Shrinkage estimation of a mean matrix of a multivariate complex normal distribution

Yoshihiko Konno *

October 20, 2010

Abstract

The problem of estimating a mean matrix of a multivariate complex normal distribution with an unknown covariance matrix is considered under an invariant loss function. By using complex versions of the Stein identity, the Stein-Haff identity, and calculus on eigenvalues, a formula is obtained for an unbiased estimate of the risk of an invariant class of estimators, from which several minimax shrinkage estimators are constructed.

1 Introduction

The multivariate complex normal and complex Wishart distributions were first explored in Goodman [8], and followed by Khatri [14]. These models play an important role in signal processing methods. See Kay [13] and Schreier and Scharf [24] for the need of complex data models with complex parameters and DoGondžić and Neborai [6] for a unified approach based on complex GMANOVA models to analyze and extend signal processing models. Lillestøl [20] first investigated Stein-like shrinkage methods on simultaneous estimation of a mean vector of the complex normal model. However, shrinkage methods for these models have received less attention so far, although it is important to develop these methods beyond the maximum likelihood estimator of estimating the unknown signals in the multivariate complex normal distribution. The goal of this paper is to show how certain decision theoretical results concerning the problem of estimating a mean matrix of the real normal distribution can be extended to the complex multivariate normal case.

In this paper, we consider the problem of estimating an $m \times p$ unknown constant complex matrix $\boldsymbol{\Xi}$ that is observed with additive complex normal random errors in a decision theoretic set-up. Our

*Faculty of Science, Japan Women's University. email: konno@fc.jwu.ac.jp

observations are an $m \times p$ data matrix $\mathbf{Z}$ and a $p \times p$ positive definite Hermitian matrix $\mathbf{S}$, which is represented as

$$\begin{aligned} &\mathbf{Z} : m \times p \sim \mathbb{C}N_{m\times p}(\boldsymbol{\Xi}, \mathbf{I}_m \otimes \boldsymbol{\Sigma}), \\ &\mathbf{S} : p \times p \sim \mathbb{C}W_p(\boldsymbol{\Sigma}, n) \qquad \text{with } \mathbf{Z} \text{ and } \mathbf{S} \text{ independent,} \end{aligned} \tag{1}$$

where $n > p$, $\boldsymbol{\Sigma}$ is a $p \times p$ positive definite Hermitian constant matrix. Here we assume that $\boldsymbol{\Xi}$ and $\boldsymbol{\Sigma}$ are unknown. Furthermore $\mathbb{C}N_{m\times p}(\boldsymbol{\Xi}, \mathbf{I}_m \otimes \boldsymbol{\Sigma})$ and $\mathbb{C}W_p(\boldsymbol{\Sigma}, n)$ stand for a matrix-variate complex normal distribution with the mean matrix $\boldsymbol{\Xi}$ and the covariance matrix $\mathbf{I}_m \otimes \boldsymbol{\Sigma}$ and a complex Wishart distribution with the degree of freedom n and the parameters $\boldsymbol{\Sigma}$, respectively. In other words, the model (1) means that the density of $\mathbf{Z}$ with respect to the Lebesgue measure on $\mathbb{C}^{m\times p}$ is given as

$$\pi^{-mp}\mathrm{Det}(\boldsymbol{\Sigma})^{-m} \exp\big\{-\mathrm{Tr}\big((\mathbf{z} - \boldsymbol{\Xi})\boldsymbol{\Sigma}^{-1}(\mathbf{z} - \boldsymbol{\Xi})^*\big)\big\}, \qquad \mathbf{z} \in \mathbb{C}^{m\times p},$$

while the density of $\mathbf{S}$ with respect to the Lebesgue measure on $\mathbb{C}_H^{p\times p}$ is given by

$$\frac{\mathrm{Det}\,(\mathbf{s})^{n-p} \exp\big(-\mathrm{Tr}\,(\mathbf{s}\boldsymbol{\Sigma}^{-1})\big)}{\mathrm{Det}\,(\boldsymbol{\Sigma})^{n}\pi^{p(p-1)/2}\Pi_{k=1}^{p}\Gamma(n+1-k)}, \qquad \mathbf{s} \in \mathbb{C}_{+}^{p\times p}. \tag{2}$$

Here $\Gamma(\,\cdot\,)$ is the usual Gamma function, $\mathrm{Tr}\,(\,\cdot\,)$ and $\mathrm{Det}\,(\,\cdot\,)$ denote the trace and determinant of a square matrix, and the superscript "$*$" means the complex conjugate transpose of a matrix. Furthermore $\mathbb{C}^{m\times p}$, $\mathbb{C}_H^{p\times p}$, and $\mathbb{C}_{+}^{p\times p}$ stand for the sets of all $m \times p$ complex matrices, of all $p \times p$ Hermitian complex matrices, and of all $p \times p$ positive definite Hermitian complex matrices, respectively.

