Computational Statistics (1994) 9:25-39 Computational Statistics

Density of the Least Squares Estimator in the Multivariate Linear Model with Arbitrarily Normal Variables

B. B. van der Genugten
Department of Econometrics, Tilburg University, Tilburg, The Netherlands

Keywords: multivariate linear model, least squares estimator, numerical evaluation of probability densities, computer package Mathematica.

Abstract

An expression is obtained for the density of the least squares estimator $\hat{B} = (X_2'X_2)^{-1}X_2'X_1$ in the multivariate linear model $X_1 = X_2B + E$ for the case that $X = [X_1\ X_2]$ has a normal distribution.

This expression is a multi-dimensional integral with a dimension independent of the number of rows of X. The univariate case with two independent variables is worked out in detail. Finally, some experiences with the computer package Mathematica are described.

AMS classification: 60E05, 62J05.

1 Introduction

Consider the multivariate linear model

$$X_1 = X_2B + E \tag{1.1}$$

with n observations, dependent variables $X_1 \in \mathbf{R}^{n\times p}$ and explanatory variables $X_2 \in \mathbf{R}^{n\times k}$ $(n \geq k)$. The least squares estimator for B is given by

$$\hat{B} = (X_2'X_2)^{-1}X_2'X_1. \tag{1.2}$$

We are interested in the numerical evaluation of the distribution of $\hat{B}$ in non-standard cases with all kind of correlation structures in (X_2, E). We restrict ourselves to normal variables. Then the most general assumption is that $X = [X_1\ X_2] \in \mathbf{R}^{n\times m}$ (with $m = p + k$) has a normal distribution, say

$$\text{vec}(X) \sim N_{nm}(\mu, \Omega), \quad \Omega > 0. \tag{1.3}$$

This covers many types of lag structures including multivariate ARMA-models and non-stationary generalizations. The assumption $\Omega = \{\omega_{ij}\} > 0$ is made for sake of convenience. Formally, this excludes any deterministic (non-constant) explanatory variables. However, such cases can easily be treated using limit arguments.

The difficulty in deriving a suitable expression for the distribution of $\hat{B}$ is the non-linearity in X_2. Only very few cases are analytically tractable. We give two examples, also used for future reference and for testing computer programs.

Example 1 ($n = k = p = 1$; $\mu = 0$; general Ω). It is a simple exercise to show that $\hat{B}$ follows a Cauchy-distribution with density g, given by

$$g(B) = \pi^{-1}|\Omega|^{\frac{1}{2}}\omega_{22}\{1 + (\omega_{21} - \omega_{22}B)^2/|\Omega|\}^{-1}. \quad \square \tag{1.4}$$

Example 2 (general n, k, p; $\mu = 0$; $\Omega = I_{nm}$). This is exactly the multivariate standard normal linear model in which all elements of X_2 and E are mutually independent and standard normal. It is well-known that $\hat{B} \sim t_{n,k,p}$, the central standardized matrix t-distribution.

We recall some facts about $t_{n,k,p}$ (see for details Johnson and Kotz (1972), p. 151).

By definition $T = J^{-1}C \sim t_{n,k,p}$ if $C \in \mathbf{R}^{k\times p}$ and $J \in \mathbf{R}^{k\times k}$ are independent with C standard normal and $JJ' = \Sigma_1^n u_j u_j' \sim \chi^2_{n,m} = \chi^2_{n,m}(I_m)$, the central standard Wishart distribution with n degrees of freedom and dimension m (i.e. $u_j \sim N_m(0,\Sigma)$ with $\Sigma = I_m$ and independent). With this definition it is easy to see that $\mathcal{L}(\hat{B}) = t_{n,k,p}$. Since $\operatorname{vec}(\hat{B}) = \operatorname{vec}((X_2'X_2)^{-1}X_2'X_1) = (I_p \otimes (X_2'X_2)^{-1}X_2')\operatorname{vec}(X_1)$ we have $\mathcal{L}(\operatorname{vec}(\hat{B})|X_2) = N_{kp}(0, I_p \otimes (X_2'X_2)^{-1})$. So $\mathcal{L}(\operatorname{vec}((X_2'X_2)^{1/2}\hat{B}|X_2) = \mathcal{L}(I_p \otimes (X_2'X_2)^{1/2}\operatorname{vec}(\hat{B})|X_2) = N_{kp}(0, I_{kp})$. Hence, $J = (X_2'X_2)^{1/2}$ and $C = (X_2'X_2)^{1/2}\hat{B}$ are independent with C standard normal and $JJ' = X_2'X_2 \sim \chi^2_{n,k}$. Then, by definition, $\hat{B} = J^{-1}C \sim t_{n,k,p}$.

