Estimation of the noncentrality matrix of a noncentral Wishart distribution with unit scale matrix.
A matrix generalization of leung’s domination result

Heinz Neudecker∗

Cesaro

Abstract

The main aim is to estimate the noncentrality matrix of a noncentral Wishart distribution. The method used is Leung’s but generalized to a matrix loss function. Parallelly Leung’s scalar noncentral Wishart identity is generalized to become a matrix identity. The concept of Löwner partial ordering of symmetric matrices is used.

MSC: 62H12, 15A24, 15A45

Keywords: Noncentral Wishart matrix identity, noncentrality matrix, decision-theoretic estimation, matrix loss function, Löwner matrix ordering, Haffians

1 Introduction

We consider $S \sim W_m(n, I_m, M'M)$. Following Leung (1994) we recall that the habitual unbiased estimator of $M'M$ is $T := S - n I_m$. Under certain conditions $T_\alpha := T + \alpha(tr S)^{-1} I_m$ dominates $T$ for a suitable choice of $\alpha$, as was shown by Leung, who used the loss function

$$A \left[(M'M)^{-1}, R\right] := \text{tr}\left\{(M'M)^{-1} R - I_m\right\}^2.$$


∗Address for correspondence: Oosterstraat, 13. 1741 GH Schagen. The Netherlands. Postal address: NL 1741 GH 13. E-mail: heinz@fee.uva.nl
Received: May 2001
Accepted: October 2003
In this article we propose to use a matrix loss function, \( L[(M'M)^{-1}, R] := \left\{ (M'M)^{-1} R - I_m \right\} \left\{ (M'M)^{-1} R - I_m \right\} \quad \) and apply the concept of Löwner partial ordering of symmetric matrices. We shall show that Leung’s result still holds approximately, the error term being of order \( o(n^{-1}) \). For accomplishing this we need a matrix version of Leung’s Identity for the noncentral Wishart distribution. This will be presented first.

A matrix version of an ancillary lemma by Leung, viz his Lemma 3.1 will next be established. The generalized domination result will then follow straightforwardly.

We shall employ an approximation of \( E(tr S^{-1}F) \), where \( E \) is the expectation operator. A lemma on the matrix Haffian \( \nabla \phi F \), where \( \phi \) and \( F \) are scalar and matrix functions of \( S \), will be proved in Appendix 1. In Appendix 2 we shall prove a lemma on the scalar Haffian \( tr \nabla F_2 A F_1 \), when \( F_1 \) and \( F_2 \) are matrix functions of \( S \) and \( A \) is a constant matrix.

2 A matrix version of Leung’s identity for the noncentral Wishart distribution

We quote Leung’s Theorem 2.1, where without loss of generality we take \( h = 1 \), \( h \) being a scalar function of \( S \) in Leung’s work:

\[
E tr \Sigma^{-1} F = 2E tr \nabla F + (n - m - 1)E tr S^{-1} F + E_1 tr \Sigma^{-1} M' MS^{-1} F, \tag{1}
\]

where \( S \sim W_m(n, \Sigma, \Sigma^{-1} M'M) \), \( E \) denotes the expectation with respect to this distribution, \( E_1 \) denotes the expectation with respect to the distribution \( W_m(n + m + 1, \Sigma, \Sigma^{-1} M'M) \), \( F = F(S) \) and \( n > m + 1 \). The matrices \( S, \Sigma, F \) and \( \nabla \) are square of dimension \( m \), whereas \( M \) has dimension \( n \times m \). It is assumed that \( M \) has full column rank. Further \( \nabla F \) is the matrix Haffian as denoted by Neudecker (2000b). Inspired by Haff (1981), who did it for the central Wishart distribution, we shall establish a matrix version of (1).

**Theorem 1**

\[
EF_1 \Sigma^{-1} F_2 = 2EF_1 \nabla F_2 + 2 \left( EF_1' \nabla F_1' \right) + (n - m - 1)EF_1 S^{-1} F_2 + E_1 F_1 \Sigma^{-1} M' MS^{-1} F_2, \tag{2}
\]

for \( F_1 \) and \( F_2 \) satisfying the conditions of Lemma 5.

