Chilean Journal of Statistics

Vol. xx, No. x, Month 20xx, 1–31

Stochastic representations and a geometric parametrization of the two-dimensional Gaussian law

Thomas Dietrich[1], Steve Kalke[1] and Wolf-Dieter Richter[1]

[1]Institute of Mathematics, University of Rostock, Rostock, Germany

(Received: 13 November 2012 · Accepted in final form: 06 August 2013)

Abstract

Using different types of polar and elliptical polar coordinates, different stochastic representations of the axis-aligned and the regular two-dimensional Gaussian distribution are derived. Advantages and disadvantages of these stochastic representations are discussed. The non-Euclidean geometric measure representation of the axis-aligned two-dimensional Gaussian distribution in Richter (2011) is taken to derive a new geometric interpretation of the correlation coefficient and to motivate a new geometric parametrization of the regular Gaussian law. Estimators of the new parameters and corresponding distributions are derived. A comparison with different approaches from the literature shows the numerical stability of our results.

Keywords: Stochastic representations · random coordinates · maximum likelihood estimation · exact distributions · correlation.

Mathematics Subject Classification: Primary 60D05 · Secondary 62E15 · 62F10 · 62H10 · 62H20.

1. Introduction

Gaussian random vectors and their distributions are among the most often used ones in probability theory and mathematical statistics, see, e.g., Anderson (2003), Cambanis et al. (1981), Fang et al. (1990), Muirhead (1982) or Tong (1990). It may be surprising therefore that even the two-dimensional case is not yet completely studied with respect to all its basic properties. Similarly, the correlation coefficient is one of the original notions expressing dependence properties not only for the components of Gaussian random vectors and plays even these days an important role in applications like finance and insurance, see, e.g., Embrechts et al. (2002). Clearly, rotations of two-dimensional Gaussian random vectors affect in general the correlation coefficient as well as the variances, see, e.g., Pyati (1993). But even in the case of two correlated random variables following a joint normal distribution it is not yet completely revealed which geometric properties of the random vector and its distribution this coefficient reflects. The non-Euclidean geometric measure representation of the two-dimensional heteroscedastic axis-aligned Gaussian law recently derived in Richter (2011) stimulates some new considerations.

The focus of the present paper is first on stochastic representations of two-dimensional Gaussian random vectors and second on a suitable parametrization of their distributions.

*Emails: thomas.dietrich@uni-rostock.de — steve.kalke@uni-rostock.de — wolf-dieter.richter@uni-rostock.de

ISSN: 0718-7912 (print)/ISSN: 0718-7920 (online)

It can be a fruitful way to analyze continuous random vectors by making use of coordinates other than Cartesian ones which reflect certain basic geometric properties of the density level sets. For example, studying spherical distributions the usual polar coordinates lead to a geometric measure representation in Richter (1991) which enabled the authors in Kalke et al. (2013) to derive the exact distribution of, e.g., the component-by-component product of a continuous two-dimensional spherical distribution and to contribute this way to studies on statistics of an elliptically contoured sample like the one in Nadarajah and Gupta (2005). Moreover, suitable generalizations of polar coordinates were introduced for studying the multivariate p-generalized homoscedastic normal (or power exponential) distribution and its $l_{n,p}$-symmetric generalizations in Richter (2007, 2009), for studying the heteroscedastic uncorrelated Gaussian case and its elliptically contoured generalizations in Richter (2011, 2013), and simplex coordinates were introduced in Henschel and Richter (2002) for studying the multivariate exponential distribution and its simplicially contoured generalizations. In case of the $l_{n,p}$-symmetric distributions, the generalized polar coordinates lead to stochastic representations which were used to extend the class of skewed distributions in Arellano-Valle and Richter (2012) and which also were used to generalize the Box-Muller method in context of a simulation of the p-generalized normal distribution in Kalke and Richter (2013). Again, it may be surprising that a standard mathematical method, which was developed in earlier centuries, may even nowadays essentially contribute to a real scientific novelty. However, one of the deep problems arizing in this area of mathematical work was described in Szablowski (1998) as to find suitable coordinates. It has been solved in Richter (2007, 2009) and developed further, e.g., in Richter (2011, 2013).

To be more concrete, imagine we are given a two-dimensional Gaussian random vector $(\xi, \eta)^T$ centered at the origin with heteroscedastic and independent components and denote its polar coordinates by $\mathcal{R}_P$ and Φ_P. It is known that $\mathcal{R}_P$ and Φ_P are not independent. The distributions of $\mathcal{R}_P$ and Φ_P may be derived explicitly, but that of $(\cos \Phi_P, \sin \Phi_P)^T$ cannot be easily interpreted. Introducing elliptical polar coordinates $\mathcal{R}_{EP}$ and Φ_{EP} instead involves the advantages that $\mathcal{R}_{EP}$ and Φ_{EP} are stochastically independent, $\mathcal{R}_{EP}$ has a more simple distribution than $\mathcal{R}_P$ and that Φ_{EP} is even uniformly distributed. However, the angle Φ_{EP} itself cannot be easily interpreted and the so called uniform basis $\mathcal{W}_{EP}$ from the stochastic representation $(\xi, \eta)^T \overset{d}{=} \mathcal{R}_{EP} \cdot \mathcal{W}_{EP}$, which takes its values on a corresponding axis-aligned ellipse $E_{(a,b)}$ with main axes of lengths $2a$ and $2b$, does not follow a uniform distribution on the ellipse with respect to the Euclidean arc length.

One can think about representing $(\xi, \eta)^T$ with the help of other kinds of coordinates possessing other stochastical properties. In this paper, we shall make use of the $E_{(a,b)}$-generalized polar coordinates $\Phi_{GEP} = \Phi_P$ and $\mathcal{R}_{GEP} = \mathcal{R}_{EP}$ having the most favorable properties among all the types of coordinates considered here: Φ_{GEP} and $\mathcal{R}_{GEP}$ are independent, their distributions can be derived explicitly and both random variables as well as the corresponding basis vector $\mathcal{U}_{GEP} = (a \cos_{(a,b)} \Phi_{GEP}, b \sin_{(a,b)} \Phi_{GEP})^T$ are suitable to describe the geometry behind the axis-aligned two-dimensional Gaussian law. Here, $\cos_{(a,b)}$ and $\sin_{(a,b)}$ denote certain generalized cosine and sine functions which have been proved to be very useful for deriving geometric measure representations on the basis of non-Euclidean geometry in Richter (2011). In this context, $\mathcal{R}_{GEP}$ can be interpreted as a certain length and Φ_{GEP} can be interpeted as a certain direction of a centered bivariate Gaussian distributed random vector with independent components. Furthermore, the associated basis vector $\mathcal{U}_{GEP}$ follows the $E_{(a,b)}$-generalized uniform distribution on $E_{(a,b)}$ introduced in Richter (2011) as well.

The paper is structured as follows. A detailed comparative study of the mentioned coordinates and their extension to the arbitrary regular Gaussian case will be made in Section 2. It will turn out in Section 3 that the same non-Euclidean geometry which plays a fundamental role in Section 2 allows to formulate a new look onto the well known notion of the correlation coefficient of a regular Gaussian distribution. All new insights gained this way will give the motivation to introduce in Section 4 a parametrization of the regular Gaussian distribution which reflects on the one hand the new non-Euclidean geometry mentioned above and which is on the other hand closely connected with the well known principal axes theory and principal component theory. For general work in the latter area we refer to Anderson (2003), Hyvärinen et al. (2001), Jolliffe (2004) and Muirhead (1982). Maximum likelihood estimates of the new parameters and some exact distributions are dealt with in Sections 5 and 6, respectively. Section 7 presents an example and an outlook on future work. The Appendix gives the proof of Theorem 6.1 which is omitted in the main text of the paper.

2. Stochastic representations

2.1 The multiple standard case

Let $(\xi, \eta)^T$ be distributed according to the two-dimensional normal distribution with expectation $0_2 = (0,0)^T$ and with covariance matrix $\sigma^2 I_2$ being a multiple of the unit matrix $I_2 = \mathrm{diag}(1,1)$, $\sigma > 0$, i.e. $(\xi, \eta)^T \sim \Phi_{0_2, \sigma^2 I_2}$. It is well known that such a vector allows the representation

$$\begin{pmatrix} \xi \\ \eta \end{pmatrix} \stackrel{d}{=} \sigma \cdot \mathcal{R} \cdot \mathcal{U}$$

where $\mathcal{R}$ follows a chi-distribution with two degrees of freedom (d.f.), $\mathcal{R} \sim \chi_2$, and where $\mathcal{U}$ is independent from $\mathcal{R}$ and follows the uniform distribution on the unit circle $C = \{(x,y)^T \in \mathrm{R}^2 : x^2 + y^2 = 1\}$, $\mathcal{U} \sim \omega_C$. Making use of random polar coordinates for $\frac{1}{\sigma}(\xi, \eta)^T$, we have

$$\sigma \cdot \mathcal{R} = \sqrt{\xi^2 + \eta^2}\,, \quad \mathcal{U} = \begin{pmatrix} \cos\Phi \\ \sin\Phi \end{pmatrix}\,,$$

where $\Phi = f_1(\xi,\eta) \arctan|\eta/\xi| + \pi f_2(\xi,\eta)$ with $f_1(\xi,\eta) = 1_{Q_1 \cup Q_3}(\xi,\eta) - 1_{Q_2 \cup Q_4}(\xi,\eta)$ and $f_2(\xi,\eta) = 1_{Q_2 \cup Q_3}(\xi,\eta) + 2 \cdot 1_{Q_4}(\xi,\eta)$. Here, $\mathcal{R}$ and Φ are independent, Q_i denotes as usual the i-th quadrant of R^2, $i \in \{1, \ldots, 4\}$, and 1_A is the indicator function of the set A.

One may consider the representation

$$\frac{1}{\sigma} \cdot \begin{pmatrix} \xi \\ \eta \end{pmatrix} \stackrel{d}{=} \mathcal{R} \cdot \begin{pmatrix} \cos\Phi \\ \sin\Phi \end{pmatrix} \tag{1}$$

as corresponding to the random polar coordinate representation of $\frac{1}{\sigma}(\xi,\eta)^T$. Here, the independent random variables $\mathcal{R}$ and Φ follow the joint density

$$f_{(\mathcal{R},\Phi)}(r,\varphi) = \frac{1}{2\pi} r\, e^{-\frac{r^2}{2}}\,, \quad r > 0,\ 0 \le \varphi < 2\pi\,.$$

It is well known that a uniformly on the interval $[0,1)$ distributed random variable U, $U \sim \omega_{[0,1)}$, and a standard Cauchy distributed random variable Z are connected by $F^{-1}(U) \stackrel{d}{=} Z$ where F is the cumulative distribution function (cdf) of the standard Cauchy distribution and F^{-1} is the corresponding quantile function. Moreover, $F(Z) \sim \omega_{[0,1]}$ because of the continuity of F. Hence, the arctan-function transforms the standard Cauchy distribution into the uniform distribution on $[-\pi/2, \pi/2)$. In context of the stochastic representation of Φ, the absolute value of a standard Cauchy distributed random variable η/ξ is transformed into U, $U \sim \omega_{[0,\pi/2)}$, by the arctan-function, see Figure 1.

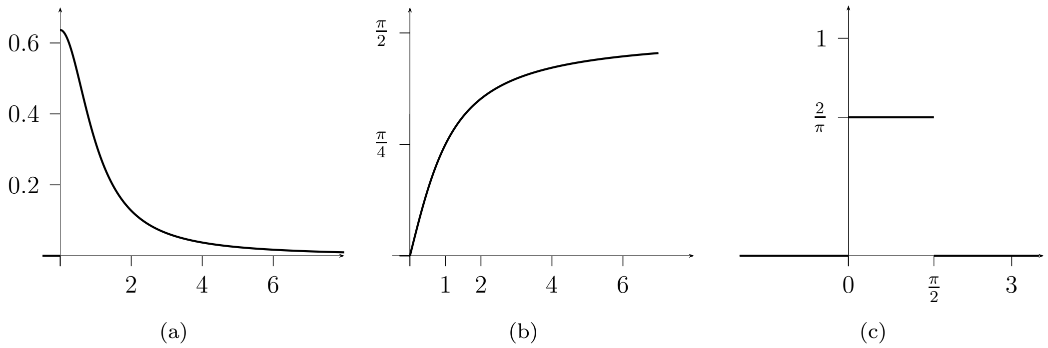


Figure 1. The arctan-transformation of $|\eta/\xi|$ yields a variable with uniform density. (a) density of $|\eta/\xi|$ (b) arctan-function (c) density of $\omega_{[0,\pi/2)}$

2.2 Notations for the axis-aligned heteroscedastic case

Here and in the following subsections, we consider vectors of uncorrelated, heteroscedastic and at the origin centered random variables, i.e. we assume that $(\xi, \eta)^T$ follows the elliptically contoured normal distribution $\Phi_{0_2,\Sigma}$ with expectation 0_2 and with regular covariance matrix $\Sigma = \mathrm{diag}(a^2, b^2)$, which generates an axis-aligned and at the origin centered ellipse

$$E_{(a,b)} = \left\{(x,y)^T \in \mathrm{R}^2 : (x,y)\Sigma^{-1}(x,y)^T = 1\right\}, \quad a, b > 0,$$

where

$$(x,y)^T \mapsto \left|(x,y)^T\right|_{(a,b)} := \left((x,y)\Sigma^{-1}(x,y)^T\right)^{1/2}$$

defines a norm. Density level sets of such a distribution are shown in Figure 2(a).

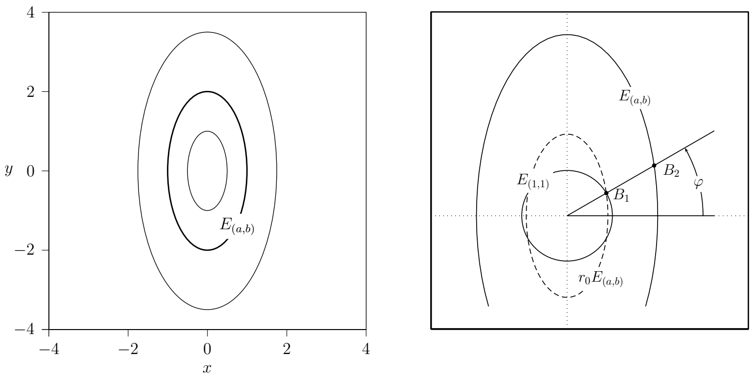
In the following subsections, we present and compare stochastic representations for $(\xi, \eta)^T$ which make use of different random coordinates. In doing so, we assume that $a \neq b$ because the case $a = b$ can be dealt with based upon the results in Subsection 2.1. The stochastic representation given there already reflects the geometry behind the two-dimensional Gaussian distribution in the multiple standard case. The reader will find some more or less known facts besides several new results. The new aspects turning out from this survey will be discussed at the end of Subsection 2.4 and 2.5, respectively. This will motivate some conclusions in the following sections.

2.3 Random polar coordinate representation

The vector $(\xi, \eta)^T \sim \Phi_{0_2,\mathrm{diag}(a^2,b^2)}$ allows in polar coordinates the stochastic representation (1) with $\sigma = 1$ and the joint density of the polar coordinates $(\mathcal{R}, \Phi)^T$ of $(\xi, \eta)^T$ is

$$f_{(\mathcal{R},\Phi)}(r, \varphi) = \frac{1}{2\pi\, a\, b}\, r\, e^{-\frac{r^2}{2} N^2_{(a,b)}(\varphi)}, \quad r > 0,\, 0 \le \varphi < 2\pi.$$

The geometric meaning of $N_{(a,b)}(\varphi) = \left|(\cos\varphi, \sin\varphi)^T\right|_{(a,b)}$ is illustrated in Figure 2(b).



(a) (b)

Figure 2. (a) Density level sets of $(\xi, \eta)^T$ in the case $a = 1$, $b = 2$.
(b) The meaning of $N_{(a,b)}(\varphi)$, $a = 2$, $b = 4$. Here, $B_1 = (\cos\varphi, \sin\varphi)^T$, $B_2 = B_1/N_{(a,b)}(\varphi)$ and $r_0 = |B_1|_{(a,b)} < 1$.

Theorem 2.1

a) The marginal density f_Φ of Φ satisfies the representation $f_\Phi(\varphi) = f^{E_{(a,b)}}(\varphi)$ where

$$f^{E_{(a,b)}}(\varphi) = \frac{1}{2\pi\, a\, b\, N^2_{(a,b)}(\varphi)}\,, \quad 0 \le \varphi < 2\pi\,. \tag{2}$$

b) The function $f^{E_{(a,b)}} : [0, 2\pi) \to (0, \infty)$ is π-periodical and its restriction to $[0, \pi)$ is symmetric with respect to (w.r.t.) $\varphi = \pi/2$. Furthermore, $f^{E_{(a,b)}}$ is strictly monotonically increasing on $(0, \pi/2)$ if $a < b$, and analogously strictly monotonically decreasing on $(0, \pi/2)$ if $a > b$, see Figure 3.

