A CLASS OF ORTHOGONALLY INVARIANT MINIMAX ESTIMATORS FOR NORMAL COVARIANCE MATRICES PARAMETRIZED BY SIMPLE JORDAN ALGEBARAS OF DEGREE 2

Yoshihiko Konno

Faculty of Science, Japan Women's University, Tokyo, Japan

Dedicated to Professor A.K. Md. E. Saleh on the occasion of his 75th birthday

Normal covariance models parametrized by simple Jordan algebras of degree 2, equivalently the Lorentz cone, were first discussed by Tolver Jensen (1988) who described, for the family of normal distributions, the structure of those statistical models which are linear in both covariance and inverse covariance. Recently Konno (2007) developed minimax estimation for normal covariance matrices parametrized by the irreducible symmetric cones which include the Lorentz cone. In this paper a new class of minimax estimators is proposed for the Lorentz Wishart models under Stein's loss function. This class includes analogue of estimators by Dey and Srinivasan (1985) and Perron (1992) for the real Wishart models.

Keywords and phrases: *Wishart distribution, Lorentz cone, Stein estimators, Jordan algebras, Symmetric cones.*

1. Introduction

James and Stein (1961) first employed the Stein loss function and considered the problem of minimax estimation of a mean matrix of the real Wishart distribution. They used a result of Kiefer (1957) to construct minimax estimators having a constant risk. Later Stein (1977) pointed out that the eigenvalues of the Wishart matrix spread out more than the eigenvalues of the expected value of the Wishart matrix. This phenomenon suggests that the eigenvalues of the Wishart matrix should be shrunk toward a middle value of the eigenvalues. Furthermore he gave an unbiased risk estimate for a class of orthogonally invariant estimators, from which he obtained minimax estimators which are uniformly better than the James and Stein estimator. Recently Konno (2007) extended these results to the estimation problem of general Wishart distributions on the symmetric cones. These general models include the Wishart models of real, complex, and quaternion entries, and the Lorentz Wishart models.

The Lorentz Wishart models were originally discussed by Tolver Jensen (1988) who describe, for the family of normal distributions, the structure of those statistical models which are linear in both covariance and inverse covariance. In this paper we focus on the problem of estimating the mean of the Lorentz Wishart distributions which are the special case of the Wishart models on the symmetric cones.

In Section 2, a brief introduction to simple Jordan algebras of degree 2 is given. In Section 3, estimation problem is introduced while in Section 4, minimax estimators are given by following the results in Konno (2007). In Section 5, a class of orthogonally invariant minimax estimators is given by using an unbiased risk estimate for the orthogonally invariant estimators given in Konno (2007). In the Appendix, technical proofs are given.

2 Simple Jordan algebras of degree 2

In this section we review some notion of simple Jordan algebra of degree 2 which is a special case of finite-dimensional Euclidean simple Jordan algebras. See Faraut and Korányi (1994) for recent results on the symmetric cones and Jordan algebras.

Let $\mathscr{W}$ be a real vector space of dimension $v-1\,(v\geq 3)$ with a symmetric bilinear form B. We denote by $\|\cdot\|_B$ a norm induced by this symmetric bilinear form B. We assume that $\mathscr{V}=\mathbb{R}\times\mathscr{W}$ has a Jordan multiplication defined for $(\alpha,\,\mathbf{a})$, $(\beta,\,\mathbf{b})$ in $\mathbb{R}\times\mathscr{W}$ by

$$(\alpha,\,\mathbf{a})\circ(\beta,\,\mathbf{b})=(\alpha\beta+B(\mathbf{a}|\,\mathbf{b}),\,\alpha\mathbf{b}+\beta\mathbf{a}). \tag{2.1}$$

Throughout the paper, the small cap alphabets are used for elements in $\mathscr{V}$, the small cap Greek letters are used for elements in $\mathbb{R}$, and bold faced small cap alphabets are used for elements in the vector space $\mathscr{W}$. For example, we write $a = (\alpha,\, \mathbf{a}) \in \mathbb{R} \times \mathscr{W}$ for $a \in \mathscr{V}$.

Note that the identity element in $\mathscr{V}$ is given by $(1,\, \mathbf{0})$ where $\mathbf{0}$ is the zero vector in $\mathscr{W}$ and that the multiplication $\circ$ on $\mathscr{V}$ satisfies for all $a,\, b \in \mathscr{V}$;

$$\begin{aligned} a \circ b &= b \circ a, \\ a \circ (a^2 \circ b) &= a^2 \circ (a \circ b), \end{aligned}$$

where $a^2 = a \circ a$. However, the multiplication $\circ$ does not satisfy the communicative law. The associated symmetric cone Ω is the well-known Lorentz cone in the enveloping space $\mathscr{V}$ defined by

$$\Omega = \{a = (\alpha,\, \mathbf{a}) \in \mathbb{R} \times \mathscr{W};\ \alpha > 0,\ \alpha^2 - B(\mathbf{a}|\, \mathbf{a}) > 0\}.$$

We define the trace and determinant of $a = (\alpha,\, \mathbf{a}) \in \mathscr{V} = \mathbb{R} \times \mathscr{W}$ by

$$\mathrm{tr}(a) = 2\alpha, \qquad \det(a) = \alpha^2 - B(\mathbf{a}|\, \mathbf{a}). \tag{2.2}$$

