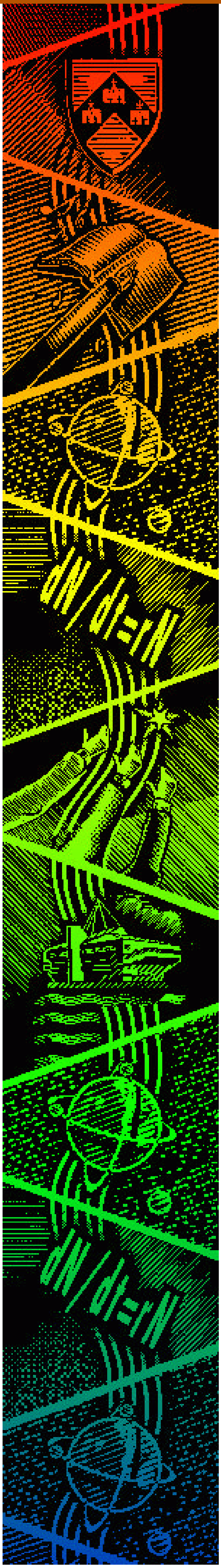

THE UNIVERSITY of York

*Discussion Papers in Economics*

No. 2001/02

# The Exact Cumulative Distribution Function of a Ratio of Quadratic Forms in Normal Variables with Application to the AR(1) Model

by

Giovanni Forchini

**Department of Economics and Related Studies**
**University of York**
**Heslington**
**York, YO10 5DD**

# The Exact Cumulative Distribution Function of a Ratio of Quadratic Forms in Normal Variables, with Application to the AR(1) Model

G. Forchini*

University of York

First version: November 1999
This version: March 2001

*I would like to thank Grant Hillier for patient help, suggestions and useful discussion. I am also grateful to the Karim Abadir, Dieter Balkenborg, Patrick Marsh, Peter Phillips, Jan Podivinsky and Richard Smith for discussions. Grants from ESRC (No. R00429424115 and No. R000236627) and the University of Southampton are gratefully acknowledged.

Proposed running head: *The CDF of a Ratio of Quadratic Forms*

Mailing Address:

Giovanni Forchini
Department of Economics and Related Studies
University of York
Heslington
York YO1 5DD
U.K.

## Abstract

Often neither the exact density nor the exact cumulative distribution function (CDF) of a statistic of interest are available in the statistics and econometrics literature (for example the maximum likelihood estimator of the autocorrelation coefficient in a simple Gaussian AR(1) model with zero start-up value). In other cases the exact CDF of a statistic of interest is very complicated despite the statistic being "simple" (for example the circular serial correlation coefficient, or a quadratic form of a vector uniformly distributed over the unit n-sphere). The first part of the paper tries to explain why this is the case by studying the analytic properties of the CDF of a statistic under very general assumptions. Differential geometric considerations show that there can be points where the CDF of a given statistic is not analytic, and such points do not depend on the parameters of the model but only on the properties of the statistic itself. The second part of the paper derives the exact CDF of a ratio of quadratic forms in normal variables, and for the first time a closed form solution is found. These results are then specialised to the maximum likelihood estimator of the autoregressive parameter in a Gaussian AR(1) model with zero start-up value, which is shown to have precisely those properties highlighted in the first part of the paper.

# 1 Introduction

The work of von Neumann (1941), Anderson (1942), Koopmans (1942), Anderson (1971) and Hillier (2001) suggests that the cumulative distribution function (CDF) of a ratio of quadratic forms might fail to be analytic at certain points of its domain. This has very important implications, for example, for the derivation of the exact distribution of a ratio of quadratic forms because the lack of analyticity in the whole domain of its CDF implies that the CDF has a different functional form over different intervals separated by these points of nonanalyticity. In general, the derivation of the exact density or CDF of a statistic of interest can be simplified (i) if we can establish whether the CDF of such a statistic is analytic everywhere in its domain or it has points where it is not analytic, and (ii) if, in the latter case, we can easily determine where such points are.

This paper addresses the problem of characterizing the existence of points where the CDF of a statistic of interest is not analytic. By generalizing some results of Mulholland (1965) and Saldanha and Tomei (1996), it will be shown in Section 2.1 that the analytic properties of the CDF of a statistic, $R$ say, depend only on the properties of the mapping $\mathrm{R}$ defining the random variable $R = \mathrm{R}(X)$ where $X$ is a random variable taking values in $\mathbb{R}^T$, and $T$ is the sample size, provided that the distribution of the original data is *smooth enough.*

These general results allow us to infer both that the CDF of ratio of quadratic forms in normal variables is not analytic at some points, and where these points are (see Section 2.2). Thus, the search for the exact CDF of such a ratio is simplified in the sense that we know that the CDF has "unusual" properties at these known points. This allows us to find an expression for the exact CDF of a ratio of quadratic forms where the quadratic form in the denominator is positive semidefinite and the covariance matrix of the normal random variables is not scalar. These results are more general than those so far available in the exact distribution theory literature, but their derivation is surprisingly simple: we write the CDF of a ratio of quadratic forms in normal variables as the probability that the difference of two independent positive definite quadratic forms in normal variables is less or equal to zero; this probability can be written in terms of a double integral (having as argument the joint density of these two independent quadratic forms) which can be evaluated in term of infinite series of zonal polynomials.

In the final part of the paper we specialise the above results to the maximum likelihood estimator of the autoregressive parameter in the simple Gaussian AR(1) model. Although the Gaussian AR(1) model has been extensively analysed for many years, and a well developed (first order and higher order) asymptotic theory (for the first see, among others, Anderson (1959), White (1959), Dickey and Fuller (1979), Evans and Savin (1981), Evans and Savin (1984), Phillips (1986b), Abadir (1993), and for the latter see for instance Phillips (1977), Phillips (1978) and Satchell (1984)) for the estimators of the autoregressive parameter and other test statistics is available, little is known about the exact (fixed $T$) distribution of the statistics that are usually of interest (with the exception of von Neumann (1941), Anderson (1942), Koopmans (1942), Anderson (1971), Hillier (2001) and Forchini (2000)).

# 2 Main results

In the first part of this section some general results concerning the analytic properties of the CDF of a statistic $\mathrm{R} : \mathbb{R}^T \to \mathbb{R}$ are obtained. The second part derives the exact CDF of a ratio of quadratic forms in normal variables.

## 2.1 Analytical properties of the CDF of a statistic $R$

It is well known that there are functional discontinuities in the density functions of the serial correlation coefficient (see L.R. Anderson (1942), and T.W. Anderson (1971)), the von Neumann ratio (von Neumann (1941)), and the sample skewness (Geary (1947)), and, in general, in the density of a quadratic form of a vector uniformly distributed on the unit $n$-sphere ( Saldanha and Tomei (1996) and Hillier (2001)).

Geary (1947) was the first to establish the link between *critical points* of a statistic (i.e. points where the gradient vanishes) and singularities in its density (see also Mulholland (1965)). The intuition for this can be easily expressed as follows. Let $\mathrm{pdf}_Y(y)$, $y \in (-1, 1)$ and $\mathrm{pdf}_X(x)$, $x \in (-1, 1)$ be the (smooth) density functions of two independent random variables $Y$ and $X$, and let $R = XY$ be the statistic of interest. The gradient of the function $\mathrm{R} : (-1, 1) \times (-1, 1) \to (-1, 1)$ defined as $(y, x) \to xy = r$ vanishes at $x = y = 0$. Some typical level surfaces for this function have been plotted in Figure 1. This shows that the level surface for $r = 0$ is the set of points on the coordinate axes. This set is not a manifold because no smooth parameterization of this set exists in a neighbourhood of the origin $(0, 0)$.

[Figure 1 approximately here]

For all $r \neq 0$ the density of $R$ is (Tjur (1980), Theorem 8.1.2)

$$\mathrm{pdf}_R(r) = \int_{\mathrm{R}^{-1}(r)} \frac{\mathrm{pdf}_X(x)\,\mathrm{pdf}_Y(y)}{\sqrt{x^2 + y^2}} dxdy$$

where $\mathrm{R}^{-1}(r) = \{(x, y) \in (-1, 1) \times (-1, 1) : xy = r\}$. The term $(x^2 + y^2)^{-\frac{1}{2}}$ is the Jacobian of the transformation $\mathrm{R} : (-1, 1) \times (-1, 1) \to (-1, 1)$, and in a neighbourhood of the origin can become arbitrarily large. If $r \neq 0$ then $\mathrm{pdf}_R(r)$ is well defined everywhere, but as soon as $r$ equals zero, the term $(x^2 + y^2)^{-\frac{1}{2}}$ tends to make the integral over $\mathrm{R}^{-1}(r)$ large, and, unless the increase of this term is compensated by an equivalent decrease of $\mathrm{pdf}_X(x)\,\mathrm{pdf}_Y(y)$ in a neighbourhood of the origin, the integral will be infinity, and the density of $R$ will have a functional discontinuity at $r = 0$. By integrating the density of $R$ to obtain the CDF the degree of smoothness is increased, so that the CDF of $R$ is continuous everywhere but it is not differentiable at $r = 0$.

In higher dimensional spaces the Jacobian of the transformation may still fail to be positive at some points. However, we can interpret the averaging over a level surface as a repeated integral, for which each integration yields a smoother function. Thus the degree of smoothness of the CDF of a statistic tends to increase with the sample size.

The next assumption and the next two theorems will formalize this intuition. The proofs of the theorems are in Appendix A.

**Assumption 1.** Let $\mathrm{pdf}_Y(y) > 0$, $y \in \mathbb{R}^T$, be the density function of a $T$-dimensional random vector $Y$ at the point $Y = y$. Suppose all derivatives of $\mathrm{pdf}_Y(y)$ exist and are continuous, and let $R = \mathrm{R}(Y)$ , where $\mathrm{R}$ is a mapping from $\mathbb{R}^T$ to the real numbers. The function $\mathrm{R}$ is assumed continuous and with continuous derivatives of all order.

The assumptions on the differentiability of $\mathrm{pdf}_Y(y)$ and $\mathrm{R}(y)$ can be weakened but the formulation of Theorem 1 and 2 below would become more complex. Also $\mathrm{pdf}_Y(y)$ and $\mathrm{R}(y)$ could be defined on a differential submanifold of $\mathbb{R}^T$ rather than on $\mathbb{R}^T$ itself and the results stated below would still apply.

The CDF of the statistic $R = \mathrm{R}(Y)$ is

$$\mathrm{F}_R(r) = \int_{\mathrm{R}^{-1}((-\infty,r))} \mathrm{pdf}_Y(y)\, dy \tag{1}$$

where

$$\mathrm{R}^{-1}((-\infty,r)) = \left\{y \in \mathbb{R}^T : \mathrm{R}(y) < r\right\}.$$

It will be shown that the only "discontinuities" in the graph of $\mathrm{F}_R(r)$ that can occur are at the points $r^* = \mathrm{R}(y^*)$, where $\nabla\,\mathrm{R}(y^*) = 0$, and $\nabla = \left(\frac{\partial}{\partial y_1}, \frac{\partial}{\partial y_2}, ..., \frac{\partial}{\partial y_T}\right)'$ denotes the gradient operator so that $\nabla\,\mathrm{R}(y)$ is a $T \times 1$ vector. A point $y^*$ where $\nabla\,\mathrm{R}(y^*) = 0$ is called a *singular* (or *critical*) *point*, and $r^* = \mathrm{R}(y^*)$ is called the *singular* (or *critical*) *value* of $\mathrm{R}(y)$ at $y^*$.

Note that $\mathrm{pdf}_Y(y)$ and thus $\mathrm{F}_\mathrm{R}(r)$ may depend on some parameters. These, however, do not play any part in the following analysis.

**Theorem 1** *If $r$ is not a singular value of $\mathrm{R}(y)$ then all the derivatives of $\mathrm{F}_R(r)$ exist and are continuous provided there is a number $r_0$ such the set*

$$\mathrm{R}^{-1}([r_0, r]) = \left\{y \in \mathbb{R}^T : r_0 \le \mathrm{R}(y) \le r\right\}$$

*is compact and does not contain any singular point of $\mathrm{R}$.*

The theorem is proved by changing coordinates to facilitate the verification of the existence of the derivatives of $\mathrm{F}_R(r)$. This is done by reparameterizing the set $\mathrm{R}^{-1}([r_0, r])$ in terms of points on a level surface

$$\mathrm{R}^{-1}(r_0) = \left\{y \in \mathbb{R}^T : \mathrm{R}(y) = r_0\right\},$$

and points on the *flow* in $\mathbb{R}^T$ generated by the vector field $\nabla\,\mathrm{R}$ (Milnor (1963) and Spivak (1970)). The assumption that the set $\mathrm{R}^{-1}([r_0, r])$ does not contain critical points of $\mathrm{R}(y)$ guarantees the existence of the flow. Compactness ensures its uniqueness on the whole set (Milnor (1963)).

