Journal of Multivariate Analysis 131 (2014) 254–264

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

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

A characterization of elliptical distributions and some optimality properties of principal components for functional data

Graciela Boente [a,*], Matías Salibián Barrera [b], David E. Tyler [c]

[a] Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires and CONICET, Argentina
[b] Department of Statistics, University of British Columbia, Canada
[c] Department of Statistics, Rutgers University, USA

ARTICLE INFO

Article history:
Received 8 February 2013
Available online 28 July 2014

AMS subject classifications:
62H25
60G07

Keywords:
Elliptical distributions
Functional data analysis
Principal components

ABSTRACT

As in the multivariate setting, the class of elliptical distributions on separable Hilbert spaces serves as an important vehicle and reference point for the development and evaluation of robust methods in functional data analysis. In this paper, we present a simple characterization of elliptical distributions on separable Hilbert spaces, namely we show that the class of elliptical distributions in the infinite-dimensional case is equivalent to the class of scale mixtures of Gaussian distributions on the space. Using this characterization, we establish a stochastic optimality property for the principal component subspaces associated with an elliptically distributed random element, which holds even when second moments do not exist. In addition, when second moments exist, we establish an optimality property regarding unitarily invariant norms of the residuals covariance operator.

1. Introduction

When considering finite-dimensional random vectors, a natural and commonly used generalization of the family of multivariate normal distributions is given by the class of elliptical distributions. This class allows for heavy tail models while preserving many of the attractive properties of the multivariate normal model, such as regression being linear, as discussed for example in [8,15,28,35]. Multivariate elliptical distributions include the t-distributions, the symmetric generalized hyperbolic distribution, the multivariate Box–Tiao or power exponential family distributions and the sub-Gaussian α-stable distributions, among others. From a practical point of view, Frahm [15] has argued that the class of heavy tailed elliptical distributions offers a good alternative for modeling financial data in which the Gaussian assumption may not be reliable. Multivariate elliptical models have also been considered extensively within the area of robust statistics as a starting point for the development of the M-estimates of multivariate location and scatter, see e.g. [26], and also as a class of models under which the asymptotic behavior and the influence functions of robust multivariate methods, such as robust principal components, can be evaluated and judged. (See, for instance, [18,20,27].)

In many areas of statistics, the collected data are more naturally represented as functions rather than finite-dimensional numerical vectors, as argued e.g. in [32]. Simplifying the functional model by discretizing the observations as sequences

* Correspondence to: Departamento de Matemáticas, FCEyN, UBA, Ciudad Universitaria, Pabellón 2, Buenos Aires, C1428EHA, Argentina.
E-mail addresses: gboente@dm.uba.ar, gboente@fibertel.com.ar (G. Boente), matias@stat.ubc.ca (M. Salibián Barrera), dtyler@rci.rutgers.edu (D.E. Tyler).

http://dx.doi.org/10.1016/j.jmva.2014.07.006
0047-259X/

of numbers can often fail to capture some of its important characteristics, such as the smoothness and continuity of the underlying functions. For this reason, in the last decades different methods have been proposed to handle this type of "functional" data, which can be viewed as instances of random elements taking values in a space of functions such as $L^2(\mathcal{I})$, with $\mathcal{I} \subset \mathbb{R}$ a finite interval. A more general and inclusive framework is to view the observations as elements in a separable Hilbert space $\mathcal{H}$, which is not necessarily finite-dimensional.

The notion of principal components analysis, which is a fundamental concept in multivariate statistics, has been extended to the functional data setting and is commonly referred to as FPCA or functional principal components analysis. The first few principal components are typically used to explore the main characteristics of the data within a reduced dimensional space. In particular, exploring this lower dimensional principal components space can be useful for detecting atypical observations or outliers in the data set. The principal components subspace has the well known property that the first q principal components associated with the distribution of a random element with finite second moment provide the best q-dimensional linear approximation to the random element in terms of mean squared error. These linear approximations also minimize unitarily invariant norms of the covariance matrix of the residuals, in the finite-dimensional setting.

As in the multivariate setting, the class of elliptical distributions on separable Hilbert spaces can serve as an important vehicle and reference point for the development and evaluation of robust methods in functional data analysis. In addition, they allow for the development of FPCA even if the random elements do not possess second moments. An extension of the class of elliptical distributions to separable Hilbert spaces is given in the relatively recent paper by Bali and Boente [2], while the Fisher-consistency of some robust estimates of principal directions for this class of elliptical distributions is established in [4]. The main purpose of the present short paper is two-fold. First, in Section 2, we present a simple characterization of elliptical distributions on separable Hilbert spaces, namely we show that the class of elliptical distributions in the infinite-dimensional case is equivalent to the class of scale mixtures of Gaussian distributions on the space, unless the random element is essentially finite-dimensional. Second, we then use this representation in Section 3.1 to establish a stochastic best lower-dimensional approximation for elliptically distributed random elements and an optimality property for the scatter operator of the associated residuals. That is, we derive two optimality properties for the eigenfunctions associated with the largest eigenvalues of the scatter operator that hold even when second moments do not exist and which recover the same best lower dimensional approximation properties mentioned above when second moments do exist.