If $\det(a) \neq 0$, then we define the inverse of a by

$$a^{-1} = \frac{1}{\det(a)}(\alpha,\, -\mathbf{a}).$$

Note that, by (2.2),

$$a \circ a^{-1} = a^{-1} \circ a = \frac{1}{\det(a)}(\alpha^2 - B(\mathbf{a}|\, \mathbf{a}),\, \mathbf{0}) = (1,\, \mathbf{0}),$$

and that, by (2.1) and (2.2), $\mathscr{V}$ is equipped with the inner product

$$(a|\, b) = \mathrm{tr}(a \circ b) = 2\{\alpha\beta + B(\mathbf{a}|\, \mathbf{b})\},$$

for $(\alpha,\, \mathbf{a})$, $(\beta,\, \mathbf{b})$ in $\mathbb{R} \times \mathscr{W}$. We denote by $\|\cdot\|$ a norm on $\mathscr{V}$ induced by the above inner product. Note that, from the definition of the norms on $\mathscr{V}$ and $\mathscr{W}$ and the multiplication rule on $\mathscr{V}$, we have

$$\|(0,\, \mathbf{a})\|^2 = \mathrm{tr}\,\{(0,\, \mathbf{a}) \circ (0,\, \mathbf{a})\} = \mathrm{tr}\,(B(\mathbf{a}|\, \mathbf{a}),\, \mathbf{0}) = 2B(\mathbf{a}|\, \mathbf{a}) = 2\,\|\mathbf{a}\|_B^2, \tag{2.3}$$

for $\mathbf{a} \in \mathscr{W}$.

We write a Jordan frame, a complete system of orthogonal primitive idempotents, as

$$c_1 = \frac{1}{2}(1,\, \mathbf{h}), \qquad c_2 = \frac{1}{2}(1,\, -\mathbf{h}),$$

where $\mathbf{h} \in \mathscr{W}$ is fixed such that $B(\mathbf{h}|\, \mathbf{h}) = 1$. Let $\mathscr{V}_1$, $\mathscr{V}_2$, and $\mathscr{V}_{12}$ be a subspace of $\mathscr{V}$ defined by

$$\mathscr{V}_1 = \mathbb{R}c_1,\ \mathscr{V}_2 = \mathbb{R}c_2,\ \mathscr{V}_{12} = \{(0,\, \mathbf{u});\ \mathbf{u} \in \mathscr{W},\ B(\mathbf{u}|\, \mathbf{h}) = 0\}.$$

Then note that $c_1^2 = c_1$, $c_2^2 = c_2$, $c_1 \circ c_2 = 0$, and $c_1 + c_2 = (1,\, \mathbf{0})$. Then we see that, for all $u \in \mathscr{V}_{12}$, $c_1 \circ u = (1/2)u$ and $c_2 \circ u = (1/2)u$. It is known that $\mathscr{V}$ is decomposed into $\mathscr{V} = \mathscr{V}_1 \oplus \mathscr{V}_{12} \oplus \mathscr{V}_2$. In fact, we have, for $a = (\alpha,\, \mathbf{a}) \in \mathscr{V}$,

$$a = (\alpha + B(\mathbf{a}|\, \mathbf{h}))c_1 \oplus (0,\, \mathbf{a} - B(\mathbf{a}|\, \mathbf{h})\mathbf{h}) \oplus (\alpha - B(\mathbf{a}|\, \mathbf{h}))c_2.$$

This decomposition is a special case of the Peirce decomposition for finite-dimensional Euclidean Jordan algebras.

For $a = (\alpha, \mathbf{a})$ and $b = (\beta, \mathbf{b})$ in $\mathscr{V}$, the quadratic representation is defined by

$$\begin{aligned} P(a)b &= 2a \circ (a \circ b) - a^2 \circ b \\ &= 2(\alpha\beta + B(\mathbf{a}|\,\mathbf{b}))(\alpha,\, \mathbf{a}) + \det(a)(-\beta,\, \mathbf{b}). \end{aligned} \tag{2.4}$$

For $a,\, b \in \mathscr{V}$, we write $a = \alpha_1 c_1 + \alpha_2 c_2 + a_{12}$ and $b = \beta_1 c_1 + \beta_2 c_2 + b_{12}$, where $\alpha_1,\, \alpha_2,\, \beta_1,\, \beta_2 \in \mathbb{R}$ and $a_{12} = (0,\, \mathbf{a}),\, b_{12} = (0,\, \mathbf{b}) \in \mathscr{V}_{12}$. From the definition of $c_1,\, c_2$, note that $\alpha = \alpha_1 + \alpha_2$ and that $\beta = \beta_1 + \beta_2$. Then a triangular subgroup transformation is defined by

$$T(a)b = \alpha_1^2 \beta_1 c_1 \oplus \big(\alpha_1\beta_1 a_{12} + \alpha_1\alpha_2 b_{12}\big) \oplus \bigg(\beta_1 \|\mathbf{a}\|_B^2 + 2\alpha_2^2 B(\mathbf{a}|\,\mathbf{b}) + \alpha_2^2\beta_2\bigg) c_2. \tag{2.5}$$

We denote by $\mathscr{T}$ the set of all triangular transformations given by (2.5).

From the fact that $(1,\, \mathbf{0}) = c_1 + c_2$ and (2.5), we have

$$T(b)(1,\, \mathbf{0}) = \beta_1^2 c_1 \oplus \beta_1(0,\, \mathbf{b}) \oplus (\|\mathbf{b}\|_B^2 + \beta_2^2) c_2,$$

where $b = \beta_1 c_1 \oplus (0,\, \mathbf{b}) \oplus \beta_2 c_2$. Then we have $T(b)(1, \mathbf{0}) = a$ if we set

$$\beta_1 = \sqrt{\alpha + B(\mathbf{a}|\,\mathbf{h})}, \quad \beta_2 = \sqrt{\frac{\alpha^2 - \|\mathbf{a}\|_B^2}{\alpha + B(\mathbf{a}|\,\mathbf{h})}}, \quad \mathbf{b} = \frac{1}{\sqrt{\alpha + B(\mathbf{a}|\,\mathbf{h})}} \bigg(\mathbf{a} - B(\mathbf{a}|\,\mathbf{h})\mathbf{h}\bigg). \tag{2.6}$$