Note that the condition that the set $\mathrm{R}^{-1}([r_0, r])$ is compact could be further relaxed by noting that in general the set $\mathrm{R}^{-1}([r_0, r])$ is a closed subset of $\mathbb{R}^T$ and it is thus a locally compact $\sigma$-compact set, i.e. it can be covered by a countable sequence of compact subsets in which the conclusion of Theorem 1 would apply. However, Theorem 1 is all we need for the study of ratios of quadratic forms.

The behaviour of $\mathrm{F}_R(r)$ in a neighbourhood of radius $\varepsilon$, of the critical level $r^*$ is more difficult to determine, because it is difficult to find local coordinates which simplify the integral in (1). If the critical points are isolated, the set

$$\mathrm{R}^{-1}([r^* - \varepsilon, r^* + \varepsilon]) = \left\{y \in \mathbb{R}^T : \mathrm{R}(y) \in [r^* - \varepsilon, r^* + \varepsilon]\right\}$$

can be written as the union of two sets, one containing a neighbourhood of the critical point $y^*$, $S_1$, and one containing no critical points, $S_2$,

$$\mathrm{R}^{-1}([r^* - \varepsilon, r^* + \varepsilon]) = S_1 \cup S_2$$

The analyticity of $\mathrm{F}_R(r)$ for $r \in [r^* - \varepsilon, r^* + \varepsilon]$ depends on the properties of $\nabla \mathrm{R}(y)$ and thus on the diffeomorphism $\varphi(t, y_0)$ defined in the proof of Theorem 1: since $\varphi(t, y_0)$ may fail to be analytic in $S_1$, the properties of $\mathrm{F}_R(r)$ could change quite drastically in a neighbourhood of $r^*$.

The study of the behaviour of $\mathrm{F}_R(r)$ in $[r^* - \varepsilon, r^* + \varepsilon]$ depends on the rank of the Hessian, $H\,\mathrm{R}(y)$, of $\mathrm{R}(y)$ at $y^*$, where the operator $H$ is defined as $H = \nabla\nabla'$. If $\mathrm{rank}(H\,\mathrm{R}(y^*)) = T$, the critical point $y^*$ is said to be *nondegenerate*. If $\mathrm{rank}(H\,\mathrm{R}(y^*)) < T$, $y^*$ is a *degenerate critical point*.

Mulholland (1965) has shown that if the critical point $y^*$ corresponding to the critical level $r^*$ is non degenerate then $\mathrm{F}_R(r)$ has continuous derivatives in a neighbourhood of $r^*$ up to order equal to the integer part of $T/2$. This is also true if $y^*$ is an isolated degenerate critical point as the following generalization of Theorem 2 of Mulholland (1965) shows.

**Theorem 2** *The derivatives of order $p$ of $\mathrm{F}_R(r)$ exist and are continuous in a neighbourhood of the critical level $r^*$ contained in the compact set $\mathrm{R}^{-1}([r^* - \varepsilon, r^* + \varepsilon])$, $\varepsilon > 0$, provided $\mathrm{R}(y)$ has isolated critical points. The number $p = [m/2]$ where $m \leq T$ is the number of nonzero eigenvalues of the Hessian of $\mathrm{R}(y)$ at $y = y^*$, and $[s]$ denotes the integer part of $s$.*

This theorem can be proved by using some differential geometric results (i.e. the *reduction lemma* for the degenerate case and the *Morse lemma* for the nondegenerate case (see, for instance, Castrigiano and Hayes (1993))) to find a change of coordinates in a neighbourhood of a critical point so that $\mathrm{R}(y)$ can be written as a difference of two quadratic forms (plus a smooth remainder in the degenerate case). This can then be used to prove the continuous differentiability of $\mathrm{F}_R(r)$ in a neighbourhood of the singular values.

In Theorem 1 the assumption that the critical points of $\mathrm{R}(y)$ are isolated is fundamental, since it implies by Sard's lemma (see for instance Milnor (1963)) that the set of critical points has measure zero. If this assumption fails then the density of the statistic of interest does not exist (see Jupp and Mardia (1978)).

Note that for both Theorem 1 and Theorem 2 the domain of $\mathrm{R}(y)$ can be replaced by a differentiable submanifold of $\mathbb{R}^T$. If at the critical points the density of $Y$ does not vanish, the results above still hold. Moreover, these results can be generalized from statistics with values in $\mathbb{R}$ to statistics with values in $\mathbb{R}^k$. The density of a statistic with values in $\mathbb{R}^k$ can also have functional discontinuities since the level sets of $\mathrm{R}(y)$ may fail to be manifolds. Note also that if $R = \mathrm{R}(Y) = (\mathrm{R}_1(Y), ..., \mathrm{R}_k(Y))$ is a $k$ dimensional statistic. The density of $R_k$ can be obtained by averaging the joint density of $R$ with respect to the $k-1$ random variables $R_1, ..., R_{k-1}$. By so doing the density of $R_k$ will be *smoother* than the density of $R$.

Finally, note that Theorems 1 and 2 have practical implications for the derivation of the exact density of statistics having critical points. They suggest that the functional form of the density of such a statistic is different on different intervals, and, thus, help explain why theoretical results for some statistics, for example ratios of quadratic forms, have been so difficult to obtain.

## 2.2 The CDF of a ratio of quadratic forms in normal variables

Most of the attempts to derive the density or the CDF of a ratio of quadratic forms are based on the inversion of the joint characteristic function of the two quadratic forms in the numerator and in the denominator (see among others Koopmans (1942), Gurland (1948), Gurland (1956), White (1959), Satchell (1984)). This, however, leads to integrals which are difficult to evaluate unless the sample size tends to infinity, and no general exact solution seems available so far (for solution to specific cases see for example von Neumann (1941), Koopmans (1942), Anderson (1942), Anderson (1971) and Hillier (2001)). Saddlepoint approximations to these integrals are given by Lieberman (1994) and Marsh (1998).

The results derived in Section 2.1 suggest that the CDF of a ratio of quadratic forms has a different functional form over different intervals delimited by the critical values of the statistic itself. We will obtain a representation of the CDF of a ratio of quadratic forms in normal variables at a particular point by writing it as the CDF evaluated at zero of the difference of two independent positive definite quadratic forms in normal variables which are constructed by separating the eigenvalues of a certain indefinite quadratic form into positive and negative (see also Johnson and Kotz (1970), Chapter 7, for similar procedures). It will be shown that, in a neighbourhood of a critical value (at least) one of these eigenvalues changes sign.

The density of a positive definite quadratic form of normal random variables is given for example by Gurland (1956), Ruben (1962), James (1964) for the central case and by Phillips (1986a) for the noncentral case. The distribution of an indefinite quadratic form is given by Gurland (1955) and Robinson (1965) for the case of central normal random variables and by Shah (1963) for the noncentral case (see also Johnson and Kotz (1970)), however, we will not use these results because they give expressions which are not convergent everywhere or contain unsolved integrals. Imhof (1961), Davies (1973) and Shively, Ansley, and Kohn (1990) give numerical algorithms for the computation of the density and CDF of a quadratic form.

Suppose that $Y$ is a $(T \times 1)$ random vector having a multivariate normal distribution with mean vector 0 and (positive definite) covariance matrix $\Omega$,

$$\mathrm{pdf}_Y(y) = (2\pi)^{-\frac{T}{2}} |\Omega|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2} y' \Omega^{-1} y\right\}. \tag{2}$$

It is required to find the CDF of

$$Q = \mathrm{Q}(Y) = \frac{Y'AY}{Y'BY}, \tag{3}$$

where $A$ and $B$ are $(T \times T)$ symmetric matrices and $B$ is positive semidefinite. Note that the critical values $q^* = \mathrm{Q}(y^*)$ satisfy $|A - q^* B| = 0$, and if $B$ is positive definite, these are the eigenvalues of $B^{-1}A$. It follows from Theorems 1 and 2 that the CDF of $Q$ is analytic everywhere apart from the points in a small neighbourhoods of the critical values. This suggests that the CDF of $Q$ has a different functional form over different intervals. In the rest of this section it will be shown that this is indeed the case. Note also that if $B$ is positive definite then $Q$ takes values between the smallest and the largest eigenvalues of $B^{-1}A$. If however $B$ is positive semidefinite the range of $Q$ is (i) the whole real line if

$A$ is indefinite, (ii) the positive part of the real line if $A$ is positive semidefinite and (iii) the negative part of the real line if $A$ is negative semidefinite.

As Hillier (2001) pointed out, the distribution of $Q$ is the same as the distribution of $(V'AV)/(V'BV)$ where $V = Y(Y'Y)^{-\frac{1}{2}}$ is a vector distributed on the unit $T$-sphere. Therefore, the results below hold for scale-mixtures of normals, and in particular for spherically symmetric distributions.

Let $\mathrm{F}_Q(q)$ be the CDF of $Q$ at the point $q$ given that $Y \sim N(0, \Omega)$,

$$\begin{aligned} \mathrm{F}_Q(q) &= \Pr\left\{\frac{Y'AY}{Y'BY} \leq q \mid Y \sim N(0,\Omega)\right\} \\ &= \Pr\left\{\frac{Y'A^*Y}{Y'B^*Y} \leq q \mid Y \sim N(0,I_T)\right\} \\ &= \Pr\left\{Y'(A^* - qB^*)Y \leq 0 \mid Y \sim N(0,I_T)\right\}, \end{aligned}$$

where $A^* = \Omega^{\frac{1}{2}} A \Omega^{\frac{1}{2}}$ and $B^* = \Omega^{\frac{1}{2}} B \Omega^{\frac{1}{2}}$. The third line of the display above is the only point where the assumption that $B$ (and thus $B^*$) is positive semidefinite is used.

Let $Y = H'X$, where $H$ is an orthogonal matrix which diagonalizes $A^* - qB^*$,

$$H'(A^* - qB^*)H = \begin{pmatrix} D_1(q) & 0 & 0 \\ 0 & -D_2(q) & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

and $D_1(q)$ and $-D_2(q)$ are diagonal matrices containing the $n_1 \geq 0$ positive and the $n_2 \geq 0$ negative eigenvalues of $A^* - qB^*$ respectively. Note that $n_1$ and $n_2$ vary as $q$ varies. By partitioning $X$ conformably to $H'(A^* - qB^*)H$, $X = (X_1', X_2', X_3')'$, we obtain

$$\mathrm{F}_Q(q) = \Pr\left\{X_1'D_1(q)X_1 - X_2'D_2(q)X_2 \leq 0 | X_1 \sim N(0, I_{n_1}), X_2 \sim N(0, I_{n_2})\right\}, \tag{4}$$

where $X_1$, $X_2$ and $X_3$ are independent, and

$$n_1 + n_2 = \operatorname{rank}(A^* - qB^*) = \operatorname{rank}(A - qB) \leq T.$$

At a critical value $q = q^*$ the rank of $A^* - qB^*$ is less or equal to $T-1$, while for $q \neq q^*$ the rank of $A^* - qB^*$ is $T$. This suggests that in a neighbourhood of a critical point $n_1$ and $n_2$ change (this will be analysed in more detail for the maximum likelihood estimator of the autoregressive parameter in a Gaussian AR(1) model in Section 3.2). Note also that the number of positive and negative eigenvalues of $A - qB$ is the same as the number of positive and negative eigenvalues of $A^* - qB^*$ by Sylvester's law of inertia.

The above results allow us to write the CDF of $Q$ as

$$\mathrm{F}_Q(q) = \Pr\{Q_1 - Q_2 \leq 0\}, \tag{5}$$

where $Q_1 = X_1' D_1(q) X_1 > 0$ and $Q_2 = X_2' D_2(q) X_2 > 0$ are independent quadratic forms in normal variables. Note that $\mathrm{F}_Q(q) = 0$ for values of $q$ for which $n_2 = 0$, and $\mathrm{F}_Q(q) = 1$ for values of $q$ for which $n_1 = 0$. If $n_1 > 0$ and $n_2 > 0$ we can find the joint density of $(Q_1, Q_2)$ as a product of the marginal densities of $Q_1$ and $Q_2$ (since they are independent). Thus the CDF of $Q$ at the point $q$ is

$$\mathrm{F}_Q(q) = \int_{q_2>0} \int_{0<q_1<q_2} \mathrm{pdf}_{Q_1}(q_2)\,\mathrm{pdf}_{Q_2}(q_2)\, dq_1 dq_2. \tag{6}$$

This integral can be evaluated by expanding the densities of $Q_1$ and $Q_2$ as infinite series and by integrating term by term. This procedure, detailed in Appendix B, leads to two expressions for the CDF of a ratio of two quadratic forms in normal variables which do not seem to have been derived before in the statistical literature.

