Neurocomputing 150 (2015) 357–366

Contents lists available at ScienceDirect

Neurocomputing

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

Bayesian approach for parameter estimation of the diagonal of the Modified Riesz Distribution

Taher Ben Arab [a,b,*], Atizez Hadrich [a,b], Afif Masmoudi [a], Mourad Zribi [b]

[a] Laboratory of Probability and Statistics, Faculty of Sciences of Sfax, Sfax University, B. P. 1171, Sfax, Tunisia
[b] Laboratoire d'Informatique Signal et Image de la Côte d'Opale (LISIC-EA 4491), ULCO, 50 Rue Ferdinand Buisson BP 719, 62228 Calais Cedex, France

ARTICLE INFO

Article history:
Received 15 December 2013
Received in revised form
31 July 2014
Accepted 27 August 2014
Available online 19 October 2014

Keywords:
Diagonal of the Modified Riesz Distribution
Maximum likelihood
Method of moments
Bayesian approach
Monte Carlo methods
Mean square error

ABSTRACT

In this paper, we examine the problem of estimating the parameters of the proposed Diagonal of the Modified Riesz Distribution (DMRD) defined in $\mathbb{R}^r$. Several methods can be used for estimating the parameters of the DMRD. In the present work, we address the issue of estimating the parameters of the DMRD through a Bayesian approach associated with Monte Carlo methods. The proposed approach is compared via a simulation study with maximum likelihood method and method of moments by calculating the Mean Square Error (MSE).

1. Introduction

In the context of statistics and related fields, the analysis of symmetric cones [8,9] has been developed in depth through the last seven decades. First, the studies were only focused on the real case. Then recently, a complex case has been performed. This analysis has been fundamental in the development of several statistical areas such as random matrix distributions. These mathematical tools have been commonly implemented in the characterization of non-central distributions, statistical theory of shape, etc. A basic source of the general theory of symmetric cones can be found in [5,10,14,15,20,23]. According to this theory, Hassairi and Lajmi [14] have introduced a family of distributions on symmetric cones called the Riesz distribution which generalizes the matrix multivariate Gamma [4] and Wishart [7,14] distributions, considering them as particular cases.

Our contribution in this paper is to introduce a new statistical model based on the Riesz distribution [19] called the Diagonal of Modified Riesz Distribution (DMRD). This statistical model represents the distribution of the random diagonal elements of the matrix ΣX, where X is a random square matrix distributed according to a Riesz distribution on the cone of (r,r)-symmetric positive definite matrices and Σ represents a symmetric positive definite matrix. Our study is based on the concept of transform Laplace. The distribution of the random vector $Y = Diag(\Sigma X)$ (*Diag* represents the diagonal elements of the matrix) is characterized by the known parameters' vector (s_1, s_2) and the unknown parameters' vector $(\Sigma_{11}, \Sigma_{22}, \Sigma_{12}^2)$. We determined the distribution of Yin two cases: $s_1 = s_2$ and $s_2 > s_1$. In the first case, $s_1 = s_2$, these distributions are reduced to Kibble bivariate Gamma distribution [18]. In the second case, $s_2 > s_1$, the transform Laplace of Yrepresents the product of the transform Laplace of a Kibble bivariate Gamma distribution by the Laplace transform of a univariate Gamma distribution. Since the support of this distribution is $]0, +\infty[^2$ and its correlation coefficient is positive, they give it an important role in the field of image processing [16,21] (i.e., image registration problem). Image registration is a fundamental task in image processing used to match two or more pictures taken, for example, at different times, from different sensors or from different viewpoints. Over the years, a broad range of techniques has been developed for the various types of data and problems [2,22].

The proposed DMRD is characterized by three parameters which must be estimated. To estimate these parameters, we apply three methods (maximum likelihood method (ML), method of moments (MOM) and Bayesian approach associated with Monte Carlo methods).

This paper is organized as follows. Section 2 presents some definitions of the Riesz distribution. Section 3 introduces some important results of the DMRD. In section 4, an estimation of the parameters of the DMRD using the ML and the MOM is given. Section 5 discusses the proposed Bayesian estimation method in which some simulation results are presented in order to illustrate the performance of the three estimators of the parameter vector of the DMRD model. Finally, Section 6 presents some concluding remarks.

* Corresponding author.
E-mail addresses: taher.ben_arab@yahoo.fr (T. Ben Arab), hadrich_atizez@yahoo.fr (A. Hadrich), Afif.Masmoudi@fss.rnu.tn (A. Masmoudi), Mourad.Zribi@lisic.univ-littoral.fr (M. Zribi).

http://dx.doi.org/10.1016/j.neucom.2014.08.078
0925-2312/

2. Riesz distribution

The study of Riesz distributions originates from the paper [23] of Marcel Riesz, where he introduces a family of convolution operators as a generalization of the classical Rieman–Liouville operators on the positive reals to the Lorentz cone in $\mathbb{R}^r$.

