Improving the sample covariance matrix for a complex elliptically contoured distribution

Yoshihiko Konno *

Faculty of Science, Japan Women’s University
2-8-1 Mejirodai, Bunkyo-ku, Tokyo 112-8691, Japan

Abstract. In this paper the problem of estimating the scale matrix in a complex elliptically contoured distribution (complex ECD) is addressed. An extended Haff-Stein identity for this model is derived. It is shown that the minimax estimators of the covariance matrix obtained under the complex normal model remain robust under the complex ECD model when the Stein loss function is employed.

Key words: covariance matrix, shrinkage estimator, complex Wishart distribution, minimax, unbiased risk of estimate, complex normal distribution, eigenvalue distribution in sample covariance matrix.

MSC: primary; 62H12; secondary: 62F10

1 Introduction

The focus of the present paper is on estimating the scale matrix $\Sigma$ that appear in a standard complex MANOVA model

$$Y = C\beta + E,$$  \hspace{1cm} (1)

in which $Y$ is an $N \times p$ matrix of response complex variables, $C$ is an $N \times m$ known matrix of complex entries, $\beta$ is an $m \times p$ unknown parameter matrix of complex entries, and $E$ is an

* Email Address: konno@fc.jwu.ac.jp

Supported in part by the Japan Society for the Promotion of Science through Grants-in-Aid for Scientific Research(C)(No.17500185).
$N \times p$ error matrix of complex entries. The following assumptions on (1) are made to hold; (i) the error matrix $E$ follows a complex elliptically contoured distribution with the density function

$$\text{Det}(\Sigma)^{-N} f(\text{Tr}(\Sigma^{-1}e^*e)) \quad e \in \mathbb{C}^{N \times p},$$

in which $\Sigma$ is a $p \times p$ unknown positive definite Hermitian matrix, and $f(\cdot)$ is a nonnegative function on the nonnegative real line such that $\int_0^\infty f(t) \, dt = 1$; (ii) $N - m - p + 1 > 0$. Here $\mathbb{C}^{N \times p}$ is the space of $N \times p$ matrices of complex entries, “$*$” stands for the conjugate transpose of a matrix of complex entries, “$\text{Tr}$” and “$\text{Det}$” mean the usual trace and the determinant of a squared matrix, respectively.

The model (1), along the previous assumptions, includes the complex MANOVA model with the complex normal error, which is extensively explored by Andersen et al. (1995). The practical significance of this model in signal processing and communications is discussed, for example, in DoGondžić and Neborai (2003). See Krishnaiah and Lin (1986) and Micheas et al. (2005) for construction of an elliptical family of complex distributions via an extension of elliptical family of real distributions.

Stein (1977) pointed out that the eigenvalues of the sample covariance matrix obtained from the multivariate real normal sample spread out more than the corresponding eigenvalues of the parameter covariance matrix, and that the problem gets more serious when the parameter covariance matrix is near spherical. Since this fact was pointed out, there has been growing literature on improved estimators of the covariance matrix of the multivariate real normal distributions. We refer to Lin and Perlman (1985), Young and Berger (1994), Kubokawa and Srivastava (1999), Daniels and Kass (1999, 2001), and the cited references therein for further information on the improved estimation of the real normal covariance matrix.

The sample covariance matrix obtained from independently and identically distributed complex random vectors may be improved by applying shrinkage-expansion techniques to its eigenvalues since the eigenvalues of the sample covariance matrix are biased away from those of the true parameter covariance matrix. Abraham and Dey (1994) demonstrated that the adaptive beamforming and detection performance loss due to the estimation of an unstructured covariance matrix is reduced by modifying the eigenvalues of the sample covariance matrix. The methods of Abraham and Dey (1994) heavily depend on the results of Dey and Srinivasan (1985) for estimating the covariance matrix of a real multivariate normal random vector. It is conjectured in the paper of Abraham and Dey (1994) that the results of Dey and Srinivasan (1985) may be extended to the complex multivariate normal
The underlying techniques in the paper of Dey and Srinivasan (1985) are twofold; one is integration-by-parts for the real Wishart distribution, which is derived by Stein (1977) and Haff (1979) independently; the other is an algebraic derivation of the eigenstructure of a positive definite matrix of real entries, which is developed by Stein (1977) and Haff (1991). Later, Maiwald and Kraus (1994) used well-known isomorphism between real and complex random matrices in order to give the real representation of integration-by-parts for the complex Wishart distribution. More recently, Svensson and Lundberg (2004) obtained an unbiased estimate of the risks for unitary invariant estimators of the scale matrix of the complex Wishart distribution and used the techniques of Haff (1991) in order to derive the variational form of the Bayes estimator.

In this paper we consider the problem of estimating the Hermitian scale matrix $\Sigma$ for the complex ECD model (1) in a decision theoretic manner. We employ the Stein loss function in order to evaluate the performance of every estimator of $\Sigma$. We first generalize the techniques of Dey and Srinivasan (1985) to the complex elliptically contoured distributions. Namely, we develop integration-by-parts for the complex ECD model (1) and some algebraic derivations which are applied to the eigenstructure of a positive definite Hermitian matrix. Using these results we show that the dominance results for the real ECD model given by Kubokawa and Srivastava (1999) are generalized to the complex setting. In other words we obtain dominance results in the complex case similar to those in the real case given by James and Stein (1961) and Dey and Srinivasan (1985).