**Theorem 3** *If $Y \sim N(0, \Omega)$, $D_1 = D_1(q)$ and $D_2 = D_2(q)$, and $q$ is in the interval for which $n_1 > 0$ and $n_2 > 0$, the CDF of $Q$, defined in (3), is*

$$\begin{aligned}\mathrm{F}_Q(q) \quad = \quad & \frac{\Gamma\left(\frac{n_1+n_2}{2}\right)}{\Gamma\left(\frac{n_2}{2}\right)\Gamma\left(\frac{n_1}{2}+1\right)|D_1|^{\frac{1}{2}}|D_2|^{\frac{1}{2}}} \\ & \sum_{a=0}^{\infty}\sum_{b=0}^{\infty}\sum_{\alpha}\sum_{\beta} \frac{\left(\frac{n_2-1}{2}\right)_\alpha \left(\frac{n_1+1}{2}\right)_\beta \left(\frac{n_1+n_2}{2}\right)_{a+b}}{\left(\frac{n_2}{2}\right)_\alpha \left(\frac{n_1}{2}+1\right)_\beta a!b!} \\ & \left[\mathrm{tr}\left(D_2^{-1}\right) + \mathrm{tr}\left(D_1^{-1}\right)\right]^{-a-b-\frac{n_1+n_2}{2}} C_\alpha\left(D_2^{-1}\right) C_\beta\left(D_1^{-1}\right),\end{aligned} \tag{7}$$

*where $C_\alpha(.)$ and $C_\beta(.)$ are zonal polynomial corresponding to the partitions $\alpha$ and $\beta$ of the integers $a$ and $b$ respectively (James (1964) or Muirhead (1982)), and $(z)_k$ denotes the quantity $(z)_0 = 1$ and $(z)_k = z(z+1)(z+2)\cdots(z+k-1)$, $k = 1, 2, \ldots$.*

An expression involving only top-order zonal polynomials (i.e. corresponding to the partition $[j] = (j, 0, 0..., 0)$ of the integer $j$), and thus easier to evaluate numerically, is given in the following theorem.

**Theorem 4** *If $Y \sim N(0,\Omega)$, $D_1 = D_1(q)$ and $D_2 = D_2(q)$, and $q$ is in the interval for which $n_1 > 0$ and $n_2 > 0$, an alternative expression for the CDF of $Q$, defined in (3), in terms of top-order zonal polynomials is*

$$
\begin{aligned}
\mathrm{F}_Q(q) &= \frac{\Gamma\left(\frac{n_1+n_2}{2}\right)}{\Gamma\left(\frac{n_2}{2}\right)\Gamma\left(\frac{n_1}{2}+1\right)|D_1|^{\frac{1}{2}}|D_2|^{\frac{1}{2}}} \\
&\quad \sum_{j=0}^{\infty}\sum_{p=0}^{\infty}\frac{\left(\frac{1}{2}\right)_j\left(\frac{1}{2}\right)_p\left(\frac{n_1+n_2}{2}\right)_{p+j}(-1)^{j+p}}{\left(\frac{n_1}{2}+1\right)_j\left(\frac{n_2}{2}\right)_p j!p!}\left[\mathrm{tr}\left(D_2^{-1}\right)+\mathrm{tr}\left(D_1^{-1}\right)\right]^{-j-p-\frac{n_1+n_2}{2}} \\
&\quad C_{[j]}\left(D_1^{-1}-I_{n_1}\mathrm{tr}\left(D_1^{-1}\right)\right)C_{[p]}\left(D_2^{-1}-I_{n_2}\mathrm{tr}\left(D_2^{-1}\right)\right) \\
&\quad {}_2F_1\left(p+j+\frac{n_1+n_2}{2},1;j+\frac{n_1}{2}+1,\frac{\mathrm{tr}\left(D_1^{-1}\right)}{\mathrm{tr}\left(D_2^{-1}\right)+\mathrm{tr}\left(D_1^{-1}\right)}\right) \qquad (8)
\end{aligned}
$$

*where the hypergeometric function has scalar argument.*

The matrices $D_1^{-1}$ and $D_2^{-1}$ can be differentiated with respect to $q$ provided $q \neq q^*$. However, calculating the density function of $Q$ by differentiating $\mathrm{F}_Q(q)$ term by term is very complicated. It is probably easier to differentiate equation (6) with respect to $q$ under the integral sign, and re-evaluate the integrals.

# 3 The autocorrelation coefficient

The simple Gaussian AR(1) model

$$Y_t = \rho Y_{t-1} + \varepsilon_t,\ \varepsilon_t \sim NID\left(0,\sigma^2\right),\ t = 1,2,...,T,\ Y_0 = 0, \qquad (9)$$

has been extensively analysed for many years, and a well developed asymptotic theory for the estimators of the autoregressive parameter and other test statistics is available (see, among others, Anderson (1959), White (1959), Dickey and Fuller (1979), Evans and Savin (1981), Evans and Savin (1984), Phillips (1986b), Abadir (1993)). The higher order asymptotic theory is also well developed (see for instance Phillips (1977), Phillips (1978) and Satchell (1984)). However, little is known about the exact (fixed $T$) distribution of the statistics that are usually of interest in (9).

Most of the known exact results for the Gaussian AR(1) model are in the book by Anderson (1971), who, among other things, derives the exact distribution of the serial correlation coefficient in the circular model (i.e. $Y_0 = Y_T$). Recently Hillier (2001) has derived the density of a quadratic form in a vector uniformly distributed on the unit $n$-sphere which can be used to calculate the density of several statistics of interest in (9) for the case $\rho = 0$. This yields as a special case the density of the serial correlation coefficient

$$\bar{R} = \frac{Y' A_T Y}{Y'Y}, \tag{10}$$

where $Y = (Y_1, Y_2, ..., Y_T)'$, $A_T = \frac{1}{2}(L_T' + L_T)$, and

$$L_T = \begin{pmatrix} 0 & 1 & \cdots & & 0 \\ 0 & 0 & \ddots & & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}, \tag{11}$$

(see also Anderson (1942) and Koopmans (1942)), and of the von Neumann ratio (von Neumann (1941)) when the autoregressive parameter is zero.

Many of the statistics of interest in (9) can be written as ratios of quadratic forms in normal variables, where the quadratic form in the denominator is, in general, positive semidefinite. For example, the maximum likelihood estimator (MLE) (which, in this case, coincides with the ordinary least squares estimator) for $\rho$ in (9) is

$$\hat{R} = \frac{Y' A_T Y}{Y' B_T Y}, \tag{12}$$

where $B_T = L_T L_T'$. The exact density function and CDF of $\hat{R}$ have not been successfully investigated yet. The only exact results so far available are the bias of $\hat{R}$ (Hurwicz (1950)), and its exact moments (Sawa (1978), Jones (1987), Nankervis and Savin (1988), Smith (1989), Roberts (1995), and Vinod and Shenton (1996)).

This Section applies the results derived in Section 2 to the autocorrelation coefficients (10) and (12) in model (9). The organization of the section is the same as that of Section 2.

## 3.1 Analyticity of the density of $\bar{R}$ and $\hat{R}$

In this section we assume that the error terms have a smooth joint density function, and the results obtained do not rely on the assumption of normal errors.

The statistics $\hat{R}$ and $\bar{R}$ are homogeneous functions of degree zero in $Y$. So by transforming to polar coordinates, $Y = Q^{1/2}V$, $Q > 0$, $V'V = 1$ (Muirhead (1982)), $\hat{R}$ and $\bar{R}$ can be thought of as defined in terms of $V$ only as

$$\hat{R} = \frac{V'A_T V}{V'B_T V}$$

and

$$\bar{R} = V'A_T V$$

where the density of $V$ is

$$pdf_V(v) = \frac{1}{2}\int_0^{\infty} pdf_Y\left(q^{1/2}v\right) q^{\frac{T}{2}-1} dq.$$

As Theorems 1 and 2 show, the analytic properties of the densities of $\hat{R}$ and $\bar{R}$ are determined uniquely by the gradient and the Hessian of $\hat{R}$ and $\bar{R}$ regarded as function of $V$.

**Corollary 1** *If $pdf_V(v)$ is continuous and has continuous partial derivatives of all orders, then the density of $\hat{R}$ is analytic everywhere apart from the points in a neighbourhood of $\hat{r}_k^* = \cos\left(\frac{k\pi}{T-1}\right)$, $k = 1, 2, ..., T-2$. In a neighbourhood of $\hat{r}_k^*$ the order of differentiability of $F_{\hat{R}}(\hat{r})$ is between 1 and $[(T-2)/2]$.*

**Corollary 2** *If $pdf_V(v)$ is continuous and has continuous partial derivatives of all orders, then the density of $\bar{R}$ is analytic everywhere apart from the points in a neighbourhood $\bar{r}_k^* = \cos\left(\frac{k\pi}{T+1}\right)$, $k = 1, 2, ..., T$. In a neighbourhood of $\bar{r}_k^*$ the order of differentiability of $F_{\bar{R}}(\bar{r})$ is between 1 and $[(T-1)/2]$ if $T$ is odd or $[T/2]$ if $T$ is even.*

**Remarks**

(i) Corollaries 1 and 2 hold if the density of $Y$ is continuous and has continuous derivatives. Thus the assumption of normality does not affect this conclusion. Note that neither $\hat{r}_k^*$ nor $\bar{r}_k^*$ depend on the autoregressive parameter. For the case $\rho = 0$ and the case of Gaussian errors the density of $\bar{R}$, which can be obtained from the results in Hillier (2001) (and for, a special case, Anderson (1942), Anderson (1971)), has different functional forms in different intervals.

(ii) The densities of $\hat{R}$ and $\bar{R}$ in model (9) are smooth everywhere apart from a neighbourhood of the critical levels, $\hat{r}_k^*$ and $\bar{r}_k^*$ respectively. This suggests that the densities of $\hat{R}$ and $\bar{R}$ might be piecewise continuous. Figures 2 and 3 obtained by numerically integrating equations (4.8) and (4.9) in Forchini (1998), for $T = 3$ and 4 and $\rho = 0$ and 1, clearly show this property of the density of $\hat{R}$. For $T = 3$, the order of differentiability of $F_{\hat{R}}(\hat{r})$ is $[1/2] = 0$, so the density of $\hat{R}$ has a discontinuity at $\hat{r} = 0$. For $T = 4$ and $T = 5$, the density of $\hat{R}$ is well defined and continuous everywhere but it is not differentiable at the critical points. For $T \geq 6$, the discontinuity involves higher order derivatives, and becomes very difficult to notice in a graph of the density.

[Figures 2 and 3 approximately here]

## 3.2 The CDF of $\hat{R}$

Since $\hat{R}$ can be written as a ratio of quadratic forms in normal variables, its CDF is given by Theorems 3 and 4 of Section 2.3 with $A = A_T$, $B = B_T$ and $\Omega = \sigma^2(I_T + \rho^2 B_T - 2\rho A_T)^{-1}$. This Section establishes a link between Sections 2.1, 3.1 and 2.2. For this reason we now study the eigenvalues of $A_T - \hat{r}B_T$, which are relevant when $\rho = 0$. This is done in a series of Lemmas. The main result, contained in Corollary 4 below, shows how the functional form of $\mathrm{F}_{\hat{R}}(\hat{r})$ changes at the critical values $\hat{r} = \hat{r}_k^*$, $k = 1, ..., T-2$ (defined in Corollary 1) in the case of normal errors.

**Lemma 1** *If* $\hat{r} = \cos\left(\frac{k\pi}{T-1}\right)$, $k = 1, 2, .., T-2$, *then* $\operatorname{rank}(A_T - \hat{r}B_T) = T - 1$, *otherwise* $\operatorname{rank}(A_T - \hat{r}B_T) = T$.