Proof a) The proof follows immediately by integrating the joint density $f_{(\mathcal{R},\Phi)}(r, \varphi)$ w.r.t. the radius $r > 0$.
b) Considering $N^2_{(a,b)}(\varphi)$ in terms of $\cos^2\varphi$, the monotonicity properties follow from those of $\cos^2\varphi$. ■

Notice that the quantity $\pi\, a\, b$ is discovered to be a suitably defined ellipse number within a certain non-Euclidean geometry in Richter (2011), $\pi\, a\, b = \pi_{(a,b)}$. Another well known representation of f_Φ is

$$f_\Phi(\varphi) = \frac{\sqrt{1 - \rho_*^2}}{2\pi\,(1 - \rho_* \cos(2\varphi))}\,, \quad \varphi \in [0, 2\pi)\,,$$

the density of the so called offset normal distribution in the special case of uncorrelated components ξ and η, see, e.g., Jammalamadaka and SenGupta (2001) or Mardia (1972). Here, $\rho_* := (a^2 - b^2)/(a^2 + b^2)$. In this regard, (2) is a new non-Euclidean representation of this special offset normal density. We shall give a probabilistic description of the quantity ρ_* in Section 3.

Let the set A be an element of the Borel σ-algebra $\mathfrak{A}_{(a,b)}$ on $E_{(a,b)}$ and denote by $\omega_{(a,b)}$ the $E_{(a,b)}$-generalized uniform probability distribution on $E_{(a,b)}$ introduced in Richter (2011). Then

$$\omega_{(a,b)}(A) = \int\limits_{Pol^{*-1}_{(a,b)}(A)} f^{E_{(a,b)}}(\varphi)\, d\varphi\,, \quad A \in \mathfrak{A}_{(a,b)}\,.$$

While $Pol_{(a,b)} : (r, \varphi) \mapsto (x, y)$ means the $E_{(a,b)}$-generalized elliptical polar coordinate transformation introduced in Richter (2011), the inverse of its restriction to $r = 1$, $Pol^*_{(a.b)}(\varphi) = Pol_{(a,b)}(1, \varphi)$, is denoted by $Pol^{*-1}_{(a,b)}$. We call therefore f_Φ the polar angle density corresponding to the $E_{(a,b)}$-generalized uniform probability distribution on $E_{(a,b)}$. The cdf of the polar angle distribution corresponding to $\omega_{(a,b)}$ will be denoted by F_Φ.

Corollary 2.2 The marginal cdf F_Φ of Φ satisfies the representation $F_\Phi = F^{E_{(a,b)}}$ where

$$F^{E_{(a,b)}}(\varphi) = \frac{1}{2\pi} \arctan\left[\frac{a}{b}\tan\varphi\right] + \frac{1_{(\pi/2,2\pi)}(\varphi) + 1_{(3\pi/2,2\pi)}(\varphi)}{2}, \quad 0 \leq \varphi < 2\pi.$$

Proof The proof of Corollary 2.2 follows by integrating f_Φ. An alternative proof may be given by using the fact that $a\eta/(b\xi)$ is distributed according to a Cauchy distribution. ■

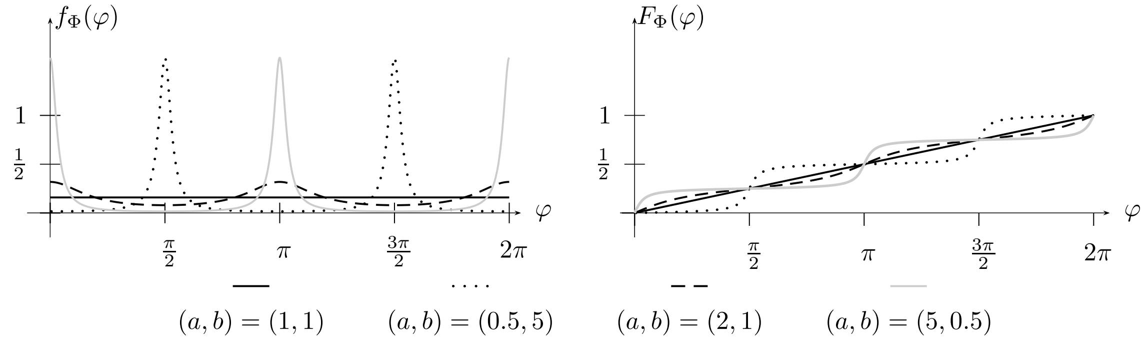


Figure 3. Polar angle density f_Φ and cdf F_Φ corresponding to $\omega_{(a,b)}$.

Notice that in contrast to the situation in Subsection 2.1, the transformation $(\xi, \eta) \mapsto f_1(\xi, \eta) \arctan|\eta/\xi| + \pi f_2(\xi, \eta)$ does not lead here to the uniform distribution on $[0, \pi/2)$. This is reflected in Figure 4.

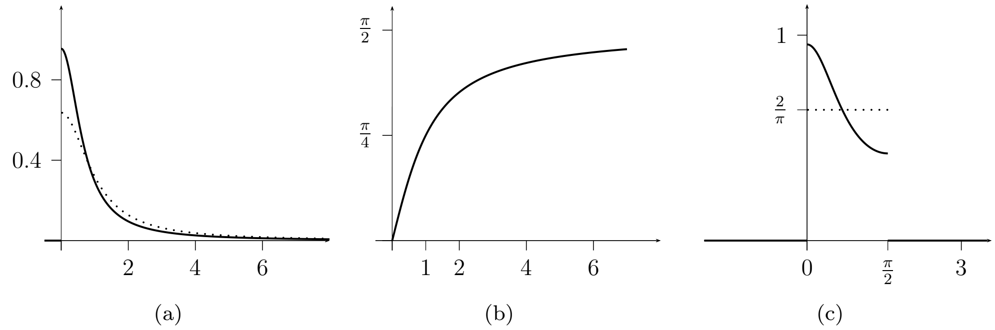


Figure 4. The arctan-transformation does not lead here to the uniform density.
(a) density of $|\eta/\xi|$, where $a = 3, b = 2$ (b) arctan-function
(c) density $1_{[0,\pi/2)}(\varphi) \cdot 4\, f^{E_{(a,b)}}(\varphi)$ of $\arctan|\eta/\xi|$
The dotted curves correspond to the limit case $a = b$ and $\omega_{[0,\pi/2)}$.

The distribution of $\mathcal{R} = \sqrt{\xi^2 + \eta^2}$ in the general case of $(\xi, \eta)^T$ following an absolutely continuous Gaussian distribution was studied extensively in the middle of the last century mostly in the context of military applications, see, e.g., the first sections in Guenther and Terragno (1964). Nowadays, however, the random variable $\mathcal{R}$ has relevance not only to military applications (Nelson, 1988) but is also of concern in industrial processing involving the drilling of holes (McCool, 2006), in the so called home range model for animal locations (Solow, 1990) or in measuring the accuracy of GPS point positioning and navigation (van

Diggelen, 2007; McEwen et al., 2005). Common synonyms for $\mathcal{R}$ are circular error (Harter, 1960) or radial error (Weil, 1954) where circular error is only used in the special case of zero expectation. The moments of $\mathcal{R}$ can be found, e.g., in Scheuer (1962) or White (1975).

COROLLARY 2.3

a) The conditional density of $\mathcal{R}$ under $\Phi = \varphi$ is the Rayleigh density with parameter $N^{-2}_{(a,b)}(\varphi)$, $\varphi \in [0, 2\pi)$.
b) The random polar coordinates $\mathcal{R}$ and Φ are not stochastically independent.

PROOF It follows immediately from the representations of $f_{(\mathcal{R},\Phi)}$ and f_Φ that

$$f_{\mathcal{R}|\Phi=\varphi}(r) = r\, N^2_{(a,b)}(\varphi)\, \exp\left(-\frac{r^2}{2}\, N^2_{(a,b)}(\varphi)\right), \quad r > 0\,.$$

This function depends on φ because of $a \neq b$, hence $\mathcal{R}$ and Φ are not independent. ∎

The result of this corollary can be found in Solow (1990) in another form but for the more general case of an arbitrary invertible covariance matrix Σ while our notation emphasizes the influence of the polar angle φ in terms of the quantity $N^2_{(a,b)}(\varphi)$.

The next theorem deals with representations for the cdf and pdf of the random variable $\mathcal{R}$. Especially, there will be given two representations for the pdf in b). While the representation b2) can be found in the literature the representations a) and b1) are based upon a geometric measure representation for the Gaussian law and seem to be new.

THEOREM 2.4

a) The cdf of $\mathcal{R}$ is equal to zero for non positive arguments. If $r > 0$, then

$$F_{\mathcal{R}}(r) = 1 - \exp\left\{-\frac{1}{2}\left(\frac{r}{\max\{a,b\}}\right)^2\right\}$$
$$+\frac{2}{\pi}\int\limits_{\frac{r}{\max\{a,b\}}}^{\frac{r}{\min\{a,b\}}} \rho \exp\left\{-\frac{\rho^2}{2}\right\} \arctan\sqrt{\frac{r^2 - \rho^2 \min\{a,b\}^2}{\rho^2 \max\{a,b\}^2 - r^2}}\, d\rho\,.$$

b) The density function of $\mathcal{R}$ vanishes for $r \leq 0$ and allows for $r > 0$ (and $a \neq b$) the two representations

$$\text{b1)} \qquad f_{\mathcal{R}}(r) = \frac{2}{\pi}\left[\int_{\frac{r}{\max\{a,b\}}}^{\frac{r}{\min\{a,b\}}} \rho \exp\left\{-\frac{\rho^2}{2}\right\} \frac{r}{\sqrt{(\rho^2 b^2 - r^2)(r^2 - \rho^2 a^2)}}\, d\rho\right],$$

$$\text{b2)} \qquad f_{\mathcal{R}}(r) = \frac{r}{ab} \exp\left\{-\frac{r^2(b^2+a^2)}{4a^2b^2}\right\} I_0\left(\frac{r^2(b^2-a^2)}{4a^2b^2}\right),$$

where I_0 denotes the modified Bessel function of the first kind and order zero. For a basic reference for I_0 and related functions, see Watson (1995).

PROOF We note that a) follows by using a geometric measure representation from Richter (1995). Alternatively, one might combine a result in Guenther and Terragno (1964) and Waugh (1961) concerning $P(\mathcal{R} < r)$ with the geometric measure representation of the noncentral chisquare distribution in Ittrich et al. (2000). We omit these straightforward but also tedious calculations and refer to Richter (2007), Kalke et al. (2013) and Günzel et al. (2012) who examplify the use of geometric measure representations. The proof of b1) follows by taking the derivative of $F_{\mathcal{R}}$ from a) according to Leibniz's integral rule, that of b2) follows by integrating the joint density of $\mathcal{R}$ and Φ w.r.t. the angle variable. These proofs may be called the integral-local and local-integral approaches to the density $f_{\mathcal{R}}$, respectively. ■

Remark 2.1 In contrast to b1) the representations a) and b2) hold also if $a = b$ and coincide then with the cdf and the pdf of a multiple of a chi-distributed random variable. In this regard, the new representation of the cdf reflects the deviation from a scaled cumulative chi-distribution function.

The quantiles of the radial errors cdf are of special interest in the aforementioned applications. The median is refered to as the circular error probability (CEP) (Nelson, 1988; Zhang and Weilian, 2012) and can either be approximated by integrating $f_{\mathcal{R}}$ numerically (Harter, 1960), by series expansions (Weil, 1954; Shnidman, 1995) or by a simple function of a and b (Nelson, 1988). If $a = b$, then $a^{-1}\mathcal{R} \sim \chi_2$, see Subsection 2.1. Therefore, the quantiles can be calculated explicitly which is why this case is often assumed for simplicity. Arguing that any correlation between ξ and η can be eliminated by a proper rotation some authors assume the components of $(\xi, \eta)^T$ to be independent without loss of generality (w.l.o.g.), see, e.g., Gillis (1991). In fact, the distribution of $\mathcal{R}$ is invariant under rotations but 'the correlation coefficient of the old variables manifests itself in the variance of the new variables', as criticised in Pyati (1993), which we will consider in detail in Subsection 2.6.

To distinguish between the polar coordinates used in this subsection and those coordinates used in the following subsections, let us finally denote the polar coordinates by $(\mathcal{R}, \Phi) = (\mathcal{R}_P, \Phi_P)$.

2.4 Random elliptical polar coordinate representation

Let $\mathcal{R}_{EP}$ and Φ_{EP} denote elliptical polar coordinates of the random vector $(\xi, \eta)^T$, i.e.

$$\mathcal{R}_{EP} = \left(\frac{\xi^2}{a^2} + \frac{\eta^2}{b^2}\right)^{\frac{1}{2}} \qquad \text{and} \qquad \Phi_{EP} = f_1(\xi, \eta) \arctan\left|\frac{a\eta}{b\xi}\right| + \pi f_2(\xi, \eta)\,.$$

Following Richter (2011), we have $\mathcal{R}_{EP} = \mathcal{R}_{(a,b)} = \left|(\xi, \eta)^T\right|_{(a,b)}$ and the stochastic representation

$$\begin{pmatrix}\xi\\ \eta\end{pmatrix} \stackrel{d}{=} \mathcal{R}_{(a,b)} \cdot \mathcal{W}_{(a,b)} \qquad \text{with} \qquad \mathcal{W}_{(a,b)} = \begin{pmatrix} a\,\cos\Phi_{EP}\\ b\,\sin\Phi_{EP}\end{pmatrix}\,. \tag{3}$$

Remark 2.2 It should be emphasized that the elliptical polar angle $\varphi_{EP} = \varphi_{EP}(x, y)$ of a point $(x, y)^T$ does a.e. not coincide with the (usual) polar angle $\varphi_P = \varphi_P(x, y)$ (unless for $a = b$), see Figures 5(a) and 5(b). Actually, we have

$$\begin{aligned}\varphi_{EP}(x, y) &= f_1(x, y) \arctan\left|a/b\,\tan\big(\varphi_P(x, y)\big)\right| + \pi f_2(x, y)\,,\\ \varphi_P(x, y) &= f_1(x, y) \arctan\left|b/a\,\tan\big(\varphi_{EP}(x, y)\big)\right| + \pi f_2(x, y)\end{aligned}$$

and $\tan\big(\varphi_P(x, y)\big) = b/a\,\tan\big(\varphi_{EP}(x, y)\big)$.

The joint density of $\mathcal{R}_{(a,b)}$ and Φ_{EP} is

$$f_{(\mathcal{R}_{(a,b)}, \Phi_{EP})}(r, \varphi) = \frac{1}{2\pi}\, r\, e^{-\frac{r^2}{2}}\,, \quad r > 0\,,\, 0 \le \varphi < 2\pi\,,$$

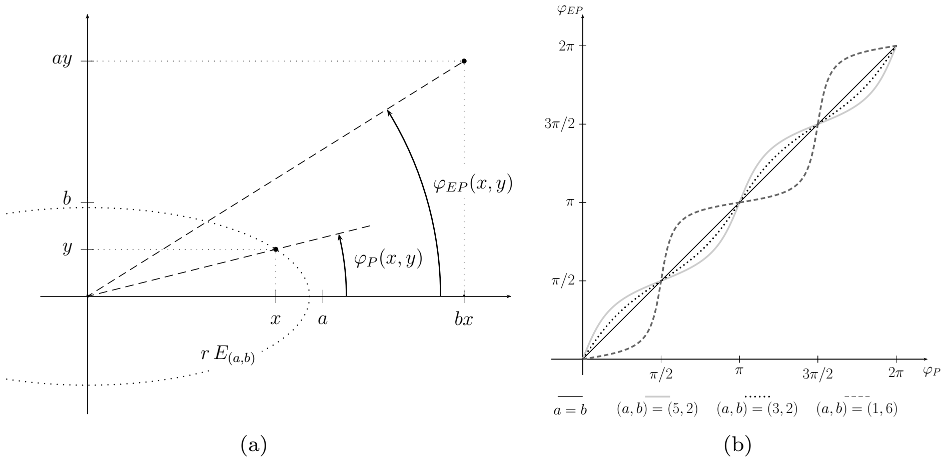


Figure 5. (a) Polar angle $\varphi_P(x,y)$ and elliptical polar angle $\varphi_{EP}(x,y)$ where $(x,y) = (4,1)$, $(a,b) = (5,2)$ and $r = \left|(x,y)^T\right|_{(a,b)}$.
(b) The angles $\varphi_{EP}(x,y)$ and $\varphi_P(x,y)$ of (x,y) running anti-clockwise oriented through $E_{(a,b)}$ and starting from $(a,0)$.

hence $\mathcal{R}_{(a,b)}$ and Φ_{EP} are independent where $\mathcal{R}_{(a,b)}$ follows the chi-distribution with two d.f., $\mathcal{R}_{(a,b)} \sim \chi_2$, and Φ_{EP} is uniformly distributed on the interval $[0, 2\pi)$, $\Phi_{EP} \sim \omega_{[0,2\pi)}$. For a deeper understanding of the quantities in representation (3), let us define the subsets

$$A_P(\varphi_1,\varphi_2) := \left\{(x,y)^T = r(\varphi_P)\,(\cos\varphi_P, \sin\varphi_P)^T : \varphi_1 \le \varphi_P(x,y) \le \varphi_2\right\},$$
$$A_{EP}(\varphi_1,\varphi_2) := \left\{(x,y)^T = (a\cos\varphi_{EP}, b\sin\varphi_{EP})^T : \varphi_1 \le \varphi_{EP}(x,y) \le \varphi_2\right\}$$

of the ellipse $E_{(a,b)}$ where $0 \le \varphi_1 < \varphi_2 < 2\pi$ and $r(\varphi_P) = N^{-1}_{(a,b)}(\varphi_P)$. Then

$$P\big(\mathcal{W}_{(a,b)} \in A_{EP}(\varphi_1,\varphi_2)\big) = P(\varphi_1 \le \Phi_{EP} \le \varphi_2) = \frac{\varphi_2 - \varphi_1}{2\pi}.$$

This does not mean, however, that $\mathcal{W}_{(a,b)}$ follows the uniform distribution on the ellipse $E_{(a,b)}$ w.r.t. the Euclidean arc length (as it would for $a = b$). To demonstrate this, let $a > b$ and denote the Euclidean arc length by AL. Then $AL\left(A_{EP}\left(0,\pi/4\right)\right) < AL\left(A_{EP}\left(\pi/4,\pi/2\right)\right)$ while $AL\left(A_P\left(0,\pi/4\right)\right) > AL\left(A_P\left(\pi/4,\pi/2\right)\right)$, see also Figure 6.