3 Estimation problem

Consider that $\mathbf{\Lambda} :\, \mathbb{R} \times \mathscr{W} \mapsto \mathbb{R}_S^{p\times p}$ is a one-to-one Jordan algebra homeomorphism, i.e., Λ is a one-to-one linear mapping such that $\mathbf{\Lambda}(xy) = (1/2)(\mathbf{\Lambda}(x)\mathbf{\Lambda}(y) + \mathbf{\Lambda}(y)\mathbf{\Lambda}(x))$ for $x,\, y \in \mathbb{R} \times \mathscr{W}$. Here $\mathbb{R}_S^{p\times p}$ is the space of $p \times p$ symmetric matrices and I_p denotes the $p \times p$ identity matrix. We denote by Det and Tr the determinant and the trace of linear transformations.

Let $X_1,\, X_2,\, \ldots,\, X_n$ be a random sample from a p-dimensional multivariate normal distribution

$$f_Z(\boldsymbol{z}) = (2\pi)^{-p/2}\, \mathrm{Det}\mathbf{\Lambda}(\sigma)^{-1/2} \exp\left\{-\frac{1}{2}\boldsymbol{z}'\mathbf{\Lambda}(\sigma)^{-1}\boldsymbol{z}\right\}, \tag{3.1}$$

with $\sigma \in \Omega$, and consider the problem of estimating σ based on $\boldsymbol{X} = (X_1,\, X_2,\, \ldots,\, X_n)$. We define a Wishart random variable $w = (\omega,\, \mathbf{w})$ in the closure of Ω as $\mathrm{Tr}(\boldsymbol{X}\boldsymbol{X}'\mathbf{\Lambda}(a)) = (a|\, w)$ for any $a = (\alpha,\, \mathbf{a})$ in $\mathbb{R} \times \mathscr{W}$. From this and the fact that $\Lambda(1,\, \mathbf{0}) = I_p$, we have

$$\alpha\mathrm{Tr}(\boldsymbol{X}\boldsymbol{X}') + \mathrm{Tr}(\boldsymbol{X}\boldsymbol{X}'\mathbf{\Lambda}(0,\, \mathbf{a})) = 2\alpha\omega + 2B(\mathbf{a}|\,\mathbf{w}) \tag{3.2}$$

for any $a = (\alpha,\, \mathbf{a})$ in $\mathbb{R} \times \mathscr{W}$. From Proposition 4 in Konno (2007), we can see that w has a density with respect to the Lebesgue measure as follows:

$$f_\Omega(w|\, n,\, v,\, p,\, \sigma) = \frac{2^{-np/2}}{\Gamma_\Omega(np/4)} \det(\sigma)^{-np/4} \det(w)^{np/4 - v/2} \exp\left\{-\frac{1}{2}(\sigma^{-1}|\, w)\right\}, \tag{3.3}$$

where

$$\Gamma_\Omega(s) = (2\pi)^{(v-2)/2} \Pi_{j=1}^2 \Gamma\bigg(s - \frac{v-2}{2}(j-1)\bigg).$$

Furthermore, it can be seen from Konno (2007) that the maximum likelihood estimator for σ is given by $\widehat{\sigma}_{mle} = (\widehat{\sigma}_0,\, \widehat{\boldsymbol{\sigma}})$, where

$$\widehat{\sigma}_0 = \frac{1}{np}\mathrm{Tr}(\boldsymbol{X}\boldsymbol{X}') \qquad \text{and} \qquad B(\mathbf{a}|\,\widehat{\boldsymbol{\sigma}}) = \frac{1}{np}\mathrm{Tr}(\boldsymbol{X}\boldsymbol{X}'\mathbf{\Lambda}(0,\, \mathbf{a}))$$

for any $\mathbf{a} \in \mathscr{W}$.

We employ a loss function

$$\mathscr{L}(\widehat{\sigma}, \sigma) = (\sigma^{-1} | \widehat{\sigma}) - \log\det(\widehat{\sigma}) + \log\det(\sigma) - 2, \tag{3.4}$$

where $\widehat{\sigma}$ is an estimator for σ. Note that $\mathscr{L}(\cdot, \cdot)$ is a strictly convex of its first argument and that it is nonnegative and minimized at $\widehat{\sigma} = \sigma$ as usual. The loss function is a counterpart of the usual Stein loss function for the problem of estimating a normal covariance matrix. The risk function is defined as

$$\mathscr{R}(\hat{\sigma}, \sigma) = \mathbb{E}[(\sigma^{-1} | \widehat{\sigma}) - \log\det(\hat{\sigma}) + \log\det(\sigma) - 2],$$

where the expectation above is taken with respect to (3.3).

4 Minimax estimation

In this section we give a brief summary of the minimax estimation theory for the mean matrix of the Wishart models on simple Jordan algebras of the degree 2. For a more detailed exposure to this theory see Konno (2007).