**Lemma 2** *The eigenvalues of* $A_T - \hat{r}B_T$ *have the form* $\lambda(\hat{r}) = -\hat{r} - \cos\theta$, *where* $\theta \in \mathbb{C}$ *solves*

$$f_T(\theta) = \frac{\sin((T+1)\theta)}{\sin(T\theta)} = -2\hat{r}, \tag{13}$$

$$\sin(\theta) \neq 0. \tag{14}$$

*Precisely, (i) if* $|2\hat{r}| \leq 1 + 1/T$ *then there are* $T$ *real solutions* $\theta_1, \theta_2, ..., \theta_T$ *to the* *equation (13) with*

$$\begin{aligned}\theta_1 &\in \left[0, \frac{\pi}{T}\right) \\ \theta_k &\in \left(\frac{k\pi}{T}, \frac{(k+1)\pi}{T}\right), \; k = 2, ..., T-1 \\ \theta_T &\in \left(\frac{(T-1)\pi}{T}, \pi\right],\end{aligned}$$

*and (ii) if* $|2\hat{r}| > 1 + 1/T$ *then (13) has* $T-1$ *real solutions and one complex solution:*

| $\hat{r} > 0$ | $\hat{r} < 0$ |
|---|---|
| $\theta_k \in \left(\frac{(k+1)\pi}{T+1}, \frac{(k+1)\pi}{T}\right), \; k = 1, ..., T-1$ | $\theta_1 = ia, \; a > 0$ |
| $\theta_T = \pi + ia, \; a > 0$ | $\theta_k \in \left(\frac{k\pi}{T}, \frac{(k+1)\pi}{T+1}\right), \; k = 2, ..., T$ |

An immediate consequence of this lemma is:

**Corollary 3** $\cos\theta_k < \cos\theta_{k+1}$, $k = 1, ..., T-1$

In general, the function $f_T(\theta)$ equals zero at $\theta = k\pi/(T+1)$, for $k, = 1, 2, ...$, and $f_T(\theta)$ has poles at $\theta = k\pi/T$, for $k, = 1, 2, ....$ The zeros and the poles of $f_T(\theta)$ cancel if both $(T+1)k/T$ and $k$ are integers. Note that equation (13) can be explicitly solved for $\theta \in \mathbb{R}$ in some special cases, such as $\hat{r} = 0, \pm 1/2$.

Note also that all eigenvalues of $A_T - \hat{r}B_T$ are real even when $\theta_T$ is complex since $\cos(\pi k + ia) = (-1)^k \cosh(a) \in \mathbb{R}$, and that the largest eigenvalue is always positive and the smallest always negative as shown in the following lemma.

**Lemma 3** *Let* $\lambda_i = -\hat{r} - \cos\theta_i$, $i = 1, 2, ..., T$, *where* $\theta_i$ *is a solution to (13). Then* $\min \lambda_i < 0$ *and* $\max \lambda_i > 0$.

Now, note that the ordered critical values of $\hat{r}$, $\hat{r}_1^* > \hat{r}_2^* > ... > \hat{r}_{T-2}^*$, divide $\mathbb{R}$ into disjoint subsets:

$$(-\infty, \hat{r}_1^*) \cup \{\hat{r}_1^*\} \cup (\hat{r}_1^*, \hat{r}_2^*) \cup \{\hat{r}_2^*\} \cup ... \cup \left(\hat{r}_{T-3}^*, \hat{r}_{T-2}^*\right) \cup \left\{\hat{r}_{T-2}^*\right\} \cup \left(\hat{r}_{T-2}^*, +\infty\right).$$

Let $I_1 = (-\infty, \hat{r}_1^*)$, $I_k = \left(\hat{r}_{k-1}^*, \hat{r}_k^*\right)$, $k = 2, ..., T-2$, and $I_{T-1} = \left(\hat{r}_{T-2}^*, +\infty\right)$. Then it follows from Corollary 3 and the fact that the eigenvalues are continuous functions of $\hat{r}$ (Lemma 2) that:

**Corollary 4** *If $\hat{r} \in I_k$, then $\lambda_1(\hat{r}) > \lambda_2(\hat{r}) > ... > \lambda_{T-k-1}(\hat{r}) > 0 > \lambda_{T-k}(\hat{r}) > ... > \lambda_T(\hat{r})$.*

Corollary 4 makes clear what happens when $\hat{r}$ varies in the interval $(\hat{r}_k^* - \varepsilon, \hat{r}_k^* + \varepsilon)$, where $\varepsilon$ is a small positive quantity: $\lambda_{T-k-1}(\hat{r})$ changes from negative to positive. Equation (7) shows that the CDF of $\hat{R}$ depends only on the nonzero eigenvalues of $A_T - \hat{r}B_T$, and the nonanalyticity at $\hat{r}_k^*$ is due to the fact that the dimensions of the matrices $D_1(\hat{r})$ and $D_2(\hat{r})$ change at such a point.

**Remarks**

(i) By the implicit function theorem, the eigenvalues of $A_T^* - \hat{r}B_T^*$, and thus the CDF of $\hat{R}$, can be differentiated with respect to $\rho$ at all points. Therefore the exact CDF does not have that discontinuity in $\rho$ which characterizes the asymptotic distribution of $\hat{R}$.

(ii) The case where $\rho \neq 0$ can be treated analogously. The only difference is that in this case we need to analyse the eigenvalues of $A_T^* - \hat{r}B_T^*$, where $A^* = \Omega^{\frac{1}{2}} A \Omega^{\frac{1}{2}}$ and $B^* = \Omega^{\frac{1}{2}} B \Omega^{\frac{1}{2}}$ and $\Omega = \sigma^2 \left(I_T + \rho^2 B_T - 2\rho A_T\right)^{-1}$. Note that the values of $\hat{r}$ where analyticity fails are again $\hat{r}_k^* = \cos\left(\frac{k\pi}{T-1}\right)$, $k = 1, 2, .., T-2$ since $\operatorname{rank}(A_T^* - \hat{r}B_T^*) = \operatorname{rank}(A_T - \hat{r}B_T)$. Note also that by Sylvester's law of inertia the number of positive and negative eigenvalues of $A_T^* - \hat{r}B_T^*$ is the same as the number of positive and negative eigenvalues of $A_T - \hat{r}B_T$.

(iii) The cumulative distribution function of $\bar{R}$ can be easily obtained from (7) by noting that the eigenvalues of $A_T - \bar{r}I_T$ are $-\bar{r} + \cos\left(\frac{k\pi}{T+1}\right)$. The comments following (5) show that $\mathrm{F}_{\bar{R}}(\bar{r}) = 0$ for $\bar{r} \leq \cos\left(\frac{T\pi}{T+1}\right)$, $\mathrm{F}_{\bar{R}}(\bar{r}) = 1$ for $\bar{r} \geq \cos\left(\frac{\pi}{T+1}\right)$. For $-1 \leq \bar{r} \leq 1$, $\mathrm{F}_{\bar{R}}(\bar{r})$ is given by (7).

(iv) The technique used in the previous Sections can be employed to show the existence and the almost everywhere analyticity of the density of the OLS estimator for the autoregressive parameter in the Gaussian AR(1) model with a random start-up value. The main difference with the zero start-up case is the existence of $T-1$ rather than $T-2$ critical levels for the OLS estimator of $\rho$.

# 4 Conclusion

This paper has dealt with two problems. The first one concerns the analyticity of the CDF of a statistic under very general assumptions. Differential geometric considerations have shown that there are points where the CDF of a given statistic may not be analytic, and such points do not depend on the parameters of the model. This suggests that some statistics might have density functions with different functional forms over different intervals, and explains why exact results have been so difficult to derive for some statistics.

The second problem considered in the paper concerns the exact CDF of a ratio of quadratic forms in normal variables. For the first time a closed form solution has been derived for the CDF of such a statistic which according to the results of the first part has point of nonanalyticity. The maximum likelihood estimator for the autoregressive parameter in a Gaussian AR(1) model with zero start-up value has been used to illustrate what happens at a point where the CDF is not analytic.

# References


ABADIR, K. M. (1993): "The Limiting Distribution of the Autocorrelation Coefficient under a Unit Root," *The Annals of Statistics*, 21, 1058–1070.

ANDERSON, L. R. (1942): "The Distribution of the Serial Correlation Coefficient," *The Annals of Mathematical Statistics*, 13, 1–13.

ANDERSON, T. W. (1959): "On Asymptotic Distributions of Estimates of Parameters of Stochastic Difference Equations," *The Annals of Mathematicals Statistics*, 30, 676–687.

——— (1971): *The Statistical Analysis of Time Series*. John Wiley and Sons, New York.

CASTRIGIANO, D. P. L., AND S. A. HAYES (1993): *Catastrophe Theory*. Addison-Wesley Publishing Company, New York.

DAVIES, R. B. (1973): "Numerical Inversion of a Characteristic Function," *Biometrika*, 60, 415–417.

DICKEY, D. A., AND W. A. FULLER (1979): "Distribution of the Estimators for Autoregressive Series with a Unit Root," *Journal of the American Statistical Association*, 74, 427–431.

EVANS, G. B. A., AND N. E. SAVIN (1981): "Testing for Unit Roots:," *Econometrica*, 49, 753–799.

——— (1984): "Testing for Unit Roots:," *Econometrica*, 52, 1241–1269.

FORCHINI, G. (1998): "Exact Distribution Theory for Some Econometric Problems," Ph.D. thesis, University of Southampton.

——— (2000): "The Density of the Sufficient Statistics for a Gaussian AR(1) Model in Terms of Generalized Functions," *Statistics and Probability Letters*, 50, 237–243.

GEARY, R. C. (1947): "The Frequency Distribution of $\sqrt{b_1}$ for Samples of All Sizes Drawn at Random from a Normal Population," *Biometrika*, 34, 68–97.

GURLAND, J. (1948): "Inversion Formulae for the Distribution of Ratios," *The Annals of Mathematical Statistics*, 19, 228–237.

——— (1955): "Distribution of Definite and Indefinite Quadratic Forms," *The Annals of Mathematical Statistics*, 26, 122–127.

——— (1956): "Quadratic Forms in Normally Distributed Random Variables," *Sankhya*, 17, 37–50.

HILLIER, G. H. (2001): "The Density of a Quadratic Form in a Vector Uniformly Distributed on the n-Sphere," *Econometric Theory*, 17.

HURWICZ, L. (1950): "Least-Squares Bias in Time Series," in *Statistical Inference in Dynamic Economic Models*, ed. by T. C. Koopmans. Cowles Commission for Research in Economics, Monograph No. 10, Chapman and Hall, London.

IMHOF, J. P. (1961): "Computing the Distribution of Quadratic Forms in Normal Variables," *Biometrika*, 48, 419–426.

JAMES, A. T. (1964): "Distributions of Matrix Variates and Latent Roots Derived from Normal Samples," *Annals of Mathematical Statistics*, 35, 475–501.

JOHNSON, N. L., AND S. KOTZ (1970): *Continuous Univariate Distribution*, vol. 2. Houghton Mifflin, New York.

Jones, M. C. (1987): "On Moments of Ratios of Quadratic Forms in Normal Variables," *Statistics and Probability Letters*, 6, 129–136.

Jupp, P. E., and K. V. Mardia (1978): "Density-Preserving Statistics and Densities for Sample Means," *The Annals of Probability*, 6, 688–694.

Koopmans, T. C. (1942): "Serial Correlation and Quadratic Forms in Normal Variables," *The Annals of Mathematical Statistics*, 12, 14–33.

Lieberman, O. (1994): "Saddlepoint Approximation for the Distribution of a Ratio of Quadratic Forms in Normal Variables," *Journal of the American Statistical Association*, 89, 924–928.

Marsh, P. W. N. (1998): "Saddlepoint Approximations for Noncentral Quadratic Forms," *Econometric Theory*, 14, 539–559.

Milnor, J. (1963): *Morse Theory.* Princeton University Press, Princeton.

Muirhead, R. J. (1982): *Aspects of Multivariate Statistical Theory.* John Wiley and Sons, New York.

Mulholland, H. P. (1965): "On the Degree of Smoothness and on Singularities in Distributions of Statistical Functions," *Proceedings of the Cambridge Philosophical Society*, 61, 721–739.

Nankervis, J. C., and N. W. Savin (1988): "The Exact Moments of the Least Squares Estimator for the Autoregressive Model: Corrections and Extensions," *Journal of Econometrics*, 37, 381–388.

Phillips, P. C. B. (1977): "Approximations to Some Finite Sample Distributions Associated with a First-Order Stochastic Difference Equation," *Econometrica*, 45, 463–485.