The density g is given by

$$g(B) = \pi^{-kp/2}\frac{\Gamma_k((n+p)/2)}{\Gamma_k(n/2)}|I_k + BB'|^{-(n+p)/2}. \tag{1.5}$$

where Γ_k stands for the multivariate gamma function

$$\Gamma_k(n/2) = \pi^{k(k-1)/4}\prod_{j=1}^{k}\Gamma((n-j+1)/2). \quad \square \tag{1.6}$$

The generalization of the examples for the non-central case $\mu \neq 0$ leads already to complicated expressions. In example 2 we are led to certain types of non-central t-distributions. Expressions for the density are usually given with the help of Zonal-polynomials.

For the general problem we follow quite another approach in deriving an expression for the density g of $\hat{B}$. For $k = p = 1$ the right hand side of (1.2) is the quotient of two random variables $C_{21} = X_2'X_1$ and $C_{22} = X_2'X_2$ with positive denominator C_{22}. So the density g is given by Cramer's inversion formula based on the characteristic function of (C_{21}, C_{22}). This formula has been generalized for arbitrary k, p by Phillips (1985). Using this result we are able to express the density g as a multiple integral with a fixed dimension $k \times p$ not depending on n:

$$g(B) = \int_{\mathbf{R}^{k\times p}}\varphi(B, U_{21})dU_{21}. \tag{1.7}$$

Section 2 gives the expression for the integrand φ. The densities (1.4) and (1.5) in the examples 1 and 2 are derived from the general result. These examples suggest strongly that a simple expression for φ cannot be obtained for arbitrary μ and Ω.
Section 3 specifies some useful results for computer programs for the univariate case $p = 1$ with $k \leq 2$.
Section 4 gives the proof of the theorem in section 2.
Finally, section 5 describes some experiences with the computer package Mathematica. The main idea is to use its computer algebra facility for obtaining an expression for φ and to use its numerical part for the calculation of the integral, thereby giving the density g.

2 Statement of the result

In the specification of the integrand φ in (1.7) the concept of the determinant-differential operator is used. Such an operator D works on functions $\psi : \mathbf{R}^{k\times k} \to \mathbf{R}$ and is defined by $D\psi(U) = \det \partial/\partial U = \{\partial/\partial u_{rs}\}$, $r, s = 1, \ldots, k$. So $D\psi(U)$ is a linear combination of $k!$ partial derivatives of ψ of order k. As usual higher orders are defined by $D^p = D(D^{p-1})$,$p = 2, 3, \ldots$. The particular value at $U = U_0$ is denoted by $D^p\psi(U_0)$. For $k = 1$ $D\psi(U)$ reduces to $(d/dU)\psi$. For $k = 2$ we get

$$D\psi(U) = \left\{ \det \begin{pmatrix} \partial/\partial u_{11} & \partial/\partial u_{12} \\ \partial/\partial u_{21} & \partial/\partial u_{22} \end{pmatrix} \right\} \psi = \left(\frac{\partial^2}{\partial u_{11}\partial u_{22}} - \frac{\partial^2}{\partial u_{12}\partial u_{21}} \right) \psi.$$

