SORT 28 (2) July-December 2004, 191-200

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 $T := S - nI_m$. Under certain conditions $T_\alpha := T + \alpha(\operatorname{tr} S)^{-1} I_m$ dominates T for a suitable choice of α, as was shown by Leung, who used the loss function

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

He extended work by Perlman & Rasmussen (1975), Saxena & Alam (1982), Chow (1987) and Leung & Muirhead (1987).

* *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, *viz* $L[(M'M)^{-1}, R] := \left\{(M'M)^{-1} R - I_m\right\}' \left\{(M'M)^{-1} R - I_m\right\}$ 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 $E(\operatorname{tr} S)^{-1}S$, where E is the expectation operator. A lemma on the matrix Haffian $\nabla\varphi F$, where φ 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 $\operatorname{tr}\nabla F_2 A F_1$, 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:

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

where $S \sim W_m\left(n, \Sigma, \Sigma^{-1}M'M\right)$, 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, Σ, F and ∇ are square of dimension m, whereas M has dimension $n \times m$. It is assumed that M has full column rank. Further ∇F is the matrix Haffian as denoted by Neudecker (2000*b*). Inspired by Haff (1981), who did it for the central Wishart distribution, we shall establish a matrix version of (1).

Theorem 1

$$EF_1\Sigma^{-1}F_2 = 2EF_1\nabla F_2 + 2\left(EF_2'\nabla F_1'\right)' + (n - m - 1)EF_1S^{-1}F_2 + E_1F_1\Sigma^{-1}M'MS^{-1}F_2, \qquad (2)$$

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:

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

with constant A. For a proof see Lemma 5.

Taking $A = e_j e_i'$ we get

$$E\,\mathrm{tr}\Sigma^{-1}F_2 e_j e_i' F_1 = 2E\,\mathrm{tr}\,(\nabla F_2)\,e_j e_i' F_1 + 2E\,\mathrm{tr}\left(\nabla F_1'\right) e_i e_j' F_2' + \\ + (n-m-1)\mathrm{tr}\,ES^{-1}F_2 e_j e_i' F_1 + E_1\,\mathrm{tr}\,\Sigma^{-1}M'MS^{-1}F_2 e_j e_i' F_1$$

or equivalently

$$\left(EF_1\Sigma^{-1}F_2\right)_{ij} = 2\,(EF_1\nabla F_2)_{ij} + 2\left(EF_2'\nabla F_1'\right)_{ji} + \\ + (n-m-1)\left(EF_1S^{-1}F_2\right)_{ij} + \left(E_1F_1\Sigma^{-1}M'MS^{-1}F_2\right)_{ij}.$$

□

Note: It was assumed that (1) holds for all $F = F_2 e_j e_i' F_1$, 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

$$E\,(tr\,S)^{-1}(M'M)^{-1}S\,(M'M)^{-1} < nE\,(trS)^{-1}(M'M)^{-2} - \\ -2(n-4)E\,(tr\,S)^{-2}(M'M)^{-2} + E_1\,(tr\,S)^{-1}(M'M)^{-1} - \\ -2E_1\,(tr\,S)^{-2}(M'M)^{-1},$$

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 $F_1 = (\mathrm{tr}\,S)^{-1}(M'M)^{-1}$ and $F_2 = S\,(M'M)^{-1}$. By Theorem 1 (with $\Sigma = I_m$):

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

$$= nE\,(\operatorname{tr} S)^{-1}\,(M'M)^{-2} - 2E\,(\operatorname{tr} S)^{-2}\,(M'M)^{-1}\,S\,(M'M)^{-1} +$$
$$+E_1\,(\operatorname{tr} S)^{-1}\,(M'M)^{-1} \qquad (i)$$