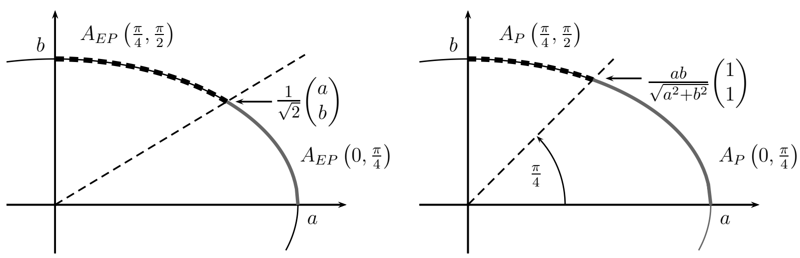


Figure 6. $\mathcal{W}_{(a,b)}$ is not uniformly distributed: The case $a > b$.

The previous results do not give a comprehensive insight into the nature of the bivariate and at the origin centered normal distribution. On the one hand, the introduction of random polar coordinates as in Subsection 2.3 does not lead to a decomposition of $(\xi,\eta)^T$ into independent factors. On the other hand, the introduction of random elliptical coordinates as in Subsection 2.4 is strongly connected with the described problems of interpreting the resulting random angle. Moreover, the uniform distribution of the angle Φ_{EP} does not reflect the non uniform distribution on $E_{(a,b)}$ of the corresponding random vector $\mathcal{W}_{(a,b)}$. Searching for a deeper understanding of the axis-aligned Gaussian law, we shall make use of coordinates in the next section which play a basic role in non-Euclidean or Minkowski geometry.

2.5 Random $E_{(a,b)}$-generalized elliptical polar coordinate representation

Let us consider now the $E_{(a,b)}$-generalized elliptical polar coordinates $(\mathcal{R}_{GEP}, \Phi_{GEP})^T$ of $(\xi, \eta)^T$,

$$\mathcal{R}_{GEP} = \mathcal{R}_{(a,b)} = \mathcal{R}_{EP} = \left| \begin{pmatrix} \xi \\ \eta \end{pmatrix} \right|_{(a,b)} \qquad \text{and} \qquad \Phi_{GEP} = \Phi_P = \Phi$$

where Φ is the usual polar angle. These coordinates were introduced in Richter (2011) and it follows immediately from the considerations in Example 15 there that $\mathcal{R}_{(a,b)}$ and Φ are independent and satisfy $\mathcal{R}_{(a,b)} \sim \chi_2$, $\Phi \sim F^{E_{(a,b)}}$, that is, the random variables $\mathcal{R}_{(a,b)}$ and Φ follow the joint density function

$$f_{(\mathcal{R}_{(a,b)},\Phi)}(r,\varphi) = r\, e^{-\frac{r^2}{2}} \frac{1}{2\pi_{(a,b)}\, N^2_{(a,b)}(\varphi)}\,, \quad r > 0\,,\ 0 \le \varphi < 2\pi\,. \tag{4}$$

Further, let $\mathcal{U}_{(a,b)} = (a\, \cos_{(a,b)}(\Phi), b\, \sin_{(a,b)}(\Phi))^T$ where

$$\cos_{(a,b)}(\varphi) = \frac{\cos\varphi}{a\, N_{(a,b)}(\varphi)} \qquad \text{and} \qquad \sin_{(a,b)}(\varphi) = \frac{\sin\varphi}{b\, N_{(a,b)}(\varphi)}$$

denote the $E_{(a,b)}$-generalized trigonometric functions, see Figures 7(a) and 7(b).

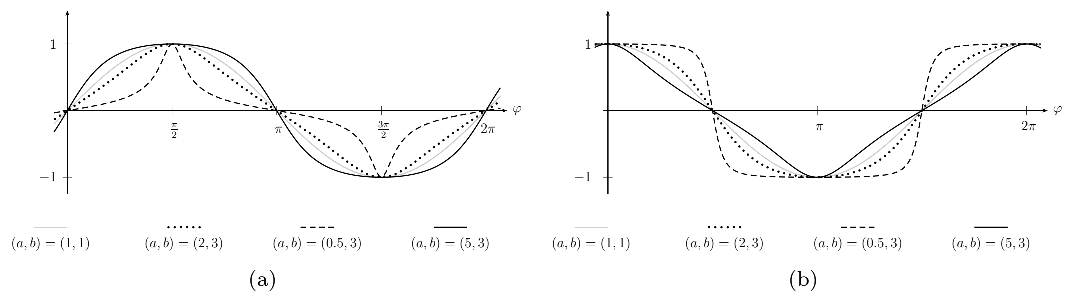


Figure 7. (a) The $E_{(a,b)}$-generalized sine function. (b) The $E_{(a,b)}$-generalized cosine function.

The random vector $(\xi, \eta)^T$ satisfies the stochastic representation

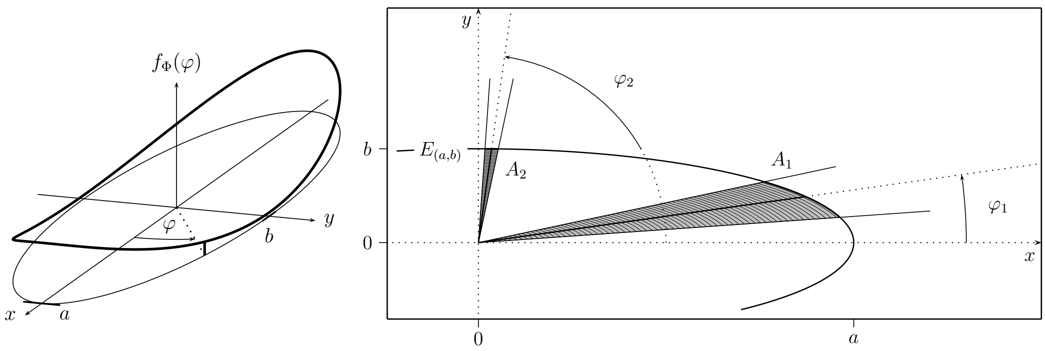
$$\begin{pmatrix} \xi \\ \eta \end{pmatrix} \stackrel{d}{=} \mathcal{R}_{(a,b)} \cdot \mathcal{U}_{(a,b)} \tag{5}$$

with $\mathcal{R}_{(a,b)}$ and $\mathcal{U}_{(a,b)}$ being independent, $\mathcal{U}_{(a,b)} \sim \omega_{(a,b)}$, $\mathcal{R}_{(a,b)} \sim \chi_2$ and also with $\mathcal{R}_{(a,b)}$ being independent of Φ. Formula (5) will be called the stochastic representation corresponding to the $E_{(a,b)}$-generalized elliptical polar coordinate representation of an axis-aligned and at the origin centered (elliptically contoured) Gaussian random vector. For emphasizing the non-standard geometric nature of $\mathcal{U}_{(a,b)}$, one might also speak about this approach as a non-Euclidean geometric-stochastic representation. Although the random vector $\mathcal{W}_{(a,b)}$

in Subsection 2.4 follows the same distribution as $\mathcal{U}_{(a,b)}$ in this subsection, the stochastic representation (5) may be preferable over the stochastic representation (3), because the polar angle used inhere describes in a more natural way than Φ_{EP} the direction of $(\xi,\eta)^T$. In this regard, the distribution of Φ described in Theorem 2.1 perfectly reflects the distribution of $\mathcal{U}_{(a,b)}$. Consequently, we restrict our considerations in the following sections to the type of coordinates used here. While the meaning of the quantity $\mathcal{R}_{(a,b)}$ is pretty clear, we want to give more explanation for the density of Φ. To this end, let us consider a closed parametric curve $\mathfrak{C}_\Phi : [0,2\pi) \to \mathrm{R}^3$ defined by $\mathfrak{C}_\Phi(\varphi) = (a \cos_{(a,b)}(\varphi), b \sin_{(a,b)}(\varphi), f_\Phi(\varphi))$. One may call this curve the $E_{(a,b)}$-generalized elliptical polar representation of f_Φ, see Figure 8(a). Moreover, the values $f_\Phi(\varphi)$ are closely connected with cone-probabilities $P(\mathcal{V} \in C(A_i)) = P(\mathcal{U}_{(a,b)} \in A_i)$ where $A_i = A_P(\varphi_i - \varepsilon, \varphi_i + \varepsilon)$ generates the cone

$$C(A_i) = \left\{ (x,y)^T \in \mathrm{R}^2 : \frac{(x,y)^T}{|(x,y)^T|_{(a,b)}} \in A_i,\ |(x,y)^T|_{(a,b)} \le 1 \right\}$$

for some $\varepsilon > 0$ and $\mathcal{V}$ follows the uniform distribution on $K_{(a,b)} = \{(x,y)^T \in \mathrm{R}^2 : |(x,y)^T|_{(a,b)} \le 1\}$, see Example 4.4. (b) in Richter (2011) and Figure 8(b). Here, large values of $f_\Phi(\varphi_i)$ correspond to large values of $P(\mathcal{U}_{(a,b)} \in A_i)$ as well as small values of $f_\Phi(\varphi_i)$ correspond to small values of $P(\mathcal{U}_{(a,b)} \in A_i)$.



(a) (b)

Figure 8. In both figures it is $a = 4$ and $b = 1$.
(a) The $E_{(a,b)}$-generalized elliptic polar representation $\mathfrak{C}_\Phi$ for f_Φ.
(b) Large values of $f_\Phi(\varphi_i)$ correspond to those of $P\left(\mathcal{U}_{(a,b)} \in A_i\right)$. Here, $\varepsilon = 4°$, $\varphi_1 = 8°$ and $\varphi_2 = 82°$.

In directional statistics the density f of a polar angle is sometimes represented by the closed curve $((1+f(\varphi))\cos\varphi, (1+f(\varphi))\sin\varphi)$, $\varphi \in [0,2\pi)$, which is called circular or polar representation of the density f (w.r.t. the unit circle C), see, e.g., Jammalamadaka and SenGupta (2001) or Mardia and Jupp (2000). A circular representation of f_Φ is shown in Figure 9.

2.6 Regularly distributed Gaussian vectors

Let now the vector $(\xi,\eta)^T$ be distributed according to a regular normal distribution, i.e. $(\xi,\eta)^T \sim \Phi_{\mu,\Sigma}$ with an arbitrary expectation vector $\mu = (\mu_1,\mu_2)^T \in \mathrm{R}^2$ and a regular covariance matrix

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\,\sigma_1\sigma_2 \\ \rho\,\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}.$$

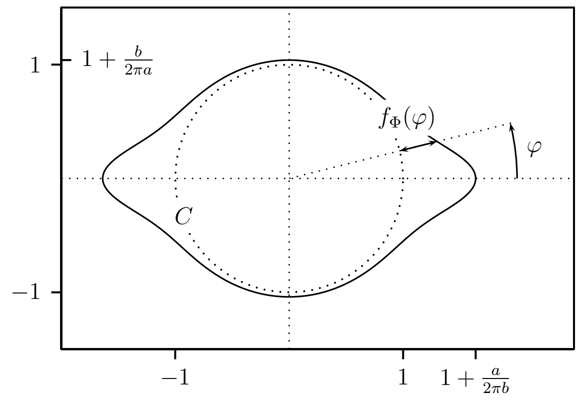


Figure 9. Circular representation of f_Φ in the case $a = 4$, $b = 1$.

With regard to the basic concept in principal component analysis, see for example Anderson (2003), the random vector $(\xi - \mu_1, \eta - \mu_2)^T$ can be orthogonally transformed to follow an axis-aligned Gaussian distribution $\Phi_{0_2,\mathrm{diag}(a^2,b^2)}$ with variances a^2 and b^2 being the eigenvalues of Σ. In other words, the ellipse E generated by the symmetric positive-definite matrix Σ and μ, $E = \big\{(x,y)^T \in \mathrm{R}^2 : (x-\mu_1, y-\mu_2)\Sigma^{-1}(x-\mu_1, y-\mu_2)^T = 1\big\}$, can be rotated clockwise with respect to μ and through the angle α, $\alpha \in [0, \pi/2)$, to an axis-aligned ellipse $E_{(a,b)} + \mu$ where a and b are the half lengths of the principal axes of E and where α is the anticlockwise-measured (acute) angle included by the positive x-axis and one of the principal axes of $E - \mu$. The half length of this principal axis associated with the rotation angle α will be denoted by a, see Figure 10. Notice that every orthogonal transformation of $E-\mu$ can be represented by a properly chosen rotation of $E-\mu$. Putting $\alpha = 0$ if both $\sigma_1 = \sigma_2$ and $\rho = 0$, there is a one-to-one map

$$T_{(AA,R)} : (a, b, \alpha) \mapsto (\sigma_1, \sigma_2, \rho)$$

from the parameters in the axis-aligned case to those in the regular case where $(\sigma_1, \sigma_2, \rho) \in (0,\infty)^2 \times (-1, 1)$ and $(a, b, \alpha) \in \big(\{(x,y) \in \mathrm{R}^2 : x > 0,\ y > 0,\ x \neq y\} \times [0, \pi/2)\big) \cup \big\{(x, x, 0) \in \mathrm{R}^3 : x > 0\big\}$. The inverse of this map, $T^{-1}_{(AA,R)} : (\sigma_1, \sigma_2, \rho) \mapsto (a, b, \alpha)$, is given by

$$\alpha = \alpha(\Sigma) = \begin{cases} 0 & \rho = 0 \quad (6.1) \\ \gamma + \begin{cases} 0 & \rho\,(\sigma_1 - \sigma_2) > 0 \\ \pi/2 & \rho\,(\sigma_1 - \sigma_2) < 0 \end{cases} & \sigma_1 \neq \sigma_2 \ \wedge\ \rho \neq 0 \quad (6.2) \\ \pi/4 & \sigma_1 = \sigma_2 \ \wedge\ \rho \neq 0 \quad (6.3) \end{cases} \tag{6}$$

where $\gamma = \frac{1}{2}\arctan\frac{2\sigma_1\sigma_2\,\rho}{\sigma_1^2 - \sigma_2^2}$ and

$$\begin{aligned} a = a(\Sigma) &= \sqrt{\sigma_1^2 \cos^2\alpha + \sigma_2^2 \sin^2\alpha + 2\,\rho\,\sigma_1\sigma_2 \sin\alpha\cos\alpha} \quad , \\ b = b(\Sigma) &= \sqrt{\sigma_2^2 \cos^2\alpha + \sigma_1^2 \sin^2\alpha - 2\,\rho\,\sigma_1\sigma_2 \sin\alpha\cos\alpha} \quad . \end{aligned} \tag{7}$$

The nontrivial cases (6.2) and (6.3) occurring in the evaluation of $\alpha(\Sigma)$ in (6) are illustrated in Figure 11 by giving the corresponding density level sets of $(\xi, \eta)^T \sim \Phi_{\mu,\Sigma}$ where w.l.o.g. it is $\mu = 0_2$. Notice that $\gamma < 0$ in Figure 11(a) and that the omitted subcase $\sigma_1 > \sigma_2$, $\rho \neq 0$ in Figure 11 was already dealt with in Figure 10.

