Star-shaped distributions: Euclidean and non-Euclidean representations

Wolf-Dieter Richter
University of Rostock, Institute of Mathematics, Rostock, Germany - wolf-dieter.richter@uni-rostock.de

Abstract

Stochastic representations of random vectors following a star-shaped distribution, and geometric measure representations of the corresponding probability laws are considered. These representations generalize those for spherical and elliptically contoured distributions and apply to various statistical and probabilistic problems. Special emphasis is on the class of p-generalized elliptically contoured distributions including big subclasses of norm and antinorm contoured distributions. Finally, an outlook is given to exact distribution theory under non-standard model assumptions.

Keywords: p-generalized elliptically contoured distributions; stochastic representations; geometric measure representations; star-shaped sample clouds; star-shaped density level sets; generalized surface measures; generalized ball numbers; circle numbers for star discs.

1. Introduction

A random vector U being uniformly distributed on the Euclidean sphere in R^n is often called the uniform basis of the family of spherically distributed random vectors. Every element X of this family can be constructed multiplying U by a suitably defined non-negative random variable R being independent of U and commonly called the generating variate of X. The corresponding distributions are studied in Kelker(1970) and Fang et al.(1990). An analogous construction for characteristic functions has been dealt with already in Schoenberg(1938).
Elliptically contoured distributed random vectors can always be derived by a matrix-transformation from a spherically distributed vector. Such constructions have been successfully used since the basic work in Cambanis et al.(1981), Fang and Zang(1990) and Anderson(1993). A recent statistical application can be found in Pynnönen(2012).
In Balkema et al.(2010), Joenssen and Vogel(2012), Mosler(2013) and in other papers the authors reveal that star-shaped sets and, correspondingly, star-shaped distributions occur in different applied disciplines. In related data situations, a statistician might be confronted with a cloud of sample points reflecting certain star-shaped contours rather than just elliptically ones.
Let us assume throughout this note that the density level sets of X are the boundaries $S(r) = rS$ of star bodies $K(r) = rK, r > 0$, with the origin being an inner point of K, and with S being different from the Euclidean sphere, general. It is the aim of this paper to discuss then a stochastic representation for X being analogous to that in the spherical case. Note that the particular case of elliptically contoured distributed vectors allows both matrix-transformed Euclidean and non-Euclidean representations, the latter being proved in Richter(2013).
The paper is organized as follows. We continue with Euclidean and non-Euclidean stochastic representations of random vectors and corresponding geometric measure representations of their probability laws in Sections 2 and 3, respectively. Thereby, an emphasis will be on the notion of generalized surface content measure. Section 4 deals with the specific class of p-generalized elliptically contoured distributions including big classes of norm and antinorm contoured distributions. In Section 5, we outline the close connection between normalizing constants of star shaped density generating functions and so called ball numbers (being circle numbers of star discs if dimension is two). We finish with an outlook to exact distribution theory under non-standard model assumptions, in Section 6.

2. Euclidean representations

It is well known that if X is a spherically distributed random vector, i.e. if K is the Euclidean unit ball then

X allows according to the results in Schoenberg(1938) and in Kelker(1970) the stochastic representation

$$X \stackrel{d}{=} R \cdot U \tag{1}$$

with a non-negativ random variable R being independent of a random vector U which follows the uniform distribution on the Euclidean unit sphere S. If X has a density f with density generating function g and normalizing constant $C(g)$,

$$f(x) = C(g)g(||x||),\ x \in R^n$$

where $||.||$ denotes the Euclidean norm, then the density level sets of X are Euclidean spheres $S(r) = rS$ of positive radius r, and the distribution Φ_g of X allows according to Richter(1985, 1991) the geometric measure representation

$$\Phi_g(A) = C(g) \int_0^\infty r^{n-1} O(A \cap S(r)) g(r) dr \tag{2}$$

where O denotes the Euclidean surface content measure. Numerous probabilistic and statistical applications of this representation are surveyed in Richter(2012, 2014) and the literature mentioned there. As just to mention a few of them we refer to the derivation of exact distributions of linear combinations, extreme values and other order statistics, products and ratios of components of $l_{n,p}$-symmetrically distributed vectors, the construction of generalized Student, Chi-square and Fisher distributions and statistics, the derivation of geometric representations for skewed elliptically contoured and $l_{n,p}$-symmetric distributions, and the derivation of numerous directional distributions being generalizations of the von Mises distribution. Some of the related work can be found in Arellano-Valle and Richter (2012), Batún-Cutz et al. (2013), Dietrich et al. (2013), Richter(2014), Richter and Venz (2014), Stehlik et al. (2014) and Müller and Richter (2015). Note that (2) reflects the classical method of indivisibles of Cavalieri and Torricelli, which was established already before calculus was established by Leibniz and Newton, in a more stringent way, and extends it. For a related review, see Richter(1985, 2012).