Based on $(\mathbf{Z}, \mathbf{S})$ we consider the problem of estimating the mean matrix $\boldsymbol{\Xi}$ with respect to a loss function

$$\mathscr{L}(\widehat{\boldsymbol{\Xi}}, (\boldsymbol{\Xi}, \boldsymbol{\Sigma})) = \mathrm{Tr}\,\{(\widehat{\boldsymbol{\Xi}} - \boldsymbol{\Xi})\boldsymbol{\Sigma}^{-1}(\widehat{\boldsymbol{\Xi}} - \boldsymbol{\Xi})^*\},$$

where an $m \times p$ random matrix $\widehat{\boldsymbol{\Xi}}$ is an estimator of $\boldsymbol{\Xi}$. The risk function corresponding to this loss function is

$$\mathscr{R}(\widehat{\boldsymbol{\Xi}}, (\boldsymbol{\Xi}, \boldsymbol{\Sigma})) = \mathbb{E}[\mathscr{L}(\widehat{\boldsymbol{\Xi}}, (\boldsymbol{\Xi}, \boldsymbol{\Sigma}))],$$

where the expectation above is taken with respect to the joint distribution of $(\mathbf{Z}, \mathbf{S})$.

This estimation problem is important since it is a prototype of estimating the regression matrix of a complex MANOVA model and of predicting multivariate responses in a linear regression complex model. We extend a large body of the results obtained by Efron and Morris [7], Bilodeau and Kariya [2], Kariya *et al.* [12], Konno [15], and van der Merwe and Zidek [30] in the multivariate real normal set-up to the complex normal set-up (1). The results in the real normal model were obtained by extensive use of the integration by parts approach, known as the Stein identity derived by Stein [26, 28], and the

Stein-Haff identity by Stein [27] and Haff [9, 10]. In addition to these identities, the eigenvalue calculus, developed by Loh [21, 22, 23], Konno [15], and Kariya *et al.* [12], is important to the development for a systematic search for shrinkage estimators. We extend these approaches to the complex normal setup. The Stein identity for the multivariate complex normal is easily derived by using an isomorphism between real and complex variables stated in Andersen *et al.* [1] while the Stein-Haff identity was extended to the complex Wishart distribution by Svensson and Lundberg [29]. These identities and the eigenvalue calculus for the complex matrix developed in this paper are exploited to establish a systematic search for shrinkage estimators for the model (1), which includes the FICYREG estimator of van der Merwe and Zidek [30].

Shrinkage methods for estimating the regression matrix in a multivariate linear regression model have been extensively investigated to overcome the shortcomings of the ordinary least squares estimator. The literature includes Brown and Zidek [4, 5] and van der Merwe and Zidek [30]. Later Breiman and Friedman [3] proposed to predict a future observation by a ridge-type shrinkage estimator in order to use information of correlated variables. See Solanky [25] for further investigation on this problem, which is application of minimax estimators to construct better predictors in order to overcome shortcomings of the predictor based on the least squares estimator. This shows that the results obtained in this paper can be immediately applied to the problem of predicting a future observation in a multivariate linear model for complex data.

The remaining parts of this papers are organized as follows. In Section 2, some notation used throughout this paper are introdouced. Next integration by parts formulae, complex versions of the Stein identity and the Stein-Haff identity, are given. These identities play vital roles in obtaining unbiased risk estimate in Section 3. In Section 3, we obtain unbiased risk estimate for invariant estimators, from which several shrinkage estimators are derived. The detailed proof for the results is available at http://mp-w3math.jwu.ac.jp/˜ konno/pdf/tr10.pdf.