4.1 Minimax estimator with constant risk

To describe the minimax risk, we first consider a class of estimators having the form

$$\widehat{\sigma}(Tw) = T\widehat{\sigma}(w) \tag{4.1}$$

for any element T in the triangular transformation group $\mathscr{T}$. Using (2.6), we decompose $w = T(b)(1, \mathbf{0})$ where $T(b) \in \mathscr{T}$ and $b = \beta_1 c_1 + \beta_2 c_2 + (0, \mathbf{b})$ with

$$\beta_1 = \sqrt{\omega + B(\mathbf{w}| \mathbf{h})}, \quad \beta_2 = \sqrt{\frac{\omega^2 - \|\mathbf{w}\|_B^2}{\omega + B(\mathbf{w}| \mathbf{h})}}, \quad \mathbf{b} = \frac{1}{\sqrt{\omega + B(\mathbf{w}| \mathbf{h})}}\Big(\mathbf{w} - B(\mathbf{w}| \mathbf{h})\mathbf{h}\Big). \tag{4.2}$$

Then a standard argument such as those in Muirhead (1982) and Eaton (1989) shows that (4.1) holds if and only if, for some $\delta_1 > 0$, $\delta_2 > 0$ and $d_{12} \in \mathscr{V}_{12}$,

$$\widehat{\sigma}(w) = T(b)\left(\delta_1 c_1 + \delta_2 c_2 + d_{12}\right). \tag{4.3}$$

Furthermore, we can see that, from Proposition 11 in Konno (2007), the estimator $T(b)(\delta_1 c_1 + \delta_2 c_2)$ with

$$\delta_1^{-1} = np/2 + (v-2), \qquad \delta_2^{-1} = np/2 - (v-2), \tag{4.4}$$

is minimax. Since the maximum likelihood estimator $\widehat{\sigma}_{mle}$ belongs to the class of estimators of the form (4.1), it is improved by the estimator $T(b)(\delta_1 c_1 + \delta_2 c_2)$ with (4.4). Furthermore, its minimax risk is given by $-\sum_{j=1}^{2}\{\log\delta_j + \mathbb{E}[\log u_j^2]\}$, where u_j^2's follow chi-squared distributions with the degree of freedom $(np/2) - (v-2)(j-1)$ $(j = 1, 2)$.

To express the minimax estimator in terms of $w = (\omega, \mathbf{w})$, set $T(b)(\delta_1 c_1 + \delta_2 c_2) = (\chi, \mathbf{x})$. Using (2.5), we have

$$\begin{aligned} T(b)(\delta_1 c_1 + \delta_2 c_2) &= \frac{\delta_1\beta_1^2}{2}(1, \mathbf{h}) \oplus \delta_1\beta_1(0, \mathbf{b}) \oplus \frac{1}{2}\left(\frac{\delta_1}{2}\|(0, \mathbf{b})\|^2 + \delta_2\beta_2^2\right)(1, -\mathbf{h}) \\ &= \left(\frac{\delta_1\beta_1^2}{2} + \frac{\delta_1\|(0, \mathbf{b})\|^2}{4} + \frac{\delta_2\beta_2^2}{2}, \, (\frac{\delta_1\beta_1^2}{2} - \frac{\delta_2\beta_2^2}{2} - \frac{\delta_1\|(0, \mathbf{b})\|^2}{4})\mathbf{h} + \delta_1\beta_1\mathbf{b}\right). \end{aligned} \tag{4.5}$$

Furthermore note that, from (4.2),

$$\|(0, \mathbf{b})\|^2 = \mathrm{tr}(\,(0, \mathbf{b}) \circ (0, \mathbf{b})) = \mathrm{tr}\,(B(\mathbf{b}| \mathbf{b}), \mathbf{0}) = 2\|\mathbf{b}\|_B^2 = 2\frac{\|\mathbf{w}\|_B^2 - B^2(\mathbf{w}| \mathbf{h})}{\omega + B(\mathbf{w}| \mathbf{h})}.$$

Putting this equation and (4.2) into (4.5), we have $T(b)(\delta_1 c_1 + \delta_2 c_2) = (\chi,\, \mathbf{x})$, where

$$\begin{aligned}
\chi &= \frac{\delta_1}{2}\left(\beta_1^2 + \|\mathbf{b}\|_B^2\right) + \frac{\delta_2\beta_2^2}{2} \\
&= \frac{\delta_1}{2}\left(\omega + B(\mathbf{w}|\,\mathbf{h}) + \frac{\|\mathbf{w}\|_B^2 - B^2(\mathbf{w}|\,\mathbf{h})}{\omega + B(\mathbf{w}|\,\mathbf{h})}\right) + \frac{\delta_2}{2}\frac{\omega^2 - \|\mathbf{w}\|_B^2}{\omega + B(\mathbf{w}|\,\mathbf{h})} \\
&= \delta_1 \frac{\omega^2 + 2\omega B(\mathbf{w}|\,\mathbf{h}) + \|\mathbf{w}\|_B^2}{2(\omega + B(\mathbf{w}|\,\mathbf{h}))} + \delta_2 \frac{\omega^2 - \|\mathbf{w}\|_B^2}{2(\omega + B(\mathbf{w}|\,\mathbf{h}))}
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{x} &= \frac{\delta_1}{2}\left(\beta_1^2 - \|\mathbf{b}\|_B^2\right)\mathbf{h} + \delta_1\beta_1\mathbf{b} - \frac{\delta_2}{2}\beta_2^2\mathbf{h} \\
&= \frac{\delta_1}{2}\left((\omega + B(\mathbf{w}|\,\mathbf{h})) - \frac{\|\mathbf{w}\|_B^2 - B^2(\mathbf{w}|\,\mathbf{h})}{\omega + B(\mathbf{w}|\,\mathbf{h})}\right)\mathbf{h} + \delta_1(\mathbf{w} - B(\mathbf{w}|\,\mathbf{h})\mathbf{h}) \\
&\quad - \frac{\delta_2}{2}\frac{\omega^2 - \|\mathbf{w}\|_B^2}{\omega + B(\mathbf{w}|\,\mathbf{h})}\mathbf{h} \\
&= \delta_1\mathbf{w} + \frac{\delta_1}{2}\frac{\omega - \|\mathbf{w}\|_B^2}{\omega + B(\mathbf{w}|\,\mathbf{h})}\mathbf{h} - \frac{\delta_2}{2}\frac{\omega^2 - \|\mathbf{w}\|_B^2}{\omega + B(\mathbf{w}|\,\mathbf{h})}\mathbf{h}.
\end{aligned}$$

4.2 Orthogonally invariant estimators and their unbiased risk estimates

To describe the orthogonally invariant estimators we need the following lemma which states the singular value decomposition of an element in Ω.