**Proof.** Take \( F = F_2 e_j e'^i F_1 \), with unit vectors \( e_i \) and \( e_j \). We then use the identity:

\[
tr \nabla F_2 A F_1 = tr (\nabla F_2) A F_1 + tr (\nabla F_1') A' F_2',
\]

with constant \( A \). For a proof see Lemma 5.
Taking \( A = e_i e'_j \) we get

\[
E \text{tr} \Sigma^{-1} F_2 e_i e'_j F_1 = 2 E \text{tr} (\nabla F_2) e_i e'_j F_1 + 2 E \text{tr} (\nabla F'_1) e_i e'_j F'_2 + (n - m - 1) \text{tr} E S^{-1} F_2 e_i e'_j F_1 + E_1 \text{tr} \Sigma^{-1} M'MS^{-1} F_2 e_i e'_j F_1
\]

or equivalently

\[
(EF_1 \Sigma^{-1} F_2)_{ij} = 2 (EF_1 \nabla F_2)_{ij} + 2 (EF'_2 \nabla F'_1)_{ji} + (n - m - 1) (EF_1 S^{-1} F_2)_{ij} + (E_1 F_1 \Sigma^{-1} M'MS^{-1} F_2)_{ij}.
\]

\[\Box\]

Note: It was assumed that (1) holds for all \( F = F_2 e_i e'_j F_1 \), which puts stronger conditions on the input matrix than was necessary for (1). By choosing \( F_1 = I_m \) and taking traces we derive (1) from (2).

For discussion of the central Wishart case we refer to Haff (1981).

3 A matrix version of Leung’s lemma 3.1

Lemma 2

\[
E (\text{tr} S)^{-1} (M'M)^{-1} S (M'M)^{-1} < n E (\text{tr} S)^{-1} (M'M)^{-2} - 2(n - 4)E (\text{tr} S)^{-2} (M'M)^{-2} + E_1 (\text{tr} S)^{-1} (M'M)^{-1} - 2E_1 (\text{tr} S)^{-2} (M'M)^{-1}.
\]

where \( S \sim W_m(n, I_m, M'M) \) and \( M'M \) is assumed to be nonsingular. The inequality \( A < B \), for symmetric \( A \) and \( B \), stands for the Löwner ordering meaning that \( B - A \) is positive definite.

Proof. Take \( F_1 = (\text{tr} S)^{-1} (M'M)^{-1} \) and \( F_2 = S (M'M)^{-1} \). By Theorem 1 (with \( \Sigma = I_m \)):

\[
E (\text{tr} S)^{-1} (M'M)^{-1} S (M'M)^{-1} = 2 E (\text{tr} S)^{-1} (M'M)^{-1} \nabla S (M'M)^{-1} + 2 \left\{ E (M'M)^{-1} S \nabla (\text{tr} S)^{-1} (M'M)^{-1} \right\}^T + (n - m - 1) E (\text{tr} S)^{-1} (M'M)^{-2} + E_1 (\text{tr} S)^{-1} (M'M)^{-1} = (m + 1) E (\text{tr} S)^{-1} (M'M)^{-2} - 2 E (\text{tr} S)^{-2} (M'M)^{-1} S (M'M)^{-1} + (n - m - 1) E (\text{tr} S)^{-1} (M'M)^{-2} + E_1 (\text{tr} S)^{-1} (M'M)^{-1} =
\]
\[ = nE \left( \text{tr} S \right)^{-1} (M'M)^{-2} - 2E \left( \text{tr} S \right)^{-2} (M'M)^{-1} S (M'M)^{-1} + \\
+E_1 \left( \text{tr} S \right)^{-1} (M'M)^{-1} \]  

(i)

Further

\[
E \left( \text{tr} S \right)^{-2} (M'M)^{-1} S (M'M)^{-1} = 2E \left( \text{tr} S \right)^{-2} (M'M)^{-1} \nabla S (M'M)^{-1} + \\
+2 \left[ E (M'M)^{-1} S \nabla (\text{tr} S)^{-2} (M'M)^{-1} \right] + (n - m - 1) E \left( \text{tr} S \right)^{-2} (M'M)^{-2} + \\
+E_1 \left( \text{tr} S \right)^{-2} (M'M)^{-1},
\]

where we applied Theorem 1 (with \( \Sigma = I_m \)) using \( F_1 = (\text{tr} S)^{-2} (M'M)^{-1} \) and \( F_2 = S (M'M)^{-1} \). Proceeding as before we get