The present paper is organized in the following way. In Section 2, we give integration-by-formula for the complex elliptically contoured distribution via an extension of a formula on the real ECD models obtained by Kubokawa and Srivastava (1999). We also give calculus lemmas on the eigenstructure of Hermitian matrices. These lemmas are similar to algebraic lemmas on the eigenstructure of the symmetric matrices given by Haff (1991). These lemmas appear in a unpublished manuscript of Svensson (2004). Utilizing the Haff-Stein identify for the complex ECD models we show that minimax estimators which are obtained in a similar way to those under the real normal distributions remain robust under the complex ECD models (1). In Section 3 the proofs of the calculus lemmas are presented since our proofs of the lemmas are more elementary.
2 Main Result

To describe a canonical form of our estimation problem, let $P$ be an $N \times N$ unitary matrix such that $PC = ((C^*C)^{1/2}, 0_{m \times (N-m)})^*$ and $PP^* = P^*P = I_N$. Also set $\mu = (C^*C)^{1/2}\beta$. Let $n = N - m$ and let $X$ and $Z$ be $m \times p$ and $n \times p$ matrices of complex entries such that $(X^*, Z^*)^* = PY$. Then the joint density function of $X$ and $Z$ has the form

$$\text{Det}(\Sigma)^{-N} f(\text{Tr}(\Sigma^{-1}(x - \mu)^*(x - \mu) + \Sigma^{-1}z^*z)), \quad x \in \mathbb{C}^{m \times p}, z \in \mathbb{C}^{n \times p}. \quad (2)$$

Set

$$S = Z^*Z.$$

Based on $S$, we consider the problem of estimating $\Sigma$ under the Stein loss function

$$\mathcal{L}(\hat{\Sigma}, \Sigma) = \text{Tr}(\hat{\Sigma}\Sigma^{-1}) - \text{Det}(\hat{\Sigma}\Sigma^{-1}) - p, \quad (3)$$

in which $\hat{\Sigma}$ is an estimator of a $p \times p$ Hermitian matrix $\Sigma$.

The loss function given by (3) has been suggested by James and Stein (1961) for the problem of estimating the real normal covariance matrix. Certainly there are many other possible loss functions which we may employ. However the above loss function has the attractive feature that it is relatively easy to work with [see Muirhead (1982)].

2.1 Integration by parts formula

Let $\mathbb{C}$ denote the field of complex numbers. The imaginary unit is denoted by $\sqrt{-1}$. For every $c$ in $\mathbb{C}$ we denote by $\text{Re}c$ and $\text{Im}c$ the real and the imaginary parts of $c$, respectively. The conjugation of a complex number $c$ is given by $\bar{c} = \text{Re}c - \sqrt{-1}\text{Im}c$. The absolute value of $c$ is given by $|c| = (c\bar{c})^{1/2}$. We denote by $\mathbb{C}^{n \times p}$ the set of all $n \times p$ matrices of complex entries.

To describe integration-by-parts for $S = Z^*Z$, we need the following notation: For a nonnegative $x$ and $f(\cdot)$ in (2), set $F(x) = \int_x^\infty f(t) \, dt$ and define

$$\mathbb{E}^F[h(X, Z)] = \mathbb{E} \left[h(X, Z) \frac{F(R)}{f(R)} \right]. \quad (4)$$

Here, $h(\cdot, \cdot)$ is a real-valued integrable function and

$$R = \text{Tr}(\Sigma^{-1}(X - \mu)^*(X - \mu) + \Sigma^{-1}Z^*Z).$$

The notation in (4) was introduced by Kubokawa and Srivastava (1999) in order to describe integration-by-parts for the real elliptically contoured distributions. Let $G(S)$ be a $p \times p$
matrix of complex entries, the \((i, j)\)-element \(g_{ij}(S)\) of which is a function of \(S = (s_{ij})\). For a \(p \times p\) Hermitian matrix \(S = (s_{jk})\), let \(D_S = (d_{jk}^S)\) be a \(p \times p\) operator matrix, the \((j, k)\)-element of which is given by

\[
d_{jk}^S = \frac{1}{2}(1 + \delta_{jk}) \left\{ \frac{\partial}{\partial (\text{Re} \ s_{jk})} + \sqrt{-1} \frac{\partial}{\partial (\text{Im} \ s_{jk})} \right\}, \quad j, k = 1, 2, \ldots, p. \quad (5)
\]

Here, \(\delta_{jk}\) is the Kronecker delta \((= 1\ if\ j = k\ and\ = 0\ if\ j \neq k)\). Thus the \((j, k)\)-element of \(D_sG(S)\) is

\[
\{D_sG(S)\}_{jk} = \sum_{l=1}^{p} d_{jl}^S g_{lk}(S) = \frac{1}{2}(1 + \delta_{jl}) \sum_{l=1}^{p} \left\{ \frac{\partial g_{lk}}{\partial (\text{Re} \ s_{jl})}(S) + \sqrt{-1} \frac{\partial g_{lk}}{\partial (\text{Im} \ s_{jl})}(S) \right\}.
\]

**Lemma 2.1** Assume that each entry of \(G(S)\) is a partially differentiable function with respect to \(\text{Re} \ s_{jk}\) and \(\text{Im} \ s_{jk}\), \(j, k = 1, 2, \ldots, p\). Also assume that, for any real \(a\),

\[
\begin{align*}
\text{(i)} & \quad \mathbb{E}[|\text{Tr}(G(S)\Sigma^{-1})|] \text{ is finite;} \\
\text{(ii)} & \quad \lim_{|z_{jk}| \to \infty} |z_{jk}|G(S)S^{-1}F(|z_{jk}|^2 + a^2) = 0,
\end{align*}
\]

for \(j = 1, 2, \ldots, n\) and \(k = 1, 2, \ldots, p\). Then we have

\[
\mathbb{E}[\text{Tr}(G(S)\Sigma^{-1})] = \mathbb{E}^F[(n - p)\text{Tr}(G(S)S^{-1}) + \text{Tr}(D_sG(S))]. \quad (6)
\]