2 Preliminaries: Notation and Basic identities

2.1 Notation

Let $\mathbb{R}$ and $\mathbb{C}$ denote the field of real and complex numbers, respectively. We represent any element $c \in \mathbb{C}$ as $c = a + \sqrt{-1}\, b$, where a, $b \in \mathbb{R}$. We also denote the real and imaginary parts of c by $\mathrm{Re}\, c$ and $\mathrm{Im}\, c$, respectively. In particular we denote by $\mathbb{R}_+$ the set of all positive real numbers. The conjugate of a complex number c is given by $\bar{c} := a - \sqrt{-1}\, b$. We define by $\mathbb{R}^p$ and $\mathbb{C}^p$ the sets of all p-tuples of real and complex numbers, respectively. We set $\mathbb{R}^p_> = \{(\ell_1, \ell_2, \ldots, \ell_p) \in \mathbb{R}^p : \ell_1 > \ell_2 > \cdots > \ell_p > 0\}$. In this paper, these tuples are represented as columns. The sets of all $m \times p$ matrices of real and complex entries are denoted by $\mathbb{R}^{m\times p}$ and $\mathbb{C}^{m\times p}$, respectively. The transpose and the conjugate of $\mathbf{C}$ is denoted by $\mathbf{C}'$ and $\overline{\mathbf{C}}$, respectively. Furthermore the conjugate transpose of an $m \times p$ matrix $\mathbf{C} \in \mathbb{C}^{m\times p}$ is denoted by $\mathbf{C}^* = \overline{\mathbf{C}'}$. The set of $p \times p$ Hermitian positive definite matrices is denoted by $\mathbb{C}^{p\times p}_+$. For any $\mathbf{c} = \mathbf{a} + \sqrt{-1}\,\mathbf{b} \in \mathbb{C}^p\,(\mathbf{a}, \mathbf{b} \in \mathbb{R}^p)$, we denote by $[\mathbf{c}]$ a $2p$-dimensional real vector $(\mathbf{a}', \mathbf{b}')'$. For a positive integer q and real numbers $a_1, a_2, \ldots, a_q$, $\mathrm{Diag}(a_1, a_2, \ldots, a_q)$ denotes a $q \times q$ diagonal matrix with the i-th diagonal element $a_i\,(i = 1, 2, \ldots, q)$. For an $m \times p$ complex matrix $\mathbf{C} = \mathbf{A} + \sqrt{-1}\,\mathbf{B}\,(\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times p})$, we denote by $\{\mathbf{C}\}$ a $2m \times 2p$ real matrix

$$\begin{pmatrix} \mathbf{A} & -\mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{pmatrix}.$$

Let $g(x, y)$ be a real-valued function on an open set $U \in \mathbb{R}^2$. We say that g is differentiable if $\partial g/\partial x$ and $\partial g/\partial y$ exist on U. Let u, v be real-valued functions on an open set $U \in \mathbb{R}^2$. A function $g := u + \sqrt{-1}\,v$ is called differentiable if u, v are differentiable. For $z = x + \sqrt{-1}\,y\,(x, y \in \mathbb{R})$ and differentiable function $g(z) = u(z) + \sqrt{-1}\,v(z)$, we define

$$\begin{aligned} \frac{\partial}{\partial z} g &= \frac{1}{2}\left(\frac{\partial}{\partial x} - \sqrt{-1}\,\frac{\partial}{\partial y}\right) g = \frac{1}{2}\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) + \frac{\sqrt{-1}}{2}\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right), \\ \frac{\partial}{\partial \bar{z}} g &= \frac{1}{2}\left(\frac{\partial}{\partial x} + \sqrt{-1}\,\frac{\partial}{\partial y}\right) g = \frac{1}{2}\left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}\right) + \frac{\sqrt{-1}}{2}\left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\right). \end{aligned}$$

It is checked directly that

$$\frac{\partial}{\partial z} z = 1, \qquad \frac{\partial}{\partial z} \bar{z} = 0, \qquad \frac{\partial}{\partial \bar{z}} z = 0, \qquad \frac{\partial}{\partial \bar{z}} \bar{z} = 1.$$

If g is differentiable, then

$$\overline{\frac{\partial}{\partial z} g} = \frac{\partial}{\partial \bar{z}} \bar{g}. \tag{3}$$