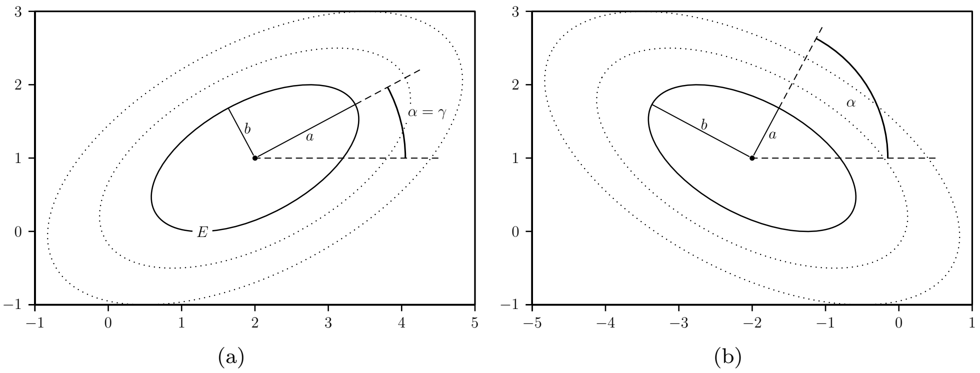


Figure 10. Density level sets of $(\xi, \eta)^T$ for $\sigma_1 > \sigma_2$.
(a) $\mu = \binom{2}{1}$, $\Sigma = \begin{pmatrix} 2 & 0.75 \\ 0.75 & 1 \end{pmatrix}$ (b) $\mu = \binom{-2}{1}$, $\Sigma = \begin{pmatrix} 2 & -0.75 \\ -0.75 & 1 \end{pmatrix}$

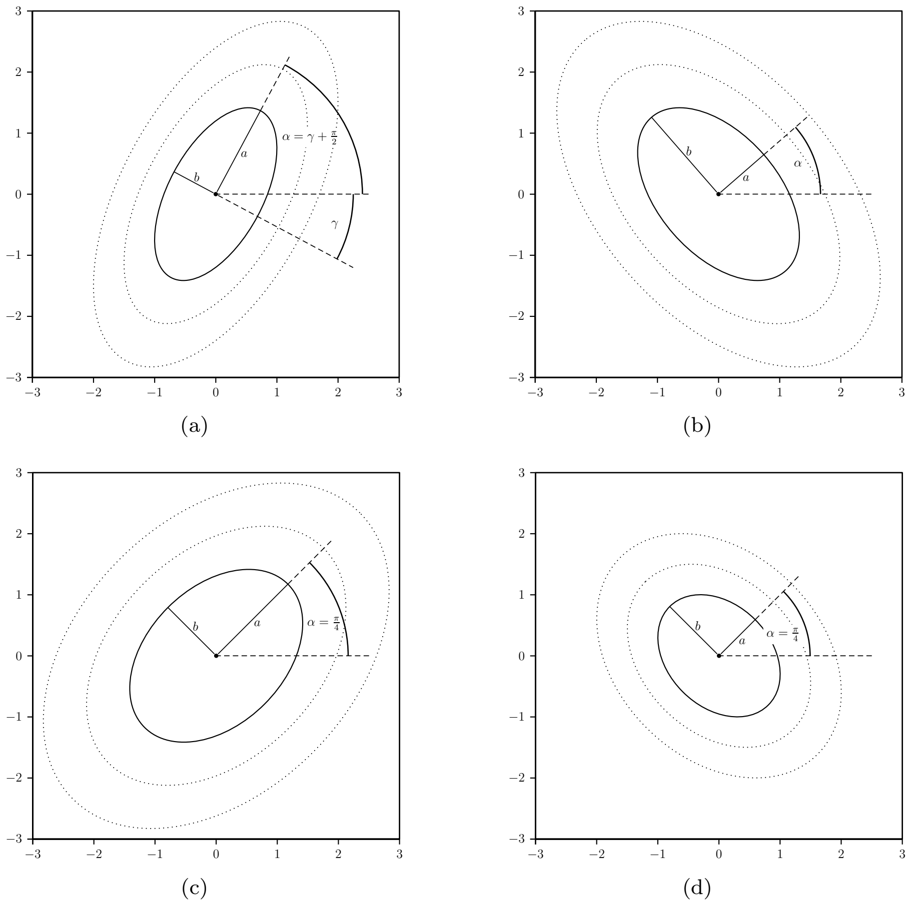


Figure 11. The case (6.2), $\sigma_1 < \sigma_2$: (a) $\Sigma = \begin{pmatrix} 1 & 0.75 \\ 0.75 & 2 \end{pmatrix}$ (b) $\Sigma = \begin{pmatrix} 1.75 & -0.9 \\ -0.9 & 2 \end{pmatrix}$
The case (6.3): (c) $\Sigma = \begin{pmatrix} 2 & 0.75 \\ 0.75 & 2 \end{pmatrix}$ (d) $\Sigma = \begin{pmatrix} 1 & -0.3 \\ -0.3 & 1 \end{pmatrix}$

Keep in mind that $D\Sigma D^T = \mathrm{diag}(a^2, b^2) = cov\left(D\left(\xi, \eta\right)^T\right)$ where

$$D = D(\alpha) = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}$$

is an orthogonal matrix defining a clockwise rotation through the uniquely defined angle α and w.r.t. the origin. That is why we refer to (6) and (7) as the geometric representation of α, a and b. The rows of D are orthonormal eigenvectors of Σ corresponding to the eigenvalues a^2 and b^2 and in addition, the rows define the principal components $(\cos\alpha, \sin\alpha)(\xi-\mu_1, \eta-\mu_2)^T$ and $(-\sin\alpha, \cos\alpha)(\xi-\mu_1, \eta-\mu_2)^T$ of $(\xi-\mu_1, \eta-\mu_2)^T$ having variances a^2 und b^2. From the latter interpretation of a^2 and b^2 as eigenvalues of Σ follow the representations

$$\begin{aligned} \max\{a,b\} = \max\{a(\Sigma), b(\Sigma)\} &= \sqrt{\frac{\sigma_1^2+\sigma_2^2+\sqrt{(\sigma_1^2-\sigma_2^2)^2+4\rho^2\sigma_1^2\sigma_2^2}}{2}}\,, \\ \min\{a,b\} = \min\{a(\Sigma), b(\Sigma)\} &= \sqrt{\frac{\sigma_1^2+\sigma_2^2-\sqrt{(\sigma_1^2-\sigma_2^2)^2+4\rho^2\sigma_1^2\sigma_2^2}}{2}}\,. \end{aligned} \tag{8}$$

We therefore refer to (8) as the algebraic representation of the eigenvalues of the covariance matrix Σ arranged in descending order. In contrast to the geometric representations in (7), it seems at the first look that there is no (explicit) use of the rotation angle $\alpha = \alpha(\Sigma)$ in (8). But notice that the correlation coefficient $\rho = \rho(\alpha)$ depends on α in a well defined but subtle way. This will be discussed in detail in the next section.

To summarize our considerations so far, the centered and rotated (principal component) vector

$$\begin{pmatrix} X \\ Y \end{pmatrix} := D(\alpha) \begin{pmatrix} \xi-\mu_1 \\ \eta-\mu_2 \end{pmatrix}$$

follows an axis-aligned Gaussian distribution that generates an ellipse $E_{(a,b)}$ with $a = a(\Sigma) > 0$, $b = b(\Sigma) > 0$. With regard to Subsection 2.5, the random vector $(\xi, \eta)^T$ satisfies therefore the representation

$$\begin{pmatrix} \xi \\ \eta \end{pmatrix} \stackrel{d}{=} \mu + D(\alpha)^T \left(\mathcal{R}_{(a,b)} \cdot \mathcal{U}_{(a,b)}\right) \tag{9}$$

where the parameter triple (a, b, α) is uniquely determined by (6) and (7), $\mathcal{R}_{(a,b)}$ and $\mathcal{U}_{(a,b)}$ are independent and $\mathcal{R}_{(a,b)} \sim \chi_2$, $\mathcal{U}_{(a,b)} \sim \omega_{(a,b)}$. Formula (9) will be called the $E_{(a,b)}$-generalized elliptical polar coordinate representation of a regularly distributed Gaussian random vector. Notice that $\mathcal{R}_{(a,b)} = |D(\alpha)(\xi-\mu_1, \eta-\mu_2)^T|_{(a,b)}$, i.e. $\mathcal{R}_{(a,b)}$ can be interpreted as the value of the Minkowski functional w.r.t. the set $D(\alpha)^T E_{(a,b)}$ at the point $(\xi-\mu_1, \eta-\mu_2)^T$. Incidentally, the chi-distribution of $\mathcal{R}_{(a,b)}$ in (3), (5) or (9), respectively, is a consequence of the well known fact that $(Z-\mu)^T\Sigma^{-1}(Z-\mu)$ is chi-square distributed with m degrees of freedom if $Z \sim \Phi_{\mu,\Sigma}$, $\mu \in \mathrm{R}^m$ and $\Sigma \in \mathrm{R}^{m\times m}$, see, e.g., Muirhead (1982). Moreover, the radius $\mathcal{R}_{(a,b)}$ and $\Phi_\alpha = \Phi + \alpha \mod 2\pi$, where Φ_α is the polar angle of $(\xi-\mu_1, \eta-\mu_2)^T$ and Φ is the polar angle of $D(\alpha)(\xi-\mu_1, \eta-\mu_2)^T$, follow the joint pdf

$$f_{(\mathcal{R}_{(a,b)}, \Phi_\alpha)}(r, \varphi) = r\, e^{-\frac{r^2}{2}} \frac{1}{2\pi_{(a,b)}\, N^2_{(a,b)}(\varphi-\alpha)}\,, \quad r > 0\,,\, 0 \le \varphi < 2\pi\,.$$

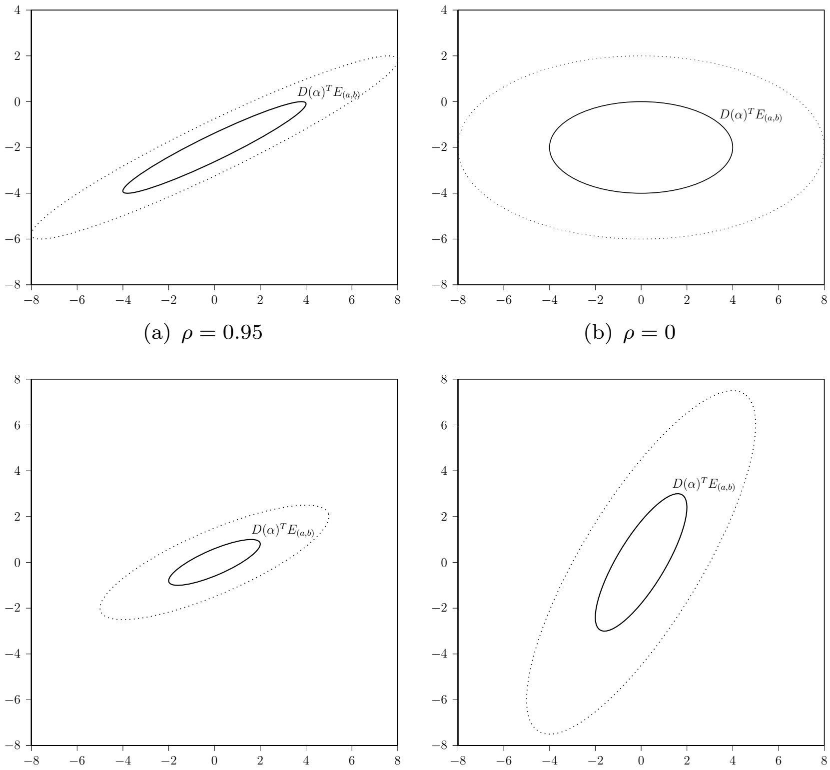
Equation (4) implies $f_{\Phi_\alpha}(\varphi) = f_\Phi((\varphi-\alpha) \mod 2\pi)$, $0 \le \varphi < 2\pi$. We shall call $D(\alpha)^T\mathcal{U}_{(a,b)}$ the generalized uniform basis of the random vector $(\xi-\mu_1, \eta-\mu_2)^T$ and $\mathcal{R}_{(a,b)}$ its generalized radius variable. This notion follows the one in Fang et al. (1990) and extends it in the spirit of Richter (2011) to non-Euclidean geometry. The representation (9), however, differs essentially from those in Cambanis et al. (1981) and Tong (1990), which refer to the uniform distribution ω_C on the unit circle (w.r.t. the Euclidean arc length). For the discussion of the possibility to choose the parameters a and b in descending order see Remark 6.2

Let us finally discuss the main new aspect turning out from the survey given in this section. As we have seen, there are several other more or less known and common possibilities to represent the random vector $(\xi, \eta)^T$ with the help of a certain angle and a certain radius variable. As indicated in Subsections 2.5 and 2.6, however, there are pretty good reasons for us to follow further here the approach in (9), which is based on non-Euclidean geometry. Moreover, this geometric-stochastic representation of a regularly distributed Gaussian random vector itself motivates us to introduce a corresponding parametrization of the Gaussian law in Section 4. Before doing this, we will analyze in Section 3 the meaning of the correlation coefficient within this non-Euclidean geometric approach.

3. The non Euclidean geometric representation of ρ

We consider again the random vector $(\xi, \eta)^T \sim \Phi_{0_2,\Sigma}$ and assume that Σ is regular and has two different eigenvalues. As it was shown in Section 2, the distribution of $(\xi, \eta)^T$ is strongly connected then with a non-Euclidean geometry. In this context, equation (9) gives a decomposition of $(\xi, \eta)^T$ into independent random elements, an $E_{(a,b)}$-generalized radius and a generalized uniform basis. The latter follows a generalized uniform distribution on the ellipse $E_{(a,b)}$ rotated anti-clockwise through the angle α, where (a, b, α) can be obtained from (6) and (7). Besides, the considerations in Subsection 2.6 make clear that the existence of both linear and stochastic dependencies between ξ and η result only from the rotation of a Gaussian random vector with independent components through the angle α. Hence, if $a \neq b$, the rotation angle α causes the existence of a nontrivial relationship between ξ and η and can be used therefore as a parameter expressing dependence properties. This raises the questions in which sense the usual stochastic parameters σ_1^2, σ_2^2 and ρ are suitable to describe the geometry behind the regular Gaussian law and especially in which sense ρ reflects the linear (geometric) dependencies between ξ and η.
As can be seen in Figures 12(a) and 12(b), the variances σ_1^2 and σ_2^2 do not indicate a lot about the shape of the density level sets $D(\alpha)^T E_{(a,b)}$ of $(\xi, \eta)^T$ if the correlation coefficient is unknown. Also, if ρ is known but one of the two variances σ_1^2 and σ_2^2 is unknown, then again the shape of the density level sets of $(\xi, \eta)^T$ may look in various different ways, see Figures 12(c) and 12(d). Consequently, one has to know the complete triple $(\sigma_1^2, \sigma_2^2, \rho)$ to understand the regular Gaussian law from the geometric point of view. In contrast, the quantities a^2 and b^2 determine already the lengths-ratio of the principle axes of the density level sets of $(\xi, \eta)^T$ and in this sense a main part of their shape even if the rotation angle α is unknown. On the other hand, the angle α already involves a certain main direction of the considered density level sets even if a^2 and b^2 are unknown. In this context, $\tan \alpha$ is the slope of the principle axis of (ξ, η)'s density level sets having half length a, and $\tan \alpha$ is also the slope of the line through the origin with minimum or maximum expected squared distance to $(\xi, \eta)^T$ if $a > b$ or $a < b$, respectively, see Anderson (2003) or Hyvärinen et al. (2001). It is common to consider ρ as a parameter expressing dependence properties of ξ and η. Let us emphasize here as a consequence of Subsection 2.6 that $\Sigma = \Sigma(\alpha)$ and especially $\rho = \rho(\alpha)$ are functions of the rotation angle α where

$$\rho(\alpha) = \frac{\sin \alpha \cos \alpha \cdot (1 - \frac{b^2}{a^2})}{\sqrt{\cos^2 \alpha + \frac{b^2}{a^2} \sin^2 \alpha} \sqrt{\frac{b^2}{a^2} \cos^2 \alpha + \sin^2 \alpha}}\,, \quad 0 \leq \alpha < \frac{\pi}{2}\,. \tag{10}$$



(a) $\rho = 0.95$ (b) $\rho = 0$

(c) $\sigma_2^2 = 1$ (d) $\sigma_2^2 = 9$

Figure 12. (a), (b) The shape of the density level sets $D(\alpha)^T E_{(a,b)}$ with $\sigma_1^2 = 16$ and $\sigma_2^2 = 4$ depends on ρ.
(c), (d) The shape of $D(\alpha)^T E_{(a,b)}$ with $\sigma_1^2 = 4$ and $\rho = 0.8$ depends on σ_2^2.