Lemma 4.1 *For any $w = (\omega,\, \mathbf{w}) \in \Omega$, set*

$$z = \frac{1}{\sqrt{4\omega^2 + 2\|\mathbf{w}\|_B B(\mathbf{w}|\,\mathbf{h}) - 2\|\mathbf{w}\|_B^2}}(2\omega,\, \mathbf{w} - \|\mathbf{w}\|_B\mathbf{h}). \tag{4.6}$$

Then we have $w = P(z)(\lambda_1 c_1 + \lambda_2 c_2)$, where

$$\lambda_1 = \omega + \|\mathbf{w}\|_B \qquad \textit{and} \qquad \lambda_2 = \omega - \|\mathbf{w}\|_B. \tag{4.7}$$

Furthermore, we have $\mathrm{Det}(P(z)) = 1$.

Proof. The first assertion can be obtained from a straightforward application of Corollary 1 in Faybusovich and Tsuchiya (2003). Then, using Lemma 1(ii) in Konno (2007) and noting that $\det z = 1$, we can complete the proof. □

We use Lemma 4.1 to decompose the element w as

$$w = P(z)\left\{\frac{\omega + \|\mathbf{w}\|_B}{2}(1,\, \mathbf{h}) + \frac{\omega - \|\mathbf{w}\|_B}{2}(1,\, -\mathbf{h})\right\} = P(z)(\omega,\, \|\mathbf{w}\|_B\mathbf{h}),$$

where z is given by (4.6). For a decomposition of w stated in Lemma 4.1, we consider orthogonally invariant estimators of the form

$$\widehat{\sigma}_\varphi = P(z)\left(\frac{\varphi_1(\lambda_1,\, \lambda_2)}{2}(1,\, \mathbf{h}) + \frac{\varphi_2(\lambda_1,\, \lambda_2)}{2}(1,\, -\mathbf{h})\right), \tag{4.8}$$

where φ_1 and φ_2 are differentiable functions from $\mathbb{R}^2$ to $\mathbb{R}$. Let

$$\mathscr{R}^{\#}(\widehat{\sigma}_\varphi,\, \sigma) = \mathbb{E}[(\sigma^{-1}|\,\widehat{\sigma}_\varphi) - \log\det(\widehat{\sigma}_\varphi)].$$

It is easily seen that comparison between two estimators of the form (4.8) in terms of the risk $\mathscr{R}$ is equivalent to that in terms of $\mathscr{R}^{\#}$. The next theorem is a generalization of Lemma 2.1 in Dey and Srinivasan (1985) to the setting of the Lorentz Wishart distributions, which is derived from a general results in Konno (2007).

Lemma 4.2 *Consider the estimators given by* (4.8). *Then an unbiased risk estimate for* $\mathscr{R}^{\#}(\widehat{\sigma}_{\varphi},\, \sigma)$ *is given as*

$$\widehat{\mathscr{R}}^{\#}(\widehat{\sigma}_{\varphi}) = \sum_{j=1}^{2}\left\{2\frac{\partial\varphi_j}{\partial\lambda_j} + \left(\frac{np}{2} - v\right)\frac{\varphi_j}{\lambda_j} + 2(v-2)\frac{\varphi_1 - \varphi_2}{\lambda_1 - \lambda_2} - \log\varphi_j\right\}, \tag{4.9}$$

i.e., we have $\mathscr{R}^{\#}(\widehat{\sigma}_{\varphi},\, \sigma) = \mathbb{E}[\widehat{\mathscr{R}}^{\#}(\widehat{\sigma}_{\varphi})]$.

5 A new class of estimators

Using Lemma 4.2, Konno (2007) showed that the estimator

$$\widehat{\sigma}_m = P(z)(\delta_1\lambda_1 c_1 + \delta_2\lambda_2 c_2) \tag{5.1}$$

is minimax, where δ_j's and λ_j's are given by (4.4) and (4.7). However, this estimator does not satisfy a natural restrictions on the estimated eigenvalues, i.e., $\varphi_1 \geq \varphi_2$ in (4.8).

We can construct explicit form of orthogonally invariant estimators which include estimators corresponding to those in Dey and Srinivasan (1985), Perron (1992), and Takemura (1984). We again decompose w into $P(z)(\lambda_1 c_1 + \lambda_2 c_2)$ where $P(z)$, λ_1, and λ_2 are defined as in Lemma 4.1, and we consider a class of estimators for σ, being of the form

$$\widehat{\sigma}_{\gamma} = P(z)\{\phi_1^{(\gamma)}\lambda_1 c_1 + \phi_2^{(\gamma)}\lambda_2 c_2\}, \tag{5.2}$$

where γ is a positive constant, and

$$\phi_1^{(\gamma)} = \left(\frac{\lambda_1^{\gamma}}{\lambda_1^{\gamma} + \lambda_2^{\gamma}}\delta_1 + \frac{\lambda_2^{\gamma}}{\lambda_1^{\gamma} + \lambda_2^{\gamma}}\delta_2\right) \quad \text{and} \quad \phi_2^{(\gamma)} = \left(\frac{\lambda_2^{\gamma}}{\lambda_1^{\gamma} + \lambda_2^{\gamma}}\delta_1 + \frac{\lambda_1^{\gamma}}{\lambda_1^{\gamma} + \lambda_2^{\gamma}}\delta_2\right).$$

Note that, if $\gamma \leq 1$, the eigenvalues of the estimator (5.2) are order-preserving. If $\gamma = 1$ then the estimator (5.2) corresponds to an analogue of the estimator for the normal covariance matrix given by Perron (1992) while it corresponds to an analogue of the estimator for the normal covariance matrix given by Takemura (1984) if $\gamma = 1/2$.