Let $\mathbf{G} = (g_{ij})_{i=1,2,\ldots,m,\, j=1,2,\ldots,p}$ be an $m \times p$ matrix, where g_{ij}'s are complex-valued differentiable functions on $\mathbb{C}^{m\times p}$. For $\mathbf{z} = (z_{ij})_{i=1,2,\ldots,m,\, j=1,2,\ldots,p} \in \mathbb{C}^{m\times p}$, we set

$$\nabla_{\mathbf{z}} = \left(\frac{\partial}{\partial z_{ij}}\right)_{i=1,2,\ldots,m,\, j=1,2,\ldots,p},$$

and we define

$$\mathrm{Re}(\mathrm{Tr}\,(\nabla_z' \mathbf{G})) = \mathrm{Tr}\,(\mathrm{Re}(\nabla_z' \mathbf{G})) = \sum_{j=1}^{p}\sum_{i=1}^{m}\left\{\frac{\partial(\mathrm{Re}\, g_{ij})}{\partial(\mathrm{Re}\, z_{ij})} + \frac{\partial(\mathrm{Im}\, g_{ij})}{\partial(\mathrm{Im}\, z_{ij})}\right\}.$$

2.2 Complex normal distributions and the Stein identity

Recall that a $p \times 1$ complex random vector Z is said to have a p-variate complex normal distribution with a mean vector $\theta \in \mathbb{C}^p$ and a covariance matrix $\mathbf{\Sigma} \in \mathbb{C}_+^{p\times p}$ if the density of Z with respect to Lebesgue measure on $\mathbb{C}^p$ is given as

$$f_Z(\mathbf{z}) = \frac{1}{\pi^p}\mathrm{Det}(\mathbf{\Sigma})^{-1}\exp\{-(\mathbf{z}-\theta)^*\mathbf{\Sigma}^{-1}(\mathbf{z}-\theta)\}, \qquad \mathbf{z} \in \mathbb{C}^p.$$

We use the notation $Z \sim \mathbb{C}N_p(\theta, \mathbf{\Sigma})$ for this.

Lemma 1. *Let Z be a $p \times 1$ complex random vector having $\mathbb{C}N_p(\theta, \mathbf{\Sigma})$ and let $\mathbf{g} = (g_1, g_2, \dots, g_p) : \mathbb{C}^p \to \mathbb{C}^p$ be differentiable with*

$$\mathbb{E}\left|\frac{\partial\,(\mathrm{Re}\, g_i)}{\partial\,(\mathrm{Re}\, z_i)}\right|_{\mathbf{z}=Z} < \infty, \qquad \mathbb{E}\left|\frac{\partial\,(\mathrm{Im}\, g_i)}{\partial\,(\mathrm{Im}\, z_i)}\right|_{\mathbf{z}=Z} < \infty, \qquad i = 1, 2, \dots, p.$$

Then we have

$$\mathbb{E}[(Z-\theta)^*\mathbf{\Sigma}^{-1}\mathbf{g}(Z) + \mathbf{g}^*(Z)\mathbf{\Sigma}^{-1}(Z-\theta)] \quad = \quad \mathbb{E}\left[\sum_{i=1}^p \left\{\frac{\partial\,(\mathrm{Re}\, g_i)}{\partial\,(\mathrm{Re}\, z_i)} + \frac{\partial\,(\mathrm{Im}\, g_i)}{\partial\,(\mathrm{Im}\, z_i)}\right\}\Bigg|_{\mathbf{z}=Z}\right].$$

2.3 Complex Wishart distributions and the Stein-Haff identity

Assume that a $p\times p$ Hermitian positive definite matrix $\mathbf{S}$ has a complex Wishart distribution $\mathbb{C}W_p(\mathbf{\Sigma}, n)$ with the density function (2). Let $\mathbf{G}(\mathbf{S})$ be a $p \times p$ matrix, the (i, j) element $g_{ij}(\mathbf{S})$ of which is a complex-valued function of $\mathbf{S} = (s_{ij})$. For a $p \times p$ Hermitian matrix $\mathbf{S} = (s_{jk})$, let $\mathbf{D}_S = (\partial/\partial s_{jk})$ be a $p \times p$ operator matrix, the (j, k) element of which is given by

$$\frac{\partial}{\partial s_{jk}} = \frac{1}{2}(1+\delta_{jk})\left\{\frac{\partial}{\partial(\mathrm{Re}\, s_{jk})} + \sqrt{-1}\frac{\partial}{\partial(\mathrm{Im}\, s_{jk})}\right\}, \qquad j, k = 1, 2, \dots, p. \tag{4}$$