Further

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

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

$$E\,(\operatorname{tr} S)^{-2}\,(M'M)^{-1}\,S\,(M'M)^{-1} = (m+1)E\,(\operatorname{tr} S)^{-2}\,(M'M)^{-2} -$$
$$-4E\,(\operatorname{tr} S)^{-3}\,(M'M)^{-1}\,S\,(M'M)^{-1} + (n-m-1)E\,(\operatorname{tr} S)^{-2}\,(M'M)^{-2} +$$
$$+E_1\,(\operatorname{tr} S)^{-2}\,(M'M)^{-1} = nE\,(\operatorname{tr} S)^{-2}\,(M'M)^{-2} -$$
$$-4E\,(\operatorname{tr} S)^{-3}\,(M'M)^{-1}\,S\,(M'M)^{-1} + E_1\,(\operatorname{tr} S)^{-2}\,(M'M)^{-1}\,.$$

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

$$(n-4)E\,(\operatorname{tr} S)^{-2}\,(M'M)^{-2} + E_1\,(\operatorname{tr} S)^{-2}\,(M'M)^{-1} < E\,(\operatorname{tr} S)^{-2}\,(M'M)^{-1}\,S\,(M'M)^{-1}\,.$$

Insertion in (i) finally yields

$$E\,(\operatorname{tr} S)^{-1}\,(M'M)^{-1}\,S\,(M'M)^{-1} < nE\,(\operatorname{tr} S)^{-1}\,(M'M)^{-2} - 2(n-4)E\,(\operatorname{tr} S)^{-2}\,(M'M)^{-2}$$
$$-2E_1\,(\operatorname{tr} S)^{-2}\,(M'M)^{-1} + E_1\,(\operatorname{tr} S)^{-1}\,(M'M)^{-1}\,.$$

□

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.
3. Taking traces in Lemma 2 yields Leung's Lemma 3.1.

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 $0 < \alpha \leqslant 4(n-4)$, *where*

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

$$T := S - nI_m \quad \textit{and} \quad T_\alpha := T + \alpha\,(tr\,S)^{-1} I_m.$$

Proof.

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

Its expected value is

$$2n\alpha E\,(\mathrm{tr}\,S)^{-1}(M'M)^{-2} - \alpha^2 E\,(\mathrm{tr}\,S)^{-2}(M'M)^{-2} - \alpha E\,(\mathrm{tr}\,S)^{-1}S\,(M'M)^{-2} -$$
$$-\alpha E\,(\mathrm{tr}\,S)^{-1}(M'M)^{-2}S + 2\alpha E\,(\mathrm{tr}\,S)^{-1}(M'M)^{-1} >$$
$$> 2\alpha E\,(\mathrm{tr}\,S)^{-1}(M'M)^{-1}S\,(M'M)^{-1} +$$
$$+4\alpha(n-4)E\,(\mathrm{tr}\,S)^{-2}(M'M)^{-2} - 2\alpha E_1(\mathrm{tr}\,S)^{-1}(M'M)^{-1} + 4\alpha E_1(\mathrm{tr}\,S)^{-2}(M'M)^{-1} -$$
$$-\alpha^2 E\,(\mathrm{tr}\,S)^{-2}(M'M)^{-2} - \alpha E\,(\mathrm{tr}\,S)^{-1}S\,(M'M)^{-2} - \alpha E\,(\mathrm{tr}\,S)^{-1}(M'M)^{-2}S +$$
$$+2\alpha E\,(\mathrm{tr}\,S)^{-1}(M'M)^{-1}$$
$$= 2\alpha E\,(\mathrm{tr}\,S)^{-1}(M'M)^{-1}S\,(M'M)^{-1} + \alpha\,[4(n-4) - \alpha]\,E\,(\mathrm{tr}\,S)^{-2}(M'M)^{-2} +$$
$$+2\alpha\left\{E\,(\mathrm{tr}\,S)^{-1}(M'M)^{-1} - E_1(\mathrm{tr}\,S)^{-1}(M'M)^{-1}\right\} +$$
$$+4\alpha E_1(\mathrm{tr}\,S)^{-2}(M'M)^{-1} - \alpha E\,(\mathrm{tr}\,S)^{-1}S\,(M'M)^{-2} - \alpha E\,(\mathrm{tr}\,S)^{-1}(M'M)^{-2}S,$$

by Lemma 2.