The operator D appears in a natural way in the differentiation of characteristic functions. We abbreviate 'exp tr' to 'etr'. For $X, U \in \mathbf{R}^{k\times k}$ we have

$$\begin{aligned} \int |X| f(X) \operatorname{etr}(iU'X) dX &= \int |X| f(X) \exp\{i \sum_r \sum_s u_{rs} x_{rs}\} dX = \\ &= i^{-k} \sum_{(i_1,\ldots,i_k)} \int f(X) (\partial^k / \partial u_{i_1} \ldots \partial u_{i_k}) \operatorname{etr}(iU'X) dX \\ &= i^{-k} D \int f(X) \operatorname{etr}(iU'X) dx \end{aligned}$$

and more general for $p = 1, 2, \ldots$:

$$\int_{\mathbf{R}^{k\times k}} |X|^p f(X)\,\mathrm{etr}(iU'X)dX = i^{-kp}D^p \int_{\mathbf{R}^{k\times k}} f(X)\,\mathrm{etr}(iU'X)dX. \qquad (2.1)$$

Theorem

Under the condition that the integral in (1.7) is absolutely convergent we have:

$$\varphi(B, U_{21}) = \left(\frac{1}{2\pi i}\right)^{kp} D_{22}^p \psi(U_{21}, -(BU_{21}' + U_{21}B')/2) \qquad (2.2)$$

with determinant-differential operator

$$D_{22} = \det(\partial/\partial U_{22}), \quad U_{22} \in \mathbf{R}^{k\times k} \qquad (2.3)$$

and with

$$\psi(U_{21}, U_{22}) = |\Omega|^{1/2} \exp\left(-\tfrac{1}{2}\mu'\Omega^{-1}\mu\right) \cdot |\Phi|^{-1/2} \exp\left(\tfrac{1}{2}\mu'\Omega^{-1}\Phi^{-1}\Omega^{-1}\mu\right) \quad (2.4)$$

$$\Phi = \Phi(U_{21}, U_{22}) = \Omega^{-1} - i(V + V') \qquad (2.5)$$

$$V = V(U_{21}, U_{22}) = U \otimes I_n \qquad (2.6)$$

$$U = \begin{bmatrix} 0 & 0 \\ U_{21} & U_{22} \end{bmatrix} \in \mathbf{R}^{m\times m}. \qquad (2.7)$$

Example 1 (continued)

In this case $m = 2$ and so with (2.6), (2.7)

$$V = U = \begin{bmatrix} 0 & 0 \\ U_{21} & U_{22} \end{bmatrix} \in \mathbf{R}^{2\times 2}.$$

This leads to

$$|I_{nm} - i\Omega(V + V')| = 1 + |\Omega|U_{21}^2 - 2i(\omega_{21}U_{21} + \omega_{22}U_{22}).$$

Hence, with (2.4), (2.5)

$$\begin{aligned}\psi(U_{21},U_{22}) &= |\Omega|\{1+|\Omega|U_{21}^2-2i(\omega_{21}U_{21}+\omega_{22}U_{22})\}^{-\frac{1}{2}}\\ \partial\psi/\partial U_{22} &= i\omega_{22}|\Omega|\{1+|\Omega|U_{21}^2-2i(\omega_{21}U_{21}+\omega_{22}U_{22})\}^{-\frac{3}{2}}\end{aligned}$$

and so with (2.2)

$$\begin{aligned}\varphi(B,U_{21}) &= (2\pi i)^{-1}D_{22}\psi(U_{21},-BU_{21}) =\\ &= (2\pi)^{-1}\omega_{22}|\Omega|\{1+|\Omega|U_{21}^2+2i(\omega_{22}B-\omega_{21})U_{21}\}^{-\frac{3}{2}}.\end{aligned}$$

Substitution into (1.7) leads with the transformation

$$W=\{|\Omega|+(\omega_{22}B-\omega_{21})^2\}^{-\frac{1}{2}}\{|\Omega|U_{21}+i(\omega_{22}B-\omega_{21})\}$$

to

$$\begin{aligned}g(B) &= \int_{-\infty}^{\infty}\varphi(B,U_{21})dU_{21}\\ &= (2\pi)^{-1}\omega_{22}|\Omega|^{\frac{1}{2}}\{1+(\omega_{22}B-\omega_{21})^2/|\Omega|\}^{-1}\int_{-\infty}^{\infty}(1+W^2)^{-\frac{3}{2}}dW\end{aligned}$$

and this equals (1.4). Note the use of Cauchy's theorem in the integral transformation. □