**Proof.** The proof is given in Section 3.

**Remark 2.1** When \(f(x) = \pi^{-np} \exp(-x)\), then the model (1) becomes the multivariate linear complex normal model and \(S\) follows the complex Wishart distribution \(CW_p(n, \Sigma)\), the density function of which is

\[
\frac{\text{Det} \ (s)^{n-p} \exp(-\text{Tr} \ (s\Sigma^{-1}))}{\text{Det} \ (\Sigma)^n \pi^{p(p-1)/2} \prod_{j=1}^{p} \Gamma(n + 1 - j)}, \quad s \in \mathbb{C}^{p \times p}_S.
\]

Here \(\Gamma(\cdot)\) indicates the usual gamma function and \(\mathbb{C}^{p \times p}_S\) stands for the space of \(p \times p\) semi-positive definite Hermitian matrices. See Goodman (1963), Khatri (1965), and Andersen et al. (1995) for further information of the complex Wishart distribution. Under the assumption of the complex normal error, \(\mathbb{E}^F\) in the right hand side of (6) becomes the usual expectation taken with respect to the above density function. Therefore the identity (6) becomes a complex extension of the Haff-Stein identity for the real Wishart distribution given by Haff (1979). This complex analogue of the Haff-Stein identity for the complex Wishart distribution is given by Svensson and Lundberg (2004).
The next two lemmas were given by Svensson (2004). Since our proofs are more elementary and self-contained, we record the proofs of the lemmas in Section 3.

**Lemma 2.2** Decompose $S = ULU^*$ in which $U = (u_{jk})$ is a $p \times p$ unitary matrix and $L = \text{Diag}(\ell_1, \ell_2, \ldots, \ell_p)$ with $\ell_1 \geq \ell_2 \geq \cdots \geq \ell_p > 0$. Then we have, for $j, k, s, t = 1, 2, \ldots, p$,

\[
\begin{align*}
    d_{st}^{jk} \ell_k &= u_{sk}u_{tk}, \\
    d_{st}^{jk} u_{jk} &= \sum_{a \neq k} \frac{u_{ja}u_{ta}u_{sk}}{\ell_k - \ell_a}, \\
    d_{st}^{jk} \overline{u_{jk}} &= \sum_{a \neq k} \frac{u_{ja}u_{sa}\overline{u_{tk}}}{\ell_k - \ell_a},
\end{align*}
\]

in which $d_{st}^{jk}$ is given by (5).

**Proof.** The proof is given in Section 3.

Let $\mathbb{R}_p^\geq = \{ (x_1, x_2, \ldots, x_p) \in \mathbb{R}^p : x_1 \geq x_2 \geq \cdots \geq x_p > 0 \}$. The next lemma is a complex analogue of Theorem 5.1 given by Haff (1991).

**Lemma 2.3** Let $\varphi_1(L), \varphi_2(L), \ldots, \varphi_p(L)$ be differentiable real-valued functions on $\mathbb{R}_p^\geq$ with $\varphi_k(L) \geq 0$, $k = 1, 2, \ldots, p$, and let $\Phi(L) = \text{Diag}(\varphi_1(L), \varphi_2(L), \ldots, \varphi_p(L))$. Then,

\[
D_S[U\Phi(L)U^*] = U\Phi^{(1)}(L)U^*,
\]

in which $\Phi^{(1)}(L) = \text{Diag}(\varphi_1^{(1)}(L), \varphi_2^{(1)}(L), \ldots, \varphi_p^{(1)}(L))$ and

\[
\varphi_k^{(1)}(L) = \sum_{a \neq k} \frac{\varphi_k(L) - \varphi_a(L)}{\ell_k - \ell_a} + \frac{\partial \varphi_k(L)}{\partial \ell_k}, \quad k = 1, 2, \ldots, p.
\]

**Proof.** The proof is given in Section 3.

**Lemma 2.4** Let $\varphi_1(L), \varphi_2(L), \ldots, \varphi_p(L)$ be differentiable real-valued functions on $\mathbb{R}_p^\geq$ with $\varphi_k(L) \geq 0$, $k = 1, 2, \ldots, p$, and let $\Phi(L) = \text{Diag}(\varphi_1(L), \varphi_2(L), \ldots, \varphi_p(L))$. Then, under suitable conditions corresponding to those of Lemma 2.1, we have

\[
E[\text{Tr}(U\Phi(L)U^*\Sigma^{-1})] = E^F \left[ \sum_{k=1}^{p} \left\{ \frac{\partial \varphi_k(L)}{\partial \ell_k} + \sum_{a \neq k} \frac{\varphi_k(L) - \varphi_a(L)}{\ell_k - \ell_a} + (n - p)\frac{\varphi_k(L)}{\ell_k} \right\} \right].
\]

**Proof.** The proof is obtained immediately from Lemmas 2.1 and 2.3. \qed

**Remark 2.2** As in Remark 2.1, set $f(x) = \pi^{-np}\exp(-x)$. Then the identity in Lemma 2.4 reduces to that appearing in Theorem 1 of Svensson and Lundberg (2004).
2.2 Alternative estimators

The next proposition is a complex analogue of Proposition 1 of Kubokawa and Srivastava (1999).

**Proposition 2.1** Decompose \( S = TT^* \) in which \( T \) is a \( p \times p \) lower triangular matrix. Also set \( V = \text{Diag}(v_1, \ldots, v_p) \) with the \( v_k \)'s, \( k = 1, \ldots, p \), being positive constants. Then we have

\[
\mathbb{E}\left[ \text{Tr} \left( S \Sigma^{-1} \right) + \text{Tr} \left( TVT^* \Sigma^{-1} \right) \right] = \mathbb{E}^F \left[ np + \sum_{k=1}^{p} (n + p + 1 - 2k) v_k \right],
\]

where \( \mathbb{E}^F \) is defined by (4).