Considering the correlation coefficient as a function defined for $\alpha \in [0, 2\pi)$, $\alpha \mapsto \rho(\alpha)$ is π-periodical and if $a > b$, ρ is strictly monotonically increasing on $(0, \pi/4) \cup (3\pi/4, \pi)$ and strictly monotonically decreasing on $(\pi/4, 3\pi/4)$. Analogously, if $a < b$, the function $\alpha \mapsto \rho(\alpha)$ is strictly monotonically increasing on $(\pi/4, 3\pi/4)$ and strictly monotonically decreasing on $(0, \pi/4) \cup (3\pi/4, \pi)$, see Figure 13. Moreover, $\alpha \mapsto \rho(\alpha)$ satisfies $\rho\,(\pi/4 - \alpha) = \rho\,(\pi/4 + \alpha)$ and $\rho\,(\pi/2 - \alpha) = -\rho\,(\pi/2 + \alpha)$. Incidentally, the latter properties of $\alpha \mapsto$

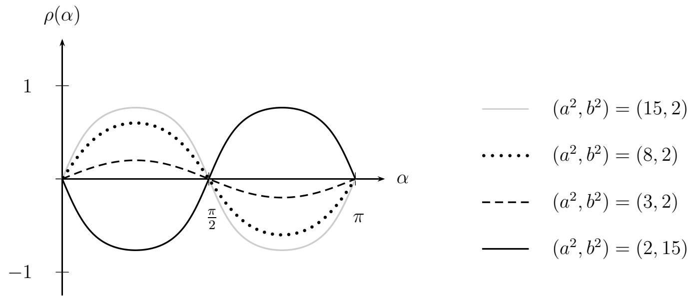


Figure 13. The correlation coefficient ρ as a function of α, $\alpha \in [0, \pi)$.

$\rho(\alpha)$, $\alpha \in [0, 2\pi)$, make clear that it suffices to consider the rotation angle as an element of $[0, \pi/2)$ to obtain any possible correlation $-1 < \rho(\alpha) < 1$. To be more concrete, for given $a > b > 0$ the maximum and minimum values of $\rho(\alpha)$ are $\rho_* = (a^2 - b^2)/(a^2 + b^2)$ and zero, respectively. On the other hand, the maximum and minimum values of $\rho(\alpha)$ are zero and $\rho_* < 0$, respectively, if $b > a > 0$. This means that the general inequalities $-1 \leq \rho \leq 1$

can be specified by the sharper inequalities

$$\min\{0,\rho_*\} \le \rho \le \max\{0,\rho_*\}\ , \tag{11}$$

where $\rho_* = \rho(\pi/4)$ and $|\rho_*| < 1$, which we have not been aware of in the literature. As $\lim_{b/a\to 0}\rho(\alpha) = 1$ and $\lim_{a/b\to 0}\rho(\alpha) = -1$, a and b may be chosen in such a way that an arbitrary prespecified value from $(-1,1)$ is attained by the correlation coefficient. For a basic study of the closely connected notion of a maximal correlation for given marginals we refer to Fréchet (1957) and Höffding (1940). Notice that in terms of the non-Euclidean norm $|\cdot|_{(a,b)}$, ρ is given by

$$\rho(\alpha) = \frac{\rho_*\sin(2\,\alpha)}{2}\cdot\left(\frac{\left|\begin{pmatrix}\cos(\alpha+\pi/2)\\ \sin(\alpha+\pi/2)\end{pmatrix}\right|_{(a,b)}}{\left|\begin{pmatrix}\cos(\alpha)\\ \sin(\alpha)\end{pmatrix}\right|_{(a,b)}} + \frac{\left|\begin{pmatrix}\cos(\alpha)\\ \sin(\alpha)\end{pmatrix}\right|_{(a,b)}}{\left|\begin{pmatrix}\cos(\alpha+\pi/2)\\ \sin(\alpha+\pi/2)\end{pmatrix}\right|_{(a,b)}}\right). \tag{12}$$

We call therefore (12) the non-Euclidean geometric representation for the correlation coefficient of the regular Gaussian law. This representation demonstrates that ρ relies on the underlying geometry of the density level sets of $(\xi,\eta)^T$. In detail, in representation (10) the correlation coefficient ρ depends not only on the rotation angle α and the ratio of the half lengths a and b of the principal axes, but even both of these quantities are weighted and mixed in a sophisticated way. In the non-Euclidean geometric representation (12), the correlation coefficient ρ depends on the rotation angle α and the ratio of the $E_{(a,b)}$-generalized lengths of the unit vectors corresponding to the polar angles α and $\alpha+\pi/2$. Recall that $\tan\alpha$ and $\tan(\alpha+\pi/2)$ are the slopes of the lines through the origin with extremal expected squared distances to $(\xi,\eta)^T$. Moreover, the quantities α, a, b are weighted and mixed by the $\pi/2$-periodical function

$$\psi(\alpha) = \frac{1}{2}\cdot\left(\frac{N_{(a,b)}(\alpha+\pi/2)}{N_{(a,b)}(\alpha)} + \frac{N_{(a,b)}(\alpha)}{N_{(a,b)}(\alpha+\pi/2)}\right),$$

which takes it's minimal value 1 at $\alpha = \pi/4$ and it's maximal value $\tilde{\psi} = (a/b + b/a)/2$ at $\alpha = 0$. In this context, the π-periodicity of the function $\rho(\alpha)$ follows from that of the function $\sin(2\alpha)$, because π is the least common multiple of the periods of both functions, see Figure 14(a). Notice that in Figure 14(b) the level l of $\tilde{\psi}$ tends to ∞ if $b/a \to 0$.

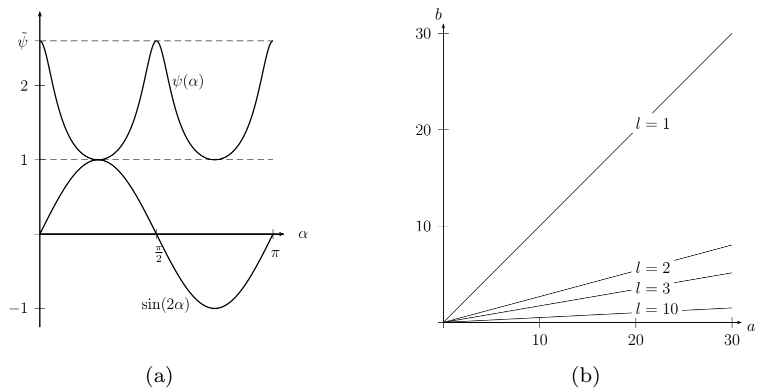


Figure 14. (a) The functions $\sin(2\alpha)$ and $\psi(\alpha)$.
(b) Level sets of $\tilde{\psi}$ considered as a function of a and b, where $a > b > 0$.

As a consequence of this discussion, we will use in what follows the rotation angle α instead of the correlation coefficient ρ to express the stochastic and linear dependencies between ξ and η, particularly with regard to the fact that every such dependency can be eliminated by a clockwise oriented rotation through the angle α. Finally, let us remark that the case $a = b$ entails $\rho \equiv 0$ for any rotation angle and $\alpha = 0$ at the same time.

To summarize this section, we point out that there are good reasons to describe the regular Gaussian distribution with the help of the parameter vector (a, b, α) instead of $(\sigma_1, \sigma_2, \rho)$. This reparametrization of the Gaussian density will be introduced in the next section.

4. The principal axes parametrization of the Gaussian law

In this section, we derive a parametrization of the regular Gaussian law which reflects on the one hand the non-Euclidean geometric nature of the stochastic representation (9) from Subsection 2.6 and which is on the other hand closely connected with a basic idea in principal component analysis. Regardless of the multiplicity of books and papers concerning the last mentioned subject, see for example Anderson (2003), Hyvärinen et al. (2001), Jolliffe (2004) and Muirhead (1982), the authors were not aware in the literature of the parametrization presented in this section.

Consider $\rho \in (0, 1)$,

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\,\sigma_1\sigma_2 \\ \rho\,\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}, \qquad |(x, y)^T|_\Sigma := (x, y)\Sigma^{-1}(x, y)^T,$$

and let the vector $(\xi, \eta)^T$ be distributed according to a regular normal distribution with expectation vector $\mu = (\mu_1, \mu_2)^T \in \mathrm{R}^2$ and regular covariance matrix Σ. Using the Σ-norm $|\cdot|_\Sigma$, the density of the random vector $(\xi, \eta)^T$ may be written as

$$f_{(\xi,\eta)}(x, y) = \frac{1}{2\pi\,(det\Sigma)^{\frac{1}{2}}} \exp\left\{ -\frac{1}{2} \left| \begin{pmatrix} x - \mu_1 \\ y - \mu_2 \end{pmatrix} \right|_\Sigma^2 \right\}, \quad (x, y)^T \in \mathrm{R}^2.$$

It follows from Subsection 2.6 that $det\Sigma = a^2b^2$ and

$$f_{(\xi,\eta)}(x, y) = \frac{1}{2\,\pi_{(a,b)}} \exp\left\{ -\frac{1}{2} \left| D(\alpha) \begin{pmatrix} x - \mu_1 \\ y - \mu_2 \end{pmatrix} \right|_{(a,b)}^2 \right\} \tag{13}$$

where the triple of parameters (a, b, α) has the meaning as in (6) and (7). This includes that α is uniquely determined in $[0, \pi/2)$, $a > 0$, $b > 0$, and that $D(\alpha)$ satisfies $D(\alpha)\Sigma^{-1}D(\alpha)^T = \mathrm{diag}\left(a^{-2}, b^{-2}\right)$. Moreover, the pdf of $(\xi, \eta)^T$ can be represented as

$$\begin{aligned} f_{(\xi,\eta)}(x, y) &= \frac{1}{2\pi_{(a,b)}} \exp\left\{ -\frac{1}{2} \left(N^2_{(a,b)}(\alpha)(x - \mu_1)^2 + N^2_{(a,b)}(\alpha + \pi/2)(y - \mu_2)^2 \right) \right\} \\ &\qquad \cdot \exp\left\{ \sin(2\alpha)\, N^2_{(a,b)}(\pi/4)\, \rho_*\, (x - \mu_1)(y - \mu_2) \right\} \qquad (14) \\ &= \frac{1}{2\,\pi_{(a,b)}} \exp\left\{ -\frac{s_1^2(x - \mu_1)^2 + s_2^2(y - \mu_2)^2 - 2c^2\,\rho_* \sin(2\alpha)(x - \mu_1)(y - \mu_2)}{2} \right\} \end{aligned}$$

where $c^2 = (a^2 + b^2)/(2\,a^2b^2)$. Here, $\rho_* = (a^2 - b^2)/(a^2 + b^2)$ is according to (11) the maximum or minimum value of the correlation coefficient $\rho(\alpha)$ for given $a > b$ or $b > a$, respectively. Recall that according to Richter (2011) the quantity $\pi_{(a,b)} = \pi\, a\, b$ denotes the ellipse number of $E_{(a,b)}$. The quantities s_1^2, s_2^2 and c^2 allow the following geometric interpretations:

$$s_1^2 = \left|\begin{pmatrix}\cos\alpha\\ \sin\alpha\end{pmatrix}\right|^2_{(a,b)}, \quad s_2^2 = \left|\begin{pmatrix}\cos\left(\alpha+\frac{\pi}{2}\right)\\ \sin\left(\alpha+\frac{\pi}{2}\right)\end{pmatrix}\right|^2_{(a,b)} \quad \text{and} \quad c^2 = \left|\begin{pmatrix}\cos\frac{\pi}{4}\\ \sin\frac{\pi}{4}\end{pmatrix}\right|^2_{(a,b)},$$

see Figure 15(a). Here, the vector $(\cos\pi/4, \sin\pi/4)^T$ corresponds to the angle $\alpha_* = \pi/4$, which is according to Section 3 connected with the extremal correlation coefficient ρ_*.

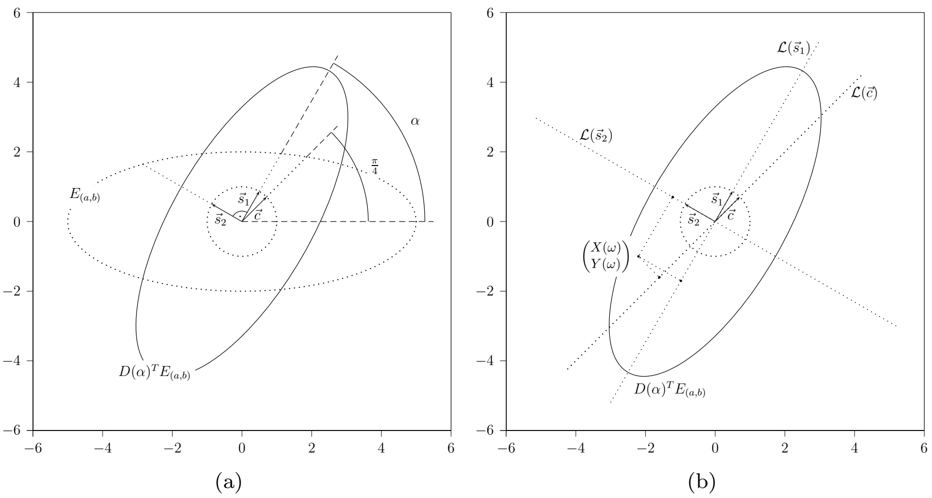


Figure 15. (a) The quantities s_1, s_2 and c are the $|\cdot|_{(a,b)}$-lengths of the vectors $\vec{s}_1 = \begin{pmatrix}\cos\alpha\\ \sin\alpha\end{pmatrix}$, $\vec{s}_2 = \begin{pmatrix}\cos(\alpha+\pi/2)\\ \sin(\alpha+\pi/2)\end{pmatrix}$ and $\vec{c} = \begin{pmatrix}\cos\pi/4\\ \sin\pi/4\end{pmatrix}$, respectively.
(b) The quantities $a^2b^2s_1^2$, $a^2b^2s_2^2$ and $a^2b^2c^2$ are variances of the coefficients corresponding to the orthogonal projections of $(X,Y)^T$ onto the one-dimensional subspaces $\mathcal{L}(\vec{s}_2)$, $\mathcal{L}(\vec{s}_1)$ and $\mathcal{L}(\vec{c})$, respectively.

The quantities s_1^2, s_2^2 and c^2 can also be considered from a stochastic point of view as

$$a^2b^2s_1^2 = V\left(\left\langle\begin{pmatrix}X\\Y\end{pmatrix}, \begin{pmatrix}\cos(\alpha+\pi/2)\\ \sin(\alpha+\pi/2)\end{pmatrix}\right\rangle\right), \quad a^2b^2s_2^2 = V\left(\left\langle\begin{pmatrix}X\\Y\end{pmatrix}, \begin{pmatrix}\cos\alpha\\ \sin\alpha\end{pmatrix}\right\rangle\right)$$

and $a^2b^2c^2 = V\left(\langle (X,Y)^T, (\cos\pi/4, \sin\pi/4)^T\rangle\right)$ where $(X,Y)^T = D(\alpha)\left((\xi,\eta)^T - \mu\right)$ and $\langle\cdot,\cdot\rangle$ denotes the Euclidean inner product in R^2, see Figure 15(b). With regard to principal component analysis, the components of $D(\beta)(X,Y)^T = D(\alpha+\beta)((\xi,\eta)^T - \mu)$ have variances lower than $\max\{a^2,b^2\}$ and greater than $\min\{a^2,b^2\}$ for any β not beeing a multiple of $\pi/2$ and $a \neq b$. In this context,

$$\min\{a^2,b^2\} < V(\xi) < \max\{a^2,b^2\} \quad \text{and} \quad \min\{a,^2 b^2\} < V(\eta) < \max\{a^2,b^2\}$$

if $\alpha \neq 0$ and $a \neq b$. The formula (13) or any of the representations in (14) will be called the principal axes representation of the two-dimensional regular Gaussian density with parameters (a,b,α) satisfying $a > 0$, $b > 0$ and $\alpha \in [0,\pi/2)$, $(\xi,\eta)^T \sim \Phi^{PA}_{a,b,\alpha}$, where $a = b$ implies $\alpha = 0$.