Theorem 5.1 *Let* $\boldsymbol{X} = (X_1,\, X_2,\, \ldots,\, X_n)$ *where* $X_1,\, X_2,\, \ldots,\, X_n$ *are independently and identically distributed as* (3.1) *for some* σ *in* Ω *and assume that* w *is an element in the closure of* Ω *such that* $\mathrm{Tr}(\boldsymbol{X}\boldsymbol{X}'\boldsymbol{\Lambda}(a)) = (a|\, w)$ *for any element* a *in* $\mathbb{R}\times\mathcal{W}$. *Then the estimators given by* (5.2) *are minimax if* $\gamma \geq 1$. *Furthermore, the eigenvalues of the estimators satisfy the natural order* $\phi_1^{(1)}\lambda_1 \geq \phi_2^{(1)}\lambda_2$ *when* $\gamma = 1$.

Proof. We apply Lemma 4.2 with $\varphi_j = \phi_j^{(\gamma)}\lambda_j$ $(j = 1,\, 2)$. From straightforward calculation we have

$$\sum_{j=1}^{2} \frac{\partial \varphi_j}{\partial \lambda_j} = \delta_1 + \delta_2 + \frac{2\gamma(\lambda_1\lambda_2)^\gamma}{(\lambda_1^\gamma + \lambda_2^\gamma)^2}(\delta_1 - \delta_2) \leq \delta_1 + \delta_2, \tag{5.3}$$

$$\begin{aligned}\frac{\varphi_1 - \varphi_2}{\lambda_1 - \lambda_2} &= \left\{\left(\frac{\lambda_1^\gamma}{\lambda_1^\gamma + \lambda_2^\gamma}\delta_1 + \frac{\lambda_2^\gamma}{\lambda_1^\gamma + \lambda_2^\gamma}\delta_2\right)\lambda_1 - \left(\frac{\lambda_2^\gamma}{\lambda_1^\gamma + \lambda_2^\gamma}\delta_1 + \frac{\lambda_1^\gamma}{\lambda_1^\gamma + \lambda_2^\gamma}\delta_1\right)\lambda_2\right\}\Big/(\lambda_1 - \lambda_2) \\ &= \frac{(\lambda_1^\gamma + \lambda_2^\gamma)(\lambda_1 - \lambda_2) + \lambda_1\lambda_2(\lambda_1^{\gamma-1} - \lambda_2^{\gamma-1})}{(\lambda_1^\gamma + \lambda_2^\gamma)(\lambda_1 - \lambda_2)}\delta_1 + \frac{\lambda_1\lambda_2(\lambda_2^{\gamma-1} - \lambda_1^{\gamma-1})}{(\lambda_1^\gamma + \lambda_2^\gamma)(\lambda_1 - \lambda_2)}\delta_2 \\ &= \delta_1 + \frac{\lambda_1\lambda_2(\lambda_1^{\gamma-1} - \lambda_2^{\gamma-1})}{(\lambda_1 - \lambda_2)(\lambda_1^\gamma + \lambda_2^\gamma)}(\delta_1 - \delta_2) \leq \delta_1, \end{aligned} \tag{5.4}$$

provided $\gamma \geq 1$. Furthermore, from the convexity of the logarithmic function and Jensen's inequality, we have

$$\log\det(\widehat{\sigma}_\gamma) = \log(\phi_1^{(\gamma)}\phi_2^{(\gamma)}) + \log(\lambda_1\lambda_2) \geq \sum_{j=1}^{2}\{\log\lambda_j + \log\delta_j\}. \tag{5.5}$$

Putting (5.3)-(5.5) into (4.9) and using Corollary 9 in Konno (2007), we have

$$\begin{aligned}\mathscr{R}(\widehat{\sigma}_\gamma,\, \sigma) &\leq 2(\delta_1 + \delta_2) + \left(\frac{np}{2} - v\right)(\delta_1 + \delta_2) + 2(v-2)\delta_1 \\ &\quad - \log(\delta_1\delta_2) - \mathbb{E}[\log(\lambda_1\lambda_2)] + \log\det\sigma - 2 \\ &= -\sum_{j=1}^{2}\{\mathbb{E}[\log u_j^2] + \log\delta_j\},\end{aligned}$$

where $\mathcal{L}(u_j^2) = \chi^2_{(np/2)-(v-2)(j-1)}$ $(j = 1,\, 2)$. This completes the proof. □

Remark 5.1 Letting γ in (5.2) go to the infinity, we can see that the estimator $\widehat{\sigma}_\gamma$ tends to (5.1), an analogue of Dey and Srinivasan's estimator for the real normal covariance matrix, i.e., $\widehat{\sigma}_\infty = \widehat{\sigma}_m$. On the other hand, if $\gamma = 1/2$, then the estimator $\widehat{\sigma}_{1/2}$ becomes an analogue of an estimator of Takemura (1984) for the real normal covariance matrix. We conject that $\widehat{\sigma}_{1/2}$ is minimax. However, we can not show that it is minimax because of the complex nature of the Jordan algebras of the degree 2.

Acknowledgements

The author would like to thank a referee for her/his careful reading of the manuscript and helpful comments. This work was in part supported by the Japan Society for the Promotion of Science through Grants-in-Aid for Scientific Research (C) (No.17500185).

References:

Dey, D,K. and Srinivasan, C. (1985): Estimation of a covariance matrix under Stein's loss, *Ann. Statist.* 13, 1581-1591.