Example 2 (continued)
Substitution of $\Omega = I_{nm} = I_m \otimes I_n$ into (2.5) gives

$$\Phi=\left(\begin{bmatrix} I_p & 0\\ 0 & I_k\end{bmatrix}-i\begin{bmatrix}0 & U_{21}'\\ U_{21} & U_{22}+U_{22}'\end{bmatrix}\right)\otimes I_n.$$

So, with (2.4)

$$\psi(U_{21},U_{22})=|\Phi|^{-1/2}=|I_k+U_{21}U_{21}'-i(U_{22}+U_{22}')|^{-n/2}. \qquad (2.8)$$

It is not easy to calculate $D_{22}^p\psi(U_{21},U_{22})$ from (2.8). Using some known results about central Wishart-distributions we can show that

$$D_{22}^p\psi(U_{21},U_{22}) = \\ (2i)^{kp}\frac{\Gamma_k(n/2+p)}{\Gamma_k(n/2)}\cdot|I_k+U_{21}U_{21}'-i(U_{22}+U_{22}')|^{-n/2-p}. \tag{2.9}$$

The derivation of (2.9) from (2.8) can be found at the end of this example.

Substitution of (2.9) into (2.2) leads with (1.7) to

$$g(B) = \\ \left(\frac{1}{\pi}\right)^{kp}\frac{\Gamma_k(n/2+p)}{\Gamma_k(n/2)}\int_{\mathbf{R}^{k\times p}}|I_k+U_{21}U_{21}'+i(BU_{21}'+U_{21}B')|^{-n/2-p}dU_{21}.$$

The integral can be reduced in the same way as in example 1 (see also in Phillips (1985), p. 160). Transform with $W=(I_k+BB')^{-1/2}(U_{21}+iB)$. Then

$$I_k+U_{21}U_{21}'+i(BU_{21}'+U_{21}B')=(I_k+BB')^{1/2}(I_k+WW')(I_k+BB')^{1/2}.$$

The transformation has Jacobian $|I_k+BB'|^{p/2}$. The domain of integration can be taken to be $\mathbf{R}^{k\times p}$ as before using standard arguments in contour-integration. This leads to

$$g(B)=\left(\frac{1}{\pi}\right)^{kp}\frac{\Gamma_k(n/2+p)}{\Gamma_k(n/2)}|I_k+BB'|^{-(n+p)/2}\int_{\mathbf{R}^{k\times p}}|I_k+WW'|^{-n/2-p}dW.$$

Since (e.g. see Dickey (1967), p. 512)

$$\int_{\mathbf{R}^{k\times p}}|I_k+WW'|^{-n/2-p}dW=\pi^{kp/2}\frac{\Gamma_k((n+p)/2)}{\Gamma_k(n/2+p)}$$

we get (1.5). This is the density of the central standardized matrix t distribution $t_{n,k,p}$.

We conclude this example with the derivation of (2.9). Let $\hat{W}\sim\chi^2_{n,k}(\Sigma)$. Then the characteristic function $\varphi(U)=E\{\mathrm{etr}(iU'\hat{W}\}$, $U\in\mathbf{R}^{k\times k}$ is given by

$$\varphi(U)=\{|\Sigma^{-1}|/|\Sigma^{-1}-i(U+U')|\}^{n/2}$$

and the density $f(W)$ for $W > 0$ by

$$f(W) = K(n, \Sigma^{-1})|W|^{(n-k-1)/2} \operatorname{etr}(-\tfrac{1}{2}\Sigma^{-1}W),$$

where

$$K(n, \Sigma^{-1}) = \{2^{nk/2}\Gamma_k(n/2)\}^{-1}|\Sigma|^{n/2}$$

(see Johnson and Kotz (1972), p. 162-163).