5. Point Estimation

5.1 Estimation of the semi-major and the semi-minor axes lengths

In this subsection, we consider a sample of i.i.d. random vectors $(\xi_1, \eta_1)^T, \ldots, (\xi_n, \eta_n)^T$ from a two-parameter Gaussian density with expectation $\mu = 0_2$, which can be written according to the principal axes representation (13) as

$$f_{(\xi,\eta)}(x,y) = \frac{1}{2\pi ab} \cdot \exp\left\{ -\frac{1}{2} \left(\frac{(x\cos\alpha + y\sin\alpha)^2}{a^2} + \frac{(-x\sin\alpha + y\cos\alpha)^2}{b^2} \right) \right\} \quad (15)$$

where $(a,b) \in \Theta_2$ is the unknown parameter vector, $\Theta_2 = \{(x,y) \in \mathrm{R}^2 : x > 0,\ y > 0,\ x \neq y\}$, and $\alpha \in [0, \pi/2)$ is assumed to be a known angle of rotation. The maximum likelihood estimators of a^2 and b^2 are a.s. given by

$$\widehat{a^2} = \frac{1}{n} \sum_{i=1}^{n} (\xi_i \cos\alpha + \eta_i \sin\alpha)^2 \qquad \text{and} \qquad \widehat{b^2} = \frac{1}{n} \sum_{i=1}^{n} (\eta_i \cos\alpha - \xi_i \sin\alpha)^2 .$$

We observe that $\widehat{a^2}$ and $\widehat{b^2}$ are scaled squares of the row-norms of the sample

$$\begin{pmatrix} \xi \\ \eta \end{pmatrix}_{(n)} := \begin{pmatrix} \xi_1 & \xi_2 & \cdots & \xi_n \\ \eta_1 & \eta_2 & \cdots & \eta_n \end{pmatrix}$$

rotated clockwise through the angle α w.r.t. the origin 0_2, i.e.

$$\widehat{a^2} = \frac{1}{n} \left\| \begin{pmatrix} \cos\alpha & \sin\alpha \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix}_{(n)} \right\|^2 \qquad \text{and} \qquad \widehat{b^2} = \frac{1}{n} \left\| \begin{pmatrix} -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix}_{(n)} \right\|^2 \quad (16)$$

where $\| \cdot \|$ denotes the Euclidean norm in R^n. The exact and asymptotic distributions of $\widehat{a^2}$ and $\widehat{b^2}$ will be determined in Section 6.

5.2 The case of three unknown parameters

Let $(\xi, \eta)^T_{(n)} = (\xi_{(n)}, \eta_{(n)})^T$ again be a sample from the Gaussian density in (15) where now $(a, b, \alpha) \in \Theta_3 = \Theta_2 \times [0, \pi/2)$ is assumed to be completely unknown. The maximum likelihood estimators of a^2, b^2 and α are a.s. given by

$$\widehat{a^2} = \frac{1}{n} \left\| \begin{pmatrix} \cos\hat\alpha & \sin\hat\alpha \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix}_{(n)} \right\|^2 , \quad \widehat{b^2} = \frac{1}{n} \left\| \begin{pmatrix} -\sin\hat\alpha & \cos\hat\alpha \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix}_{(n)} \right\|^2$$

and

$$\widehat{\alpha} = \begin{cases} \frac{1}{2} \arctan \frac{Z}{N} & , \text{ if } Z \cdot N > 0 \\ \frac{1}{2} \arctan \frac{Z}{N} + \frac{\pi}{2} & , \text{ if } Z \cdot N < 0 \end{cases}$$

where $Z = 2 \langle \xi_{(n)}, \eta_{(n)} \rangle$, $N = \left\| \xi_{(n)} \right\|^2 - \left\| \eta_{(n)} \right\|^2$ and where $\langle \cdot , \cdot \rangle$ denotes from now on the Euclidean inner product in R^n. Notice that the maximum likelihood approach for estimating (a, b, α) is in this case equivalent to applying $T^{-1}_{(AA,R)}$ to the suitably transformed components of

$$\widehat{\Sigma} = \frac{1}{n} \begin{pmatrix} \sum_{i=1}^{n} \xi_i^2 & \sum_{i=1}^{n} \xi_i \eta_i \\ \sum_{i=1}^{n} \xi_i \eta_i & \sum_{i=1}^{n} \eta_i^2 \end{pmatrix},$$

i.e.

$$T^{-1}_{(AA,R)} \left(\sqrt{\frac{1}{n} \sum_{i=1}^{n} \xi_i^2}, \sqrt{\frac{1}{n} \sum_{i=1}^{n} \eta_i^2}, \frac{\sum_{i=1}^{n} \xi_i \eta_i}{\sqrt{\sum_{i=1}^{n} \xi_i^2} \sqrt{\sum_{i=1}^{n} \eta_i^2}} \right) = \left(\sqrt{\widehat{a^2}}, \sqrt{\widehat{b^2}}, \widehat{\alpha} \right).$$

Moreover, the result of this approach is also equal to that of the method of moments approach taking into account the three moments of second order $\mathrm{E}\xi_1^2$, $\mathrm{E}\eta_1^2$ and $\mathrm{E}\xi_1\eta_1$. In this context, the diagonalisation of $\widehat{\Sigma}$ leads according to equations (8) to the representations

$$\max\left\{\widehat{a^2}, \widehat{b^2}\right\} = \frac{1}{2n} \left(\|\xi_{(n)}\|^2 + \|\eta_{(n)}\|^2 + \sqrt{\left(\|\xi_{(n)}\|^2 - \|\eta_{(n)}\|^2 \right)^2 + 4 \left\langle \xi_{(n)}, \eta_{(n)} \right\rangle^2} \right) \quad (17)$$

and

$$\min\left\{\widehat{a^2}, \widehat{b^2}\right\} = \frac{1}{2n} \left(\|\xi_{(n)}\|^2 + \|\eta_{(n)}\|^2 - \sqrt{\left(\|\xi_{(n)}\|^2 - \|\eta_{(n)}\|^2 \right)^2 + 4 \left\langle \xi_{(n)}, \eta_{(n)} \right\rangle^2} \right) \quad (18)$$

for the maximum likelihood estimators of $\max\{a^2, b^2\}$ and $\min\{a^2, b^2\}$.

6. Distribution of Point Estimators

6.1 The MLEs of a^2 and b^2 in case of a known rotation angle α

Consider the maximum likelihood estimators $\widehat{a^2}$ and $\widehat{b^2}$ in Subsection 5.1, i.e. in the case of a known angle α. As $D(\alpha) \left(\xi_i \; \eta_i\right)^T \sim \Phi_{0_2, \mathrm{diag}(a^2, b^2)}$ according to Subsection 2.6, it follows immediately from the representations in (16) that $n\, \widehat{a^2}/a^2 \sim \chi_n^2$ and $n\, \widehat{b^2}/b^2 \sim \chi_n^2$. For any finite sample size $\widehat{a^2}$ and $\widehat{b^2}$ are unbiased, stochastically independent and have variances $2a^4/n$ and $2b^4/n$, respectively. For increasing sample sizes it follows from the law of large numbers that a.s.

$$\begin{pmatrix} \widehat{a^2} \\ \widehat{b^2} \end{pmatrix} \underset{n\to\infty}{\longrightarrow} \begin{pmatrix} a^2 \\ b^2 \end{pmatrix}.$$

Also notice that maximum likelihood estimators in exponential families are consistent and asymptotically normal, see for example Lehmann (1983). Hence,

$$\mathfrak{L}\left(\sqrt{n} \begin{pmatrix} \widehat{a^2} - a^2 \\ \widehat{b^2} - b^2 \end{pmatrix} \right) \Longrightarrow \Phi_{0_2, \left(\begin{smallmatrix} 2a^4 & 0 \\ 0 & 2b^4 \end{smallmatrix}\right)}, \quad n \to \infty.$$

6.2 The three parameter case

The distributions of the maximum likelihood estimators of $\lambda_1 := \max\{a^2, b^2\}$ and $\lambda_2 := \min\{a^2, b^2\}$ from Subsection 5.2, that is in case of an unknown rotation angle α, have been studied, e.g., in Anderson (2003) and Muirhead (1975, 1982). After a brief review of their results we will give representations for the cdfs of the maximum likelihhood estimators $\widehat{\lambda_1}$ and $\widehat{\lambda_2}$ which seem to be new, see Theorem 6.1.

Consider the maximum likelihood estimator $\widehat{\lambda_1}$ of $\max\{a^2, b^2\}$, which is the greatest eigenvalue of $\widehat{\Sigma}$, see Subsection 5.2. It is a well known fact that $n\,\widehat{\Sigma} \sim W_2(\Sigma, n)$ where $W_2(\Sigma, n)$ denotes the two-dimensional Wishart distribution with n degrees of freedom, i.e. $n\,\widehat{\lambda_1}$ is the largest eigenvalue of a $W_2(\Sigma, n)$-distributed random matrix. In Muirhead (1982) a representation for the distribution function of the largest eigenvalue of a $W_2(\Sigma, n)$-distributed random matrix is given and it follows immediately from there that

$$P\left(\widehat{\lambda_1} < x\right) = \frac{\Gamma\left(\sqrt{\pi}\right)}{2\Gamma\left(\frac{n+3}{2}\right)\Gamma\left(\frac{n+2}{2}\right)}\left(\frac{nx}{2ab}\right)^n {}_1F_1\left(\frac{n}{2}; \frac{n+3}{2}; -\frac{nx}{2}\,\mathrm{diag}\left(a^{-2}, b^{-2}\right)\right) \tag{19}$$

where

$${}_1F_1(c; d; M) = \sum_{k=0}^{\infty}\sum_{\kappa(k)} \frac{(c)_{(k)}}{(d)_{(k)}}\frac{C_\kappa(M)}{k!}$$

is the hypergeometric function with matrix argument $M \in \mathrm{R}^{2\times 2}$ and parameters $c > 0$, $d > 0$. For a precise definition of ${}_pF_q(c_1, \ldots, c_p; d_1, \ldots, d_q; M)$, $M \in \mathrm{R}^{m\times m}$, $m \in \mathrm{N}$, a partition $\kappa(k) = (k_1, \ldots, k_m)$ of k, the zonal polynomial C_κ of order κ and the notation $(c)_{(k)}$, we refer to Muirhead (1982). The hypergeometric function with matrix argument $M \in \mathrm{R}^{2\times 2}$ was expressed in Muirhead (1975) as an infinite sum involving hypergeometric functions with scalar arguments. From there, we have

$$P\left(\widehat{\lambda_1} < x\right) = \tau\,(nx)^n \exp\left(-\frac{nx}{2}\left(a^{-2} + b^{-2}\right)\right)[B - A]\,, \quad x > 0\,, \tag{20}$$

with

$$\tau = 2^{-n-1}\sqrt{\pi}\,(ab)^{-n}\,\Gamma\left(\frac{n+3}{2}\right)\Gamma\left(\frac{n+2}{2}\right),$$

$$A = n\sum_{k=0}^{\infty}\frac{\left(\frac{1}{2}\right)_{(k)}}{\left(\frac{n+3}{2}\right)_{(k)}}\frac{\left(\frac{nx}{2}\left(a^{-2} + b^{-2}\right)\right)^k}{k!}\,{}_2F_1\left(-\frac{k}{2}, \frac{1}{2} - \frac{k}{2}; \frac{1}{2} - k; 4\left(\frac{ab}{a^2 + b^2}\right)^2\right)$$

and

$$\frac{B}{n+1} = \sum_{k=0}^{\infty}\frac{\left(\frac{1}{2}\right)_{(k)}}{\left(\frac{n+1}{2}\right)_{(k)}}\frac{\left(\frac{nx}{2}\left(a^{-2} + b^{-2}\right)\right)^k}{k!}\,{}_3F_2\left(-\frac{k}{2}, \frac{1}{2} - \frac{k}{2}, \frac{n}{2}; \frac{1}{2} - k, \frac{n}{2} + 1; 4\left(\frac{ab}{a^2 + b^2}\right)^2\right).$$

Alternatively to the representations (19) and (20), which involve series of zonal polynomials, we present in the following theorem a new integral representation for the cdf $F_{\widehat{\lambda_1}}$ of $\widehat{\lambda_1}$ and additionally a new integral representation for the cdf $F_{\widehat{\lambda_2}}$, i.e. the cdf of the smallest eigenvalue of the $W_2(\Sigma, n)$-distributed random matrix $n\,\widehat{\Sigma}$. It will turn out that these integral representations are particularly useful for the numerical evaluation of $F_{\widehat{\lambda_1}}$ and $F_{\widehat{\lambda_2}}$, see Section 7.

THEOREM 6.1 The cdf of $\widehat{\lambda_1}$ satisfies the representation

$$F_{\widehat{\lambda_1}}(x) = \Bigg(\frac{\omega_n^2}{(2\pi_{(a,b)})^n} \int_{r=0}^{\sqrt{nx}} \int_{s=0}^{\sqrt{nx-r^2}} (rs)^{n-1} e^{-\frac{r^2}{2a^2}-\frac{s^2}{2b^2}}\, ds\, dr$$

$$+ \frac{2\omega_n\omega_{n-1}}{(2\,\pi_{(a,b)})^n} \int_{r=0}^{\sqrt{nx}} \int_{s=\sqrt{nx-r^2}}^{\sqrt{nx}} (rs)^{n-1} e^{-\frac{r^2}{2a^2}-\frac{s^2}{2b^2}} \int_{\arccos\sqrt{g}}^{\frac{\pi}{2}} \sin^{n-2}(\psi_1)\, d\psi\, ds\, dr \Bigg) 1_{(0,\infty)}(x)$$

and the cdf $F_{\widehat{\lambda_2}}$ satisfies the representation

$$F_{\widehat{\lambda_2}}(x) = \Bigg(\frac{\omega_n^2}{(2\,\pi_{(a,b)})^n} \Bigg[\int_{r=0}^{\sqrt{nx}} \int_{s=\sqrt{nx}}^{\infty} (rs)^{n-1} e^{-\frac{r^2}{2a^2}-\frac{s^2}{2b^2}}\, ds\, dr$$

$$+ \int_{r=\sqrt{nx}}^{\infty} \int_{s=0}^{\sqrt{nx}} (rs)^{n-1} e^{-\frac{r^2}{2a^2}-\frac{s^2}{2b^2}}\, ds\, dr \Bigg]$$

$$+ \frac{2\omega_n\omega_{n-1}}{(2\,\pi_{(a,b)})^n} \int_{r=\sqrt{nx}}^{\infty} \int_{s=\sqrt{nx}}^{\infty} (rs)^{n-1} e^{-\frac{r^2}{2a^2}-\frac{s^2}{2b^2}} \int_{0}^{\arccos\sqrt{g}} \sin^{n-2}(\psi_1)\, d\psi\, ds\, dr \Bigg) 1_{(0,\infty)}(x)$$

where $g := g(r,s;n,x) := 1 + nx\,(nx - r^2 - s^2)/(r^2 s^2)$ and $\omega_k = 2\pi^{k/2}/\Gamma\,(k/2)$ denotes the Euclidean surface area of the k-dimensional unit sphere S_k, $k \geq 1$.

PROOF There will be given a scetch of the proof in the Appendix. ∎

REMARK 6.1 The representations in Theorem 6.1 can be reformulated on using

$$\int_{\arccos\sqrt{g}}^{\frac{\pi}{2}} \sin^{n-2}(\psi)\, d\psi = \begin{cases} \sum\limits_{k=1}^{\frac{n}{2}-1} \frac{\sqrt{g}(1-g)^{\frac{n-1}{2}-k}}{n-2} \cdot C_k + \frac{\left(\frac{\pi}{2}-\arccos\sqrt{g}\right)}{n-2} \cdot C_{\frac{n}{2}-1} & \text{, if } n \text{ is even} \\ \sum\limits_{k=1}^{\frac{n-1}{2}} \frac{\sqrt{g}(1-g)^{\frac{n-1}{2}-k}}{n-2} \cdot C_k & \text{, if } n \text{ is odd} \end{cases},$$

where $n \geq 3$,

$$C_k = \prod_{i=1}^{k-1} \frac{n-2i-1}{n-2i-2} = \frac{\left(\frac{3-n}{2}\right)_{(k-1)}}{\left(\frac{4-n}{2}\right)_{(k-1)}}, \quad k \in \mathrm{N},\ 1 \leq k \leq \frac{n-1}{2},$$

and $\prod_{i=1}^{0} a_i := 1,\ a_i \in \mathrm{R}$.

PROOF Let $n \geq 3$. Integration by parts yields

$$\int_{\arccos\sqrt{g}}^{\frac{\pi}{2}} \sin^{n-2}(\psi)\, d\psi = \frac{\sqrt{g}\cdot(1-g)^{\frac{n-3}{2}}}{n-2} + \frac{n-3}{n-2} \int_{\arccos\sqrt{g}}^{\frac{\pi}{2}} \sin^{n-4}(\psi)\, d\psi\,,$$

which can be used to prove the representation of $\int_{\arccos\sqrt{g}}^{\frac{\pi}{2}} \sin^{n-2}(\psi)\, d\psi$ by induction on

n. Notice that

$$\int_0^{\frac{\pi}{2}} \sin^{n-2}(x)\ dx = \frac{\sqrt{\pi}\cdot\Gamma\left(\frac{n-1}{2}\right)}{2\cdot\Gamma\left(\frac{n}{2}\right)}\,, \quad n \geq 2\,.$$

Further details of the proof are left to the reader. ■

There are different reasons for the usefulness of Theorem 6.1. First of all, the integral representations of the cdfs $F_{\widehat{\lambda_1}}$ and $F_{\widehat{\lambda_2}}$ in Theorem 6.1 can be easier implemented than the infinite series representation in (19). Notice here that neither the hypergeometric function with matrix argument M nor the involved zonal polynomials $C_\kappa(M)$ are available as basic functions of common computing environments. A second reason is the numerical stability of the involved integrals for large arguments x of the cdfs and small values of the sample size n. Using an adaptive quadrature with fault tolerance ε, $F_{\widehat{\lambda_1}}(x)$ can be approximated very accurately even in this case, see Section 7.

