GEOMETRIC STRUCTURES ON A STATISTICAL MANIFOLD AND GEOMETRY OF ESTIMATION

A Thesis submitted

in partial fulfillment for the Degree of

Doctor of Philosophy

by

HARSHA K. V.

Department of Mathematics

INDIAN INSTITUTE OF SPACE SCIENCE AND TECHNOLOGY

THIRUVANANTHAPURAM

AUGUST 2015

ABSTRACT

The main objective of this thesis is to study the various geometric structures on a statistical manifold and the geometry of parameter estimation. This study comes under the area of Information Geometry which is the geometric study of a statistical model of probability distributions. A statistical model equipped with a Riemannian metric and a pair of dual affine connections is called a statistical manifold. Amari's α-geometry is an important geometric structure on a statistical manifold which plays a major role in the asymptotic theory of estimation.

In Chapter 2 we introduce a generalized class of geometric structures on a statistical manifold called the (F, G)-geometry using a general embedding function F and a positive smooth function G. In Section 2.2 the Fisher information metric and the α-connections are computed for a statistical manifold defined on finite sets. In Theorem 2.3.5 we prove a necessary and sufficient condition for two (F, G)-connections to be dual with respect to the G-metric. In Theorem 2.3.6 we show that the α-geometry is a special case of the (F, G)-geometry. Thus we obtain a generalized dualistic structure on a statistical manifold which includes the α-geometry as a special case. Further the G-metric and the (F, G)-connections are computed for statistical manifold defined on finite sets in Section 2.3.

In Chapter 3 we study the invariance properties of various geometric structures on a statistical manifold and classify them into invariant and non-invariant classes. The covariance under reparametrization of the (F, G)-geometric structures are shown in Theorems 3.2.3 and 3.2.4. Then in Theorem 3.2.5 we prove that the (F, G)-geometry is not invariant under smooth one to one transformations of the random variable in general. In Corollary 3.2.6 we prove that the α-geometry is the only (F, G)-geometry which is invariant under smooth one to one transformations of the random variable. In Theorems 3.2.7 and 3.2.8 we show that the (α, ρ, τ)-geometry is covariant under reparametrization and is not invariant under smooth one to one transformations of the random variable in general. Also the α-geometry is the only (α, ρ, τ)-geometry which is invariant under smooth one to one transformations of the random variable. Further the relation between

the (F, G) and (α, ρ, τ)-geometries are given in Theorem 3.2.11.

In Chapter 4 first we give the (± 1)-conformal equivalence of the α-geometry and the geometry induced from the conformal transformation of the α-divergence in Propositions 4.2.3 and 4.2.4. In Corollary 4.2.6 we prove that the q-structure is the conformal flattening of the α-geometry. Then we discuss the importance of non-invariant (F, G)-geometry in the study of the dually flat geometries of the deformed exponential family. There are two dually flat geometries on a deformed exponential family, the U-geometry and the χ-geometry. In Theorem 4.3.4 we show that the U-geometry is the (F, G)-geometry for suitable choices of F and G. Further we prove that the χ-geometry is the conformal flattening of the (F, G)-geometry for suitable choices of F and G in Theorems 4.3.16, 4.3.17 and 4.3.18.

In Chapter 5 we consider the parameter estimation problem based on a mismatched model. In Theorems 5.3.1 and 5.3.2 we prove a necessary and sufficient condition for the estimator based on a mismatched model to be consistent and first order efficient. Further a theoretical formulation of the maximum likelihood estimation problem based on a mismatched model in an exponential family is given. We prove a necessary and sufficient condition for an MLE based on a mismatched model to be consistent and efficient in Theorems 5.3.8 and 5.3.9.

In Chapter 6 we define certain generalized notions like F-product, F- independence of random variables and maximum F-likelihood estimator (F-MLE) in Section 6.1. In Theorem 6.1.6 we show that the F-MLE is a MAP estimator with a prior. Then using the F-escort probability distribution we define two generalized notions of MLE, the $\mathbf{x}_N$-based F-escort MLE and the F-escort MLE based on the product of F-escort distribution of the marginal probability density of single observations in Section 6.2. In Theorem 6.2.3 we give a characterization of the q-escort MLE among the $\mathbf{x}_N$ based F-escort MLE as a Bayesian MAP estimator with a prior. Further an analytic proof of the F-version of the maximum entropy theorem is given in Theorem 6.2.5. In Theorem 6.3.2 a proof of the generalized Cramer-Rao bound defined by Naudts is given. Further we show that the U-estimator for the dual coordinate in the U-geometry of the deformed exponential family is optimal with respect to this bound in Theorem 6.3.3. This chapter ends with an open problem regarding the properties of the F-MLE in a deformed exponential family.