It follows from (2.5) that $|\Sigma|^{n/2}\psi(U_{21}, U_{22})$ is the characteristic function of $\chi^2_{n,k}(\Sigma)$ at U_{22} if $\Sigma = (I_k + U_{21}U'_{21})^{-1}$. Hence,

$$|\Sigma|^{n/2}\psi(U_{21}, U_{22}) = \int K(n, \Sigma^{-1})|W|^{(n-k-1)/2} \operatorname{etr}\left(-\tfrac{1}{2}\Sigma^{-1}W + iU'_{22}W\right) dW$$

and so, using (2.1),

$$\begin{aligned} &|\Sigma|^{n/2}D^p_{22}\psi(U_{21}, U_{22}) = \\ &= i^{kp}\int K(n, \Sigma^{-1})|W|^{(n-k-1)/2+p} \operatorname{etr}\left(-\tfrac{1}{2}\Sigma^{-1}W + iU'_{22}W\right) dW = \\ &= i^{kp}K(n, \Sigma^{-1})/K(n+2p, \Sigma^{-1})|\Sigma|^{n/2+p}|\Sigma^{-1} - i(U_{22} + U'_{22})|^{-n/2-p} \end{aligned}$$

or

$$D^p_{22}\psi(U_{21}, U_{22}) = (2i)^{kp}\frac{\Gamma_k(n/2+p)}{\Gamma_k(n/2)}\left|\Sigma^{-1} - i(U_{22} + U'_{22})\right|^{-n/2-p}$$

This proves (2.9). □

The foregoing examples make clear that there is no hope for getting simple expressions for g if $\mu \neq 0$ or if Ω has no particular structure. The keypoints are the differentiation in $D^p_{22}\psi$ and the numerical integration of (1.7).

Some useful formula are given in the next section 3 for the differentiation in univariate case $p = 1$ and $k \leq 2$ explanatory variables. Then only the numeric calculation of (1.7) remains.

For general k, p and small n it is wise to leave the differentiation process to an (alpha-numeric) computer package. We refer to section 5.

3 The univariate case $p = 1$ with $k \leq 2$

For these simple cases explicit expressions for $D_{22}^{p}\psi = D_{22}\psi$ are managerable. Write $U_{22} = \{u_{rs}\}$, $r, s = 1, 2$. We need explicit expressions for $\partial\psi/\partial u_{rs}$ and $\partial^2\psi/\partial u_{rs}\partial u_{pq}$.

Lemma

Let

$$J_{rs} = \tfrac{1}{2}\partial/\partial u_{rs}(V + V'), \tag{3.1}$$

then

$$\partial\psi/\partial u_{rs} = i a_{rs}\psi \tag{3.2}$$

with

$$a_{rs} = \mathrm{tr}(\Phi^{-1}J_{rs}) + 2\mu'\Omega^{-1}\Phi^{-1}J_{rs}\Phi^{-1}\Omega^{-1}\mu. \tag{3.3}$$

Furthermore, let

$$\Phi_{rspq} = J_{rs}\Phi^{-1}J_{pq} + J_{pq}\Phi^{-1}J_{rs}, \tag{3.4}$$

then

$$\partial^2\psi/\partial u_{rs}\partial u_{pq} = -(a_{rspq} + a_{rs}a_{pq})\psi \tag{3.5}$$

with

$$a_{rspq} = \mathrm{tr}(\Phi^{-1}\Phi_{rspq}) + 4\mu'\Omega^{-1}\Phi^{-1}\Phi_{rspq}\Phi^{-1}\Omega^{-1}\mu. \tag{3.6}$$

Proof
We have

$$\begin{aligned}\partial|\Phi|^{-1/2}/\partial u_{rs} &= -\tfrac{1}{2}|\Phi|^{-3/2}\partial|\Phi|/\partial u_{rs} = \\ &= -\tfrac{1}{2}|\Phi|^{-3/2}\cdot|\Phi|\,\mathrm{tr}\{\Phi^{-1}\partial\Phi/\partial u_{rs}\} = i|\Phi|^{-1/2}\,\mathrm{tr}(\Phi^{-1}J_{rs})\end{aligned}$$

$$\partial\Phi^{-1}/\partial u_{rs} = -\Phi^{-1}(\partial\Phi/\partial u_{rs})\Phi^{-1} = 2i\Phi^{-1}J_{rs}\Phi^{-1}.$$

