Nonparametric Information Geometry

Giovanni Pistone

Collegio Carlo Alberto, Via Real Collegio 30, 10024 Moncalieri, Italy
giovanni.pistone@carloalberto.org
http://www.giannidiorestino.it/index.html

Abstract. The differential-geometric structure of the set of positive densities on a given measure space has raised the interest of many mathematicians after the discovery by C.R. Rao of the geometric meaning of the Fisher information. Most of the research is focused on parametric statistical models. In series of papers by author and coworkers a particular version of the nonparametric case has been discussed. It consists of a minimalistic structure modeled according the theory of exponential families: given a reference density other densities are represented by the centered log likelihood which is an element of an Orlicz space. This mappings give a system of charts of a Banach manifold. It has been observed that, while the construction is natural, the practical applicability is limited by the technical difficulty to deal with such a class of Banach spaces. It has been suggested recently to replace the exponential function with other functions with similar behavior but polynomial growth at infinity in order to obtain more tractable Banach spaces, e.g. Hilbert spaces. We give first a review of our theory with special emphasis on the specific issues of the infinite dimensional setting. In a second part we discuss two specific topics, differential equations and the metric connection. The position of this line of research with respect to other approaches is briefly discussed.

Keywords: Information Geometry, Banach Manifold.

1 Introduction

In the present paper we follow closely the presentation of Information Geometry developed by S.-I. Amari and coworkers, see e.g. in [1], [2], [3], [4], with the specification that we want to construct a Banach manifold structure in the classical sense, see e.g. [5] or [6], without any restriction to parametric models. We feel that the non parametric approach is of interest even in the case of a finite state space. We build upon our previous work in this field, namely [7], [8], [9], [10], [11], [12], [13], [14], [15], [16], [17], [18], [19], [20], [21], [22]. Other contributions are referred to in the text below. We do not discuss here the non commutative/quantum case as developed e.g., in [23], [24] and the review in [25].

The rest of this introductory section contains a review of relevant facts related with the topology of Orlicz spaces which are the model spaces in our manifold structure. The review part is based on previous joint work with M. P. Rogantin
and A. Cena, but a number of examples and remarks are added in order to clarify potential issues and possible applications. The exponential manifold (originally introduced in the joint work with C. Sempi) is critically reviewed in Sec. 2, together with applications. Differential equations are discussed in Sec. 3 with examples. Sec. 4 deals with the Hilbert bundle of the exponential manifold and the computation of the metric derivative. It builds upon previous work on non parametric connections with P. Gibilisco. A variation on exponential manifolds is introduced in Sec. 5 to show it could be developed along the lines previously discussed.

1.1 Model Spaces

In this paper we consider a fixed $\sigma$-finite measure space $(\Omega, \mathcal{F}, \mu)$ and denote by $P_\triangledown$ the set of all densities which are positive $\mu$-a.s. The set of densities, without any further restriction is $P_\geq$, while $P_1$ is the set of measurable functions $f$ with $\int f \, d\mu = 1$. In the finite state space case, i.e. $\#\Omega < \infty$, $P_1$ is a plane, $P_\geq$ is the simplex, $P_\triangledown$ its topological interior. In the infinite case, the setting is much more difficult: we concentrate here mainly on strictly positive densities and we construct its geometry by taking as a guiding model the theory of exponential families, see [26], [27], [28], [29]. A non parametric approach we use was initially suggested by P. Dawid [30,31]. A geometry derived from exponential families is intrinsically bases on the positivity of the densities, see [32,33].

At each $p \in P_\triangledown$ we associate a set of densities of the form $q = e^{u-K} \cdot p$, where $u$ belongs to a suitable Banach space $B_p$ and $K$ is a constant depending on $p$ and $u$. The mapping $u \mapsto q$ will be one-to-one and its inverse $s_p: q \mapsto u$ will be a chart of our exponential manifold $eP = (P_\triangledown, \{s_p\})$. As we do not have manifold structures on the set of positive densities other that the exponential one, in the following the manifold and the set are both denoted by $P_\triangledown$.

We refer to [5, §5-7] and [6] for the theory of manifolds modeled on Banach spaces. According to this definition, a manifold is a set $P_\triangledown$ together with a collection or atlas of charts $s: \mathcal{U} \to B$ from a subset $\mathcal{U} \subset P_\triangledown$ to a Banach space $B$ such that for each couple of charts the transition maps $s' \circ s^{-1}: s(\mathcal{U} \cap \mathcal{U}') \to s'(\mathcal{U} \cap \mathcal{U}')$ are smooth functions from an open set of $B$ into an open set of $B'$. In this geometric approach, $P_\triangledown$ is a set, while all structure is in model spaces $B$.

It should be noted that the Banach spaces are not required to be equal, as the finite dimensional case seems to suggest, but they should be isomorphic when connected by a chart. Actually this freedom is of much use in our application to statistical model, but requires a careful discussion of the isomorphism. Precisely, at each $p \in P_\triangledown$, the model space $B_p$ for our manifold is an Orlicz space of centered random variables, see [34], [35, Chapter II], [36], [37, Ch 8]. We review briefly our notations and recall some basic facts from these references.

If both $\phi$ and $\phi^{-1} = \phi_*$ are monotone, continuous functions on $\mathbb{R}_{\geq 0}$ onto itself, we call the pair

$$
\Phi(x) = \int_0^{|x|} \phi(u) \, du, \quad \Phi_*(y) = \int_0^{|y|} \phi^{-1}(v) \, dv,
$$

...