Eaton, M.L. (1989): *Group Invariance Application in Statistics*, Regional conference series in Probability and Statistics Vol. 1, Institute of Mathematical Statistics.

Faraut, J. and Korányi, K. (1994): *Analysis on Symmetric Cones*, Oxford Science Publications.

Faybusovich, L. and Tsuchiya, T. (2003): Primal-dual algorithms and infinite-dimensional Jordan algebras of finite rank, *Math. Program* Ser. B 97, 471-493.

Haff, L.R. (1992): The variational form of certain Bayes estimators, *Ann. Statist.* 19, 1163-1190.

James, W. and Stein, C. (1961): Estimation with quadratic loss, in *Proc. Fourth Berkeley Symp. Math. Statist. Prob.* 1 361-380, Univ. California Press.

Kiefer, J (1957): Invariance, minimax and sequential estimation and continuous time processes, *Ann. Math. Statist.* 28, 573-601.

Konno, Y. (2007): Estimation of normal covariance matrices parametrized by irreducible symmetric cones under Stein's loss, *J. Multivariate Anal.* 98, 295-316.

Muirhead, R.J. (1982): *Aspects of Multivariate Statistical Analysis*, John Wiley & Sons, Inc.

Perron, F. (1992): Minimax estimators of a covariance matrix, *J. Multivariate Anal.* 43, 6-28.

Sheena, Y. and Takemura, A. (1992): Inadmissibility of non-order-preserving orthogonally invariant estimators of the covariance matrix in the case of Stein's loss, *J. Multivariate Anal.* 41, 117-131.

Stein, C. (1977): 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.).

Takemura, A. (1984): An orthogonally invariant minimax estimators of the covariance matrix of a multivariate normal populations, *Tsukuba J. Math.* 8, 365-376.

Tolver Jensen, S. (1988): Covariance hypotheses which are linear in both the covariance and the inverse covariance, *Ann. Statist.* 16, 302-322.

6 Appendix

6.1 Proof of (2.4)

Since $a\circ b=(\alpha\beta+B(\mathbf{a}|\,\mathbf{b}),\,\alpha\mathbf{b}+\beta\mathbf{a})$ and $a^2=(\alpha^2+\|\mathbf{a}|\|_B^2,\,2\alpha\mathbf{a})$, we have

$$\begin{aligned}
a\circ(a\circ b) &= (\alpha,\,\mathbf{a})\circ(\alpha\beta+B(\mathbf{a}|\,\mathbf{b}),\,\alpha\mathbf{b}+\beta\mathbf{a})\\
&= (\alpha^2\beta+2\alpha B(\mathbf{a}|\,\mathbf{b})+\beta\|\mathbf{a}\|_B^2,\,(2\alpha\beta+B(\mathbf{a}|\,\mathbf{b}))\mathbf{a}+\alpha^2\mathbf{b})),\\
a^2\circ b &= (\alpha^2+\|\mathbf{a}\|_B^2,\,2\alpha\mathbf{a})\circ(\beta,\,\mathbf{b})\\
&= (\alpha^2\beta+\beta\|\mathbf{a}\|_B^2+2\alpha B(\mathbf{a}|\,\mathbf{b}),\,(\alpha^2+\|\mathbf{a}\|_B^2)\mathbf{b}+2\alpha\beta\mathbf{a}).
\end{aligned}$$

From these equations, we have

$$\begin{aligned}
P(a)b &= 2(\alpha^2\beta+2\alpha B(\mathbf{a}|\,\mathbf{b})+\beta\|\mathbf{a}\|_B^2,\,(2\alpha\beta+B(\mathbf{a}|\,\mathbf{b}))\mathbf{a}+\alpha^2\mathbf{b})\\
&\quad -(\alpha^2\beta+\beta\|\mathbf{a}\|_B^2+2\alpha B(\mathbf{a}|\,\mathbf{b}),\,(\alpha^2+\|\mathbf{a}\|_B^2)\mathbf{b}+2\alpha\beta\mathbf{a})\\
&= (\alpha^2\beta+2\alpha B(\mathbf{a}|\,\mathbf{b})+\beta\|\mathbf{a}\|_B^2,\,(2\alpha\beta+B(\mathbf{a}|\,\mathbf{b}))\mathbf{a}+(\alpha^2-\|\mathbf{a}\|_B^2)\mathbf{b})\\
&= (2\alpha\beta+B(\mathbf{a}|\,\mathbf{b}))(\alpha,\,\mathbf{a})+(\alpha^2-\|\mathbf{a}\|_B^2)(-\beta,\,\mathbf{b}),
\end{aligned}$$

which completes the proof of (2.4). $\square$

6.2 Proof of (2.5)

For elements $x,\,y$ in $\mathscr{V}$, L is a map from $\mathscr{V}$ to $\mathscr{V}$ defined by $L(x)y=x\circ y$. For $z\in\mathscr{V}_{12}$, let τ_{c_1} is a map from $\mathscr{V}$ to $\mathscr{V}$ defined by

$$\tau_{c_1}(z)(x)=x_1\oplus(2L(z)x_1+x_{12})\oplus(2L(c_2)L(z)^2x_1+2L(c_2)L(z)x_{12}+x_2), \tag{6.1}$$

where $x=x_1\oplus x_{12}\oplus x_2$ is the Peirce decomposition with respect to the idempotent c_1 such that $x_1\in\{y\in\mathscr{V}|\,c_1\circ y=y\}$, $x_2\in\{y\in\mathcal{V}|\,c_2\circ y=y\}$, and $x_{12}\in\mathscr{V}_{12}=\{y\in\mathscr{V}|\,c_1\circ y=c_2\circ y=(1/2)y\}$. From Faraut and Korányi (1994), it is seen that the map from Ω to the triangular subgroup of the general linear group over $\mathscr{V}$ is given by, for $a=\alpha_1c_1+\alpha_2c_2+(0,\,\mathbf{a})\in\Omega$,

$$T(a)=P(a_1)\tau_{c_1}(a_{12})P(a_2),$$

where $a_1 = \alpha_1 c_1 + c_2$, $a_2 = c_1 + \alpha_2 c_2$, and $a_{12} = (0,\, \mathbf{a})$.

