Journal of Multivariate Analysis 143 (2016) 32–55

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

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

Estimation of the inverse scatter matrix of an elliptically symmetric distribution

Dominique Fourdrinier[a], Fatiha Mezoued[b], Martin T. Wells[c,*]

[a] Normandie Université, Université de Rouen, LITIS EA 4108, Avenue de l'Université, BP 12, 76801 Saint-Étienne-du-Rouvray, France
[b] École Nationale Supérieure de Statistique et d'Économie Appliquée (ENSSEA ex-INPS), Algiers, Algeria
[c] Cornell University, Department of Statistical Science, 1190 Comstock Hall, Ithaca, NY 14853, USA

ARTICLE INFO

Article history:
Received 16 May 2015
Available online 3 September 2015

AMS 2010 subject classifications:
62H12
62F10
62C99

Keywords:
Elliptically symmetric distributions
High-dimensional statistics
Moore–Penrose inverse
Inverse scatter matrix
Quadratic loss
Singular sample covariance matrix
Sample eigenvalues
Stein–Haff identity

ABSTRACT

We consider estimation of the inverse scatter matrices Σ^{-1} for high-dimensional elliptically symmetric distributions. In high-dimensional settings the sample covariance matrix S may be singular. Depending on the singularity of S, natural estimators of Σ^{-1} are of the form $a\,S^{-1}$ or $a\,S^{+}$ where a is a positive constant and S^{-1} and S^{+} are, respectively, the inverse and the Moore–Penrose inverse of S. We propose a unified estimation approach for these two cases and provide improved estimators under the quadratic loss $\operatorname{tr}(\hat{\Sigma}^{-1} - \Sigma^{-1})^2$. To this end, a new and general Stein–Haff identity is derived for the high-dimensional elliptically symmetric distribution setting.

1. Introduction

The estimation of covariance and inverse covariance matrices in a high-dimensional framework has seen a surge of interest in the past years. Of these, estimates of the inverse covariance matrix are required in many multivariate inference procedures including the Fisher linear discriminant analysis, confidence intervals based on the Mahalanobis distance, optimal portfolio selection, graphical models, and weighted least squares estimator in multivariate linear regression models. Estimation of the precision matrix in the classical multivariate setting has been studied by Efron and Morris [12], Haff [21], Dey [9], Krishnamoorthy and Gupta [24], Dey et al. [10], Zhou et al. [43], and Tsukuma and Konno [42].

The natural estimator of the inverse covariance matrix, based on the sample covariance matrix, is well known to be inadequate in the high-dimensional context. When the dimension is of the same order of the sample size the sample covariance matrix becomes unstable and has large estimation error. It is also well known that the eigenvalues of sample covariance matrix are over-dispersed, that is, the eigenvalues of sample covariance matrix are not good estimators of their population counterpart Marčenko and Pastur [33]. Additionally, in the setting where the dimension of the sample covariance matrix is larger than the sample size, the inverse of the sample covariance matrix does not exist. An estimator of the precision

* Corresponding author.
E-mail addresses: Dominique.Fourdrinier@univ-rouen.fr (D. Fourdrinier), mezoued.fatiha@enssea.dz (F. Mezoued), mtw1@cornell.edu (M.T. Wells).

http://dx.doi.org/10.1016/j.jmva.2015.08.012

matrix for the multivariate normal distribution based on the Moore–Penrose generalized inverse of the sample covariance matrix was developed in Kubokawa and Srivastava [27]. Kubokawa and Inoue [25] consider general types of ridge estimators for covariance and precision matrices, and derive asymptotic expansions of their risk functions. More generally, the idea to correct (shrink) the eigenvalues of the sample covariance matrix is also found in previous work by Ledoit and Wolf [29], El Karoui [14], Ledoit and Wolf [30] and Donoho [11]. The problem has been examined under many sparsity scenarios, for example, zero elements of the matrix [2,13,38,6] or its inverse [34,20,37,28,7,36], bandedness [3,4] among others.