——— (1978): "Edgeworth and Saddlepoint Approximations in the First-Order Noncircular Autoregression," *Biometrika*, 65, 91–98.

——— (1986a): "The Exact Distribution of the Wald Statistic," *Econometrica*, 54, 881–895.

——— (1986b): "Time Series Regression with a Unit Root," *Econometrica*, 55, 277–301.

Roberts, L. A. (1995): "On the Existence of Moments of Ratios of Quadratic Forms," *Econometric Theory*, 11, 759–774.

Robinson, J. (1965): "The Distribution of a General Quadratic Form in Normal Variates," *Australian Journal of Statistics*, 7, 110–114.

Ruben, H. (1962): "Probability Content of Regions under Spherical Normal Ditributions, IV: The Distribution of Homogeneous and Non-Homogeneous Quadratic Functions of Normal Variables," *The Annals of Mathematical Statistics*, 33, 542–570.

Saldanha, N. C., and C. Tomei (1996): "The Accumulated Distribution of Quadratic Forms on the Sphere," *Linear Algebra and its Applications*, 245, 335–351.

Satchell, S. (1984): "Approximation to the Finite Sample Distribution for Nonstable First Order Stochastic Difference Equations," *Econometrica*, 52, 1271–1289.

Sawa, T. (1978): "The Exact Moments of the Least Squares Estimator for the Autoregressive Model," *Journal of Econometrics*, 31, 244–257.

Shah, S. B. (1963): "Distribution of Definite and of Indefinite Quadratic Forms for a Non-Central Normal Distribution," *The Annals of Mathematical Statistics*, 34, 186–190.

Shively, T. S., C. F. Ansley, and R. Kohn (1990): "Fast Evaluation of the Distribution of the Durbin-Watson and Other Invariant Test Statistics in Time Series Regression," *Journal of the American Statistical Association*, 85, 676–685.

Smith, M. (1989): "On the Expectation of a Ratio of Quadratic Forms in Normal Variables," *Journal of Multivariate Analysis*, 31, 244–257.

Spivak, M. (1970): *A Comprehensive Introduction to Differential Geometry.* Publish and Perish, Boston.

Titchmarsh, E. C. (1993): *The Theory of Functions.* Oxford University Press, Oxford.

Tjur, T. (1980): *Probability Based on Radon Measures.* John Wiley and Sons, New York.

Vinod, H. D., and L. R. Shenton (1996): "Exact Moments for Autoregressive and Random Walk Models for a Zero or Stationary Initial Value," *Econometric Theory*, 12, 481499.

von Neumann, J. (1941): "Distribution of the Ratio of the Mean Square Successive Difference to the Variance," *The Annals of Mathematical Statistics*, 12, 367–395.

White, J. S. (1959): "The Limiting Distribution of the Serial Correlation Coefficient in the Explosive Case," *The Annals of Mathematical Statistics*, 29, 1188–1197.

# Appendix A: Proofs of results in Section 2.1

**Proof of Theorem 1.** Following Saldanha and Tomei (1996), $r$ and $r_0$ are chosen so that there are no singular points in the set $\mathrm{R}^{-1}([r_0, r]) = \{y \in \mathbb{R}^T : r_0 \leq \mathrm{R}(y) \leq r\}$. Then

$$\mathrm{F}_R(r) = \mathrm{F}_R(r_0) + \int_{\mathrm{R}^{-1}([r_0,r])} \mathrm{pdf}_Y(y)\, dy.$$

If $\mathrm{R}^{-1}([r_0, r])$ is compact, lemma 2.4 in Milnor (1963) guarantees the existence of a 1-parameter group of diffeomorphisms of $\mathrm{R}^{-1}([r_0, r])$ (i.e. a one-to-one differentiable mapping) $\varphi : \mathbb{R} \times \mathrm{R}^{-1}([r_0, r]) \to \mathrm{R}^{-1}([r_0, r])$ such that: a) $\varphi(0, y) = y$, b) $\partial\varphi(t, y)/\partial t = \nabla(\mathrm{R}(\varphi(t, y)))$, and c) $\varphi(t, \varphi(s, y)) = \varphi(t + s, y)$.

Thus the set $\mathrm{R}^{-1}([r_0, r])$ can be reparameterized is such a way that for any $y \in \mathrm{R}^{-1}([r_0, r])$, $y = \varphi(t, y_0)$, $t \in (r_0, r)$, $y_0 \in \mathrm{R}^{-1}(r_0) = \{y \in \mathbb{R}^T : \mathrm{R}(y) = r_0\}$, and

$$\mathrm{F}_R(r) = \mathrm{F}_R(r_0) + \int_{r_0}^{r} \int_{\mathrm{R}^{-1}(r_0)} \mathrm{pdf}_Y(\varphi(t, y_0))\, J_\varphi(t, y_0)\, dy_0 dt$$

where $J_\varphi(t, y_0)$ denotes the Jacobian of the transformation $x = \varphi(t, y_0)$ evaluated at the point $(t, y_0)$. The first derivative of $\mathrm{F}_R(r)$ exists, and equals,

$$\mathrm{F}'_R(r) = \mathrm{pdf}_R(r) = \int_{\mathrm{R}^{-1}(r_0)} \mathrm{pdf}_Y(\varphi(r, y_0))\, J_\varphi(r, y_0)\, dy_0.$$

Since the derivatives of $\mathrm{pdf}_Y(\varphi(r, y_0))$ and $J_\varphi(r, y_0)$ exist and are continuous, the higher order derivatives of $\mathrm{F}_R(r)$ exist and are continuous.

**Proof of Theorem 2.** The proof follows that of Theorem 2 in Mulholland (1965). Since the critical points are isolated, there is no loss of generality in assuming that there is only a critical point $y^*$ corresponding to the critical level $r^* = \mathrm{R}(y)$. The main idea of the proof is that of splitting the region $\mathrm{R}^{-1}([r^* - \varepsilon, r^* + \varepsilon])$ containing $y^*$ into two parts, a small region, $S_1$, containing the critical point $y^*$, and a region, $S_2$, which does not contain any critical point. Theorem 1 can be applied to region $S_2$, so that the integral over $S_2$ of $\mathrm{pdf}_Y(y)$ is continuously differentiable. The problems arise from the region $S_1$, which is our main concern here.

To analyse the region $\mathrm{R}^{-1}([r_0, r])$ we first parameterize it in a convenient way. Let

$$g(y) = \mathrm{R}(y) - r^*.$$

From the *reduction lemma* (Castrigiano and Hayes (1993), page 64) it follows that there exists a diffeomorphism $\varphi$ from a neighbourhood, $U_\delta(y^*)$, of radius $\delta$, of $y^*$ to a neighbourhood of $0 \subset \mathbb{R}^T$, such that

$$0 = \varphi(0)$$

and

$$g(y^* + \varphi(x_1, x_2, x_3)) = x_1' x_1 - x_2' x_2 + h(x_3),$$

where $h(.)$ is a smooth function having the following properties:
(a) $h(0) = 0$,
(b) $\operatorname{rank}(\mathrm{H}\, h(0)) = 0$, and $p$ and $m - p$ are respectively the number of positive and negative eigenvalues of $\mathrm{H}\, R(y^*)$, so that $x_1 \in \mathbb{R}^p$, $x_2 \in \mathbb{R}^{m-p}$, $x_3 \in \mathbb{R}^{T-m}$.

In order to simplify the analysis we define the region $S_1 = \varphi(S_1')$ in terms of the new coordinates as

$$S_1' = \left\{x \in \mathbb{R}^T : x_1' x_1 + x_2' x_2 < 2\delta', x_3' x_3 < \delta''\right\},$$

for suitable small $\delta'$, $\delta'' > 0$. For $r$ close enough to $r^*$ set

$$\begin{aligned} E_1(r) &= \varphi(S_1') \cap \mathrm{R}^{-1}((-\infty, r)) \cap R^{-1}((r^* - \varepsilon, r^* + \varepsilon)) \\ E_2(r) &= \left[\mathrm{R}^{-1}((r^* - \varepsilon, r^* + \varepsilon)) \setminus \varphi(S_1')\right] \cap \mathrm{R}^{-1}((-\infty, r)) \end{aligned}$$

$(\varepsilon > 0)$ then

$$F_R(r) = F_R(r^* - \varepsilon) + \int_{E_1(r)} \mathrm{pdf}_Y(y)\, dy + \int_{E_2(r)} \mathrm{pdf}_Y(y)\, dy.$$

Since there are no critical points in $S_2$, the integral over $E_2(r)$ can be shown to be continuously differentiable as in the proof of Theorem 1. The integral over $E_1(r)$ only needs to be studied.

Transforming $y$ to $(x_1, x_2, x_3)$, $(y = y^* + \varphi(x_1, x_2, x_3))$

$$\int_{E_1(r)} \mathrm{pdf}_Y(y)\, dy = \int_{E_1'(r)} \mathrm{pdf}_Y(y^* + \varphi(x_1, x_2, x_3))\, J_\varphi(x_1, x_2, x_3)\, dx_1 dx_2 dx_3$$

where $J_\varphi(x_1, x_2, x_3)$ denotes the Jacobian of the transformation and without loss of generality we can take

$$\begin{aligned} E_1'(r) &= \{x_1 \in \mathbb{R}^p, x_2 \in \mathbb{R}^{m-p}, x_3 \in \mathbb{R}^{T-m} : -\varepsilon < x_1'x_1 - x_2'x_2 + h(x_3) < \varepsilon, \\ &\qquad x_1'x_1 - x_2'x_2 + h(x_3) < r - r^*, x_1'x_1 + x_2'x_2 < 2\delta', x_3'x_3 < \delta''\}. \end{aligned}$$

Now, choose $\delta'$, $\delta''$ and $r$ so that $|h(x_3)| < \rho''$, $\rho'' < 2\delta' < \varepsilon - \rho''$, and choose $r$ so that $\rho'' < |r - r^*| < 2\delta' - \rho''$.

By transforming $x_1$ and $x_2$ to polar coordinates, $x_1 = q_1^{1/2} v_1$, $x_2 = q_2^{1/2} v_2$ with $q_1 > 0$, $q_2 > 0$, $v_1'v_1 = 1$, and $v_2'v_2 = 1$. The Jacobian of the transformation is $dx_1 dx_2 = \frac{1}{4} q_1^{\frac{p}{2}-1} q_2^{\frac{m-p}{2}-1} dq_1 dq_2 (dv_1)(dv_2)$, where for all vectors $v$ the quantity $(dv)$ denotes the unnormalized measure over the unit sphere $v'v = 1$. Then,

$$\int_{E_1(r)} \mathrm{pdf}_Y(y)\, dy = \frac{1}{4} \int_{B_1} \int_{B_2(r)} q_1^{\frac{p}{2}-1} q_2^{\frac{m-p}{2}-1} k(q_1, q_2, x_3)\, dq_1 dq_2 dx_3,$$

where

$$\begin{aligned} \mathrm{k}(q_1, q_2, x_3) &= \int_{v_1'v_1=1} \int_{v_2'v_2=1} \mathrm{pdf}_Y\left(y^* + \varphi\left(q_1^{1/2} v_1, q_2^{1/2} v_2, x_3\right)\right) \\ &\qquad J_\varphi\left(q_1^{1/2} v_1, q_2^{1/2} v_2, x_3\right) (dv_1)(dv_2) \end{aligned}$$

is continuous and has continuous derivatives, and

$$\begin{aligned} B_1 &= \{x_3 \in \mathbb{R}^{T-m} : x_3'x_3 < \delta''\} \\ B_2(r) &= \{q_1 > 0, q_2 > 0 : -\varepsilon < q_1 - q_2 + h(x_3) < \varepsilon, \\ &\qquad q_1 - q_2 + h(x_3) < r - r^*, q_1 + q_2 < 2\delta'\}. \end{aligned}$$

Now set $w = \frac{1}{2}(q_1 + q_2)$ and $z = \frac{1}{2}(q_1 - q_2)$, then the Jacobian is 2, and the region $B_2(r)$ has the form indicated in Figure 4.

