Journal of Multivariate Analysis 100 (2009) 777–793

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

On the geometry of generalized Gaussian distributions

Attila Andai

RIKEN, BSI, Amari Research Unit, 2–1, Hirosawa, Wako, Saitama 351-0198, Japan

ARTICLE INFO

Article history:
Received 21 June 2007
Available online 27 August 2008

AMS subject classifications:
94A17
53B21

Keywords:
Gaussian distribution
Differential geometry

ABSTRACT

In this paper we consider the space of those probability distributions which maximize the q-Rényi entropy. These distributions have the same parameter space for every q, and in the $q = 1$ case these are the normal distributions. Some methods to endow this parameter space with a Riemannian metric is presented: the second derivative of the q-Rényi entropy, the Tsallis entropy, and the relative entropy give rise to a Riemannian metric, the Fisher information matrix is a natural Riemannian metric, and there are some geometrically motivated metrics which were studied by Siegel, Calvo and Oller, Lovrić, Min-Oo and Ruh. These metrics are different; therefore, our differential geometrical calculations are based on a new metric with parameters, which covers all the above-mentioned metrics for special values of the parameters, among others. We also compute the geometrical properties of this metric, the equation of the geodesic line with some special solutions, the Riemann and Ricci curvature tensors, and the scalar curvature. Using the correspondence between the volume of the geodesic ball and the scalar curvature we show how the parameter q modulates the statistical distinguishability of close points. We show that some frequently used metrics in quantum information geometry can be easily recovered from classical metrics.

1. Introduction

In theoretical statistics, and in applications, the distance functions between probability distributions play an important role. The construction of a proper distance function has been considered by several authors. But even the same statistical model with different mathematical frameworks can lead to different distance functions. To narrow the family of potential distance functions we consider those which are natural from a differential geometrical point of view.

Historically, the pioneering work of Mahalanobis [24] was generalized by Rao [31], who first suggested the idea of considering the Fisher information [14] as a Riemannian metric on the space of probability distributions. Cencov [8] was the first to study monotone metrics on statistical manifolds. He proved that, up to a normalization, there exists a unique monotone metric, the Fisher information. Amari [3] and Amari and Nagaoka [4] provide modern account of the general differential geometry that arises from the Fisher information metric. The Fisher metric was studied further by Akin [1], James [16], Burbea [6], Mitchell [23], Atkinson and Mitchell [5], Skovgaard [35], Oller [26], Oller and Cuadras [28], and Oller and Corcuera [27], among other researchers. The combination of differential geometrical and statistical studies helped to find the statistical interpretation of geometrical quantities. For example, the geodesic distance between probability distributions, which is usually known as the Rao distance, is a natural distance function between probability distributions; the statistical meaning of the so-called e-curvature was first clarified by Efron [12]; the normalized volume measure of the manifold is called Jeffreys' prior [17] within the field of Bayesian statistics.

In this paper we consider the space of those probability distributions which maximize the q-Rényi entropy. These distributions have the same parameter space for every q, and in the $q = 1$ case these are the normal distributions. The first results about the geometrical properties of these spaces are due to Amari [3,2]. He considered the Fisher information

E-mail address: andaia@math.bme.hu.

0047-259X/$ – see front matter
doi:10.1016/j.jmva.2008.08.007

metric on these manifolds and computed some geometrical invariants. Some methods to endow the parameter space with a Riemannian metric are presented: the second derivative of the q-Rényi entropy [32], the Tsallis entropy [36], and the relative entropy give rise to a Riemannian metric, the Fisher information matrix is a natural Riemannian metric, and there are some geometrically motivated metrics which were studied by Siegel [34], Calvo and Oller [7] and Lovrić, Min-Oo and Ruh [33]. These metrics are different; therefore, our differential geometrical calculations are based on a new metric with parameters, which covers all the above-mentioned metrics for special values of the parameters, among others. We also compute the geometrical properties of this metric, the equation of the geodesic line with some special solutions, the Riemann and Ricci curvature tensors, and the scalar curvature. Using the correspondence between the volume of the geodesic ball and the scalar curvature we show how the parameter q modulates the statistical distinguishability of close points. We show that some frequently used metrics in quantum information geometry can be easily recovered from classical metrics.

The proofs of the presented theorems are in the Appendix.

2. q-Rényi entropy maximizing distributions

The normal distributions can be introduced as a result of the maximum entropy principle. Consider the family of density functions which are continuous and supported on the real line with given expectation value $\mu \in \mathbb{R}$ and variance $\sigma^2 \in \mathbb{R}^+$. Introducing the Lagrange multipliers a, b, c, we have the following functional on the family of probability distributions:

$$S(\rho) = -\int \rho(x) \log \rho(x)\, \mathrm{d}x - a\left(\int \rho(x)\, \mathrm{d}x - 1\right) - b\left(\int \rho(x) x\, \mathrm{d}x - \mu\right) - c\left(\int \rho(x)(x-\mu)^2 \mathrm{d}x - \sigma^2\right).$$

The variation of the functional is

$$\delta S = \int \left(-\log \rho(x) - 1 - a - bx - c(x-\mu)^2\right) \delta p(x)\, \mathrm{d}x.$$

The functional has extremal point at ρ if its variation is zero. One can show that the entropy functional has a local maximum at the point

$$\rho(x) = \exp\left(-a - bx - c(x-\mu)^2\right)$$

for appropriate parameters $a, b, c \in \mathbb{R}$.

The family of one-dimensional normal distributions S_1 can be parameterized by the expectation value $u \in \mathbb{R}$ and the parameter $d \in \mathbb{R}^+$ as

$$f(d, u, x) = \frac{\sqrt{d}}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2} d(x-u)^2}.$$

This means that S_1 can be identified with a two-dimensional space $\Xi_1 = \mathbb{R}^+ \times \mathbb{R}$. The statistical properties of the distributions lead us to define the Riemannian metric on the space Ξ_1.

In general, the family of n-dimensional normal distributions S_n can be parameterized by the expectation vector $\underline{u} \in \mathbb{R}^n$ and the inverse of the covariance matrix D. Let us denote the set of real symmetric strictly positive $n \times n$ matrices by $\mathcal{M}_n$. Then we can identify the sets S_n and $\Xi_n = \mathcal{M}_n \times \mathbb{R}^n$ using the following one-to-one map:

$$\Xi_n \to S_n \qquad (D, \underline{u}) \mapsto f(D, \underline{u}, \cdot),$$

where

$$f(D, \underline{u}, \cdot) : \mathbb{R}^n \to \mathbb{R} \qquad \underline{x} \mapsto \frac{\sqrt{\det D}}{\sqrt{(2\pi)^n}} \exp\left(-\frac{1}{2}\langle \underline{x} - \underline{u}, D(\underline{x} - \underline{u})\rangle\right).$$

Normal distributions with zero expectation will be said to be special normal distributions. The parameter space of the n-dimensional special normal distribution is $\Xi_n^{(s)} = \mathcal{M}_n$.

One can generalize the above-mentioned procedure to extend the notion of Gaussian distributions using the q-Rényi entropy [32].

Definition 2.1. Let us fix a parameter $q \in \mathbb{R}^+ \setminus \{1\}$ and consider a density function ρ. The q-Rényi entropy of the distribution ρ is

$$S_q(\rho) = \frac{1}{1-q} \log \int_{\mathbb{R}} \rho(x)^q\, \mathrm{d}x$$

if the integral exists.