Writing $c = |\Omega|^{1/2}\exp(-\frac{1}{2}\mu'\Omega^{-1}\mu)$ this gives

$$\begin{aligned}\partial\psi/\partial u_{rs} &= ci\{|\Phi|^{-1/2}\,\mathrm{tr}(\Phi^{-1}J_{rs})\cdot\exp(\mu'\Omega^{-1}\Phi^{-1}\Omega^{-1}\mu)+ \\ &+ |\Phi|^{-1/2}(2\mu'\Omega^{-1}\Phi^{-1}J_{rs}\Phi^{-1}\Omega^{-1}\mu)\cdot\exp(\frac{1}{2}\mu'\Omega^{-1}\Phi^{-1}\Omega^{-1}\mu)\} = ia_{rs}\psi,\end{aligned}$$

proving (3.3). Furthermore,

$$\begin{aligned}\partial\,\mathrm{tr}(\Phi^{-1}J_{rs})/\partial u_{pq} &= \mathrm{tr}\{(\partial\Phi^{-1}/\partial u_{pq})J_{rs}\} = \\ &= 2i\,\mathrm{tr}(\Phi^{-1}J_{pq}\Phi^{-1}J_{rs}) = i\,\mathrm{tr}(\Phi^{-1}\Phi_{rspq})\end{aligned}$$

$$\begin{aligned}\partial\Phi^{-1}J_{rs}\Phi^{-1}/\partial u_{pq} &= 2i\Phi^{-1}J_{pq}\Phi^{-1}J_{rs}\Phi^{-1} + 2i\Phi^{-1}J_{rs}\Phi^{-1}J_{pq}\Phi^{-1} = \\ &= 2i\Phi^{-1}\Phi_{rspq}\Phi^{-1}\end{aligned}$$

Therefore, using (3.3),

$$\begin{aligned}\partial^2\psi/\partial u_{rs}\partial u_{pq} &= i\{(\partial a_{rs}/\partial u_{pq})\psi + a_{rs}(\partial\psi/\partial u_{pq})\} = \\ &= i\{i\,\mathrm{tr}(\Phi^{-1}\Phi_{rspq}) + 2\mu'\Omega^{-1}(2i\Phi^{-1}\Phi_{rspq}\Phi^{-1})\Omega^{-1}\mu) + ia_{rs}a_{pq}\}\psi = \\ &= -(a_{rspq} + a_{rs}a_{pq})\psi,\end{aligned}$$

proving (3.6).

4 Proof of the theorem

The proof is given with some preparatory lemma's.

Lemma 1

Let $X \sim N_n(\mu, \Omega)$, $\Omega > 0$. Set $Y_j = X'A_jX$ with symmetric $A_j \in \mathbf{R}^{n\times n}$, $j = 1, \ldots, s$ and $Y = (Y_1, \ldots, Y_s)'$. Then for $u = (u_1, \ldots, u_s)'$:

$$\begin{aligned}\varphi(u) &= E\{\exp(iu'Y)\} = \\ &= |\Omega|^{-1/2} \exp(-\tfrac{1}{2}\mu'\Omega^{-1}\mu) \cdot |\Phi|^{-1/2} \exp(\tfrac{1}{2}\mu'\Omega^{-1}\Phi^{-1}\Omega^{-1}\mu) \end{aligned} \tag{4.1}$$

with

$$\Phi = \Phi(u) = \Omega^{-1} - 2i \sum_{j=1}^{s} u_j A_j. \tag{4.2}$$

Proof

Follows after a reformulation of Magnus (1986), Lemma 5, p. 102.