Here δ_{jk} is the Kronecker delta ($=1$ if $j=k$ and $=0$ if $j \neq k$). Thus the (j, k) element of $\mathbf{D}_S\mathbf{G}(\mathbf{S})$ is

$$\{\mathbf{D}_S\mathbf{G}(\mathbf{S})\}_{jk} = \sum_{l=1}^p \frac{\partial g_{lk}}{\partial s_{jl}}(\mathbf{S}) = \frac{1}{2}(1+\delta_{jl})\sum_{l=1}^p\left\{\frac{\partial g_{lk}}{\partial(\mathrm{Re}\, s_{jl})}(\mathbf{S}) + \sqrt{-1}\frac{\partial g_{lk}}{\partial(\mathrm{Im}\, s_{jl})}(\mathbf{S})\right\}.$$

Lemma 2. *Assume that each entry of $\mathbf{G}(\mathbf{S})$ is a partially differentiable function with respect to $\mathrm{Re}\, s_{jk}$ and $\mathrm{Im}\, s_{jk}$, $j, k = 1, 2, \dots, p$. Under conditions on $\mathbf{G}(\mathbf{S})$ specified in* Konno [17]*, the following identity holds:*

$$\mathbb{E}[\mathrm{Tr}(\mathbf{G}(\mathbf{S})\mathbf{\Sigma}^{-1})] = \mathbb{E}[(n-p)\mathrm{Tr}(\mathbf{G}(\mathbf{S})\mathbf{S}^{-1}) + \mathrm{Tr}(\mathbf{D}_S\mathbf{G}(\mathbf{S}))]. \tag{5}$$

Remark 1. The Stein-Haff identity was extended to an elliptically contoured complex distribution by Konno [17]. Hence, if we know the improved estimators for the normal case, we can establish the robustness of improvement for the elliptically contoured complex distribution in a manner similar to that demonstrated in Kubokawa and Srivastava [18, 19].

3 Unknown case and invariant loss

3.1 Unbiased risk estimate for a class of invariant estimators

Consider a class of estimators of the form $\mathbf{Z}+\mathbf{G}(\mathbf{Z},\,\mathbf{S})$, where $\mathbf{G} := \mathbf{G}(\mathbf{Z},\,\mathbf{S})$ is an $m\times p$ matrix whose $(i,\,j)$ element $g_{ij}\,(i = 1,\,2,\,\ldots,\,m,\,j = 1,\,2,\ldots,\,p)$ is a complex-valued function based on $(\mathbf{Z},\,\mathbf{S})$.

Lemma 3. *Assume that all elements of* $\mathbf{G}(\mathbf{Z},\,\mathbf{S})$ *are absolutely continuous functions of* $\mathbf{Z}$ *and* $\mathbf{S}$*. Then we have*

$$\begin{aligned}\mathscr{R}(\mathbf{Z}+\mathbf{G}(\mathbf{Z},\,\mathbf{S}),\,(\boldsymbol{\Xi},\,\boldsymbol{\Sigma})) \;&=\; mp+\mathbb{E}\big[2\,\mathrm{Tr}\,\{\mathrm{Re}\,(\nabla'_Z\mathbf{G}(\mathbf{Z},\,\mathbf{S}))\}+\mathrm{Tr}\,\{\mathbf{D}_S\mathbf{G}^*(\mathbf{Z},\,\mathbf{S})\mathbf{G}(\mathbf{Z},\,\mathbf{S})\}\\ &\quad+(n-p)\mathrm{Tr}\,\{\mathbf{G}^*(\mathbf{Z},\,\mathbf{S})\mathbf{G}(\mathbf{Z},\,\mathbf{S})\mathbf{S}^{-1}\}\big]. \end{aligned}\tag{6}$$