Remark 6.2 Instead of the one-to-one map $T_{(AA,R)}$ in Subsection 2.6 we could have introduced a quite similar map $T^*_{(AA,R)}$ where the parameters a and b, however, are assumed to be arranged in descending order, $a \geq b$, and where α has domain $[0,\pi)$. Considering $T^*_{(AA,R)}$ would involve the advantages that differently from (11) any correlation $-\rho_* \leq \rho \leq \rho_*$ could be attained by a suitable specification of the rotation angle $\alpha \in [0,\pi)$ and that the cdfs of the maximum-likelihood estimators of a^2 and b^2 could be computed, e.g., on the basis of Theorem 6.1. At the same time, however, the evaluation of the rotation angle α would become more complicated than in (6). The same would hold for the maximum-likelihood estimators of a^2 and b^2 in (16) if $\alpha \in [0,\pi)$.

7. Example and Outlook

Consider a sample of i.i.d. random vectors $(\xi_1,\eta_1)^T,\ldots,(\xi_n,\eta_n)^T$ from the Gaussian distribution $\Phi_{0_2,\Sigma}$ with

$$\Sigma = \frac{1}{4}\begin{pmatrix} 5 & \sqrt{3} \\ \sqrt{3} & 7 \end{pmatrix}$$

and let ρ denote the correlation coefficient of ξ_1 and η_1. Then $\rho = \sqrt{3/35} \approx 0.29277$ and $T^{-1}_{(AA,R)}(\sigma_1,\sigma_2,\rho) = \left(\sqrt{2},1,\pi/3\right)$. Therefore, Σ has principal axes of length $2a = 2\sqrt{2}$, $2b = 2$ and $(\xi_1,\eta_1)^T$ is distributed as the product of the generalized radius $\mathcal{R}_{(\sqrt{2},1)}$ and the generalized uniform basis vector $D(\pi/3)^T\,\mathcal{U}_{(\sqrt{2},1)}$, $(\xi_1,\eta_1)^T \sim \Phi^{PA}_{\sqrt{2},1,\pi/3}$, where $\mathcal{R}_{(\sqrt{2},1)} \sim \chi_2$ and $\mathcal{U}_{(\sqrt{2},1)} \sim \omega_{(\sqrt{2},1)}$ are independent. The maximal correlation attainable (by rotation) in this configuration of a and b is given by $\rho_* = (a^2-b^2)/(a^2+b^2) = 1/3$, see (11). Hence, ξ_1 and η_1 can be regarded as kind of nearly maximally correlated in this configuration as ρ is already pretty close to ρ_*. A simulation of the sample considered above for different values of n on the basis of the statistical software R gives realisations of the MLEs from Section 5.2, see Table 1. As can be seen there, the introduced estimates approximate the corresponding parameters very accurately if the sample size is large enough. In this context, the theoretical investigation of sufficient sample sizes for the considered estimators is an interesting open problem, which may be focussed in a future work. Regardless of this aspect, it is helpful to know the distributions of the point estimators in order to use the principal axes parametrization of the Gaussian law in statistical modelling. On the one hand, statistical inference for large sample sizes n can be based on the theory of exponential

Table 1. Realisations of $\widehat{\lambda_1} = \max\left\{\widehat{a^2}, \widehat{b^2}\right\}$, $\widehat{\lambda_2} = \min\left\{\widehat{a^2}, \widehat{b^2}\right\}$ and $\widehat{\alpha}$ where $a^2 = 2$, $b^2 = 1$ and $\alpha = \pi/3 \approx 1.0472$.

MLE \ n	100	1000	10000	100000	1000000
$\widehat{\lambda_1}$	1.917	2.045	2.025	2.005	2.001
$\widehat{\lambda_2}$	1.095	0.977	1.019	1.003	0.999
$\widehat{\alpha}$	1.195	1.117	1.039	1.041	1.046

families. Here,

$$\mathcal{L}\left(\sqrt{n}\begin{pmatrix}\widehat{a^2} - 2\\ \widehat{b^2} - 1\\ \widehat{\alpha} - \pi/3\end{pmatrix}\right) \Longrightarrow \Phi_{0_3,\Psi}\,, \quad n \to \infty\,,$$

where $\Psi = \operatorname{diag}\left(2a^4, 2b^4, a^2b^2/\left((a^2-b^2)^2\right)\right)\big|_{a^2=2,b^2=1} = \operatorname{diag}(8,2,2)$, see, e.g., Lehmann (1983). On the other hand, dealing with a small sample size, the cumulative distribution functions $F_{\widehat{\lambda_1}}$ and $F_{\widehat{\lambda_2}}$ can be approximated very accurately on the basis of Theorem 6.1, even if the corresponding argument x of the cdf is very large. In this regard, a numerical integration using the adaptive Lobatto-rule w.r.t. an estimated absolute error tolerance $\varepsilon = 10^{-9}$ leads to an approximated cdf $\widetilde{F}_{\widehat{\lambda_1}}$ and to the quantile approximations presented in Table 2. For a review of the basic principles of adaptive quadrature see, e.g., Gander and Gautschi (2000). Ensuring

$$F_{\widehat{\lambda_1}}\left(\widetilde{F}^{-1}_{\widehat{\lambda_1}}(q) + 10^{-4}\right) \geq \tilde{F}_{\widehat{\lambda_1}}\left(\widetilde{F}^{-1}_{\widehat{\lambda_1}}(q) + 10^{-4}\right) - \varepsilon > q\,,$$
$$F_{\widehat{\lambda_1}}\left(\widetilde{F}^{-1}_{\widehat{\lambda_1}}(q) - 10^{-4}\right) \leq \tilde{F}_{\widehat{\lambda_1}}\left(\widetilde{F}^{-1}_{\widehat{\lambda_1}}(q) - 10^{-4}\right) + \varepsilon < q\,,$$

for $q \in \{0.95, 0.99, 0.999, 0.9999\}$ where the respective first inequality is caused by the adaptive Lobatto-rule, we say that these quantile approximations are at least correct to the first 4 decimals. In contrast, the approximation of $F_{\widehat{\lambda_1}}(x)$ using representation (19) becomes difficult if x is large, although the authors in Koev and Edelman (2006) made an implementation of the right side in (19) available. Their implementation is very fast and effective for moderate values of x but also numerically instable for large values of x and could not be used to get quantile approximations for $F_{\widehat{\lambda_1}}$. Even the representation in (20) seems to be less appropriate than the one in Theorem 6.1 if x is sufficiently large, although the approximation of $F_{\widehat{\lambda_1}}(x)$ on the basis of (20) is also stable in this case. However, the approximations of $F_{\widehat{\lambda_1}}(x)$ determined on the basis of (20) are less accurate due to the very slow convergence of the involved infinite series. The corresponding quantile approximations of $F_{\widehat{\lambda_1}}$ differ significantly from those determined with the help of Theorem 6.1, see Table 2. Let us finally state that it is still an open problem to determine the cdf of $\widehat{\alpha}$ and to discuss the dependence-independence properties of all estimators.

Acknowledgements

The authors are grateful to the Referee and Editor-in-Chief whose valuable hints led to a better presentation of the material.

Table 2. Quantile approximations of $F_{\widehat{\lambda_1}}$ for $a^2 = 2$, $b^2 = 1$ and $n = 5$ on the basis of representation (20) using the first 80 summands and Theorem 6.1, respectively. Notice that the quantile approximations according to Theorem 6.1 are at least correct to the first 4 decimals.

Method	$\widetilde{F}^{-1}_{\widehat{\lambda_1}}(0.95)$	$\widetilde{F}^{-1}_{\widehat{\lambda_1}}(0.99)$	$\widetilde{F}^{-1}_{\widehat{\lambda_1}}(0.999)$	$\widetilde{F}^{-1}_{\widehat{\lambda_1}}(0.9999)$
Theorem 6.1	4.9479	6.6903	9.0472	11.2825
Representation (20)	4.7587	6.3471	8.5074	10.5938

Appendix A. Proof of Theorem 6.1

We consider a sample $(\xi_i, \eta_i)^T$, $i = 1, \ldots, n$ of i.i.d. $(\xi_i, \eta_i)^T$, $(\xi_i, \eta_i)^T \sim \Phi_{0_2,\Sigma}$, with regular covariance matrix $\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$, $\rho \neq 0$ and $\xi_{(n)}^T = (\xi_1, \ldots, \xi_n)$, $\eta_{(n)}^T = (\eta_1, \ldots, \eta_n)$. Consequently, $\begin{pmatrix} \xi_{(n)} \\ \eta_{(n)} \end{pmatrix} \sim \Phi_{0_{2n},\Sigma_{2n\times 2n}}$ with $\Sigma_{2n\times 2n} := \begin{pmatrix} \Sigma_1 & \Sigma_{12} \\ \Sigma_{12} & \Sigma_2 \end{pmatrix}$, $\Sigma_i = \sigma_i^2 I_n$, $i = 1, 2$, $\Sigma_{12} = \rho\sigma_1\sigma_2 I_n$ and I_n being the $n \times n$ identity matrix. Let (a, b, α) be the uniquely determined triple corresponding to $(\sigma_1, \sigma_2, \rho)$. It follows from (17) and (18) in Subsection 5.2 that

$$P\left(\widehat{\lambda_1} < t\right) = P\left(\begin{pmatrix} \xi_{(n)} \\ \eta_{(n)} \end{pmatrix} \in A_1(t)\right) = \Phi_{0_{2n},\Sigma_{2n\times 2n}}\big(A_1(t)\big)$$

and that $P\left(\widehat{\lambda_2} < t\right) = \Phi_{0_{2n},\Sigma_{2n\times 2n}}\big(A_2(t)\big)$ where $t > 0$, $\widehat{\lambda_1} := \max\left\{\widehat{a^2}, \widehat{b^2}\right\}$, $\widehat{\lambda_2} := \min\left\{\widehat{a^2}, \widehat{b^2}\right\}$,

$$A_1(t) = \left\{ \begin{pmatrix} v_{(n)} \\ w_{(n)} \end{pmatrix} \,\middle|\, \|v_{(n)}\|^2 + \|w_{(n)}\|^2 + \sqrt{\left(\|v_{(n)}\|^2 - \|w_{(n)}\|^2\right)^2 + 4\left\langle v_{(n)}, w_{(n)} \right\rangle^2} < 2nt \right\},$$

$$A_2(t) = \left\{ \begin{pmatrix} v_{(n)} \\ w_{(n)} \end{pmatrix} \,\middle|\, \|v_{(n)}\|^2 + \|w_{(n)}\|^2 - \sqrt{\left(\|v_{(n)}\|^2 - \|w_{(n)}\|^2\right)^2 + 4\left\langle v_{(n)}, w_{(n)} \right\rangle^2} < 2nt \right\},$$

and where $v_{(n)}, w_{(n)}$ are n-dimensional Euclidean vectors. The Borel sets $A_i(t)$, $i = 1, 2$, possess some kind of invariance property. Let $D(\beta) = \begin{pmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{pmatrix}$ for arbitrary $\beta \in [0, 2\pi)$ and consider for any $2n$-dimensional vector $\begin{pmatrix} v_{(n)} \\ w_{(n)} \end{pmatrix} = \begin{pmatrix} v_1 & \ldots & v_n & w_1 & \ldots & w_n \end{pmatrix}^T$ its transformation to $\begin{pmatrix} x_{(n)} \\ y_{(n)} \end{pmatrix} := (x_1, \ldots, x_n, y_1, \ldots, y_n)^T$ where

$$\begin{pmatrix} x_i \\ y_i \end{pmatrix} := D(\beta) \cdot \begin{pmatrix} v_i \\ w_i \end{pmatrix} = \begin{pmatrix} v_i \cos\beta + w_i \sin\beta \\ w_i \cos\beta - v_i \sin\beta \end{pmatrix}, \quad i = 1, \ldots, n.$$

Direct calculations show that $\|v_{(n)}\|^2 + \|w_{(n)}\|^2 = \|x_{(n)}\|^2 + \|y_{(n)}\|^2$ and

$$\sqrt{\left(\|v_{(n)}\|^2 - \|w_{(n)}\|^2\right)^2 + 4\left\langle v_{(n)}, w_{(n)} \right\rangle^2} = \sqrt{\left(\|x_{(n)}\|^2 - \|y_{(n)}\|^2\right)^2 + 4\left\langle x_{(n)}, y_{(n)} \right\rangle^2},$$

which implies $\begin{pmatrix} x_{(n)} \\ y_{(n)} \end{pmatrix} \in A_i(t)$ if and only if $\begin{pmatrix} v_{(n)} \\ w_{(n)} \end{pmatrix} \in A_i(t)$, $i = 1, 2$. Hence, for $\beta = \alpha$ it follows

$$P\left(\widehat{\lambda_1} < t\right) = \Phi_{0_{2n},\Psi_{2n\times 2n}}(A_1(t)) \qquad \text{and} \qquad P\left(\widehat{\lambda_2} < t\right) = \Phi_{0_{2n},\Psi_{2n\times 2n}}(A_2(t))$$

where $\Psi_{2n\times 2n} = \begin{pmatrix} a^2 I_n & 0_{n\times n} \\ 0_{n\times n} & b^2 I_n \end{pmatrix}$ and $0_{n\times n}$ denotes the $n\times n$ zero matrix. Consider

$$\Phi_{0_{2n},\Psi_{2n\times 2n}}(A_i(t)) = \int_{A_i(t)} \frac{1}{(2\pi)^n (ab)^n} \cdot \exp\left\{ -\frac{\|x_{(n)}\|^2}{2a^2} - \frac{\|y_{(n)}\|^2}{2b^2} \right\} d(x_{(n)}, y_{(n)}) \,, \quad \text{(A1)}$$

$i = 1, 2$, and denote by $SPH_n : M_n \to \mathrm{R}^n$, $M_n := [0,\infty) \times M_n^*$, $M_n^* := [0,\pi)^{n-2} \times [0, 2\pi)$ the n-dimensional spherical coordinate transformation. We introduce n-dimensional spherical coordinates in the two n-dimensional subspaces $\mathcal{L}_1 = span\{e_1, \ldots, e_n\}$ and $\mathcal{L}_2 = span\{e_{n+1}, \ldots, e_{2n}\}$ of R^{2n}, where $\{e_1, \ldots, e_{2n}\}$ denotes the canonical basis of R^{2n}. The almost one-to-one transformation $T : M_n \times M_n \to \mathrm{R}^{2n}$ defined by

$$T(r, \overbrace{\phi_1, \ldots, \phi_{n-1}}^{=:\phi}, s, \overbrace{\psi_1, \ldots, \psi_{n-1}}^{=:\psi}) = \begin{pmatrix} SPH_n(r, \phi_1, \ldots, \phi_{n-1}) \\ SPH_n(s, \psi_1, \ldots, \psi_{n-1}) \end{pmatrix}$$

has the Jacobian

$$J_T(r, \phi, s, \psi) = J_{SPH_n}(r, \phi) \cdot J_{SPH_n}(s, \psi) = (rs)^{n-1} \prod_{k=1}^{n-1} (\sin\phi_k \sin\psi_k)^{n-1-k} \,.$$

Equation (A1) can now be written as

$$\Phi_{0_{2n},\Psi_{2n\times 2n}}(A_i(t)) = \int_{\{T^{-1}[A_i(t)]\}} \frac{J_T(r, \phi, s, \psi)}{(2\pi)^n (ab)^n} \cdot \exp\left\{ -\frac{r^2}{2a^2} - \frac{s^2}{2b^2} \right\} d(r, s, \phi, \psi) \,, \quad i = 1, 2 \,.$$

Notice that any vector $(x_{(n)}^T, y_{(n)}^T)^T \in \mathrm{R}^{2n}$ satisfies the inequalities

$$\|x_{(n)}\|^2 + \|y_{(n)}\|^2 + \sqrt{(\|x_{(n)}\|^2 - \|y_{(n)}\|^2)^2 + 4\left\langle x_{(n)}, y_{(n)} \right\rangle^2} \geq 2 \cdot \max\{\|x_{(n)}\|^2, \|y_{(n)}\|^2\} \,, \quad \text{(A2)}$$

$$\|x_{(n)}\|^2 + \|y_{(n)}\|^2 - \sqrt{(\|x_{(n)}\|^2 - \|y_{(n)}\|^2)^2 + 4\left\langle x_{(n)}, y_{(n)} \right\rangle^2} \leq 2 \cdot \min\{\|x_{(n)}\|^2, \|y_{(n)}\|^2\} \,. \quad \text{(A3)}$$