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

$$\mu\,(n\,I_m + M'M) - 2\mu^2(n\,I_m + 2M'M) +$$
$$+2\mu^3(mn + 2\,\mathrm{tr}\,M'M)(n\,I_m + M'M),$$

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

Insertion yields

$$2\alpha E\,(\mathrm{tr}\,S)^{-1}\,(M'M)^{-1}\,S\,(M'M)^{-1} - \alpha E\,(\mathrm{tr}\,S)^{-1}\,S\,(M'M)^{-2} - \\ -\alpha E\,(\mathrm{tr}\,S)^{-1}\,(M'M)^{-2}\,S = O + o(n^{-1}).$$

Hence to the order of approximation

$$EL[\,(M'M)^{-1}, T\,] - EL[\,(M'M)^{-1}, T_\alpha] > \alpha\,[4(n-4) - \alpha]\,E\,(\mathrm{tr}\,S)^{-2}\,(M'M)^{-2} + \\ +2\alpha\left[E\,(\mathrm{tr}\,S)^{-1}\,(M'M)^{-1} - E_1\,(\mathrm{tr}\,S)^{-1}\,(M'M)^{-1}\right] + 4\alpha E_1\,(\mathrm{tr}\,S)^{-2}\,(M'M)^{-1} > O,$$

as $E\,(\mathrm{tr}\,S)^{-1} \geqslant E_1\,(\mathrm{tr}\,S)^{-1}$.

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

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 φ is a scalar function of the symmetric matrix variable X and F is a matrix function thereof. Further

$$\frac{\partial\varphi}{\partial X} := \tfrac{1}{2}\sum_{ij}\frac{\partial\varphi}{\partial x_{ij}}\left(E_{ij} + E_{ji}\right), \quad \textit{where } E_{ij} := e_i\,e_j'$$

Proof.

$$(\nabla\varphi F)_{ik} = \sum_j d_{ij}\,(\varphi F)_{jk} = \sum_j d_{ij}\varphi f_{jk} = \tfrac{1}{2}\sum_j\left(1 + \delta_{ij}\right)\frac{\partial\varphi f_{jk}}{\partial x_{ij}} =$$

$$= \frac{\partial\varphi f_{ik}}{\partial x_{ii}} + \tfrac{1}{2}\sum_{j\neq i}\frac{\partial\varphi f_{jk}}{\partial x_{ij}} = \varphi\left(\frac{\partial f_{ik}}{\partial x_{ii}} + \tfrac{1}{2}\sum_{j\neq i}\frac{\partial f_{jk}}{\partial x_{ij}}\right) + \left(\frac{\partial\varphi}{\partial x_{ii}}\right)f_{ik} +$$

$$+ \tfrac{1}{2}\sum_{j\neq i}\frac{\partial\varphi}{\partial x_{ij}}\,f_{jk} = \varphi\,(\nabla F)_{ik} + \left(\frac{\partial\varphi}{\partial X}\right)_{i.}F_{.k}, \quad \text{hence}$$

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

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

For more details see Neudecker (2000*b*). □

Appendix 2: a lemma on the scalar Haffian $\mathrm{tr}\,\nabla F_2 A F_1$

Lemma 5

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

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 $dF = \sum_l P_l(dX)Q_l'$ 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

$$\begin{aligned} d\,(F_2 A F_1) &= (dF_2)\,A F_1 + F_2 A dF_1 \\ &= P(dX)Q' A F_1 + (d\varphi) F_2 A C \end{aligned}$$