**Proof.** The above equality follows from two equalities, i.e., \( \mathbb{E}\left[ \text{Tr} \left( S \Sigma^{-1} \right) \right] = \mathbb{E}^F [np] \) and \( \mathbb{E}\left[ \text{Tr} \left( TVT^* \Sigma^{-1} \right) \right] = \mathbb{E}^F \left[ \sum_{k=1}^{p} (n + p + 1 - 2k) v_k \right] \). The first equality is derived from Lemma 2.4 with \( \varphi_k(L) = \ell_k, k = 1, 2, \ldots, p \). By an invariance argument it suffices to prove that the second inequality holds in the case that \( \Sigma = I_p \). From an application of Lemma 2.1 and the fact that \( \text{Tr} \left( D S S \right) = p^2 \), we have

\[
\mathbb{E}\left[ \text{Tr} \left( S \right) \right] = \mathbb{E}^F [p(n - p) + \text{Tr} \left( D S S \right)] = np \mathbb{E}^F [1].
\]

Furthermore, since the distribution of \( S \) is invariant under the permutation of the coordinates, we see that

\[
\mathbb{E}[s_{kk}] = n \mathbb{E}^F [1] \tag{7}
\]

for \( k = 1, 2, \ldots, p \). Also, note that the Jacobian of the transformation from \( S \) to \( T \) depends only on the diagonal elements of \( T \). Thus, it suffices to show that \( \mathbb{E}[|t_{21}|^2] = \mathbb{E}^F [1] \) in the case that \( p = 2 \) in order to prove that, for \( k \neq l \),

\[
\mathbb{E}[|t_{kl}|^2] = \mathbb{E}^F [1]. \tag{8}
\]

To prove that \( \mathbb{E}[|t_{21}|^2] = \mathbb{E}^F [1] \), let \( Z = (z_1, z_2) \) with an \( n \times 1 \) vector \( z_2 \). Set \( v = Qz_2 \) such that \( Qz_1 = (t_{11}, 0)' \). We decompose \( v = (v_1, v_2)' \) with a scalar \( v_1 \). Recall that \( S = Z^*Z \) to see that \( s_{21} = z_2^*z_1 = u_2^*Q^*Qz_1 = \bar{v}_1 t_{11} \). Furthermore the distribution of \( (X, z_1, v_1, v_2) \) is

\[
\text{Det} \left( \Sigma \right)^{-m} f \left( \text{Tr} \{ \Sigma^{-1} (x - \theta)^*(x - \theta) \} + \text{Tr} \left( u_1^*u_1 \right) + |v_1|^2 + v_2^*v_2 \right).
\]

Proceed in a manner similar to that used in the derivation of (17) in Section 3 to see that \( \mathbb{E}[|t_{21}|^2] = \mathbb{E}[|v_1|^2] = \mathbb{E}^F [1] \). By using (7) and (8) some algebraic computations complete the proof of the second equality. \( \square \)
Remark 2.3 If $S$ follows the complex Wishart distribution with the degrees of freedom $n$ and a scale matrix $\Sigma$, i.e., $\mathcal{CW}_p(n, \Sigma)$, then we have $F = f$ in (4) from which it follows that $\mathbb{E}^F[1] = 1$. Thus, Proposition 2.1 results in the equations $\mathbb{E} [\text{Tr} (TT^*\Sigma^{-1})] = \sum_{k=1}^p (n + p + 1 - 2k) v_k$ and $\mathbb{E} [\text{Tr} (SS^{-1})] = np$.

Consider estimators of the form 

$$\hat{\Sigma}^{(JS)} = T \text{Diag}(\tilde{v}_1, \ldots, \tilde{v}_p) T^*,$$  \hspace{1cm} (9) 

in which $T$ is defined in Proposition 2.1 and $\tilde{v}_k = (n + p + 1 - 2k)^{-1}$, $k = 1, 2, \ldots, p$.

**Proposition 2.2** Using the loss function $L(\hat{\Sigma}, \Sigma)$ given by (3) the James-Stein type estimator $\hat{\Sigma}^{(JS)}$ given by (9) is better than the usual estimator $n^{-1}S$ uniformly for every function $f(\cdot)$ in (2).

**Proof.** The proof is obtained immediately from an application of Proposition 2.1. \hfill $\square$

Next we consider a class of unitary invariant estimators of the form 

$$\hat{\Sigma}^{(\Phi)} = U \text{Diag}(\varphi_1(L), \ldots, \varphi_p(L)) U^*,$$  \hspace{1cm} (10) 

in which $U$ is defined in Lemma 2.2 and the $\varphi_k(L)$'s, $k = 1, 2, \ldots, p$, are differentiable functions from $\mathbb{R}^{p}_{\geq}$ to $\mathbb{R}$. This class of estimators has been suggested by Abraham and Dey (1994) via an extension of orthogonally invariant estimators for the real normal covariance matrix.