Let E be the Euclidean space of (r,r) real symmetric matrices equipped with the scalar product $\langle x,y\rangle = \mathrm{tr}(x\cdot y)$ and where Ω denotes the cone of (r,r)-symmetric positive definite matrices. For $x=(x_{ij})_{1\le i,j\le r}$ in E and $1\le k\le r$, we define the sub-matrices $P_k(x)=(x_{ij})_{1\le i,j\le k}$ and $\Delta_k(x)$ denotes the determinant of the (k,k) matrix $P_k(x)$. Then the generalized power of x in the cone Ω of positive definite elements of E is defined, for $s_*=(s_1,s_2,\dots,s_r)\in\mathbb{R}^r$, by

$$\Delta_{s_*}(x)=\Delta_1(x)^{s_1-s_2}\Delta_2(x)^{s_2-s_3}\dots\dots\Delta_r(x)^{s_r}. \quad (1)$$

For $s_*=(s_1,s_2,\dots,s_r)$ satisfying the conditions $s_i>(i-1)/2$, the absolutely continuous Riesz measure is defined by

$$R_{s*}(dx)=\frac{\Delta_{s_*-n/r}(x)}{\Gamma_\Omega(s_*)}\mathbf{1}_\Omega(x)\,dx, \quad (2)$$

where $\mathbf{1}_\Omega(x)$ represents the indicator function defined on Ω, $n=r(r+1)/2$ is the dimension of E, $\Gamma_\Omega(p)=(2\pi)^{r(r-1)/4}\prod_{i=1}^r \Gamma(s_i-\frac{1}{2}(i-1))$ and $\Gamma(a)=\int_0^{+\infty}e^{-x}x^{a-1}\,dx$, $a>0$.

Then, a result due to Gindikin [12] says that, for all $\theta\in-\Omega$, the Laplace transform of measure R_{s*} is defined by

$$L_{R_{s_*}}(\theta)=\int_E e^{\langle\theta,x\rangle}R_{s_*}(dx)=\Delta_{s_*}(-\theta^{-1}), \quad (3)$$

and, for s_* satisfying the conditions $s_i>(i-1)/2$, the Riesz distribution with a shape parameter s_* and a scale parameter $\sigma\in\Omega$ is given by

$$R(s_*,\sigma)(dx)=\frac{e^{-\langle\sigma,x\rangle}\Delta_{s_*-n/r}(x)}{\Gamma_\Omega(s_*)\Delta_{s_*}(\sigma^{-1})}\mathbf{1}_\Omega(x)\,dx. \quad (4)$$

When $s_1=s_2=\dots=s_r=p>(r-1)/2$, $R(s_*,\sigma)$ is reduced to the Wishart distribution

$$W(p,\sigma)(dx)=\frac{1}{\Gamma_\Omega(p)\ \det(\sigma^{-p})}e^{-\langle\sigma,x\rangle}\det(x)^{p-n/r}\mathbf{1}_\Omega(x)\,dx. \quad (5)$$

3. The proposed DMRD

In this section, we introduce the model of the DMRD. This model is based on Laplace transform of Riesz distribution. In our work, we study the case where $r=2$.

Let X be a random Riesz (2,2)-matrix with parameters $s_*=(s_1,s_2)\in]0,+\infty[\times]\frac{1}{2},+\infty[$, and θ in $I_2-\Omega$, where I_2 is the identity matrix of order 2. Then, according to Eq. (2), the Laplace transform of X is defined by

$$L_X(\theta)=\int_E e^{\langle\theta,x\rangle}R_{s_*}(dx)=\Delta_{s_*}((I_2-\theta)^{-1}). \quad (6)$$

In order to obtain our model, we use an affine transformation of the random X given by $\tilde{X}=\Sigma X$ where

$$\Sigma=\begin{pmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{12} & \Sigma_{22}\end{pmatrix},\quad \Sigma_{11}>0,\quad \Sigma_{22}>0$$

and $\Sigma_{11}\Sigma_{22}-\Sigma_{12}^2>0$.

The proposed model of the Diagonal of the Modified Riesz Distribution is given by $Y=(Y_1,Y_2)=diag\left(\tilde{X}\right)$ (the diagonal elements of the matrix $\tilde{X}$).

In order to define the Laplace transform $L_Y(\theta)$ of Y, we introduce the following proposition.