To describe our class of estimators, let $\mathbf{F} = \mathrm{Diag}\,(f_1,\,f_2,\,\ldots,\,f_{\min(m,\,p)})$ be the eigenvalues of $\mathbf{Z}^*\mathbf{Z}\mathbf{S}^{-1}$. For $p>m$ decompose $\mathbf{Z}\mathbf{S}^{-1}\mathbf{Z}^* = \mathbf{U}\mathbf{F}\mathbf{U}^*$, where $\mathbf{U}$ is an $m\times m$ unitary matrix. For $m>p$ we decompose $\mathbf{S} = (\mathbf{A}^*)^{-1}\mathbf{A}^{-1}$ and $\mathbf{Z}^*\mathbf{Z} = (\mathbf{A}^*)^{-1}\mathbf{F}\mathbf{A}^{-1}$, where $\mathbf{A}$ is a $p\times p$ non-singular matrix. We consider a class of estimators of the form

$$\widehat{\boldsymbol{\Xi}}_H := \widehat{\boldsymbol{\Xi}}_H(\mathbf{Z},\,\mathbf{S}) = \begin{cases}\mathbf{Z}\{\mathbf{I}_p+\mathbf{A}\mathbf{H}(\mathbf{F})\mathbf{A}^{-1}\} & \text{if } m>p\\ \{\mathbf{I}_m+\mathbf{U}\mathbf{H}(\mathbf{F})\mathbf{U}^*\}\mathbf{Z} & \text{if } p>m\end{cases}, \tag{7}$$

where $\mathbf{H} := \mathbf{H}(\mathbf{F}) = \mathrm{Diag}\,(h_1(\mathbf{F}),\,h_2(\mathbf{F}),\,\ldots,\,h_{\min(m,\,p)}(\mathbf{F}))$ whose i-th element $h_i := h_i(\mathbf{F})$, $i = 1,\,2,\,\ldots,\,\min(m,\,p)$, is a real-valued function on $\mathbb{R}_{>}^{\min(m,\,p)}$.

Let

$$\begin{aligned}\widehat{\Delta}(n,\,m,\,p;\,\mathbf{H}) \;&=\; \sum_{k=1}^{p}\Big\{2(m-p+1)h_k(\mathbf{F})+2f_kh_{kk}(\mathbf{F})+4\sum_{b>k}\frac{f_kh_k(\mathbf{F})-f_bh_b(\mathbf{F})}{f_k-f_b}\\ &\quad+(n+p-2)f_kh_k^2(\mathbf{F})-2f_k^2h_{kk}(\mathbf{F})h_k(\mathbf{F})-2\sum_{b>k}\frac{f_k^2h_k^2(\mathbf{F})-f_b^2h_b^2(\mathbf{F})}{f_k-f_b}\Big\}\\ &\quad\times\mathbb{1}\{f_1>f_2>\cdots>f_p>0\}, \end{aligned}\tag{8}$$

where $h_{kk}(\mathbf{F}) = (\partial h_k/\partial f_k)(\mathbf{F})$, $k = 1,\,2,\,\ldots,\,p$, and $\mathbb{1}\{f_1>f_2>\cdots>f_p>0\}$ is an indicator function of the set $\{(f_1,\,\ldots,\,f_p)\in\mathbb{R}^p : f_1>f_2>\cdots>f_p>0\}$.

Proposition 1. *Under the suitable conditions, we have*

$$\mathscr{R}(\widehat{\boldsymbol{\Xi}}_H,\, (\boldsymbol{\Xi},\, \boldsymbol{\Sigma})) \quad = \quad \begin{cases} mp + \mathbb{E}\big[\widehat{\Delta}(n,\, m,\, p;\, \mathbf{H})\big] & \text{if } m > p \\ mp + \mathbb{E}\big[\widehat{\Delta}(n+m-p,\, p,\, m;\, \mathbf{H})\big] & \text{if } p > m \end{cases}.$$

Remark 2. Assume that $\boldsymbol{\Xi} = \mathbf{0}$ in (1). From [14], the joint distribution of the eigenvalues of $\mathbf{Z}^*\mathbf{Z}\mathbf{S}^{-1}$ is, aparting from normalizing constants,

$$\Pi_{k=1}^{p} \frac{f_k^{m-p}}{(1+f_k)^{n+m}} \Pi_{k=1}^{p-1} \Pi_{j=k+1}^{p} (f_k - f_j)^2 \Pi_{k=1}^{p} \mathrm{d} f_k$$

if $m > p$ while it is

$$\Pi_{k=1}^{m} \frac{f_k^{p-m}}{(1+f_k)^{n+m}} \Pi_{k=1}^{m-1} \Pi_{j=k+1}^{m} (f_k - f_j)^2 \Pi_{k=1}^{m} \mathrm{d} f_k$$