**Proposition 2.3** Using the loss function $L(\hat{\Sigma}, \Sigma)$ given by (3) the estimators $\hat{\Sigma}^{(\Phi)}$ given by (10) are better than the James-Stein type estimator $\hat{\Sigma}^{(JS)}$ given by (9) uniformly for every function $f(\cdot)$ in (2), if the following inequalities hold almost surely;

$$p - \sum_{k=1}^p \left\{ \frac{\partial \varphi_k(L)}{\ell_k} + \sum_{j \neq k} \frac{\varphi_k(L) - \varphi_j(L)}{\ell_k - \ell_j} + (n - p) \frac{\varphi_k(L)}{\ell_k} \right\} \geq 0,$$  \hspace{1cm} (11) 

$$\sum_{k=1}^p \left\{ \log(n + p + 1 - 2k) + \log \frac{\varphi_k(L)}{\ell_k} \right\} \geq 0.$$  \hspace{1cm} (12) 

**Proof.** For simplicity, we omit the argument $L$ in $\varphi_k(L)$ as $\varphi_k$ for $k = 1, 2, \ldots, p$. Using Lemma 2.4 and Proposition 2.1 the risk difference between the estimators $\hat{\Sigma}^{(JS)}$ and $\hat{\Sigma}^{(\Phi)}$ is

$$L(\hat{\Sigma}^{(JS)}, \Sigma) - L(\hat{\Sigma}^{(\Phi)}, \Sigma) = \mathbb{E}^F[p] - \mathbb{E} \left[ \sum_{k=1}^p \left\{ - \log(n + p + 1 - 2k) + \log \ell_k \right\} \right]$$

$$- \mathbb{E} F \left[ \sum_{k=1}^p \left\{ \frac{\partial \varphi_k}{\ell_k} + \sum_{j \neq k} \frac{\varphi_k - \varphi_j}{\ell_k - \ell_j} + (n - p) \frac{\varphi_k}{\ell_k} \right\} \right]$$

$$+ \mathbb{E} \left[ \sum_{k=1}^p \log \varphi_k \right].$$
Rearranging the terms inside each of two different types of the expectations separately we obtain the desired result.

**Theorem 2.1** Let $\tilde{v}_k = (n + p + 1 - 2k)^{-1}$, $k = 1, 2, \ldots, p$, and decompose $S = ULU^*$ in which $U = (u_{jk})$ is a $p \times p$ unitary matrix and $L = \text{Diag}(\ell_1, \ell_2, \ldots, \ell_p)$ with $\ell_1 \geq \ell_2 \geq \cdots \geq \ell_p > 0$. Using the loss function $\mathcal{L}(\hat{\Sigma}, \Sigma)$ given by (3) the Dey-Srinivasan type estimator $\hat{\Sigma}^{(DS)} = U \text{Diag}(\tilde{v}_1\ell_1, \ldots, \tilde{v}_p\ell_p)U^*$ is better than the James-Stein type estimator $\hat{\Sigma}^{(JS)}$ given by (9) uniformly for every function $f(\cdot)$ in (2).

**Proof.** The proof follows immediately from the verification of the inequalities (11) and (12) with $\varphi_k = \tilde{v}_k\ell_k$, $k = 1, 2, \ldots, p$. □

**Remark 2.4** Sheena and Takemura (1992) called (10) order-preserving if it satisfies the condition $\varphi_1(L) \geq \varphi_2(L) \geq \cdots \geq \varphi_p(L)$. The estimator $\hat{\Sigma}^{(DS)}$ is further improved upon by a corresponding order-preserving estimators. This result follows from the fact that Lemma 1 of Sheena and Takemura holds for nonincreasing functions $f(\cdot)$ in (2). It can be also seen that an estimator similar to that given by Perron (1992) for the real normal covariance matrix improves upon the James-Stein type estimator $\hat{\Sigma}^{(JS)}$ uniformly for every function $f(\cdot)$.

### 3 Proofs of Lemmas

Before we get into proving the lemmas presented in the previous section, we state three approaches to prove the Wishart identity for the real case.

Firstly, Stein (1977) used integration-by-parts for the multivariate real normal distribution in which the mean vector is zero and the covariance matrix is identity. Then Stein (1977) used an algebraic lemma on the eigenstructure of a symmetric matrix in order to obtain the result for the real case similar to that presented in Lemma 2.4. This approach was taken by Kubokawa and Srivastava (1999) in order to extend the Wishart identity to real elliptically contoured distributions. This approach is also taken in the present paper.

Secondly, Haff used integration-by-parts for the real Wishart distribution and applied the same algebraic lemma as in Stein (1977). This approach was taken by Svensson (2004) in order to obtain the Haff-Stein identity for the complex Wishart distribution.

Thirdly, Sheena (1995) used integration-by-parts for the distribution of the eigenvalues of the real Wishart distribution. The third approach gives a direct proof of the result for the real case similar to that presented in Lemma 2.4. Also, this approach was taken by Konno (2005)
in order to obtain integration-by-parts for the Wishart distribution on irreducible symmetric cones.

The next lemma describes the chain rule for the differentiation which is useful for the proof of Lemma 2.1.

**Lemma 3.1** Let \( Y = (y_{jk}) \) be an \( n \times p \) matrix of complex entries and set \( W = Y^*Y \). Furthermore, let

\[
\frac{\partial}{\partial y_{jk}} = \frac{\partial}{\partial (\text{Re} \, y_{jk})} + \sqrt{-1} \frac{\partial}{\partial (\text{Im} \, y_{jk})}, \quad \frac{\partial}{\partial w_{kl}} = \frac{\partial}{\partial (\text{Re} \, w_{kl})} + \sqrt{-1} \frac{\partial}{\partial (\text{Im} \, w_{kl})},
\]

for \( k, l = 1, 2, \ldots, p \) and \( j = 1, 2, \ldots, n \). Then we have

\[
\frac{\partial}{\partial y_{jl}} = 2 \sum_{k=1}^{p} y_{jk} \frac{1 + \delta_{kl}}{2} \frac{\partial}{\partial w_{kl}}.
\]