[Figure 4 approximately here]

So,

$$\begin{aligned}
&\int_{E_1(r)} \mathrm{pdf}_Y\left(y\right) dy \\
= \; &\frac{1}{2}\int_{B_1}\int_{-\delta'}^{\frac{r-r^*-h(x_3)}{2}}\int_{|z|}^{\delta'} \left(z+w\right)^{\frac{p}{2}-1}\left(w-z\right)^{\frac{m-p}{2}-1}\mathrm{k}\left(z+w,w-z,x_3\right)dwdzdx_3,
\end{aligned}$$

and

$$\begin{aligned}
&\frac{\partial}{\partial r}\int_{-\delta'}^{\frac{r-r^*-h(x_3)}{2}}\int_{|z|}^{\delta'} \left(z+w\right)^{\frac{p}{2}-1}\left(w-z\right)^{\frac{m-p}{2}-1}\mathrm{k}\left(z+w,w-z,x_3\right)dwdz \\
= \; &\frac{1}{2}\int_{\left|\frac{r-r^*-h(x_3)}{2}\right|}^{\delta'}\left(\frac{r-r^*-h\left(x_3\right)}{2}+w\right)^{\frac{p}{2}-1}\left(w-\frac{r-r^*-h\left(x_3\right)}{2}\right)^{\frac{m-p}{2}-1} \\
&\mathrm{k}\left(\frac{r-r^*-h\left(x_3\right)}{2}+w,w-\frac{r-r^*-h\left(x_3\right)}{2},x_3\right)dw.
\end{aligned}$$

Let $s=\frac{r-r^*-h(x_3)}{2}$, and note that

$$0<\left|\left|r-r^*\right|-\rho''\right|\leq\left|\left|r-r^*\right|-\left|h\left(x_3\right)\right|\right|\leq 2\left|s\right|\leq\left|r-r^*\right|+\left|h\left(x_3\right)\right|<\left|r-r^*\right|+\rho''.$$

The last integral can be written as

$$\begin{aligned}
I^{\mathrm{k}}_{\frac{p}{2}-1,\frac{m-p}{2}-1}\left(s\right) &= \int_{|s|}^{\delta'}\left(s+w\right)^{\frac{p}{2}-1}\left(w-s\right)^{\frac{m-p}{2}-1}\mathrm{k}\left(s+w,w-s,x_3\right)dw, \\
&= \int_{|s|-s}^{\delta'-s}\left(w+2s\right)^{\frac{p}{2}-1}w^{\frac{m-p}{2}-1}\mathrm{k}\left(w+2s,w,x_3\right)dw
\end{aligned}$$

Since $\mathrm{k}\left(s+w,w-s,x_3\right)$ is continuous, the integral above is also continuous for (i) $p\geq 1$ and $m-p\geq 2$ or (ii) $p\geq 2$ and $m-p\geq 1$.

Assume $0<s<\delta'$ then

$$I^{\mathrm{k}}_{\frac{p}{2}-1,\frac{m-p}{2}-1}\left(s\right)=\int_{0}^{\delta'-s}\left(w+2s\right)^{\frac{p}{2}-1}w^{\frac{m-p}{2}-1}\mathrm{k}\left(w+2s,w,x_3\right)dw$$

so

$$\begin{aligned}
\frac{dI^{\mathrm{k}}_{\frac{p}{2}-1,\frac{m-p}{2}-1}\left(s\right)}{ds} &= \left(\delta'+s\right)^{\frac{p}{2}-1}\left(\delta'-s\right)^{\frac{m-p}{2}-1}\mathrm{k}\left(\delta'+s,\delta'-s,x_3\right) \\
&+\left(p-2\right)\int_0^{\delta'-s}\left(w+2s\right)^{\frac{p}{2}-2}w^{\frac{m-p}{2}-1}\mathrm{k}\left(w+2s,w,x_3\right)dw \\
&+2\int_0^{\delta'-s}\left(w+2s\right)^{\frac{p}{2}-1}w^{\frac{m-p}{2}-1}\left.\frac{\partial\,\mathrm{k}\left(v,w,x_3\right)}{\partial v}\right|_{v=w+2s}dw
\end{aligned}$$

and the integrals are convergent provided the integrands are continuous functions of $w$, and $s$, i.e. $\frac{p}{2}-2\geq 0$, and $\frac{m-p}{2}-1\geq 0$.

Now, let $-\delta'<s<0$, and $s=-t$, $0<t<\delta'$,

$$I^{\mathrm{k}}_{\frac{p}{2}-1,\frac{m-p}{2}-1}(s)=\int_{2t}^{\delta'+t}(w-2t)^{\frac{p}{2}-1}w^{\frac{m-p}{2}-1}\mathrm{k}(w-2t,w,x_3)\,dw,$$

then

$$\begin{aligned}\frac{dI^{\mathrm{k}}_{\frac{p}{2}-1,\frac{m-p}{2}-1}(s)}{ds} &= (\delta'+s)^{\frac{p}{2}-1}(\delta'-s)^{\frac{m-p}{2}-1}\mathrm{k}(\delta'+s,w,x_3)\\ &\quad-(p-2)\int_{2t}^{\delta'+t}(w+2s)^{\frac{p}{2}-2}w^{\frac{m-p}{2}-1}\mathrm{k}(w+2s,w,x_3)\,dw\\ &\quad-2\int_{2t}^{\delta'+t}(w+2s)^{\frac{p}{2}-1}w^{\frac{m-p}{2}-1}\left.\frac{\partial\,\mathrm{k}(v,w,x_3)}{\partial v}\right|_{v=y+2s}dw.\end{aligned}$$

so for $0<|s|<\delta'$,

$$\begin{aligned}\frac{dI^{\mathrm{k}}_{\frac{p}{2}-1,\frac{m-p}{2}-1}(s)}{ds} &= (\delta'+s)^{\frac{p}{2}-1}(\delta'-s)^{\frac{m-p}{2}-1}\mathrm{k}(\delta'+s,y,x_3)\\ &\quad+(p-2)\,\mathrm{sign}(s)\,I^{\mathrm{k}}_{\frac{p}{2}-2,\frac{m-p}{2}-1}(s)\\ &\quad+2\,\mathrm{sign}(s)\,I^{\mathrm{k}_{1,0}}_{\frac{p}{2}-1,\frac{m-p}{2}-1}(s),\end{aligned}$$

where

$$\mathrm{k}_{i,j}(s+w,w-s,x_3)=\left.\frac{\partial^{i+j}\mathrm{k}(t_1,t_2,x_3)}{\partial t_1^i\partial t_2^j}\right|_{\substack{t_1=s+w\\ t_2=w-s}}.$$

Note that

$$\lim_{s\to 0^+}I^{\mathrm{k}}_{\frac{p}{2}-1,\frac{m-p}{2}-1}(s)=\lim_{s\to 0^-}I^{\mathrm{k}}_{\frac{p}{2}-1,\frac{m-p}{2}-1}(s).$$

So $\frac{dI^{\mathrm{k}}_{\frac{p}{2}-1,\frac{m-p}{2}-1}(s)}{ds}$ can be differentiated again. What is left is a linear combination of differentiable terms and integrals still of the form

$$\begin{aligned}I^{\mathrm{k}_{i,0}}_{a,\frac{m-p}{2}-1}(s) &= \int_{|s|}^{\delta'}(s+w)^a(w-s)^{\frac{m-p}{2}-1}\mathrm{k}_{i,0}(s+w,w-s,x_3)\,dw,\\ &= \int_{|s|+s}^{\delta'+s}w^a(w-2s)^{\frac{m-p}{2}-1}\mathrm{k}_{i,0}(w,w-2s,x_3)\,dw\end{aligned}$$

where the smallest possible $a$ is 0 or $1/2$. Each of these terms can be differentiated again in the same way as above. So the $I^{\mathrm{k}}_{\frac{p}{2}-1,\frac{m-p}{2}-1}(s)$ can be differentiated $i+j$

times, where $i$ and $j$ satisfy (i) $p/2-1-j \geq 0$ and $\frac{m-p}{2}-1-i \geq -1/2$, or (ii) $p/2-1-j \geq -1/2$ and $\frac{m-p}{2}-1-i \geq 0$. This gives in the first case $i+j=[p/2-1]+[(m-p-1)/2]=[(m-1)/2]$, and in the second case $i+j=[(p-1)/2]+[(m-p)/2]=[(m-1)/2]$.

# Appendix B: Proofs of results in Section 2.2

**Proof of Theorem 3** The density of a positive definite quadratic form in normal variables, $Q_i$, can be found in James (1964) (page 494),

$$\mathrm{pdf}_{Q_i}(q_i) = \left[2^{\frac{n_i}{2}}\Gamma\left(\frac{n_i}{2}\right)|D_i|^{\frac{1}{2}}\right]^{-1} q_i^{\frac{n_i}{2}-1} {}_1F_1\left(\frac{1}{2};\frac{n_i}{2};-\frac{1}{2}q_iD_i^{-1}\right), \qquad i=1,2, \tag{15}$$

where the hypergeometric function has matrix argument (Muirhead (1982)).

Using this result the CDF of a ratio of quadratic forms in normal variables can be written as

$$\begin{aligned} \mathrm{F}_Q(q) &= c_1 \int_{0<q_1<q_2}\int_{q_2>0} q_1^{\frac{n_1}{2}-1} {}_1F_1\left(\frac{1}{2};\frac{n_1}{2};-\frac{1}{2}q_1D_1^{-1}\right) \\ &\quad q_2^{\frac{n_2}{2}-1} {}_1F_1\left(\frac{1}{2};\frac{n_2}{2};-\frac{1}{2}q_2D_2^{-1}\right) dq_1dq_2, \end{aligned}$$

where

$$c_1 = \left[2^{\frac{n_1+n_2}{2}}\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)|D_1|^{\frac{1}{2}}|D_2|^{\frac{1}{2}}\right]^{-1}.$$

Transforming $q_1$ to $q_1 = xq_2$, the variable $x$ varies in $(0,1)$ and $q_2 > 0$:

$$\begin{aligned} \mathrm{F}_Q(q) &= c_1 \int_{q_2>0} {}_1F_1\left(\frac{1}{2};\frac{n_2}{2};-\frac{1}{2}q_2D_2^{-1}\right) q_2^{\frac{n_1+n_2}{2}-1} \\ &\quad \int_{0<x<1} x^{\frac{n_1}{2}-1} {}_1F_1\left(\frac{1}{2};\frac{n_1}{2};-\frac{1}{2}q_2xD_1^{-1}\right) dxdq_2. \end{aligned} \tag{16}$$

The integral over $0<x<1$ can be evaluated by expanding the hypergeometric function and integrating term by term. This yields

$$\frac{\Gamma\left(\frac{n_1}{2}\right)}{\Gamma\left(\frac{n_1}{2}+1\right)} {}_1F_1\left(\frac{1}{2};\frac{n_1}{2}+1;-\frac{1}{2}q_2D_1^{-1}\right).$$

Using this in (16) and transforming the hypergeometric functions using the Kummer transformation delivers

$$\begin{aligned} \mathrm{F}_Q(q) &= \frac{c_1\Gamma\left(\frac{n_1}{2}\right)}{\Gamma\left(\frac{n_1}{2}+1\right)} \int_{q_2>0} \exp\left\{-\frac{1}{2}q_2\left[\mathrm{tr}\left(D_2^{-1}\right)+\mathrm{tr}\left(D_1^{-1}\right)\right]\right\} q_2^{\frac{n_1+n_2}{2}-1} \\ &\quad {}_1F_1\left(\frac{n_2-1}{2};\frac{n_2}{2};\frac{1}{2}q_2D_2^{-1}\right) {}_1F_1\left(\frac{n_1+1}{2};\frac{n_1}{2}+1;\frac{1}{2}q_2D_1^{-1}\right) dq_2. \end{aligned} \tag{17}$$

The final integral can be evaluated by expanding the hypergeometric functions and integrating term by term. Term by term integration of the infinite series can be justified by repeated use of Hardy's theorem (see for example Titchmarsh (1993), page 47). Note that the series in the resulting expression (7) is convergent because

$$\begin{aligned}\frac{D_2^{-1}}{\operatorname{tr}\left(D_2^{-1}\right)+\operatorname{tr}\left(D_1^{-1}\right)} &< I_{n_2} \\ \frac{D_1^{-1}}{\operatorname{tr}\left(D_2^{-1}\right)+\operatorname{tr}\left(D_1^{-1}\right)} &< I_{n_1}.\end{aligned}$$

**Proof of Theorem 4** An alternative expression for the density of a positive definite quadratic form in normal variables is

$$\begin{aligned}\operatorname{pdf}\left(q_i\right) &= \frac{1}{2^{\frac{n_1}{2}}\Gamma\left(\frac{n_1}{2}\right)\left|D_i\right|^{\frac{1}{2}}}\exp\left\{-\frac{1}{2}q_i\operatorname{tr}\left(D_i^{-1}\right)\right\} \\ & q_i^{\frac{n_1}{2}-1}{}_1F_1\left(\frac{1}{2};\frac{n_i}{2};-\frac{1}{2}q_i\left(D_i^{-1}-I_{n_i}\operatorname{tr}\left(D_i^{-1}\right)\right)\right) \quad i=1,2.\end{aligned}$$

The derivation of this expression is the same as the derivation of the formula given by James (1964). The difference comes from writing $D_i^{-1}=\left(\operatorname{tr}D_i^{-1}\right)I_{n_i}-\left[\left(\operatorname{tr}D_i^{-1}\right)I_{n_i}-D_i^{-1}\right]$ before integrating over the unit $n_i$-sphere.