if $p > m$. Note that the substitution rule to get the second distribution from the first distribution, i.e.,

$$(p,\, m,\, n) \to (m,\, p,\, n+m-p)$$

is valid to obtain the second assertion of Proposition 1 from the first assertion of Proposition 1. Hence, if we know the estimator of the form $\mathbf{Z}\{\mathbf{I}_p + \mathbf{A}\mathbf{H}\mathbf{A}^{-1}\}$ when $m > p$, we can easily write down estimators of the form $\{\mathbf{I}_m + \mathbf{U}\mathbf{H}\mathbf{U}^*\}\mathbf{Z}$ when $p > m$ by using the above substitution rule.

3.2 Alternative estimators

Proposition 2. *Let* $\gamma_1(\mathbf{F}),\, \gamma_2(\mathbf{F}),\, \ldots,\, \gamma_{\min(m,\, p)}(\mathbf{F})$ *be functions satisfying*
(i) $0 \le \gamma_k(\mathbf{F}) \le \max\{2(m-p)/(n+p),\, 2(p-m)/(n+2m-p)\}$*;*
(ii) $(\partial \gamma_k / \partial f_k)(\mathbf{F}) \ge 0$ *for* $k = 1,\, 2,\, \ldots,\, \min(m,\, p)$*;*
(iii) $\gamma_1(\mathbf{F}) \ge \gamma_2(\mathbf{F}) \ge \cdots \ge \gamma_{\min(m,\, p)}(\mathbf{F})$.
Then the estimator (7) *with*

$$\mathbf{H}(\mathbf{F}) = -\mathrm{Diag}\left(\frac{\gamma_1(\mathbf{F})}{f_1},\, \frac{\gamma_2(\mathbf{F})}{f_2},\, \ldots,\, \frac{\gamma_{\min(m,\, p)}(\mathbf{F})}{f_{\min(m,\, p)}}\right)$$

is minimax.

Corollary 1. *The* Efron-Morris *estimator*

$$\widehat{\boldsymbol{\Xi}}^{(EM)} = \begin{cases} \mathbf{Z}\{\mathbf{I}_p - \dfrac{m-p}{n+p}(\mathbf{Z}\mathbf{Z}^*)^{-1}\mathbf{S}\} & \text{if } m > p \\ \{\mathbf{I}_m - \dfrac{p-m}{n+2m-p}(\mathbf{Z}\mathbf{S}^{-1}\mathbf{Z}^*)^{-1}\}\mathbf{Z} & \text{if } p > m \end{cases}$$

is minimax.

Proposition 3. *For $k = 1, 2, \ldots, \min(m, p)$, let*

$$c_k^{(AS)} = \frac{m+p-2k}{n-p+2k}, \qquad \mathbf{H}^{(AS)}(\mathbf{F}) = -\mathrm{Diag}\left(\frac{c_1^{(AS)}}{f_1}, \frac{c_2^{(AS)}}{f_2}, \ldots, \frac{c_{\min(m,p)}^{(AS)}}{f_{\min(m,p)}}\right).$$

Then the estimator

$$\widehat{\mathbf{\Xi}}^{(AS)} = \begin{cases} \mathbf{Z}\{\mathbf{I}_p + \mathbf{A}\mathbf{H}^{(AS)}(\mathbf{F})\mathbf{A}^{-1}\} & \text{if } m > p \\ \{\mathbf{I}_m + \mathbf{U}\mathbf{H}^{(AS)}(\mathbf{F})\mathbf{U}^*\}\mathbf{Z} & \text{if } p > m \end{cases}$$

is minimax.

References

[1] H.H. Andersen, M. Højbjerre, D. Sørensen, P.S. Eriksen, Linear and Graphical Models, Springer-Verlag, New York (1995).

[2] M. Bilodeau, T. Kariya, Minimax estimators in the normal MANOVA model, J. Multivariate Anal. **28** (1989) 260–270.

[3] L. Breiman, J.H. Friedman, Predicting multivariate responses in multiple regression, J. Roy. Statist. Soc. Ser. B **59** (1998) 3–54.

[4] P.J. Brown, J.V. Zidek, Adaptive multivariate ridge regression, Ann. Statist. **8** (1980) 64–74.