**Proof.** The real and the imaginary parts of the \((c_1, c_2)\)-element of the matrix \( W \) are, respectively,

\[
\text{Re} \, w_{c_1c_2} = \sum_{c_3} \{\text{Re} \, y_{c_3c_1} \text{Re} \, y_{c_3c_2} + \text{Im} \, y_{c_3c_1} \text{Im} \, y_{c_3c_2}\},
\]

\[
\text{Im} \, w_{c_1c_2} = \sum_{c_3} \{\text{Re} \, y_{c_3c_1} \text{Im} \, y_{c_3c_2} - \text{Im} \, y_{c_3c_1} \text{Re} \, y_{c_3c_2}\}.
\]

From these two equations the real part of \( \partial/\partial y_{jl} \) is

\[
\frac{\partial}{\partial (\text{Re} \, y_{jl})} = \sum_{c_1 \ge c_2} \left\{ \frac{\partial (\text{Re} \, w_{c_1c_2})}{\partial (\text{Re} \, y_{jl})} \frac{\partial}{\partial (\text{Re} \, w_{c_1c_2})} + \frac{\partial (\text{Im} \, w_{c_1c_2})}{\partial (\text{Re} \, y_{jl})} \frac{\partial}{\partial (\text{Im} \, w_{c_1c_2})} \right\}
\]

\[
= \sum_{c_1 \ge c_2, c_3} \left[ (\delta_{c_3j} \delta_{c_1l} \text{Re} \, y_{lc_2} + \delta_{c_3j} \delta_{c_2l} \text{Re} \, y_{c_3c_1} \frac{\partial}{\partial (\text{Re} \, w_{c_1c_2})}) \right.
\]

\[
+ \left. (\delta_{c_3j} \delta_{c_1l} \text{Im} \, y_{c_3c_2} - \delta_{c_3j} \delta_{c_2l} \text{Im} \, y_{c_3c_1} \frac{\partial}{\partial (\text{Im} \, w_{c_1c_2})} \right],
\]

while the imaginary part is

\[
\frac{\partial}{\partial (\text{Im} \, y_{jl})} = \sum_{c_1 \ge c_2} \left\{ \frac{\partial (\text{Re} \, w_{c_1c_2})}{\partial (\text{Im} \, y_{jl})} \frac{\partial}{\partial (\text{Re} \, w_{c_1c_2})} + \frac{\partial (\text{Im} \, w_{c_1c_2})}{\partial (\text{Im} \, y_{jl})} \frac{\partial}{\partial (\text{Im} \, w_{c_1c_2})} \right\}
\]

\[
= \sum_{c_1 \ge c_2, c_3} \left[ (\delta_{c_3j} \delta_{c_1l} \text{Im} \, y_{c_3c_2} + \delta_{c_3j} \delta_{c_2l} \text{Im} \, y_{c_3c_1} \frac{\partial}{\partial (\text{Re} \, w_{c_1c_2})}) \right.
\]

\[
+ \left. (\delta_{c_3j} \delta_{c_1l} \text{Re} \, y_{c_3c_2} - \delta_{c_3j} \delta_{c_2l} \text{Re} \, y_{c_3c_1} \frac{\partial}{\partial (\text{Im} \, w_{c_1c_2})} \right].
\]
Therefore we have
\[\frac{\partial}{\partial (\text{Re } y_{jl})} + \sqrt{-1} \frac{\partial}{\partial (\text{Im } y_{jl})} \]
\[= \sum_{c_1 \geq c_2} \frac{1 + \delta_{c_1c_2}}{2} \left[ (\delta_{c_1l} \text{Re } y_{jc_2} + \delta_{c_2l} \text{Re } y_{jc_1}) \frac{\partial}{\partial (\text{Re } w_{c_1c_2})} \right. \]
\[+ (\delta_{c_1l} \text{Im } y_{jc_2} - \delta_{c_2l} \text{Im } y_{jc_1}) \frac{\partial}{\partial (\text{Im } w_{c_1c_2})} \]
\[+ \sqrt{-1} (\delta_{c_1l} \text{Im } y_{jc_2} + \delta_{c_2l} \text{Im } y_{jc_1}) \frac{\partial}{\partial (\text{Re } w_{c_1c_2})} \]
\[+ \sqrt{-1} (\delta_{c_2l} \text{Re } y_{jc_1} - \delta_{c_1l} \text{Im } y_{jc_2}) \frac{\partial}{\partial (\text{Im } w_{c_1c_2})} \right] \]
\[= \sum_{c_2} \frac{1 + \delta_{jc_2}}{2} \left[ \text{Re } y_{jc_2} \frac{1}{\partial (\text{Re } w_{lc_2})} + \text{Im } y_{jc_2} \frac{1}{\partial (\text{Im } w_{lc_2})} \right. \]
\[+ \sqrt{-1} \left( \text{Im } y_{jc_2} \frac{1}{\partial (\text{Re } w_{lc_2})} - \text{Re } y_{jc_2} \frac{1}{\partial (\text{Im } w_{lc_2})} \right) \]
\[+ \sum_{c_1} \frac{1 + \delta_{jc_1}}{2} \left[ (\text{Re } y_{jc_1} \frac{1}{\partial (\text{Re } w_{cl})} - \text{Im } y_{jc_1} \frac{1}{\partial (\text{Im } w_{cl})} \right. \]
\[+ \sqrt{-1} \left( \text{Im } y_{jc_1} \frac{1}{\partial (\text{Re } w_{cl})} + \text{Re } y_{jc_1} \frac{1}{\partial (\text{Im } w_{cl})} \right) \right]. \]

Since the real part of \(W\) is symmetric and the imaginary part of \(W\) is skew-symmetric, this completes the proof. \(\square\)