Most of the results for improved estimation for covariance and inverse covariance matrices have been developed in the context of the multivariate normal distribution. In this article we consider a large subclass of the elliptically contoured distributions. Let $(X, U) = (X, U_1, \ldots, U_n)$ be $n + 1$ p-dimensional random vectors having an elliptically symmetric distribution with joint density of the form

$$
\begin{aligned}
(x, u) \mapsto |\Sigma|^{-(n+1)/2} f\left((x-\theta)^\top \Sigma^{-1}(x-\theta) + \sum_{i=1}^n u_i^\top \Sigma^{-1} u_i\right) \\
= |\Sigma|^{-(n+1)/2} f\left(\mathrm{tr}\left[\Sigma^{-1}(x-\theta)(x-\theta)^\top + \Sigma^{-1} s\right]\right), \qquad (1.1)
\end{aligned}
$$

where X and the U_i's are $p \times 1$ vectors, θ is a $p \times 1$ unknown location vector, $S = UU^\top$ is a $p \times p$ matrix and Σ is a $p \times p$ unknown scatter matrix proportional to the covariance matrix. In the following, $E_{\theta,\Sigma}$ will denote the expectation with respect to the density in (1.1) and $E^*_{\theta,\Sigma}$ the expectation with respect to the density

$$
(x, u) \mapsto \frac{1}{K} |\Sigma|^{-(n+1)/2} F\left(\mathrm{tr}\left[\Sigma^{-1}(x-\theta)(x-\theta)^\top + \Sigma^{-1} s\right]\right), \qquad (1.2)
$$

where

$$
F(t) = \frac{1}{2} \int_t^\infty f(u) du \qquad (1.3)
$$

and

$$
K = \int_{\mathbb{R}^{p+k}} |\Sigma|^{-(n+1)/2} F\left(\mathrm{tr}\left[\Sigma^{-1}(x-\theta)(x-\theta)^\top + \Sigma^{-1} s\right]\right) dx\, du \qquad (1.4)
$$

is the normalizing constant which is assumed to be finite. Note that these two expectations are related since, for any integrable function $H(X, U)$, we have

$$
K\, E^*_{\theta,\Sigma}[H(X, U)] = E_{\theta,\Sigma}[\varphi_{\theta,\Sigma}(X, U)\, H(X, U)] \qquad (1.5)
$$

where

$$
\varphi_{\theta,\Sigma}(X, U) = \frac{F\left((X-\theta)^\top \Sigma^{-1}(X-\theta) + \sum_{i=1}^n U_i^\top \Sigma^{-1} U_i\right)}{f\left((X-\theta)^\top \Sigma^{-1}(X-\theta) + \sum_{i=1}^n U_i^\top \Sigma^{-1} U_i\right)}.
$$

The general model in (1.1) has been considered by various authors, for more details, see Fourdrinier, Strawderman and Wells [19] where the model is viewed as the canonical form of the general linear model. For more on elliptically symmetric distributions and the various choices of $f(\cdot)$ in (1.1) see Bilodeau and Brenner [5] and Fang, Kotz, and Ng [15]. The class in (1.1) contains models such as the multivariate normal, t-, and Kotz-type distributions. In the setting of the multivariate normal distribution, since $F = f$, we have $E_{\theta,\Sigma} = E^*_{\theta,\Sigma}$. Improved estimation of the scatter matrix for elliptical distribution models, from a decision theoretic point of view, has been considered by Fang and Li [16], Fang and Li [32], Leung and Ng [31], and Tsukuma [41].

In this article, we consider estimation of the inverse scatter matrix Σ^{-1} in (1.1) under the quadratic loss

$$
L(\hat{\Sigma}^{-1}, \Sigma^{-1}) = \mathrm{tr}\big((\hat{\Sigma}^{-1} - \Sigma^{-1})^2\big), \qquad (1.6)
$$

where $\hat{\Sigma}^{-1}$ estimates Σ^{-1} and $\mathrm{tr}(M)$ denotes the trace of a matrix M. By definition, the risk of $\hat{\Sigma}^{-1}$ is

$$
R(\hat{\Sigma}^{-1}, \Sigma^{-1}) = E_{\theta,\Sigma}[L(\hat{\Sigma}^{-1}, \Sigma^{-1})]. \qquad (1.7)
$$

When S is invertible ($p \le n$), the "usual" estimators are of the form $a\, S^{-1}$ for some positive constant a. Tsukuma [41] showed that there exists a_* such that, $a_* S^{-1}$ is unbiased where

$$
a_* = a_0 \left(\frac{2\,\pi^{p/2}}{\Gamma(p/2)} \int_0^\infty r^{p-1}\,(-2 f'(r^2))\, dr\right)^{-1}, \qquad (1.8)
$$

where $a_0 = n - p - 1$ and $S = UU^\top$. Note that for the normal distribution, $a_* = a_0$.