqrt{t + k_4} \end{cases}$$

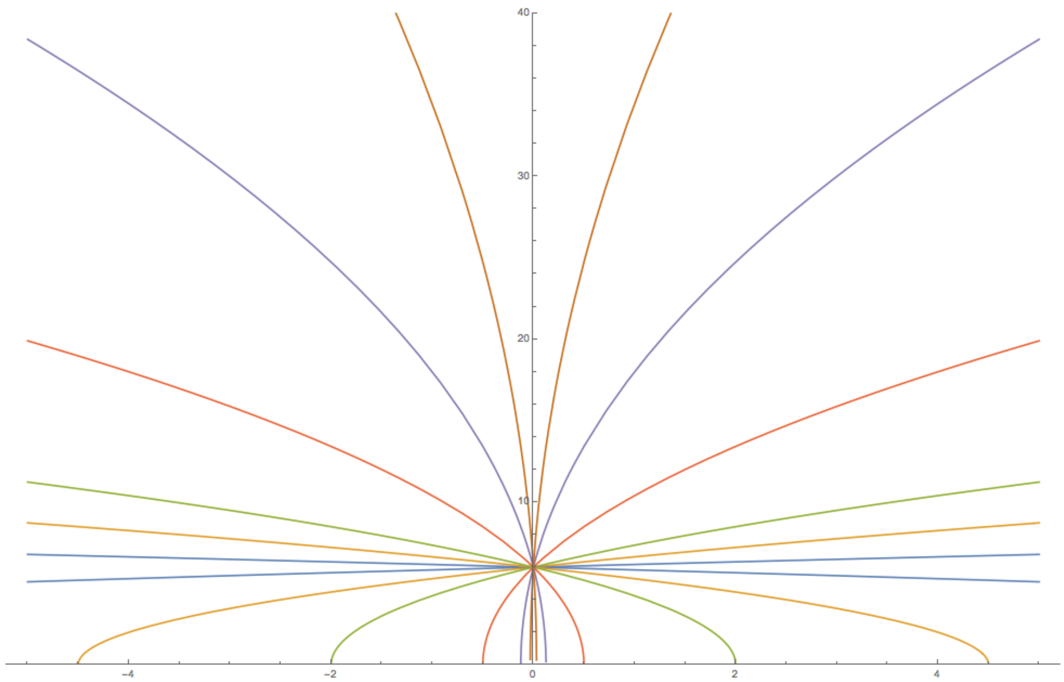

Figure 1: Examples of geodesics on the statistical manifold.

6 Conclusion

We proved that the Fisher-information metric gives a Riemannian metric on the 2-dimensional statistical manifold of Cauchy Distribution (identified as the upper-half plane). We showed the Gaussian curvature is everywhere zero, and characterized the geodesics on the manifold.

The normal distribution is the simplest motivating example of a statistical manifold, where the two parameters used to construct the 2-dimensional manifold are mean and standard deviation. The Cauchy distribution is often the canonical example of a pathological distribution since both its mean and its variance are undefined. Nevertheless, the definition of statistical manifolds and the Fisher-information metric are flexible enough to construct a manifold with any parameters of the distribution, not necessarily requiring a well-defined mean or variance.

The statistical manifold of the Cauchy Distribution is the upper half plane because its two parameters are restricted in the upper half plane. There are other probability distributions with the upper half plane as parameter space, but with different metrics on the upper half plane. It would be interesting to study and compare the geometric structures of these probability distributions.

In our case of the Cauchy Distribution, it is rather surprising that the complicated formula for Fisher-information metric results in a very simple expression for the metric. The Cauchy distribution is linked to the solution to the differential equation describing forced resonance. Because of its physical importance, it would perhaps be meaningful to consider whether the metric, curvature, and geodesics of the statistical manifold have any physical meaning in the context of forced resonance.

Information Geometry has the potential to be an important field moving forward, as it unifies statistics and geometric structure in a way that applications to Quantum Information, Mathematical Finance and even in Machine Learning. It would also be interesting to see whether the geometric structure of the Cauchy Distribution has any applications in these fields.

References

[1] S.I. Amari, H. Nagaoka, *Methods of information geometry*, Transl. Math. Monographs, Vol.191, American Math. Soc. 2000.