For $b = \beta_1 c_1 + \beta_2 c_2 + b_{12}$ with $b_{12} = (0,\, \mathbf{b})$, we first compute $P(a_2)b$ as

$$\begin{aligned}
P(a_2)b &= \{2L(a_2)^2 - L(a_2^2)\}(\beta_1 c_1 + \beta_2 c_2 + b_{12}) \\
&= 2L(a_2)(c_1 + \alpha_2 c_2) \circ (\beta_1 c_1 + \beta_2 c_2 + b_{12}) - (c_1 + \alpha_2^2 c_2) \circ (\beta_1 c_1 + \beta_2 c_2 + b_{12}) \\
&= 2L(a_2)\left(\beta_1 c_1 + \alpha_2 \beta_2 c_2 + \frac{1}{2}(1+\alpha_2) b_{12}\right) - \left(\beta_1 c_1 + \alpha_2^2 \beta_2 c_2 + \frac{1}{2}(1+\alpha_2^2) b_{12}\right) \\
&= \left(2\beta_1 c_1 + 2\alpha_2^2 \beta_2 c_2 + \frac{1}{2}(1+\alpha_2)^2 b_{12}\right) - \left(\beta_1 c_1 + \alpha_2^2 \beta_2 c_2 + \frac{1}{2}(1+\alpha_2^2) b_{12}\right) \\
&= \beta_1 c_1 + \alpha_2^2 \beta_2 c_2 + \alpha_2 b_{12}.
\end{aligned}$$

Next we use (6.1) and Lemma 1(vii) in Konno (2007) in order to compute

$$\tau_{c_1}(a_{12})(\beta_1 c_1 + \alpha_2^2 \beta_2 c_2 + \alpha_2 b_{12})$$

as

$$\begin{aligned}
&\tau_{c_1}(a_{12})(\beta_1 c_1 + \alpha_2 b_{12} + \alpha_2^2 \beta_2 c_2) \\
&\quad = \beta_1 c_1 \oplus \{2L(a_{12})\beta_1 c_1 + \alpha_2 b_{12}\} \oplus \{2L(c_2)L(a_{12})^2 \beta_1 c_1 + 2L(c_2)L(a_{12})\alpha_2 b_{12} + \alpha_2^2 \beta_2 c_2\}.
\end{aligned}$$

Furthermore we have

$$2L(c_2)L(a_{12})^2 \beta_1 c_1 = \frac{\beta_1}{2} L(c_2)\|a_{12}\|^2 (c_1 + c_2) = \frac{\beta_1}{2}\|a_{12}\|^2 c_2,$$

and

$$\begin{aligned}
2L(c_2)L(a_{12})\alpha_2 b_{12} &= 2\alpha_2 L(c_2)(a_{12} \circ b_{12}) \\
&= 2\alpha_2 L(c_2)(B(\mathbf{a}|\,\mathbf{b}),\, \mathbf{0}) \\
&= 2\alpha_2^2 B(\mathbf{a}|\,\mathbf{b}) c_2,
\end{aligned}$$

from which it follows that

$$\begin{aligned}
&\tau_{c_1}(a_{12})(\beta_1 c_1 + \alpha_2^2 \beta_2 c_2 + \alpha_2 b_{12}) \\
&\quad = \beta_1 c_1 \oplus \{\beta_1 a_{12} + \alpha_2 b_{12}\} \oplus \left(\frac{\beta_1}{2}\|a_{12}\|^2 + \alpha_2^2 \beta_2 + 2\alpha_2^2 B(\mathbf{a}|\,\mathbf{b})\right) c_2 =: \tau.
\end{aligned}$$

Finally,

$$\begin{aligned}
T(a)b &= 2L(a_1)\{(\alpha_1 c_1 + c_2) \circ \tau\} - (\alpha_1^2 c_1 + c_2) \circ \tau \\
&= 2L(a_1)\Bigg\{\alpha_1 \beta_1 c_1 \oplus \left(\frac{\alpha_1 \beta_1 + \beta_1}{2} a_{12} + \frac{\alpha_1 \alpha_2 + \alpha_2}{2} b_{12}\right) \\
&\qquad \oplus \left(\frac{\beta_1}{2}\|a_{12}\|^2 + 2\alpha_2^2 B(\mathbf{a}|\,\mathbf{b}) + \alpha_2^2 \beta_2\right) c_2\Bigg\} \\
&\qquad -\alpha_1^2 \beta_1 c_1 \oplus \left(\frac{\alpha_1^2 \beta_1 + \beta_1}{2} a_{12} + \frac{\alpha_1^2 \alpha_2 + \alpha_2}{2} b_{12}\right) \oplus \left(\frac{\beta_1}{2}\|a_{12}\|^2 + 2\alpha_2^2 B(\mathbf{a}|\,\mathbf{b}) + \alpha_2^2 \beta_2\right) c_2 \\
&= \alpha_1^2 \beta_1 c_1 \oplus \{\alpha_1 \beta_1 a_{12} + \alpha_1 \alpha_2 b_{12}\} \oplus \left(\frac{\beta_1}{2}\|a_{12}\|^2 + 2\alpha_2^2 B(\mathbf{a}|\,\mathbf{b}) + \alpha_2^2 \beta_2\right),
\end{aligned}$$

which complete the verification of (2.5). □