[5] P.J. Brown, J.V. Zidek, Multivariate regression shrinkage estimators with unknown covariance matrix, Scand. J. Statist. **9** (1982) 209–215.

[6] A. DoGondžić, A. Neborai, Generalized multivariate analysis of variance, IEEE Signal Processing Magazin, September (2003) 39-54.

[7] B. Efron, C. Morris, Families of minimax estimators of the mean of a multivariate normal distribution, Ann. Statist. **4** 11–21.

[8] N.R. Goodman, Statistical analysis based on a certain multivariate complex Gaussian distribution (An introduction), Ann. Math. Statist. **34** (1963) 152-176.

[9] L.R. Haff, An identity for the Wishart distribution with application, J. Multivariate Anal. **9** (1979) 531–542.

[10] L.R. Haff, Estimation of the inverse covariance matrix: random mixtures of the inverse Wishart matrix and the identity, Ann. Statist. **7** (1979) 1264-1276.

[11] T. Kariya, Y. Konno, W.E. Strawderman, Double shrinkage estimators in the GMANOVA model, J. Multivariate Anal. **56** (1996) 245–258.

[12] T. Kariya, Y. Konno, W.E. Strawderman, Construction of shrinkage estimators for the regression coefficient matrix in GMANOVA model, Commn. Statist. Theory Methods **28** (1999) 597–611.

[13] S.M. Kay, Fundamentals of statistical signal processing: Estimation theory, Prentice Hall PTR (1993).

[14] C.G. Khatri, Classical statistical analysis based on a certain multivariate complex Gaussian distribution, Ann. Math. Statist. **36** (1965) 98-114.

[15] Y. Konno, On estimation of a matrix of normal means with unknown covariance matrix. J. Multivariate Anal. **36** (1991) 44–55.

[16] Y. Konno, Improved estimation of matrix of normal mean and eigenvalues in the multivariate F-distributions, Doctoral dissertation, University of Tsukuba (1992).

[17] Y. Konno, Improving on the sample covariance matrix for a complex elliptically contoured distribution, J. Statist. Plan. Infer. **107** (2007) 67-75.

[18] T. Kubokawa, M.S. Srivastava, Robust improvement in estimation of a covariance matrix in an elliptically contoured distribution, Ann. Statist. **27** (1999) 600–609.

[19] T. Kubokawa, M.S. Srivastava, Robust improvement in estimation of a mean matrix in an elliptically contoured distribution, J. Multivariate Anal. **76** (2001) 138–152.

[20] J. Lillstøl, Improved estimators for multivariate complex-normal regression with application to analysis of linear time-invariant relationships, J. Multivariate Anal. **7** (1977) 512–524.

[21] W.L. Loh, Estimating covariance matrices, Ph.D. thesis (1988), Stanford University.

[22] W.L. Loh, Estimating covariance matrices, Ann. Statist. **19** (1991) 283–296.

[23] W.L. Loh, Estimating covariance matrices II, J. Multivariate Anal. **36** (1991) 263–174.

[24] P.J. Schreier, and L.L. Scharf, Statistical Signal Processing of Complex-valued Data, Cambridge Univ. Press (2010).

[25] M.S. Srivastava, and T.K.S. Solanky, Predicting multivariate response in linear regression model, Comm. Statist. Simulation Comput. **32** (2003) 389–409.

[26] C. Stein, Estimation of the mean of a multivariate normal distribution, in Proc. Prague Symp. Asymptotic Statist. (1973) 35–381.

[27] C. Stein, Lectures on the theory of estimation of many parameters, in Studies in the Statistical Theory of Estimation I(I. A. Ibragimov and M. S. Nikulin, eds.) (1977).

[28] C. Stein, Estimation of the mean of a multivariate normal distribution, Ann. Statist. **9** (1981) 1135–1151.

[29] L. Svensson, M. Lundberg, Estimating complex covariance matrix, Signals, Systems and Computers, Conference Record of the Thirty-Eighth Asilomar Conference on (2004) 7-10, 2151 - 2154.

[30] A. van der Merwe, J.V. Zidek, Multivariate regression analysis and canonical variates, Can. J. Statist. **8** (1980), 27–39.

[31] J.V. Zidek, Deriving unbiased risk estimators of multinormal mean and regression coefficient estimators using zonal polynomials, Selecting a minimax estimator doing well at a point, Ann. Statist. **6** (1978) 679–782.