which implies

$$\begin{aligned} \nabla F_2 A F_1 &= \tfrac{1}{2} P' Q' A F_1 + \tfrac{1}{2}(\mathrm{tr}\,P) Q' A F_1 + \frac{\partial\varphi}{\partial X} F_2 A C, \\ \mathrm{tr}\,\nabla F_2 A F_1 &= \tfrac{1}{2}\,\mathrm{tr}\,P' Q' A F_1 + \tfrac{1}{2}(\mathrm{tr}\,P)\,\mathrm{tr}\,Q' A F_1 + \mathrm{tr}\frac{\partial\varphi}{\partial X} F_2 A C; \\ (\nabla F_2)\,A F_1 &= \tfrac{1}{2} P' Q' A F_1 + \tfrac{1}{2}(\mathrm{tr}\,P) Q' A F_1, \\ \mathrm{tr}\,(\nabla F_2)\,A F_1 &= \tfrac{1}{2}\,\mathrm{tr}\,P' Q' A F_1 + \tfrac{1}{2}(\mathrm{tr}\,P)\,\mathrm{tr}\,Q' A F_1; \\ \left(\nabla F_1'\right) A' F_2' &= \frac{\partial\varphi}{\partial X} C' A' F_2', \\ \mathrm{tr}\left(\nabla F_1'\right) A' F_2' &= \mathrm{tr}\frac{\partial\varphi}{\partial X} C' A' F_2' = \mathrm{tr}\frac{\partial\varphi}{\partial X} F_2 A C. \end{aligned}$$

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

$$\mathrm{tr}\,\nabla F_1' A' F_2' = \mathrm{tr}\,(\nabla F_1')\,A' F_2' + \mathrm{tr}\,(\nabla F_2)\,A F_1.$$

Using $\mathrm{tr}\,\nabla F' = \mathrm{tr}\,\nabla F$ we get

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

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

$$\begin{aligned} 2\nabla F_1 &= P'Q' + (\mathrm{tr}\,P)\,Q' \\ 2\nabla F_2 &= R'T' + (\mathrm{tr}\,R)\,T' \\ 2\nabla F_1' &= Q'P' + (\mathrm{tr}\,Q)\,P' \end{aligned}$$

by the Theorem in Neudecker (2000*b*).
Further

$$\begin{aligned} dF_2AF_1 &= (dF_2)AF_1 + F_2AdF_1, \\ &= R(dX)T'AF_1 + F_2AP(dX)Q', \end{aligned}$$

which implies

$$\begin{aligned} 2\nabla F_2AF_1 &= R'T'AF_1 + P'A'F_2'Q' + \\ &\quad + (\mathrm{tr}\,R)\,T'AF_1 + (\mathrm{tr}\,F_2AP)\,Q' \\ &= 2\,(\nabla F_2)\,AF_1 + P'A'F_2'Q' + (\mathrm{tr}\,F_2AP)\,Q' \end{aligned}$$

and hence

$$\begin{aligned} 2\,\mathrm{tr}\,\nabla F_2AF_1 &= 2\,\mathrm{tr}\,(\nabla F_2)\,AF_1 + \mathrm{tr}\,[Q'P' + (\mathrm{tr}\,Q)\,P']\,A'F_2' \\ &= 2\,\mathrm{tr}\,(\nabla F_2)\,AF_1 + 2\,\mathrm{tr}\left(\nabla F_1'\right)A'F_2'. \end{aligned}$$

□

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

5 References

Chow, M. S. (1987). A complete class theorem for estimating a noncentrality parameter. *Ann. Statist.*, 15, 800-4.

Haff, L. R. (1981). Further identities for the Wishart distribution with applications in regression. *Canadian J. Statist.*, 215-24.

Leung, P. L. (1994). An identity for the noncentral Wishart distribution with application. *J. Multivariate Anal.*, 48, 107-14.

Leung, P. L. & Muirhead, R. J. (1987). Estimation of parameter matrices and eigenvalues in MANOVA and canonical correlation analysis. *Ann. Statist.*, 15, 1651-66.

Neudecker, H. (2000*a*). A note on the scalar Haffian. *Qüestiió*, 24(2), 243-9.

Neudecker, H. (2000*b*). A note on the matrix Haffian. *Qüestiió*, 24(3), 419-24.

Perlman, M. D. & Rasmussen, U. A. (1975). Some remarks on estimating a noncentrality parameter. *Comm. Statist.*, 4, 455-68.

Saxena, K. M. L. & Alam, K. (1982). Estimation of the noncentrality parameter of a chi-squared distribution. *Ann. Statist.*, 10, 1012-6.