Journal of Multivariate Analysis 124 (2014) 237–246

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

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

Invariance properties of the likelihood ratio for covariance matrix estimation in some complex elliptically contoured distributions

Olivier Besson [a,*], Yuri I. Abramovich [b]

[a] University of Toulouse, ISAE, Department of Electronics, Optronics and Signal, 10 Avenue Edouard Belin, 31055 Toulouse, France

[b] W R Systems Ltd., 11351 Random Hills Road, Suite 400, Fairfax, VA 22030, USA

HIGHLIGHTS

- We consider a class of complex elliptically contoured matrix distributions (ECD).
- We investigate properties of the likelihood ratio (LR).
- We derive stochastic representations of the LR for covariance matrix estimation (CME).
- Its p.d.f. evaluated at the true CM $\boldsymbol{R}_0$ does not depend on the latter.
- This extends the expected likelihood approach for regularized CME.

ARTICLE INFO

Article history:
Received 26 December 2012
Available online 11 November 2013

AMS subject classifications:
62H10
62H15

Keywords:
Covariance matrix estimation
Elliptically contoured distribution
Expected likelihood
Likelihood ratio

ABSTRACT

The likelihood ratio (LR) for testing if the covariance matrix of the observation matrix $\boldsymbol{X}$ is $\boldsymbol{R}$ has some invariance properties that can be exploited for covariance matrix estimation purposes. More precisely, it was shown in Abramovich et al. (2004, 2007, 2007) that, in the Gaussian case, $LR(\boldsymbol{R}_0|\boldsymbol{X})$, where $\boldsymbol{R}_0$ stands for the true covariance matrix of the observations $\boldsymbol{X}$, has a distribution which does not depend on $\boldsymbol{R}_0$ but only on known parameters. This paved the way to the expected likelihood (EL) approach, which aims at assessing and possibly enhancing the quality of any covariance matrix estimate (CME) by comparing its LR to that of $\boldsymbol{R}_0$. Such invariance properties of $LR(\boldsymbol{R}_0|\boldsymbol{X})$ were recently proven for a class of elliptically contoured distributions (ECD) in Abramovich and Besson (2013) and Besson and Abramovich (2013) where regularized CME were also presented. The aim of this paper is to derive the distribution of $LR(\boldsymbol{R}_0|\boldsymbol{X})$ for other classes of ECD not covered yet, so as to make the EL approach feasible for a larger class of distributions.

1. Introduction and problem statement

The expected likelihood (EL) approach was introduced and developed in [3–5] as a statistical tool to assess the quality of a covariance matrix estimate $\hat{\boldsymbol{R}}$ from observation of a $M \times T$ matrix variate $\boldsymbol{X}$. The EL approach relies on some invariance properties of the likelihood ratio (LR) for testing $H_0 : \mathcal{E}\{\boldsymbol{X}\boldsymbol{X}^H\} = \boldsymbol{R}$ against the alternative $\mathcal{E}\{\boldsymbol{X}\boldsymbol{X}^H\} \neq \boldsymbol{R}$. More precisely, the LR is given by

$$LR(\boldsymbol{R}|\boldsymbol{X}) = \frac{p(\boldsymbol{X}|\boldsymbol{R})}{\max_{\boldsymbol{R}} p(\boldsymbol{X}|\boldsymbol{R})} = \frac{p(\boldsymbol{X}|\boldsymbol{R})}{p(\boldsymbol{X}|\boldsymbol{R}_{\mathrm{ML}})}, \tag{1}$$

* Corresponding author.
E-mail addresses: olivier.besson@isae.fr (O. Besson), yabramovich@wrsystems.com (Y.I. Abramovich).

0047-259X/$ – see front matter
http://dx.doi.org/10.1016/j.jmva.2013.10.024

where $p(\boldsymbol{X}|\boldsymbol{R})$ stands for the probability density function (p.d.f.) of the observations (which are assumed to be zero-mean) and $\boldsymbol{R}_{\text{ML}}$ denotes the maximum likelihood estimator (MLE) of $\boldsymbol{R}$. As demonstrated in [3–5] for Gaussian distributed data, the p.d.f of $LR(\boldsymbol{R}_0|\boldsymbol{X})$, where $\boldsymbol{R}_0$ is the *true* covariance matrix of $\boldsymbol{X}$, does not depend on $\boldsymbol{R}_0$ but is fully determined by M and T. Moreover, the effective support of this p.d.f. lies on an interval whose values are much below $1 = LR(\boldsymbol{R}_{\text{ML}}|\boldsymbol{X})$, see [3–5] for illustrative examples. In other words, the LR evaluated at the true covariance matrix is much lower than the LR evaluated at the MLE. This naturally raises the question of whether it would not make more sense that an estimate $\boldsymbol{R}(\hat{\boldsymbol{\beta}})$ of $\boldsymbol{R}_0$, where $\boldsymbol{R}(\boldsymbol{\beta})$ is either a parameterized model for the covariance matrix or a regularized estimate (e.g., shrinkage of the MLE to some target covariance matrix), results in a LR which is commensurate with that of $\boldsymbol{R}_0$. This is the gist of the EL approach which estimates $\boldsymbol{\beta}$ by enforcing that $LR(\boldsymbol{R}(\hat{\boldsymbol{\beta}})|\boldsymbol{X})$ takes values which are compatible with the support of the p.d.f. of $LR(\boldsymbol{R}_0|\boldsymbol{X})$. To be more specific, let us consider a classical regularized covariance matrix estimate (CME) based on shrinkage of the MLE to a target matrix $\boldsymbol{R}_t$, i.e.,

$$\boldsymbol{R}(\beta) = (1-\beta)\boldsymbol{R}_{\text{ML}} + \beta\boldsymbol{R}_t.$$