Thus, $\binom{x_{(n)}}{y_{(n)}} \in A_1(t)$ implies $\|x_{(n)}\| \leq \sqrt{nt}$ and $\|y_{(n)}\| \leq \sqrt{nt}$. The inequalities turn to equations whenever $x_{(n)} \perp y_{(n)}$. The left side of (A3) equals zero iff $y_{(n)} = q\, x_{(n)}$. Therefore, the set $A_2(t)$, in contrast to $A_1(t)$, does not entail general conditions on the Euclidean lengths of $x_{(n)}$ and $y_{(n)}$. In consequence, we have

$$\Phi_{0_{2n},\Psi_{2n\times 2n}}(A_1(t)) = \int_{r=0}^{\sqrt{nt}} \int_{s=0}^{\sqrt{nt}} \int_{\left\{(\phi,\psi) \,|\, \binom{x(r,\phi)}{y(s,\psi)} \in A_1(t)\right\}} \frac{J_T(r, \phi, s, \psi)}{(2\pi)^n (ab)^n} \cdot e^{-\frac{r^2}{2a^2} - \frac{s^2}{2b^2}} \, d(\phi, \psi)\, ds\, dr \,, \quad \text{(A4)}$$

$$\Phi_{0_{2n},\Psi_{2n\times 2n}}(A_2(t)) = \int_{r=0}^{\infty} \int_{s=0}^{\infty} \int_{\left\{(\phi,\psi) \,|\, \binom{x(r,\phi)}{y(s,\psi)} \in A_2(t)\right\}} \frac{J_T(r, \phi, s, \psi)}{(2\pi)^n (ab)^n} \cdot e^{-\frac{r^2}{2a^2} - \frac{s^2}{2b^2}} \, d(\phi, \psi)\, ds\, dr \,. \quad \text{(A5)}$$

We also observe for fixed r, s and $\phi = (\phi_1, \ldots, \phi_{n-1})$ that

$$\int_{\left\{\psi \,\mid\, \binom{x(r,\phi)}{y(s,\psi)} \in A_i(t)\right\}} J_{SPH_n}(s,\psi)\, d(\psi) = \int_{\left\{\psi \,\mid\, \binom{x(r,\phi)}{y(s,\psi)} \in A_i(t)\right\}} s^{n-1} \prod_{k=1}^{n-1} (\sin \psi_k)^{n-1-k}\; d(\psi)$$

is the Euclidean surface area in R^n of the Borel set $M_i(x_{(n)}, s) := \left\{y_{(n)} \in S_n(s) \mid \binom{x_{(n)}}{y_{(n)}} \in A_i(t)\right\}$ where $x_{(n)} = x(r, \phi)$ and where $S_n(s) \subseteq \mathrm{R}^n$ denotes the n-dimensional unit sphere with radius s, $s > 0$. For any orthogonal matrix $O \in \mathrm{R}^{n \times n}$ we have

$$\begin{aligned} O \cdot M_i(x_{(n)}, s) = \left\{O y_{(n)} \in S_n(s) \mid \binom{x_{(n)}}{y_{(n)}} \in A_i(t)\right\} &= \left\{z_{(n)} \in S_n(s) \mid \binom{x_{(n)}}{O^T z_{(n)}} \in A_i(t)\right\} \\ &\overset{(*)}{=} \left\{z_{(n)} \in S_n(s) \mid \binom{O x_{(n)}}{z_{(n)}} \in A_i(t)\right\} \\ &= M_i(O x_{(n)}, s) \end{aligned}$$

where $(*)$ means the usage of the invariance of the Euclidean inner product in R^n under orthogonal transformations. Moreover, the Euclidean surface area $\mathrm{SA}_n(B)$ of a Borel set $B \subseteq \mathrm{R}^n$ is invariant under orthogonal transformations. Hence, $O \cdot M_i(x_{(n)}, s) = M_i(O x_{(n)}, s)$ has the same surface area as $M_i(x_{(n)}, s)$. Notice that $\|O x_{(n)}\| = \|x_{(n)}\|$, i.e. the surface area of the set $M_i(x_{(n)}, s)$ depends on the radius r but not on the angle vector ϕ of $x_{(n)} = x(r, \phi)$ and will therefore be denoted by $\mathrm{SA}_{n,i}(r, s)$, $i = 1, 2$.

We consider $\mathrm{SA}_{n,i}(r, s) = \mathrm{SA}_n(M_i(x^*_{(n)}, s))$ where $x^*_{(n)} := (r, 0, \ldots, 0)^T \in \mathrm{R}^n$. Let us furthermore introduce n-dimensional spherical coordinates $(s, \psi) = (s, \psi_1, \ldots, \psi_{n-1})$ for arbitrary $y_{(n)} \in \mathrm{R}^n$, i.e. in particular $y_1 = s \cos \psi_1$. It follows

$$\begin{aligned} &\|x^*_{(n)}\|^2 + \|y_{(n)}\|^2 \pm \sqrt{(\|x^*_{(n)}\|^2 - \|y_{(n)}\|^2)^2 + 4 \left\langle x^*_{(n)}, y_{(n)} \right\rangle^2} \\ &\qquad = r^2 + s^2 \pm \sqrt{(r^2 - s^2)^2 + 4 r^2 s^2 \cos^2 \psi_1}\,. \end{aligned}$$

Considering the sets $A_1(t)$ and $A_2(t)$ we have the inequalities

$$r^2 + s^2 + \sqrt{(r^2 - s^2)^2 + 4 r^2 s^2 \cos^2 \psi_1} < 2nt\,, \tag{A6}$$

$$r^2 + s^2 - \sqrt{(r^2 - s^2)^2 + 4 r^2 s^2 \cos^2 \psi_1} < 2nt\,, \tag{A7}$$

respectively. For fixed values of $r > 0$ and $s > 0$ inequality (A6) as well as (A7) is a condition on the angle ψ_1 only. For solving these inequalities we introduce the function

$$g = g(r, s) = g(r, s; n, t) := \frac{n^2 t^2 - nt(r^2 + s^2) + r^2 s^2}{r^2 s^2}\,,$$

see Figure A1. Therefore, the initially $(n-1)$-dimensional problem of evaluating $\mathrm{SA}_{n,i}(r, s)$ for fixed $r > 0$, $\phi \in M^*_n$, $s > 0$, is reduced to a one dimensional problem concerning ψ_1,

$$\begin{aligned} \mathrm{SA}_{n,i}(r, s) &= \mathrm{SA}_n\left(M_i(x_{(n)}, s)\right) = \mathrm{SA}_n\left(\left\{y_{(n)} \in S_n(s) \mid \binom{x_{(n)}}{y_{(n)}} \in A_i(t)\right\}\right) \\ &= \mathrm{SA}_n\left(\left\{y_{(n)} \in S_n(s) \mid \binom{x^*_{(n)}}{y_{(n)}} \in A_i(t)\right\}\right) \\ &= \mathrm{SA}_n\left(\left\{y_{(n)}(s, \psi) \mid \psi \in M^*_n,\ \psi_1 \text{ satisfies (A6) or (A7), respectively}\right\}\right)\,. \end{aligned}$$

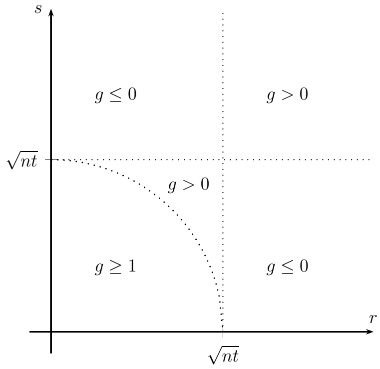


Figure A1. Behaviour of the function g according to (r, s) for fixed $n \geq 2, t > 0$.

Note that ψ_1 is restricted to $[0, 2\pi)$ if $n = 2$ and that ψ_1 ist restricted to $[0, \pi)$ if $n > 2$. Exploiting the properties of the function g it follows

$$\mathrm{SA}_{n,1}(r,s) = \begin{cases} s^{n-1}\omega_n & \text{, if } r^2 + s^2 < nt \\ 2s^{n-1}\omega_{n-1} \int_{\arccos\sqrt{g}}^{\frac{\pi}{2}} \sin^{n-2}(\psi_1)\, d\psi_1 & \text{, if } (r \leq \sqrt{nt}) \text{ and} \\ & \quad (nt - r^2 \leq s^2 \leq nt) \\ 0 & \text{, otherwise} \end{cases} \tag{A8}$$

and

$$\mathrm{SA}_{n,2}(r,s) = \begin{cases} 2s^{n-1}\omega_{n-1} \int_0^{\arccos\sqrt{g}} \sin^{n-2}(\psi_1)\, d\psi_1 & \text{, if } (r > \sqrt{nt}) \text{ and } (s > \sqrt{nt}) \\ s^{n-1}\omega_n & \text{, otherwise} \end{cases} \tag{A9}$$

where $r > 0$, $s > 0$, $g := g(r, s; n, t)$ and where ω_k, $k \geq 1$, denotes the Euclidean surface area of the k-dimensional unit sphere. Finally, we have

$$\begin{aligned} &\int_{\left\{(\phi,\psi) \,\mid\, \binom{x(r,\phi)}{y(s,\psi)} \in A_i(t)\right\}} J_T(r,\phi,s,\psi)\, d(\phi,\psi) \\ &= \int_{\phi \in M_n^*} J_{SPH_n}(r,\phi) \int_{\left\{\psi \,\mid\, \binom{x(r,\phi)}{y(s,\psi)} \in A_i(t)\right\}} J_{SPH_n}(s,\psi)\, d\psi\, d\phi \\ &= \int_{\phi \in M_n^*} J_{SPH_n}(r,\phi) \cdot \mathrm{SA}_{n,i}(r,s)\, d\phi \\ &= r^{n-1}\omega_n \cdot \mathrm{SA}_{n,i}(r,s) \end{aligned} \tag{A10}$$

where $\mathrm{SA}_{n,i}(r,s)$ can be replaced by (A8) or (A9), respectively. Using (A4) and (A5) as well as (A10) in $P\left(\widehat{\lambda}_i < t\right) = \Phi_{0_{2n}, \Psi_{2n\times 2n}}(A_i(t))$, $i = 1, 2$, completes the proof of Theorem 6.1.

References

Anderson, T.W., 2003. An Introduction to Multivariate Statistical Analysis, 3rd edition. John Wiley & Sons, Hoboken, New Jersey.

Arellano-Valle, R.B., Richter, W.-D., 2012. On skewed continuous $l_{n,p}$-symmetric distributions. Chilean Journal of Statistics, 3, 193-212.

Cambanis, S., Huang, S., Simons, G., 1981. On the Theory of Elliptically Contoured Distributions, Journal of Multivariate Analysis, 11, 368-385

Embrechts, P., McNeil, A., Straumann, D., 2002. Correlation and dependence in risk management: properties and pitfalls. In: Risk Management: Value at Risk and Beyond, ed. Dempster, M.A.H., Cambridge University Press, Cambridge, 176-223

Fang, K.T., Kotz, S., Ng, K.W., 1990. Symmetric Multivariate and Related Distributions. Chapman and Hall, New York.

Fréchet, M., 1957. Les tableaux de corréllation dont les marges sont données. Annales de l'Université de Lyon, Sciences Mathématiques et Astronomi, Série A, 4, 13-31.

Gander, W., Gautschi, W., 2000. Adaptive quadrature – Revisited. BIT Numerical Mathematics, 40, 84-101.

Gillis, J.T., 1991. Computation of the Circular Error Probability Integral. IEEE Transactions on Aerospace and Electronic Systems, 27, 906-910.

van Diggelen, F., 2007. GNSS Accuracy. Lies, Damn Lies, and Statistics. GPS World, January 2007, 26-32.

Guenther, W.C., Terragno, P.J., 1964. A Review of the Literature on a Class of Coverage Problems. The Annals of Mathematical Statistics, 35, 232-260.

Günzel, T., Richter, W.-D., Scheutzow, S., Schicker, K., Venz, J., 2012. Geometric approach to the skewed normal distribution. Journal of Statistical Planning and Inference, 142, 3209-3224.

Harter, H.L., 1960. Circular Error Probabilities. Journal of the American Statistical Association, 55, 723-731.

Henschel, V., Richter, W.-D., 2002. Geometric Generalization of the Exponential Law. Journal of Multivariate Analysis, 81, 189-204.

Höffding, W., 1940. Maßstabsinvariante Korrelationstheorie. Schriften des Mathematischen Seminars und des Instituts für angewandte Mathematik der Universität Berlin, 5, 181-233.

Hyvärinen, A., Karhunen, J., Oja, E., 2001. Independent Component Analysis. John Wiley & Sons.

Ittrich, C., Krause, D., Richter, W.-D., 2000. Probabilities and large quantiles of noncentral generalized chi-square distributions. Statistics, 34, 53-101.

Jammalamadaka, S.R., SenGupta, A., 2001. Topics in circular statistics. World Scientific, Singapore.

Jolliffe, I.T., 2004. Principal Component Analysis. Springer, New York.

Kalke, S., Richter, W.-D., 2013. Simulation of the p-generalized Gaussian distribution. Journal of Statistical Computation and Simulation, 83, 641-667.

Kalke, S., Richter, W.-D., Thauer, F., 2013. Linear combinations, products and ratios of simplicial or spherical variates. Communications in Statistics: Theory and Methods, 42, 505-527.

Koev, P., Edelman, A., 2006. The efficient evaluation of the hypergeometric function of a matrix argument. Mathematics of Computation, 75, 833-846.

Lehmann, E.L., 1983. Theory of Point Estimation. John Wiley & Sons, New York.

Mardia, K.V., 1972. Statistics of Directinal Data. Academic Press, London.

Mardia, K.V., Jupp, P.E., 2000. Directional Statistics. John Wiley & Sons, Chichester.

McCool, J. I., 2006. Control Charts for Radial Error. Quality Technology and Quantitative Management, 3, 283-293.

McEwen, R., Thomas, H., Weber, D., Psota, F., 2005. Performance of an AUV Navigation System at Arctic Latitudes. IEEE Journal of Oceanic Engineering, 30, 443-454.

Muirhead, R.J., 1975. Expressions for Some Hypergeometric Functions of Matrix Argu-

ment with Applications. Journal of Multivariate Analysis, 5, 283-293.
Muirhead, R.J., 1982. Aspects of Multivariate Statistical Theory. John Wiley & Sons, New York.
Nadarajah, S., Gupta, A.K., 2005. On the product and ratio for the elliptically symmetric Pearson type Vii distribution. Random Operators and Stochastic Equations, 13, 139-146.
Nelson, W., 1988. Use of Circular Error Probability in Target Detection. United States Air Force Technical Report.
Pyati, V.P., 1993. Computation of the Circular Error Probability (CEP) Integral. IEEE Transactions on Automatic Control, 29, 1023-1024.
Richter, W.-D., 1991. Eine geometrische Methode in der Stochastik. Rostocker Mathematisches Kolloquium, 44, 63-72.
Richter, W.-D., 1995. A geometric approach to the Gaussian law. In: Symposia Gaussiana, Conference B: Statistical Sciences, eds. Mammitzsch, V., Schneeweiss, H., 25-45.
Richter, W.-D., 2007. Generalized spherical and simplicial coordinates. Journal of Mathematical Analysis and Applications, 336, 1187-1202.
Richter, W.-D., 2009. Continuous $l_{n,p}$-symmetric distributions. Lithuanian Mathematical Journal, 49, 93-108.
Richter, W.-D., 2011. Ellipses numbers and geometric measure representations. Journal of Applied Analysis, 17, 165-179.
Richter, W.-D., 2013. Geometric and stochastic representations for elliptically contoured distributions. Communications in Statistics: Theory and Methods, 42, 579-602.
Scheuer, E.M., 1962. Moments of the Radial Error. Journal of the American Statistical Association, 57, 187-190.
Shnidman, D.A., 1995. Efficient Computation of the Circular Error Probability (CEP) Integral. IEEE Transactions on Automatic Control, 40, 1472-1474.
Solow, A.R., 1990. A note on the statistical properties of animal locations. Journal of Mathematical Biology, 29, 189-193.
Szablowski, P.J., 1998. Uniform Distributions on Spheres in Finite Dimensional L_α and Their Generalizations. Journal of Multivariate Analysis, 64, 103-117.
Tong, Y.-L., 1990. The Multivariate Normal Distribution. Springer, New York.
Watson, G.N., 1995. A Treatise on the Theory of Bessel Functions. Cambridge Mathematical Library.
Waugh, J.G., 1961. Evaluation of Integral of Elliptical Gaussian Distribution over a centered Ellipse. U.S. Naval Ordnance Test Station.
Weil, H., 1954. The Distribution of Radial Error. Annals of Mathematical Statistics, 25, 168-170.
White, R.G., 1975. Distribution and moments of radial error. NASA Technical Memorandum.
Zhang, J., Weilian, A., 2012. Assessing circular error probable when the errors are elliptical normal. Journal of Statistical Computation and Simulation, 82, 565-586.