\[
E \left( \text{tr} S \right)^{-2} (M'M)^{-1} S (M'M)^{-1} = (m + 1) E \left( \text{tr} S \right)^{-2} (M'M)^{-2} - \\
-4E \left( \text{tr} S \right)^{-3} (M'M)^{-1} S (M'M)^{-1} + (n - m - 1) E \left( \text{tr} S \right)^{-2} (M'M)^{-2} + \\
+E_1 \left( \text{tr} S \right)^{-2} (M'M)^{-1} = nE \left( \text{tr} S \right)^{-2} (M'M)^{-2} - \\
-4E \left( \text{tr} S \right)^{-3} (M'M)^{-1} S (M'M)^{-1} + E_1 \left( \text{tr} S \right)^{-2} (M'M)^{-1}.
\]

We use the Löwner ordering \( S < (\text{tr} S) I_m \), which yields \( (M'M)^{-1} S (M'M)^{-1} < (\text{tr} S) (M'M)^{-2} \). Hence we get

\[
(n - 4) E \left( \text{tr} S \right)^{-2} (M'M)^{-2} + E_1 \left( \text{tr} S \right)^{-2} (M'M)^{-1} < E \left( \text{tr} S \right)^{-2} (M'M)^{-1} S (M'M)^{-1}.
\]

Insertion in (i) finally yields

\[
E \left( \text{tr} S \right)^{-1} (M'M)^{-1} S (M'M)^{-1} < nE \left( \text{tr} S \right)^{-1} (M'M)^{-2} - 2(n - 4) E \left( \text{tr} S \right)^{-2} (M'M)^{-2} \\
-2E_1 \left( \text{tr} S \right)^{-2} (M'M)^{-1} + E_1 \left( \text{tr} S \right)^{-1} (M'M)^{-1}.
\]

Notes:

1. Nonsingularity of \( M'M \) is not trivial. A case of singularity is \( M' = \mu l \), where the \( n \) means are proportional.
2. Leung assumes \( n > 4 \). There is no need for it.
4 A matrix version of Leung’s domination result

We shall now prove the main result of this paper.

**Theorem 3**

\[ EL[ (M'M)^{-1}, T] > EL[ (M'M)^{-1}, T_\alpha] \]

for \( 0 < \alpha \leq 4 (n - 4) \), where

\[ L[ (M'M)^{-1}, R] := \left\{ (M'M)^{-1} R - I_m \right\}^\prime \left\{ (M'M)^{-1} R - I_m \right\}, \]

\[ T := S - nI_m \quad \text{and} \quad T_\alpha := T + \alpha (tr S)^{-1} I_m. \]

**Proof.**

\[ L[ (M'M)^{-1}, T] - L[ (M'M)^{-1}, T_\alpha] = \left\{ (M'M)^{-1} T - I_m \right\}^\prime \left\{ (M'M)^{-1} T - I_m \right\} - \left\{ (M'M)^{-1} T_\alpha - I_m \right\}^\prime \left\{ (M'M)^{-1} T_\alpha - I_m \right\} = 2n\alpha (tr S)^{-1} (M'M)^{-2} - a^2 (tr S)^{-2} (M'M)^{-2} - \alpha (tr S)^{-1} S (M'M)^{-2} - \alpha (tr S)^{-1} (M'M)^{-2} S + +2\alpha (tr S)^{-1} (M'M)^{-1}. \]

Its expected value is

\[ 2naE (tr S)^{-1} (M'M)^{-2} - a^2 E (tr S)^{-2} (M'M)^{-2} - \alpha E (tr S)^{-1} (M'M)^{-1} > \]

\[ > 2\alpha E (tr S)^{-1} (M'M)^{-1} S (M'M)^{-1} + \]

\[ +4\alpha(n - 4)E (tr S)^{-2} (M'M)^{-2} - 2\alpha E_1 (tr S)^{-1} (M'M)^{-1} + 4\alpha E_1 (tr S)^{-2} (M'M)^{-1} - \]

\[ -a^2 E (tr S)^{-2} (M'M)^{-2} - \alpha E (tr S)^{-1} S (M'M)^{-2} - \alpha E (tr S)^{-1} (M'M)^{-2} S + \]

\[ +2\alpha E (tr S)^{-1} (M'M)^{-1} = 2\alpha E (tr S)^{-1} (M'M)^{-1} S (M'M)^{-1} + \alpha [4(n - 4) - \alpha ] E (tr S)^{-2} (M'M)^{-2} + \]

\[ +2\alpha \left\{ E (tr S)^{-1} (M'M)^{-1} - E_1 (tr S)^{-1} (M'M)^{-1} \right\} + \]

\[ +4\alpha E_1 (tr S)^{-2} (M'M)^{-1} - \alpha E (tr S)^{-1} S (M'M)^{-2} - \alpha E (tr S)^{-1} (M'M)^{-2} S, \]

by Lemma 2.