Lemma 2

Let $X \in \mathbf{R}^{n\times m}$ with $\text{vec}(X) \sim N_{nm}(\mu, \Omega)$, $\Omega > 0$. Set $Y = X'AX$ with symmetric $A \in \mathbf{R}^{n\times n}$. Then for $U \in \mathbf{R}^{m\times m}$:

$$\begin{aligned}\varphi(U) &= E\{\text{etr}(iU'Y)\} = \\ &= |\Omega|^{-1/2} \exp(-\tfrac{1}{2}\mu'\Omega^{-1}\mu) \cdot |\Phi|^{-1/2} \exp(\tfrac{1}{2}\mu'\Omega^{-1}\Phi^{-1}\Omega^{-1}\mu) \end{aligned} \tag{4.3}$$

with

$$\Phi = \Phi(U) = \Omega^{-1} - i(U + U') \otimes A \tag{4.4}$$

Proof

Write $X = [x_1 \ldots x_m]$, $x = \text{vec}(X) = [x_1' \ldots x_m']'$, $Y = (Y_{ij})$. Then $Y_{ij} = (X'AX)_{ij} = x'A_{ij}x$ with $A_{ij} \in \mathbf{R}^{nm\times nm}$ a matrix with A for block (i,j) and zeros elsewhere. Since for $U = (u_{ij})$:

$$\text{tr}(U'Y) = \Sigma\Sigma u_{ij} Y_{ij} = \Sigma\Sigma u_{ij} x' A_{ij} x = \tfrac{1}{2}\Sigma\Sigma u_{ij} x'(A_{ij} + A_{ji})x$$

it follows from (4.1), (4.2) with $s = m^2$ that $\varphi(u)$ is given by (4.3) with

$$\Phi = \Phi(U) = \Omega^{-1} - i\{\Sigma\Sigma u_{ij}(A_{ij} + A_{ji})\} =$$
$$= \Omega^{-1} - i\{\Sigma\Sigma(u_{ij} + u_{ji})A_{ij}\} = \Omega^{-1} - i(U + U') \otimes A,$$

proving (4.4).

Lemma 3

Let $C_{21} \in \mathbf{R}^{k\times p}$, $C_{22} \in \mathbf{R}^{k\times k}$ with $C_{22} > 0$ a.s. If $E|C_{22}|^p < \infty$ and if $[C_{21}\ C_{22}]$ is absolutely continuous (with respect to C_{22} in the obvious sense) then $\hat{B} = C_{22}^{-1}C_{21}$ has a continuous density g, for $B \in \mathbf{R}^{k\times p}$ given by

$$g(B) = \left(\frac{1}{2\pi i}\right)^{kp} \int_{\mathbf{R}^{k\times p}} D_{22}^p \psi(U_{21}, -(BU_{21}' + U_{21}B')/2)dU_{21} \tag{4.5}$$

provided that the integral is absolutely convergent.

Proof

The lemma is a slight correction of Phillips (1985), theorem, p. 185.

Proof of the theorem

Follows from Lemma 2 and Lemma 3. Take $A = I_n$ and

$$U = \begin{bmatrix} 0 & 0 \\ U_{21} & U_{22} \end{bmatrix}.$$

Then combination of (4.3)-(4.5) leads to (2.4)-(2.7).

5 Some experiences with Mathematica[1]

For a recent overview of the use of computer algebra in probability theory and statistics we refer to Kendall (1993). We describe some experiences with Mathematica for our particular problem. Mathematica is a computer package for doing mathematics. It does not only contain strong numerical procedures but has also the facility of computer algebra. Its strength is the combination of both features.

[1]The programs in Mathematica in this section have been developed by Anton Markink.

We use the computer algebra part of Mathematica for obtaining an expression for φ. The programming is rather straightforward although care should be taking to prevent Mathematica to evaluate algebraic expressions too early. However, there is one exception: the procedure for the determinant-differential operator in (2.2) or (2.3). The listing of the source below gives our implementation.