The CDF of $Q$ is

$$\begin{aligned}\mathrm{F}_Q\left(q\right) &= c_1\int_{0<q_1<q_2}\int_{q_2>0}\exp\left\{-\frac{1}{2}q_1\operatorname{tr}\left(D_1^{-1}\right)-\frac{1}{2}q_2\operatorname{tr}\left(D_2^{-1}\right)\right\}q_2^{\frac{n_2}{2}-1}q_1^{\frac{n_1}{2}-1} \\ & {}_1F_1\left(\frac{1}{2};\frac{n_1}{2};-\frac{1}{2}q_1\left(D_1^{-1}-I_{n_1}\operatorname{tr}\left(D_1^{-1}\right)\right)\right) \\ & {}_1F_1\left(\frac{1}{2};\frac{n_2}{2};-\frac{1}{2}q_2\left(D_2^{-1}-I_{n_2}\operatorname{tr}\left(D_2^{-1}\right)\right)\right)dq_1dq_2,\end{aligned}$$

where

$$c_1=\left[2^{\frac{n_1+n_2}{2}}\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)\left|D_1\right|^{\frac{1}{2}}\left|D_2\right|^{\frac{1}{2}}\right]^{-1}$$

Let $q_1 = q_2 x$, then

$$\begin{aligned}
\mathrm{F}_Q(q) &= c_1 \int_{0<x<1} \int_{q_2>0} \exp\left\{-\frac{1}{2} x q_2 \operatorname{tr}\left(D_1^{-1}\right)\right\} \\
& \quad x^{\frac{n_1}{2}-1} {}_1F_1\left(\frac{1}{2}; \frac{n_1}{2}; -\frac{1}{2} x q_2 \left(D_1^{-1} - I_{n_1} \operatorname{tr}\left(D_1^{-1}\right)\right)\right) \\
& \quad \exp\left\{-\frac{1}{2} q_2 \operatorname{tr}\left(D_2^{-1}\right)\right\} q_2^{\frac{n_1+n_2}{2}-1} \\
& \quad {}_1F_1\left(\frac{1}{2}; \frac{n_2}{2}; -\frac{1}{2} q_2 \left(D_2^{-1} - I_{n_2} \operatorname{tr}\left(D_2^{-1}\right)\right)\right) dq_2 dx.
\end{aligned}$$

The integral over $0 < x < 1$ can be evaluated by expanding the first hypergeometric function:

$$\begin{aligned}
& \int_{0<x<1} \exp\left\{-\frac{1}{2} x q_2 \operatorname{tr}\left(D_1^{-1}\right)\right\} x^{\frac{n_1}{2}-1} \\
& {}_1F_1\left(\frac{1}{2}; \frac{n_1}{2}; -\frac{1}{2} x q_2 \left(D_1^{-1} - I_{n_1} \operatorname{tr}\left(D_1^{-1}\right)\right)\right) dx \\
= \; & \frac{\Gamma\left(\frac{n_1}{2}\right)}{\Gamma\left(\frac{n_1}{2}+1\right)} \sum_{j=0}^{\infty} \frac{\left(\frac{1}{2}\right)_j}{\left(\frac{n_1}{2}+1\right)_j j!} q_2^j C_{[j]}\left(-\frac{1}{2}\left(D_1^{-1} - I_{n_1} \operatorname{tr}\left(D_1^{-1}\right)\right)\right) \\
& {}_1F_1\left(j + \frac{n_1}{2}; j + \frac{n_1}{2} + 1, -\frac{1}{2} q_2 \operatorname{tr}\left(D_1^{-1}\right)\right)
\end{aligned}$$

and using Kummer transformation this becomes

$$\begin{aligned}
& \frac{\Gamma\left(\frac{n_1}{2}\right)}{\Gamma\left(\frac{n_1}{2}+1\right)} \sum_{j=0}^{\infty} \frac{\left(\frac{1}{2}\right)_j}{\left(\frac{n_1}{2}+1\right)_j j!} C_{[j]}\left(-\frac{1}{2}\left(D_1^{-1} - I_{n_1} \operatorname{tr}\left(D_1^{-1}\right)\right)\right) \\
& \exp\left\{-\frac{1}{2} q_2 \operatorname{tr}\left(D_1^{-1}\right)\right\} q_2^j \, {}_1F_1\left(1; j + \frac{n_1}{2} + 1, \frac{1}{2} q_2 \operatorname{tr}\left(D_1^{-1}\right)\right)
\end{aligned}$$

Substituting back into the CDF

$$\begin{aligned}
\mathrm{F}_Q(q) &= \frac{1}{2^{\frac{n_1+n_2}{2}}\Gamma\left(\frac{n_2}{2}\right)\Gamma\left(\frac{n_1}{2}+1\right)|D_1|^{\frac{1}{2}}|D_2|^{\frac{1}{2}}}\int_{q_2>0}\sum_{j=0}^{\infty}\frac{\left(\frac{1}{2}\right)_j}{\left(\frac{n_1}{2}+1\right)_j j!} \\
& \quad C_{[j]}\left(-\frac{1}{2}\left(D_1^{-1}-I_{n_1}\operatorname{tr}\left(D_1^{-1}\right)\right)\right) \\
& \quad \exp\left\{-\frac{1}{2}q_2\left[\operatorname{tr}\left(D_2^{-1}\right)+\operatorname{tr}\left(D_1^{-1}\right)\right]\right\}q_2^{\frac{n_1+n_2}{2}+j-1} \\
& \quad {}_1F_1\left(j+\frac{n_1}{2};j+\frac{n_1}{2}+1,\frac{1}{2}q_2\operatorname{tr}\left(D_1^{-1}\right)\right) \\
& \quad {}_1F_1\left(\frac{1}{2};\frac{n_2}{2};-\frac{1}{2}q_2\left(D_2^{-1}-I_{n_2}\operatorname{tr}\left(D_2^{-1}\right)\right)\right)dq_2.
\end{aligned}$$

Expanding the second hypergeometric function and integrating term by term we have

$$\begin{aligned}
\mathrm{F}_Q(q) &= \frac{1}{2^{\frac{n_1+n_2}{2}}\Gamma\left(\frac{n_1}{2}+1\right)\Gamma\left(\frac{n_2}{2}\right)|D_1|^{\frac{1}{2}}|D_2|^{\frac{1}{2}}}\sum_{j=0}^{\infty}\sum_{p=0}^{\infty}\frac{\left(\frac{1}{2}\right)_j\left(\frac{1}{2}\right)_p}{\left(\frac{n_1}{2}+1\right)_j\left(\frac{n_2}{2}\right)_p j!p!} \\
& \quad C_{[j]}\left(-\frac{1}{2}\left(D_1^{-1}-I_{n_1}\operatorname{tr}\left(D_1^{-1}\right)\right)\right)C_{[p]}\left(-\frac{1}{2}\left(D_2^{-1}-I_{n_2}\operatorname{tr}\left(D_2^{-1}\right)\right)\right) \\
& \quad \int_{q_2>0}\exp\left\{-\frac{1}{2}q_2\left[\operatorname{tr}\left(D_2^{-1}\right)+\operatorname{tr}\left(D_1^{-1}\right)\right]\right\} \\
& \quad q_2^{p+j+\frac{n_1+n_2}{2}-1}\,{}_1F_1\left(1;j+\frac{n_1}{2}+1,\frac{1}{2}q_2\operatorname{tr}\left(D_1^{-1}\right)\right)dq_2.
\end{aligned}$$

Term by term integration of the infinite series is possible since the resulting series is convergent. Evaluating the last integral we have the expression given in Theorem 4.

# Appendix C: Proofs of results in Section 3.1

**Proof of Corollary 1.** Let $e_T$ be an $T$-dimensional vector having all components equal to zero apart from the last one which is one. Note that $\hat{r} = \frac{v' A_T v}{v' B_T v}$. Therefore transforming $v$ to $v = e_T \cos\theta + \Lambda_1 v_1 \sin\theta$, where $\Lambda_1 = (I_{T-1}, 0)'$ is a $T \times T-1$ matrix; $0 < \theta < \pi$ and $v_1$ is an $(T-1)$-dimensional vector satisfying $v_1' v_1 = 1$. Then, in terms of the new coordinates $\hat{r}$ is

$$\hat{r} = v_1' A_{T-1} v_1 + e_{T-1}' v_1 \cot\theta \tag{18}$$

The critical points of $\hat{R}$ satisfies

$$\begin{aligned} A_{T-1} v_1 + e_{T-1} \cot\theta - \lambda v_1 &= 0 \\ e_{T-1}' v_1 &= 0 \\ v_1' v_1 &= 1. \end{aligned}$$

The second condition implies that $v_1$ has the form

$$v_1 = \begin{pmatrix} I_{T-1} \\ 0 \end{pmatrix} v_2$$

where $v_2$ is a $(T-2)$-dimensional vector such that $v_2' v_2 = 1$. Imposing this restriction, the first condition becomes

$$A_{T-1} \begin{pmatrix} I_{T-1} \\ 0 \end{pmatrix} v_2 - \lambda \begin{pmatrix} I_{T-1} \\ 0 \end{pmatrix} v_2 = 0.$$

This implies that

$$A_{T-2} v_2 - \lambda v_2 = 0$$

so that $\lambda$ is an eigenvalue of $A_{T-2}$ and $v_2$ is the corresponding normalized eigenvector. Substituting these values back into (18), and imposing the condition $e_{T-1}' v_1 = 0$, we obtain

$$\hat{r} = v_2' A_{T-2} v_2$$

so that the critical values are the eigenvalues of $A_{T-2}$, i.e. $\hat{r}_k^* = \cos\left(\frac{k\pi}{T-1}\right)$, $k = 1, 2..., T-2$.

Now, notice that since $\hat{R}$ is defined on the set

$$\begin{aligned} M &= \left\{ v \in \mathbb{R}^T : -\infty < \frac{v' A_T v}{v' B_T v} < \infty \cap v'v = 1 \right\} \\ &= \left\{ v \in \mathbb{R}^T : v'v = 1 \cap v \neq \pm e_T \right\} \end{aligned}$$

which is an open submanifold of $\mathbb{R}^T$. We need to prove that the set

$$\mathrm{R}^{-1}\left([r_0, r]\right) = \left\{ v \in M : r_0 \leq \frac{v' A_T v}{v' B_T v} \leq r \right\}$$

is a compact subset of $M$. To do this note that $M$ is a submanifold of $\mathbb{R}^T$ and that all compact subsets of $M$ are precisely the sets of the form $K \cap M$, for $K$ compact in $\mathbb{R}^T$. Note that $\mathrm{R}^{-1}\left([r_0, r]\right) \cup \{\pm e_T\}$ is a compact subset of $\mathbb{R}^T$, and that

$$\mathrm{R}^{-1}\left([r_0, r]\right) = \left[ R^{-1}\left([r_0, r]\right) \cup \{\pm e_T\} \right] \cap M.$$

So $\mathrm{R}^{-1}\left([r_0, r]\right)$ is compact. It then follows from Theorem 1 that the density of $\hat{R}$ is analytic everywhere apart from a neighbourhood of the critical points.

It follows from Theorem 2 that the order of differentiability of the cumulative distribution function of $\hat{R}$ is between 1 and $[(T-2)/2]$.

**Proof of Corollary 2.** The proof of Corollary 2 is similar to the proof of Corollary 1.

# Appendix D: Proofs of results in Section 3.2

**Proof of Lemma 1** It can be easily seen that

$$|A_T - \hat{r}B_T| = \frac{1}{4}\left|A_{T-2} - \hat{r}I_{T-2}\right|$$

which is zero for $\hat{r} = \hat{r}_k^* = \cos\left(\frac{k\pi}{T-1}\right)$, $k = 1, 2, ..., T-2$. So rank $(A_T - \hat{r}B_T) = T$ for $\hat{r} \neq \cos\left(\frac{k\pi}{T-1}\right)$, $k = 1, 2, ..., T-2$. Now, rank $\left(A_T - \cos\left(\frac{k\pi}{T-1}\right)B_T\right) = T-1$ because the submatrix obtained from $A_T - \cos\left(\frac{k\pi}{T-1}\right)B_T$ by deleting the last row and the last column is

$$\left|A_{T-1} - \cos\left(\frac{k\pi}{T-1}\right)I_{T-1}\right|$$

and this is not zero, because the polynomials $p_T(\lambda) = |A_T - \lambda I_T|$ have the Sturm sequence property, so that $p_T(\lambda)$ and $p_{T+1}(\lambda)$ do not have any zero in common.