We approximate \( E (tr S)^{-1} S \) by

\[ \mu (n I_m + M'M) - 2\mu^2 (n I_m + 2M'M) + +2\mu^3 (mn + 2 tr M'M) (n I_m + M'M), \]

with \( \mu^{-1} := tr (n I_m + M'M) \), the remainder being of order \( o(n^{-1}) \).
Insertion yields

\[ 2\alpha E (\text{tr} S)^{-1} (M^\prime M)^{-1} S (M^\prime M)^{-1} - \alpha E (\text{tr} S)^{-1} S (M^\prime M)^{-2} - \]

\[ -\alpha E (\text{tr} S)^{-1} (M^\prime M)^{-2} S = O + o(n^{-1}). \]

Hence to the order of approximation

\[ EL[ (M^\prime M)^{-1}, T] - EL[ (M^\prime M)^{-1}, T_o] > \alpha [4(n - 4) - \alpha] E (\text{tr} S)^{-2} (M^\prime M)^{-2} + \]

\[ + 2\alpha \left[ E (\text{tr} S)^{-1} (M^\prime M)^{-1} - E_1 (\text{tr} S)^{-1} (M^\prime M)^{-1} \right] + 4\alpha E_1 (\text{tr} S)^{-2} (M^\prime M)^{-1} > O, \]

as \( E (\text{tr} S)^{-1} \geq E_1 (\text{tr} S)^{-1} \).

For the auxiliary inequality see Leung (1994, p. 112).

\[ \square \]

**Appendix 1: a lemma on the matrix Haffian \( \nabla \varphi F \)**

**Lemma 4**

\[ \nabla \varphi F = \varphi \nabla F + \frac{\partial \varphi}{\partial X} F, \]

where \( \varphi \) is a scalar function of the symmetric matrix variable \( X \) and \( F \) is a matrix function thereof. Further

\[ \frac{\partial \varphi}{\partial X} := \frac{1}{2} \sum_{ij} \frac{\partial \varphi}{\partial X_{ij}} (E_{ij} + E_{ji}), \text{ where } E_{ij} := e_i e'_j \]

**Proof.**

\[ (\nabla \varphi F)_{ik} = \sum_{j} d_{ij} (\varphi F)_{jk} = \sum_{j} d_{ij} \varphi f_{jk} = \frac{1}{2} \sum_{j} \left( 1 + \delta_{ij} \right) \frac{\partial \varphi f_{jk}}{\partial X_{ij}} = \]

\[ = \frac{\partial \varphi f_{ik}}{\partial X_{ii}} + \frac{1}{2} \sum_{j \neq i} \frac{\partial \varphi f_{jk}}{\partial X_{ij}} = \varphi \left( \frac{\partial f_{ik}}{\partial X_{ii}} + \frac{1}{2} \sum_{j \neq i} \frac{\partial f_{jk}}{\partial X_{ij}} \right) + \left( \frac{\partial \varphi}{\partial X} \right) f_{ik} + \]

\[ + \frac{1}{2} \sum_{j \neq i} \frac{\partial \varphi}{\partial X_{ij}} f_{jk} = \varphi (\nabla F)_{ik} + \left( \frac{\partial \varphi}{\partial X} \right) F_{ik}, \text{ hence} \]

\[ \nabla \varphi F = \varphi \nabla F + \frac{\partial \varphi}{\partial X} F. \]

Here \( f_{jk} \) and \((F)_{jk}\) are the \( jk^{\text{th}} \) element of \( F \), \( F_{ik} \) is the \( i^{\text{th}} \) row of \( F \) and \( F_{ij} \) is the \( k^{\text{th}} \) column of \( F \).

For more details see Neudecker (2000b).  

\[ \square \]
Appendix 2: a lemma on the scalar Haffian $\text{tr } \nabla F_2 AF_1$

**Lemma 5**

$$\text{tr } \nabla F_2 AF_1 = \text{tr } (\nabla F_2) AF_1 + \text{tr } (\nabla F_1') A'F_2',$$

where $F_2$ and $F_1$ are functions of the symmetric matrix variable $X$ and $A$ is a constant matrix.