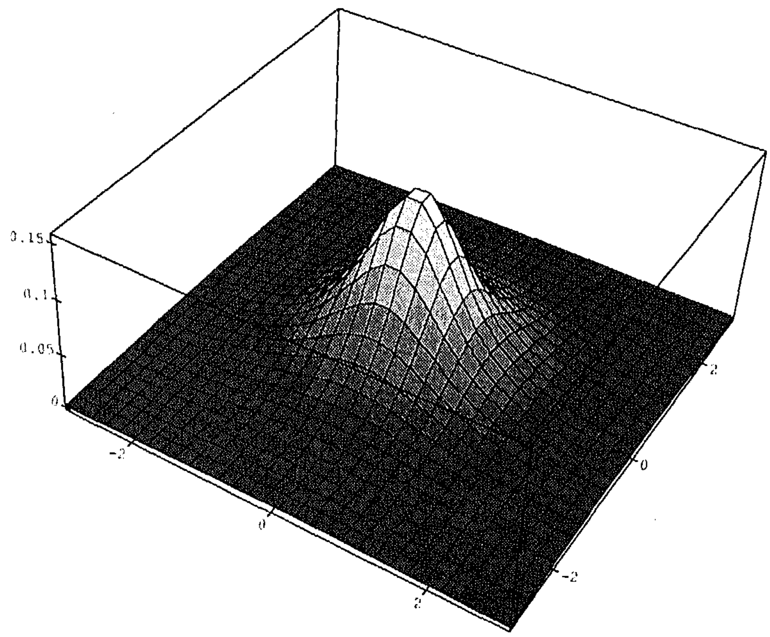
```
d22 [psi ,u22_List ] :=
(* this is comment :
   d22 is determinant-differential operator
   psi=psi(u21,u22)
   u22 : k*k matrix
*)
Block
[ { k=Length[u22] ,s ,i ,i1 ,i2 },
  (* body of procedure *)
  Catch
  [ If [k==1 , (* then *) Return [ D[psi,u22[[1,1]] ] ] ,
        (* else *)
        s=0.0 ;
        (* develop along column 1 *)
        For[i=1 ,i<=k ,i=i+1 ,
            mr=Table[0,{k-1},{k-1}] ;
            (* mr = minor of u22[i,1] *)
            For [i1=1 ,i1<=i-1 ,i1=i1+1 ,
              For [i2=1 ,i2<=k-1 ,i2=i2+1 ,
                 mr[[i1,i2]]=u22[[i1,i2+1]]
                ]] ;
            For [i1=i+1 ,i1<=k ,i1=i1+1 ,
              For [i2=1 ,i2<=k-1 ,i2=i2+1 ,
                 mr[[i1-1,i2]]=u22[[i1,i2+1]]
                ]] ;
            (* recursive call of d22 *)
            s=s + (-1)^(i+1) D[ d22[psi,mr],u22[[i,1]] ]
           ] ;
        Return [s]
       ]
  ]
 ] ;
```

The numerical part of Mathematica can be used to calculate the integral with respect to φ in order to obtain the density g.

Just in order to show that both parts of Mathematica work well together we worked out a particular case

$$p = 1, \quad k = 2, \quad n = 2 \quad \text{and} \quad \mu = 0, \quad \Omega = I_5.$$

The picture below gives the graphical representation of $g(B)$ as a function of B.



For analytically intractable cases ($\mu \neq 0$, general Ω) Mathematica uses a lot of internal computer memory for the computer algebra part. For large n the d22-procedure will not work since the number of terms in the evaluation of the determinant-differential operator will become too large. However, for such n we can rely on asymptotic expressions. For small n this is not possible and then the method provides the solution.

Of course, quite a different method for obtaining the density is simulation. Multivariate density estimating methods become available now. One could use such a method to estimate the density on the base of a large simulation run. It would be interesting to compare the efficiency of this simulation method with that of the numerical method described in this paper.

References

Dickey, J.M. (1967) - Matrixvariate generalizations of the multivariate t-distribution and the inverted multivariate t-distribution - Ann. Math. Stat. 38, 511-518.

Johnson, N.L. and Kotz, S. (1972) - Distributions in Statistics: Continuous Multivariate Distributions - Wiley, New York.

Kendall, W.S. (1993) - Computer algebra in probability and statistics - Statistica Neerlandica 47, 9-25.

Magnus, J.R. (1986) - The exact moments of a ratio of quadratic forms in normal variables - Annales d'Economie et de Statistique 4, 95-109.

Phillips, P.C.B. (1985) - The distribution of Matrix Quotients - J. Multiv. Analysis 16, 157-161.