**Proof of Lemma 2** Let $P_T = 2(A_T - \hat{r}B_T)$ and $p_T(\lambda) = |P_T - \lambda I_T|$ then, $p_T(\lambda)$ satisfies the recursion:

$$p_T(\lambda) = -(2\hat{r} + \lambda)\, p_{T-1}(\lambda) - p_{T-2}(\lambda).$$

Setting $-(2\hat{r} + \lambda) = 2\cos\theta$, and solving the recursion

$$p_T(\lambda) = 2(\cos\theta)\, p_{T-1}(\lambda) - p_{T-2}(\lambda)$$

subject to the initial conditions

$$\begin{aligned} p_0(\lambda) &= 1 \\ p_1(\lambda) &= -\lambda = 2\hat{r} + 2\cos\theta, \end{aligned}$$

requires finding the roots of the second order equation

$$z^2 - 2(\cos\theta)\, z + 1 = 0.$$

This gives:

$$\begin{aligned} z_1 &= \cos\theta + i\sin\theta = e^{i\theta} \\ z_2 &= \cos\theta - i\sin\theta = e^{-i\theta} \end{aligned}$$

for $\theta \neq k\pi$, $k = ... - 2, -1, 0, 1, 2, .....$ Thus

$$p_T(\lambda) = \frac{\sin((T+1)\theta) + 2\hat{r}\sin(T\theta)}{\sin\theta},$$

so that $\theta$ may be found by finding the zeros of $p_T(\lambda)$.

The statement of the lemma follows by noting that the eigenvalues of $A_T - \hat{r}B_T$ are $\lambda/2$. Note that $\theta \neq k\pi$.

The second part of the lemma follows from noting that $\frac{\sin((T+1)\theta)+2\hat{r}\sin(T\theta)}{\sin\theta} = 0$ has $T$ real roots if $|2\hat{r}| \leq 1 + 1/T$. If $|2\hat{r}| > 1 + 1/T$, there are only T-1 real roots. The other root can be found by setting $\theta = k\pi + ia$, $x, a \in \mathbb{R}$, $k = ..., -1, 0, 1, ...,$ and writing $p_T(\lambda)$ as:

$$\frac{(-1)^{k(T+1)}\sinh((T+1)a) + 2\hat{r}(-1)^{Tk}\sinh(Ta)}{\sinh a},$$

This can be rearranged to give

$$\frac{\sinh((T+1)a)}{\sinh(Ta)} = \operatorname{sign}(\hat{r})\, 2\hat{r}.$$

Note that

$$\frac{\sinh((T+1)a)}{\sinh(Ta)} \geq 1 + \frac{1}{T}$$

and that it is symmetric around $a = 0$.

**Proof of Lemma 3** The polynomials $p_T(\lambda)$ have the Sturm sequence property, so that the zeros of $p_T(\lambda)$ separate those of $p_{T+1}(\lambda)$. Thus we can focus on $T = 2$ and show that $p_2(\lambda)$ has one positive and one negative zero.

Only the case where $\hat{r}$ is positive will be considered, since the proof is the same for the case where $\hat{r} < 0$.

For $T = 2$ the eigenvalues of $A_2 - \hat{r}B_2$ are $\lambda = -\hat{r} - \cos\theta$, where $\theta$ solves

$$\frac{\sin(3\theta)}{\sin(2\theta)} = \frac{1}{2}(3\cos\theta - \sin\theta\tan\theta) = -2\hat{r}.$$

If $\hat{r} \geq 0$, then the smallest eigenvalue will be obtained for $\theta \in \left(\frac{\pi}{3}, \frac{\pi}{2}\right)$. In this range of values for $\theta$ the inequalities $\sin\theta\tan\theta \geq 0$ and $\cos\theta > 0$ are verified. So

$$2\cos\theta > \frac{1}{2}3\cos\theta \geq -2\hat{r}$$

which give $\lambda = -\hat{r} - \cos\theta \leq 0$.

To show that if $\hat{r} \geq 0$, then the largest eigenvalue will be nonnegative we need to consider two cases.

Case 1: $2\hat{r} \leq 1 + 1/T$. The largest eigenvalues is obtained by taking $\theta$ in the interval $\left[\frac{2\pi}{3}, \pi\right]$. In this range $\sin\theta\tan\theta \leq 0$, so that

$$\frac{1}{2}3\cos\theta \leq -2\hat{r},$$

and using the fact that $\cos\theta \leq 0$, it follows that $\lambda = -\hat{r} - \cos\theta \geq 0$.

Case 2: $2\hat{r} > 1 + 1/T$, then $\theta$ has the form $\theta = \pi + ix$. Thus $\theta$ solves

$$\frac{1}{2}\left(3\cosh x + \sinh x \tanh x\right) = 2\hat{r}$$

and $\lambda = -\hat{r} + \cosh(x)$. Noting that $\sinh x \tanh x \leq \cosh x$ gives $\lambda \geq 0$, and the lemma follows.

Figures

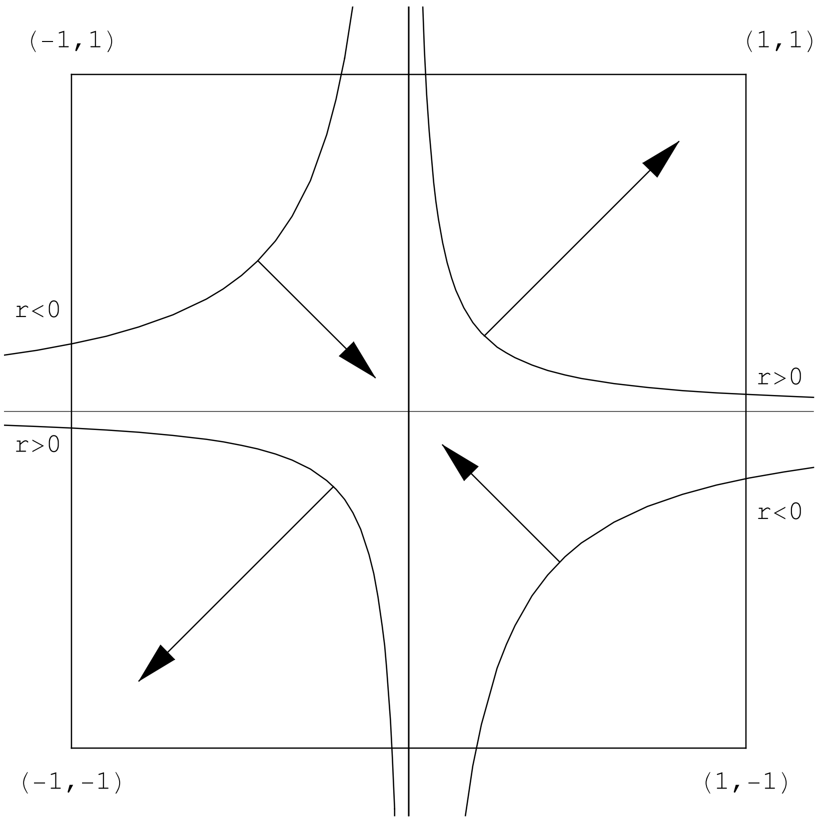


**Figure 1.** Level surfaces of the function $r = xy$. The arrows point in the direction at which $r$ increases.

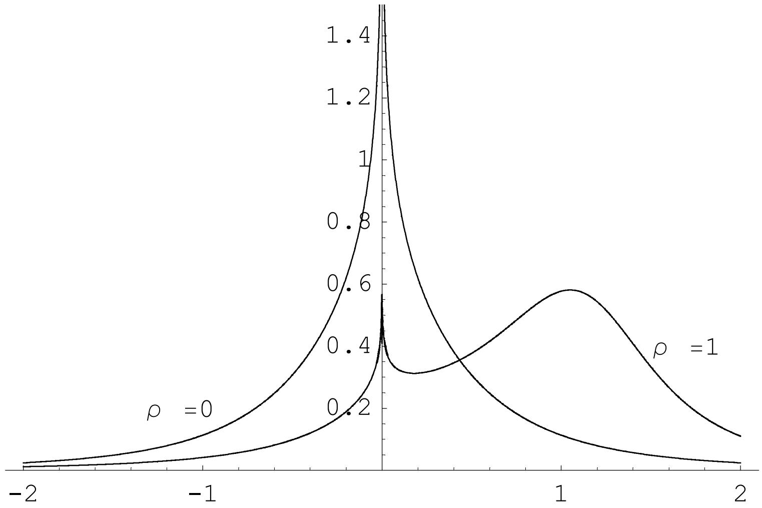


**Figure 2**. Density of $\hat{R}$ or $T = 3$ and $\rho = 0$ and $\rho = 1$. The critical level is at $\hat{r} = 0$.

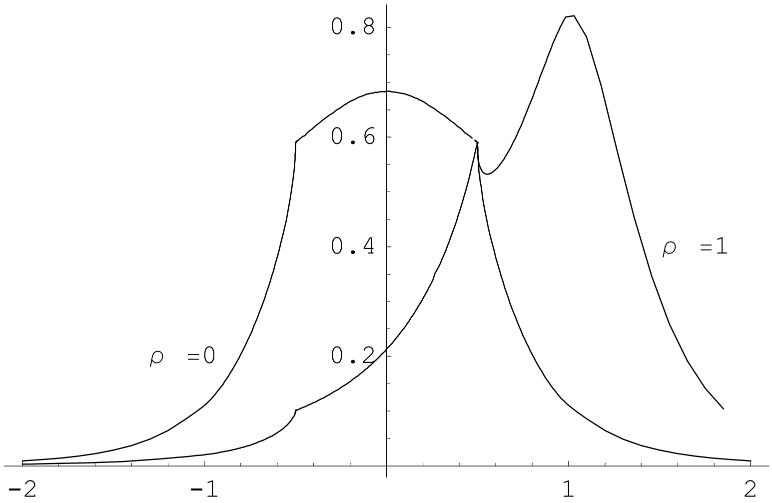


**Figure 3**. Density of $\hat{R}$ for $T = 4$ and $\rho = 0$ and $\rho = 1$. The critical level is at $\hat{r} = \frac{1}{2}, \frac{1}{2}$.

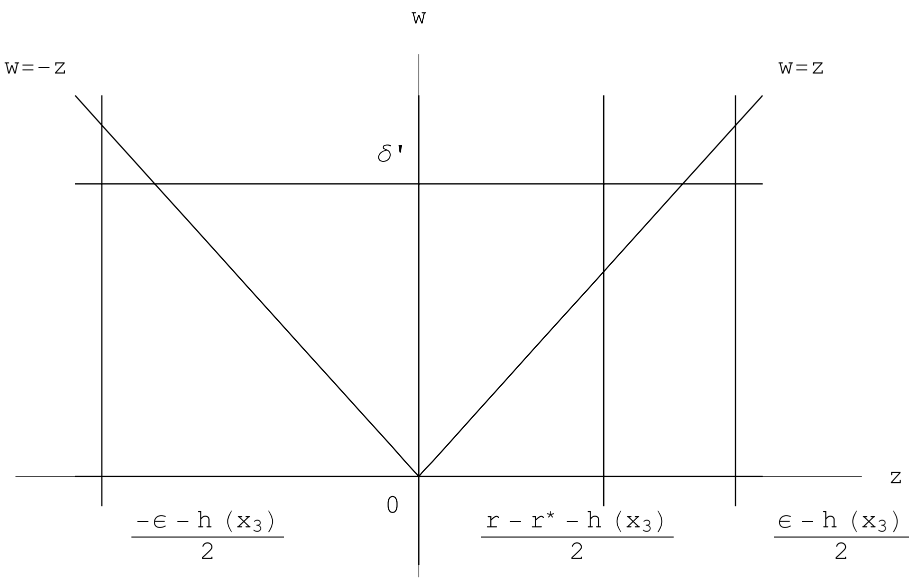


**Figure 4.** The region $B_2(r)$.