### 3.1 Proof of Lemma 2.1

To reduce the equation (9) to the case when \(\Sigma = I_p\), we first observe that

\[E[\text{Tr } (D_S G S)] = E^F[\text{Tr } (SD_S G) + p \text{ Tr } G],\]  
(13)

since the \((i, i)\)-element of \(D_S G S\) can be written as

\[(D_S G S)_{ii} = \sum_{k=1}^p \left\{ (D_S G)_{ik} s_{ki} + \sum_{j=1}^p g_{jk} \frac{1 + \delta_{ij}}{2} \left( \frac{\partial}{\partial (\text{Re } s_{ij})} + \sqrt{-1} \frac{\partial}{\partial (\text{Im } s_{ij})} \right) \right. \]
\[\times (\text{Re } s_{ki} + \sqrt{-1} \text{Im } s_{ki}) \right\} \]
\[= \sum_{k=1}^p \{ s_{ki} (D_S G)_{ik} + \sum_{j=1}^p g_{jk} (\delta_{ij} \delta_{ki} + \delta_{kj} (1 - \delta_{ij})) \} \]
\[= \sum_{k=1}^p s_{ki} (D_S G)_{ik} + \sum_{j=1}^p g_{jj}. \]
The second last equality in the above equations follows from the fact that the real part of $S$ is symmetric and the imaginary part is skew-symmetric. Replacing $G$ with $GS$ in (2), we get

$$
E[\text{Tr}(GS\Sigma^{-1})] = E^F[(n-p)\text{Tr}(G) + \text{Tr}(D_SGS)].
$$

Substitute (13) into the above equation to get that

$$
E[\text{Tr}(GS\Sigma^{-1})] = E^F[n\text{Tr}(G) + \text{Tr}(SD_SH(S))].
$$

Hence it suffices to prove that, for a complex-valued function $h(S)$,

$$
E[h(S)S\Sigma^{-1}] = E^F[nh(S)I_p + SD_SH(S)].
$$

(14)

To this end, write $\Sigma = AA^*$ and $S = AWAW^*$ in which $W = (w_{ij})$ is a $p \times p$ symmetric matrix and $A = (a_{ij})$ is a $p \times p$ nonsingular matrix with its inverse $A^{-1} = (a^{ij})$. Then

$$
\frac{1+\delta_{ij}}{2} \frac{\partial}{\partial s_{ij}} = \frac{1+\delta_{ij}}{2} \sum_{c_1 \geq c_4} \left\{ \frac{\partial}{\partial (\text{Re} s_{ij})} + \sqrt{-1} \frac{\partial}{\partial (\text{Im} s_{ij})} \right\}
$$

$$
\times \sum_{c_2, c_3} a^{c_1c_2} \left( \text{Re} s_{c_2c_3} + \sqrt{-1} \text{Im} s_{c_2c_3} \right) a^{c_4c_3} \frac{\partial}{\partial w_{c_1c_4}}
$$

$$
= \sum_{c_1 \geq c_4} \delta_{ic_3} \delta_{jc_2} a^{c_1c_2} a^{c_4c_3} \frac{\partial}{\partial w_{c_1c_4}}
$$

$$
= \sum_{c_1 \geq c_4} a^{c_1c_3} a^{c_4c_4} \frac{\partial}{\partial w_{c_1c_4}} = \{(A^*)^{-1}D_W^tA^{-1}\}_{ij},
$$

which is the $(i, j)$-element of $(A^*)^{-1}D_W^tA^{-1}$. Here $D_W^t$ denotes the transpose of the matrix $D_W$. Using this equation, we can reduce (14) to

$$
E[W\text{Tr}(S)] = E^F[nh(S)I_p + WD_W^tH(S)].
$$

(15)

Furthermore, let $Z = AY$ with $Y = (y_{ij})$. Since $W = Y^*Y$, it follows from Lemma 3.1 that

$$
\sum_{c_3} w_{ic_3} \frac{1+\delta_{jc_3}}{2} \frac{\partial h}{\partial w_{jc_3}}(S) = \sum_{k, c_3} \bar{y}_{ki}y_{kc_3} \frac{1+\delta_{jc_3}}{2} \frac{\partial h}{\partial w_{jc_3}}(S) = \frac{1}{2} \sum_k \bar{y}_{ki} \frac{\partial h}{\partial y_{kj}}(S).
$$

(16)

Furthermore, note that

$$
\frac{\partial}{\partial y_{kj}} \{\bar{y}_{ki}h(S)\} = 2\delta_{ij} + \bar{y}_{ki} \frac{\partial h}{\partial y_{kj}}(S).
$$

Thus, it suffices to prove that

$$
E[w_{ij}h(S)] = \sum_{k=1}^n E[\bar{y}_{ki}y_{kj}h(S)] = E^F \left[ \delta_{ij} + \frac{1}{2} \bar{y}_{ki} \frac{\partial h}{\partial y_{kj}}(S) \right],
$$

(17)
in which the expectation is taken with respect to a density function $\text{Det}(\Sigma)^{-m} f(\tilde{R})$ with
\[
\tilde{R} = \text{Tr}(\Sigma^{-1}x^*x) + \sum_{(c_1, c_2) \neq (j, k)} (\tilde{R}) y_{c_1c_2}^2.
\]

But, from integration-by-parts, we note that
\[
\int y_{ki} y_{kj} h(S) f(\tilde{R}) \ dx \ dy_{kj} \prod_{(c_1, c_2) \neq (j, k)} dy_{c_1c_2}
= \frac{1}{2} \int \frac{\partial}{\partial y_{kj}} \{ y_{ki} h(S) \} F(\tilde{R}) \ dx \ dy_{kj} \prod_{(c_1, c_2) \neq (j, k)} dy_{c_1c_2}
\]
and that
\[
\frac{\partial}{\partial y_{kj}} \{ y_{ki} h(S) \} = 2 \delta_{ij} + \sum_{c_1c_2} (c_1 \neq c_2) | y_{c_1c_2} |^2.
\]
Here, $dx$ and $dy_{c_1c_2}$ indicate $\wedge^m_j \wedge^p_k (\text{Re} x_{jk}) d (\text{Im} x_{jk})$ and $d (\text{Re} y_{c_1c_2}) d (\text{Im} y_{c_1c_2})$, respectively. These two equations give the equation (17), which completes the proof of the lemma.