In Section 3.2 we extend another known optimality property of principal components from Euclidean spaces to general Hilbert spaces. This optimality property holds not only for elliptical distributions, but for any distribution with finite second moments. As in the finite-dimensional case, when second moments exist a measure of closeness between a random element X and a predictor is the norm of the residuals covariance operator. Although many operator norms can be defined, in the principal components setting a reasonable requirement is that the operator norm be unitarily invariant. Under this assumption, we show that the optimal linear predictors are those obtained through the linear space spanned by the first q principal components. In Section 4 we use the theory developed in this paper to show that the spherical principal components of Locantore et al. [25] and Gervini [16] are Fisher consistent for elliptically distributed random elements. This result extends previous results obtained for random elements with a finite Karhunen–Loève expansion. Some mathematical concepts and proofs are presented in the Appendix.

2. Elliptical distributions over Hilbert spaces

There are a number of ways to define the class of elliptical distributions in the multivariate setting. An attractive constructive definition is to define them as the class of distributions generated by applying affine transformations to the class of spherical distributions. The properties of elliptical distributions then follow readily from the simpler class of spherical distributions.

Recall that a random vector $\mathbf{Z} \in \mathbb{R}^d$ is said to have a d-dimensional spherical distribution if its distribution is invariant under orthogonal transformations, i.e., if $\mathbf{QZ} \sim \mathbf{Z}$ for any $d \times d$ orthogonal matrix $\mathbf{Q}$. The classic example of a spherical distribution is the multivariate standard normal distribution. In general, if $\mathbf{Z}$ has a spherical distribution in $\mathbb{R}^d$ then $R = \|\mathbf{Z}\|_d$ and $\mathbf{U} = \mathbf{Z}/\|\mathbf{Z}\|_d$ are independent with $\mathbf{U}$ having a uniform distribution on the d-dimensional unit sphere. Here $\|\cdot\|_d$ refers to the Euclidean norm in $\mathbb{R}^d$. If $\mathbf{Z}$ is also absolutely continuous in $\mathbb{R}^d$, then it has a density of the form $f(\mathbf{z}) = g_d(\mathbf{z}^\mathrm{T}\mathbf{z})$ for some function $g_d(s) \geq 0$, i.e., it has spherical contours. The marginal density of a subset of the components of $\mathbf{Z}$ also has spherical contours, with the relationship between g_d and the k-dimensional density generator g_k, for $k < d$ being somewhat complicated. It turns out to be more convenient to denote a spherical distribution by its characteristic function. In general, the characteristic function of a spherically distributed $\mathbf{Z} \in \mathbb{R}^d$ is of the form $\psi_\mathbf{Z}(\mathbf{t}_d) = \varphi(\mathbf{t}_d^\mathrm{T}\mathbf{t}_d)$ for $\mathbf{t}_d \in \mathbb{R}^d$, and any distribution in $\mathbb{R}^d$ having a characteristic function of this form is a spherical distribution. Consequently, we express $\mathbf{Z} \sim \mathcal{S}_d(\varphi)$. This notation is convenient since, for $\mathbf{Z}^\mathrm{T} = (\mathbf{Z}_1^\mathrm{T}, \mathbf{Z}_2^\mathrm{T})$ with $\mathbf{Z}_1 \in \mathbb{R}^k$, the marginal $\mathbf{Z}_1$ is such that $\mathbf{Z}_1 \sim \mathcal{S}_k(\varphi)$. More generally, for any $k \times d$ matrix $\mathbf{Q}_k$ such that $\mathbf{Q}_k^\mathrm{T}\mathbf{Q}_k = I$, we have $\mathbf{Q}_k\mathbf{Z} \sim \mathcal{S}_k(\varphi)$. Note that if $\varphi(\mathbf{t}_d^\mathrm{T}\mathbf{t}_d)$ is a valid characteristic function in $\mathbb{R}^d$ then $\varphi(\mathbf{t}_k^\mathrm{T}\mathbf{t}_k)$, where $\mathbf{t}_k = (t_1, \dots, t_k)$, is also a valid characteristic function in $\mathbb{R}^k$ for any $k < d$. For some families of spherical distributions defined across different dimensions, such as the multivariate power exponential family considered by Kuwana and Kariya [24], the function φ may depend upon the dimension d. In such cases, the marginal distributions are not elements of the same family.

As already noted, the elliptical distributions in $\mathbb{R}^d$ correspond to those distributions arising from affine transformations of spherically distributed random vectors in $\mathbb{R}^d$. For a $d \times d$ matrix $\mathbf{B}$ and a vector $\boldsymbol{\mu} \in \mathbb{R}^d$, the distribution of $\mathbf{X} = \mathbf{BZ} + \boldsymbol{\mu}$