Proposition 3.1. *Let $s_*=(s_1,s_2)\in]0,+\infty[\times]\frac{1}{2},+\infty[$ and $\theta\Sigma\in I_2-\Omega$. Then*

$$L_Y(\theta)=[(1-\theta_{11}\Sigma_{11})(1-\theta_{22}\Sigma_{22})-\theta_{11}\theta_{22}\Sigma_{12}^2]^{-s_1}[1-\theta_{22}\Sigma_{22}]^{-(s_2-s_1)}. \quad (7)$$

Proof. Firstly, let $Y=Diag(\tilde{X})$ be a random vector of $\mathbb{R}^2$ whose elements constitute the diagonal elements of the matrix $\tilde{X}$. Then, for all $\theta\in I_2-\Omega$, we have

$$L_Y(\theta)=E\left(e^{\langle\theta,\tilde{X}\rangle}\right)=E\left(e^{tr(\theta\tilde{X})}\right)=E\left(e^{tr(\theta\Sigma X)}\right)=L_X(\theta\Sigma).$$

Secondly, if $\theta_{ij}=0$, for $i\ne j$, then

$$L_Y(\theta)=\Delta_{s_*}((I_2-\theta\Sigma)^{-1})$$
$$=[1-\theta_{22}\Sigma_{22}]^{s_1-s_2}[(1-\theta_{11}\Sigma_{11})(1-\theta_{22}\Sigma_{22})-\theta_{11}\theta_{22}\Sigma_{12}^2]^{-s_1}. \quad \square$$

In the remainder of this section, we will define the distribution of Y in two cases $s_1=s_2$ and $s_2>s_1$. In both cases, we define, in Propositions 3.2 and 3.5, the probability density function (pdf) f_Y of Y.

3.1. Case 1: $s_1=s_2=s$

Proposition 3.2. *Let $s_1=s_2=s$, $\theta\in-\Omega$ and Σ is a positive definite matrix. The distribution of the random variable Y is a Bivariate Gamma Distribution (BGD) with a pdf defined by*

$$f_Y(y_1,y_2)=\exp\left(-\frac{\Sigma_{22}y_1+\Sigma_{11}y_2}{\Sigma_{11}\Sigma_{22}-\Sigma_{12}^2}\right)\times\frac{(y_1y_2)^{s-1}}{(\Sigma_{11}\Sigma_{22}-\Sigma_{12}^2)^s\Gamma(s)}$$
$$\times f_s(\delta y_1y_2)\mathbf{1}_{[0,+\infty[^2}(y_1,y_2), \quad (8)$$

where $\mathbf{1}_{[0,+\infty[^2}(y_1,y_2)$ represents the indicator function defined on $[0,+\infty[^2$, $\delta=\Sigma_{12}^2/(\Sigma_{11}\Sigma_{22}-\Sigma_{12}^2)^2$, and $f_s(x)$ is given by $f_s(x)=\sum_{k=0}^\infty x^k/\Gamma(s+k)k!$, $\forall x\in\mathbb{R}$.

Proof. If $s_1=s_2=s$, $\Sigma_{11}>0$, $\Sigma_{22}>0$ and $\Sigma_{11}\Sigma_{22}-\Sigma_{12}^2>0$, then the Laplace transform of Y is given by

$$L_Y(\theta)=[(1-\theta_{11}\Sigma_{11})(1-\theta_{22}\Sigma_{22})-\theta_{11}\theta_{22}\Sigma_{12}^2]^{-s}. \quad (9)$$

According to [6], if $\theta\in I_2-\Omega$, then this Laplace transform is related to a BGD with a pdf defined in Eq. (8). $\square$

In Fig. 1, we show a BGD for different values of parameters.

Proposition 3.3. *The marginal distribution Y_i, $i=1,2$, is distributed according to a univariate gamma distribution with a pdf defined by*

$$f_{Y_i}(y_i)=\left(\frac{y_i}{\Sigma_{ii}}\right)^{s-1}\frac{\exp\left(-\frac{y_i}{\Sigma_{ii}}\right)}{\Sigma_{ii}\Gamma(s)}\mathbf{1}_{[0,+\infty[}(y_i). \quad (10)$$

where $s>0$ represents the shape parameter and $\Sigma_{ii}>0$, $i=1,2$, is called the scale parameter.

In the following, we denote the univariate gamma distribution by $\gamma(s,\Sigma_{ii})$ with a shape parameter s and a scale parameter $\Sigma_{ii}>0$, $i=1,2$.

Proof. The Laplace transform of Y is given by

$$L_Y(\theta)=[(1-\theta_{11}\Sigma_{11})(1-\theta_{22}\Sigma_{22})-\theta_{11}\theta_{22}\Sigma_{12}^2]^{-s}.$$

The Laplace transforms of Y_1 and Y_2 are defined by

$$L_{Y_1}(\theta_{11})=L_Y(\theta_{11},0)=(1-\theta_{11}\Sigma_{11})^{-s},\quad\text{and} \quad (11)$$

$$L_{Y_2}(\theta_{22})=L_Y(0,\theta_{22})=(1-\theta_{22}\Sigma_{22})^{-s}. \quad (12)$$

These Laplace transforms are related to a Gamma distribution $\gamma(s,\Sigma_{ii})$, $i=1,2$. $\square$