The EL approach for selection of the shrinkage factor β could possibly take the following form [4,1]:

$$\beta_{\text{EL}} = \arg\min_{\beta} \left| LR^{1/T}(\boldsymbol{R}(\beta)|\boldsymbol{X}) - \text{med}\left[\omega(LR|M,T)\right] \right|,$$

where $\omega(LR|M,T)$ is the true p.d.f. of $LR^{1/T}(\boldsymbol{R}_0|\boldsymbol{X})$ and $\text{med}\left[\omega(LR|M,T)\right]$ stands for the median value. In other words, the shrinkage factor is chosen such that the resulting LR of $\boldsymbol{R}(\beta_{\text{EL}})$ is comparable with that of $\boldsymbol{R}_0$. It is well known that regularization is particularly effective in low sample support and the EL principle was shown in [4,1] to provide a quite efficient mechanism to tune the regularization parameters. Various uses of the EL approach are possible and their effectiveness has been illustrated in different applications. For instance, it has been used successfully to detect severely erroneous MUSIC-based DOA estimates in low signal to noise ratio and it provided a mechanism to rectify the set of these estimates to meet the expected likelihood ratio values [3,5]. Accordingly, the EL approach was proven to be instrumental in designing efficient adaptive detectors in low sample support [4].

In [1,8] we extended the EL approach to a class of complex elliptically contoured distributions (ECD) (namely the $\mathcal{EMS}_{M,T}(\mathbf{0}, \boldsymbol{R}, \phi)$ type of distributions, as referred to in this paper) and we provided regularization schemes for covariance matrix estimation. Regularized covariance matrix estimation has been studied extensively in the literature, see e.g. [18,12,19,20,27] for a few examples within the framework of elliptically contoured distributions. In the latter references, the regularization parameters are selected with a view to minimize either the mean-square error or Stein loss. Our goal in this paper is not to derive and compare new covariance estimation schemes, as in [1]. Rather we focus herein in deriving invariance properties of the LR for other classes of complex elliptically contoured distributions, so as to extend the class of distributions for which the EL approach of covariance matrix estimation is feasible. How the EL approach will be used in this framework is beyond the scope of the present paper. The starting point of the present study is the following. While there is a general agreement and usually no ambiguity for defining vector elliptically contoured distributions, when it comes to extending ECD to matrix-variate, a certain number of options are possible [11]. Indeed, Fang and Zhang distinguish four classes of matrix-variate ECD whose p.d.f. and stochastic representations are different. As we shall see shortly, considering as in [1] the columns of $\boldsymbol{X}$ as independent and identically distributed (i.i.d.) elliptically distributed random vectors (r.v.) results in $\boldsymbol{X} \sim \mathcal{EMS}_{M,T}(\mathbf{0}, \boldsymbol{R}, \phi)$ (obtained from a multivariate spherical distribution in the terminology of [11]). On the other hand, the ECD considered, e.g., in [23,24] are obtained assuming that $\text{vec}(\boldsymbol{X}) \in \mathbb{C}^{MT\times 1}$ follows a vector ECD, which we will denote as $\boldsymbol{X} \sim \mathcal{EVS}_{M,T}(\mathbf{0}, \boldsymbol{R}, \phi)$.

In this paper we shall examine the p.d.f. of the likelihood ratio for two classes of complex ECD not covered in [1], namely $\boldsymbol{X} \sim \mathcal{ESS}_{M,T}(\mathbf{0}, \boldsymbol{R}, \phi)$ and $\boldsymbol{X} \sim \mathcal{EVS}_{M,T}(\mathbf{0}, \boldsymbol{R}, \phi)$. For the former, we will pay special attention to the matrix-variate Student distribution. The latter category was considered in [23,24] where Richmond proved the quite remarkable result that Kelly's generalized likelihood ratio test (GLRT) for Gaussian distributed data [17] was also the GLRT for this class of ECD. A main result of this paper includes stochastic representations of the likelihood ratio (and proof of invariance) in both the over-sampled case ($T \geq M$) and the under-sampled scenario where the number of available samples in less than the size of the observation space ($T \leq M$), in which case regularization is mandatory. Note that invariance properties of some likelihood ratios for elliptically contoured distributions (mostly EVS) have been studied, e.g., in [16,15,10,7,6], but the likelihood ratios are somewhat different from what we consider here and they serve different purposes.

2. A brief review of elliptically contoured distributions

In this section, we provide a brief summary of ECD with the only purpose of providing sufficient background for derivation and analysis of the LR in the next sections. We refer the reader to [11,6] for details that are skipped here and for an exhaustive analysis: our presentation here will follow the terminology of [11]. We also point to the recent paper [22] for an excellent comprehensive overview and applications to array processing. A vector $\boldsymbol{x} \in \mathbb{C}^M$ is said to be spherically distributed if its characteristic function $\mathcal{E}\{e^{i\text{Re}\{\boldsymbol{t}^H\boldsymbol{x}\}}\} = \phi(\boldsymbol{t}^H\boldsymbol{t})$: we will denote it as $\boldsymbol{x} \sim \mathcal{S}_M(\phi)$. Assuming that $\boldsymbol{x}$ has a density (which we will do through this document) the latter only depends on $\boldsymbol{x}^H\boldsymbol{x}$. A vector $\boldsymbol{x} \in \mathbb{C}^M$ is said to follow an elliptically contoured distribution if

$$\mathcal{E}\{e^{i\text{Re}\{\boldsymbol{t}^H\boldsymbol{x}\}}\} = \mathcal{E}\{e^{i\text{Re}\{\boldsymbol{t}^H\boldsymbol{m}\}}\}\phi(\boldsymbol{t}^H\boldsymbol{R}\boldsymbol{t}). \quad (2)$$