\[\square\]

3.2 Proof of Lemma 2.2

Since the real and the imaginary parts of $S$ are symmetric and skew-symmetric matrices respectively, we have
\[
ds_{ab}(d^*_S) = \frac{1 + \delta_{st}}{2} (d(\text{Re} s_{ab}) + \sqrt{-1} d(\text{Im} s_{ab})) \left( \frac{\partial}{\partial (\text{Re} s_{st})} + \sqrt{-1} \frac{\partial}{\partial (\text{Im} s_{st})} \right)
= \frac{1 + \delta_{st}}{2} \left\{ d(\text{Re} s_{ab}) \left( \frac{\partial}{\partial (\text{Re} s_{st})} \right) - d(\text{Im} s_{ab}) \left( \frac{\partial}{\partial (\text{Im} s_{st})} \right) \right\}
= \frac{1 + \delta_{st}}{2} \left( \delta_{as} \delta_{bt} + \delta_{at} \delta_{bs} \right) \frac{1}{1 + \delta_{st}} - (\delta_{as} \delta_{bt} - \delta_{at} \delta_{bs}) (1 - \delta_{st})
= \delta_{at} \delta_{bs}.
\]

Now taking the differentials of $S = ULU^*$ and multiplying on the left by $U^*$ and on the right by $U$, we get
\[
U^* dSU = (U^* dU) L + L(U^* dU)^* + dL. \tag{19}
\]

But, from the differentials of $U^* U = I_p$, we can see that $U^* dU$ is skew-symmetric. Expressing the diagonal and the off-diagonal elements of (19) we get
\[
d\ell_k = \{U^* dSU\}_{kk}, \tag{20}
\]
\[
\{U^* dU\}_{ak} = \begin{cases} \frac{1}{\ell_k - \ell_a} \{U^* dSU\}_{ak} & (a \neq k), \\ 0 & (a = k). \end{cases} \tag{21}
\]
Using (18), (20), and (21), we get
\[ d_{S}^{st} \ell_k = \sum_{a,b} \overline{u}_{ak} d_{ab}(d_{S}^{st})u_{bk} = \sum_{a,b} \overline{u}_{ak}\delta_{at}\delta_{bs}u_{bk} = \overline{u}_{tk}u_{sk}, \]
\[ d_{S}^{st}(u_{jk}) = \sum_{a\neq k} u_{ja}(U^{*}dU)_{ak}(d_{S}^{st}) = \sum_{a\neq k,c_1,c_2} \frac{1}{\ell_k - \ell_a} \overline{u}_{c_1a} d_{S_{c_1c_2}}(d_{S}^{st})u_{c_2k} = \sum_{a\neq k} \frac{u_{ja}u_{ta}u_{sk}}{\ell_k - \ell_a}, \]
which completes the proof of the first and the second equalities in this lemma. The third equality in this lemma is obtained by taking the complex conjugate of (19) and proceeding via a similar calculation to that presented in the derivation of the second equality in this lemma.

\[ \square \]

### 3.3 Proof of Lemma 2.3

Use Lemma 2.2 to see that the \((i, j)\)-element of \(D_{S}[U\Phi(L)U^{*}]\) is
\[ \sum_{k_1,k_2} d_{S}^{ik_1}u_{k_1k_2}\varphi_{k_2}u_{jk_2} = \sum_{k_1,k_2} \left[ \varphi_{k_2}u_{jk_2}d_{S}^{ik_1}u_{k_1k_2} + \varphi_{k_2}u_{k_1k_2}d_{S}^{ik_1}u_{jk_2} + u_{k_1k_2}u_{jk_2}u_{k_1k_2} \sum_{k_3} \frac{\partial\varphi_{k_2}}{\partial l_{k_3}} \overline{d_{S}^{ik_1}u_{jk_3}} \right] \]
\[ = \sum_{k_1,k_2} \left[ \varphi_{k_2}u_{jk_2} \sum_{a\neq k_2} \frac{u_{k_1a}u_{k_1a}u_{jk_2}}{\ell_{k_2} - \ell_a} + \varphi_{k_2}u_{k_1k_2} \sum_{a\neq k_2} \frac{u_{ja}u_{ta}u_{k_1k_2}}{\ell_{k_2} - \ell_a} \right. \]
\[ + u_{k_1k_2}u_{jk_2} \sum_{k_3} \frac{\partial\varphi_{k_2}}{\partial l_{k_3}} u_{ik_3}u_{k_1k_3} \left. \right] \]
\[ = \sum_{k} u_{ik} \left( \frac{\partial\varphi_{k}}{\partial l_{k}} + \sum_{a\neq k} \frac{\varphi_{k} - \varphi_{a}}{\ell_{k} - \ell_a} \right) \overline{u_{jk}}, \]
which completes the proof. \( \square \)

### Acknowledgements

I am grateful to the anonymous referee and the advisory editor, Professor Michael D. Perlman, for their detailed and helpful comments and suggestions which enabled me to improve the presentation of the paper. I would like to thank Dr. Lennart Svensson for providing me with his unpublished manuscript.

### References

Abraham, D.A., Dey, D.K., 1994. The application of improved covariance estimation to


Lin, S.P., Perlman, M.D., 1985. A Monte Carlo comparison of four estimators for a covariance
matrix. In Multivariate Analysis VI (P.R. Krishnaiah, ed.), North-Holland, Amsterdam, 411-426,