Each $F$ satisfies $F = \sum_k \varphi_k C_k$ or $dF = \sum P_l(dX)Q_l'$ with constant $C_k$, $P_l$ and $Q_l$.

We consider three cases. The first comprises $F_1 = \varphi C$ and $dF_2 = P(dX)Q'$, the second comprises $F_2 = \varphi C$ and $dF_1 = P(dX)Q'$, the third comprises $dF_1 = P(dX)Q'$ and $dF_2 = R(dX)T'$. The fourth case with $F_1 = \varphi_1 C_1$ and $F_2 = \varphi_2 C_2$ follows easily. Without loss of generality the summation signs were dropped.

**Proof.**

**Case 1.** We have $dF_1 = (d\varphi)C$, hence by Lemma 4 $\nabla F_1' = \frac{\partial \varphi}{\partial X}C'$. Further

$$d (F_2 AF_1) = (dF_2) AF_1 + F_2 AdF_1$$

$$= P(dX)Q'AF_1 + (d\varphi)F_2 AC$$

which implies

$$\nabla F_2 AF_1 = \frac{1}{2}P'Q'AF_1 + \frac{1}{2}(tr P)Q'AF_1 + \frac{\partial \varphi}{\partial X}F_2 AC,$$

$$\text{tr } \nabla F_2 AF_1 = \frac{1}{2}tr P'Q'AF_1 + \frac{1}{2}(tr P)tr Q'AF_1 + \text{tr } \frac{\partial \varphi}{\partial X}F_2 AC;$$

$$(\nabla F_2) AF_1 = \frac{1}{2}P'Q'AF_1 + \frac{1}{2}(tr P)Q'AF_1,$$

$$\text{tr } (\nabla F_2) AF_1 = \frac{1}{2}tr P'Q'AF_1 + \frac{1}{2}(tr P)tr Q'AF_1;$$

$$(\nabla F_1') A'F_2' = \frac{\partial \varphi}{\partial X}C' A'F_2',$$

$$\text{tr } (\nabla F_1') A'F_2' = \text{tr } \frac{\partial \varphi}{\partial X}C' A'F_2' = \text{tr } \frac{\partial \varphi}{\partial X}F_2 AC.$$

This yields the result.

**Case 2.** We replace $F_1$ by $F_2'$, $A$ by $A'$ and $F_2$ by $F_1'$ in the first result. This leads to

$$\text{tr } \nabla F_2' A'F_2' = \text{tr } (\nabla F_1') A'F_2' + \text{tr } (\nabla F_2) AF_1.$$
Using $\text{tr } \nabla F' = \text{tr } \nabla F$ we get

$$\text{tr } \nabla F_2 AF_1 = \text{tr } (\nabla F_1') A' F_2' + \text{tr } (\nabla F_2) AF_1.$$  

**Case 3.** Now $dF_1 = P(dX)Q'$ and $dF_2 = R(dX)T'$. Then

$$2\nabla F_1 = P' Q' + (\text{tr } P') Q'$$

$$2\nabla F_2 = R' T' + (\text{tr } R') T'$$

$$2\nabla F_1' = Q' P' + (\text{tr } Q') P'$$

by the Theorem in Neudecker (2000b).

Further

$$dF_2 AF_1 = (dF_2) AF_1 + F_2 AdF_1,$$

$$= R(dX)T' AF_1 + F_2 AP(dX)Q',$$

which implies

$$2\nabla F_2 AF_1 = R'T' AF_1 + P' A' F_2' Q' +$$

$$+ (\text{tr } R') T' AF_1 + (\text{tr } F_2 AP) Q'$$

$$= 2 (\nabla F_2) AF_1 + P' A' F_2' Q' + (\text{tr } F_2 AP) Q'$$

and hence

$$2 \text{tr } \nabla F_2 AF_1 = 2 \text{tr } (\nabla F_2) AF_1 + \text{tr } [Q' P' + (\text{tr } Q) P'] A' F_2'$$

$$= 2 \text{tr } (\nabla F_2) AF_1 + 2 \text{tr } (\nabla F_1') A' F_2'.$$

□

**Note:** For an introduction to the scalar Haffian see Neudecker (2000a).

### 5 References


Hola