As to prepare for the next section, let us recall an analytical definition of the notion of surface content $O(A)$ of a measurable subset A of the Euclidean sphere $S(r), r > 0$. For simplicity, let A be a part of the upper half sphere $S^+(r)$ where $x_n \geq 0$. It is well known that

$$O(A) = \int_{G(A)} h_K(N(x)) dx \tag{3}$$

where $G(A) = \{\vartheta \in R^{n-1} : \exists \eta = \eta(\vartheta) \text{ with } (\vartheta, \eta(\vartheta))^T \in A\}$, $N(\vartheta)$ denotes the outer normal vector to $S(r)$ at the point $(\vartheta, \eta(\vartheta)) \in A$, and $h_K : R^n \to R^+$ means the Minkowski functional of the Euclidean unit ball,

$$h_K(x) = \inf\{\lambda > 0 : x \in \lambda K\},\ x \in R^n.$$

Note that the integral, or differential geometric, definition of the surface measure O allows the local representation

$$O(A) = \frac{d}{dr} f_A(r) \tag{4}$$

where, with μ standing for the Lebesgue measure in R^n,

$$f_A(r) = \mu(sector(A, r)), \quad sector(A, r) = CPC(A) \cap K(r),$$

and

$$CPC(A) = \{x \in R^n : \frac{x}{h_K(x)} \in A\}$$

is the central projection cone of the set A. Finally, the distribution of U allows the geometric representation

$$P(U \in A) = \frac{O(A)}{O(S)},\ A \in \mathcal{B}_S^+ = \mathcal{B}^{(n)} \cap S^+.$$

3. Non-Euclidean representations

If K is no longer the Euclidean unit ball then, according to Richter(2014), O has to be replaced in (2) by a suitably defined non-Euclidean surface content measure, and the uniform distribution of U in (1) is defined then with respect to this non-Euclidean surface measure. The aim of this section is to shortly discuss such non-Euclidean representations. If K denotes again a star body as in Section 1, and S its boundary, we define the star-generalized surface content measure O_S on $S(r)$ by

$$O_S(A) = \frac{d}{dr} f_A(r) \tag{5}$$

where f_A is for general S defined as in the formula following (4). Note that in (5) and throughout this section the variable r plays the role of a generalized radius.
It has been discussed in Richter(2014) under which additional assumptions upon K the local definition (5) allows an integral, or differential geometric, representation similar to that in (3),

$$O_S(A) = \int\limits_{G(A)} h_{\widehat{K}}(N(x))dx \tag{6}$$

where $\widehat{K}$ has to be specified. Note that the positive homogeneous functional $h_{\widehat{K}}$ is not necessarily a norm.

Example 1

If $n = 2$ and $K = K^{||\cdot||} = \{x \in R^2 : ||x|| \leq 1\}$ for any norm $||.||$ then, in (6), $\widehat{K} = O(\pi/4)K^*$ where K^* is the unit ball with respect to the dual norm $||.||^*$ of $||.||$. Thus, $h_{\widehat{K}}$ is always a norm.

A more complex example is discussed in Section 4. Mathematicians searching for a geometry suitable for proving an analogous representation as in (3) and (6) but in the case of a more general star body K may feel themselves encouraged by corresponding remarks in Hilbert (1900). With the definition

$$\omega_S(A) = \frac{O_S(A)}{O_S(S)},\ A \in \mathcal{B}_S,$$

ω_S is called the star-generalized uniform probability distribution on the Borel sigma field $\mathcal{B}_S$ on the star sphere S.
Without giving the present geometric explanation of ω_S, a correspondingly defined distribution is often called, e.g. in Racev(1991), Song and Gupta(1997), Szablowski(1998) and numerous other papers, just the uniform distribution on $\mathcal{B}_S$ which should, however, not be confused with the one defined using the Euclidean surface measure.
Note that convex bodies having an ellipsoidal boundary generate specific norms. Two-dimensional elliptically contoured distributions satisfy therefore according to Example 1 a stochastic representation as in (1) and a geometric measure representation as in (2) where $n = 2$, S is an ellipse and O is changed with a non-Euclidean circumference measure. Corresponding n-dimensional results are derived in Richter(2013) and generalized in Richter(2014), for details see the next section.

4. The p-generalized elliptically contoured distributions

Basics of the theory of elliptically contoured distributions are to be found in Cambanis et al. (1981) and, e.g., in Fang et al.(1990). Non-Euclidean geometric and stochastic representations for elliptically contoured distributions were derived in Richter(2013) and the literature mentioned there. A stochastic vector and a geometric measure representation for the class of p-generalized elliptically contoured distributions was derived in Richter(2014). The case $p = 1$ was dealt with in Henschel and Richter(2002) if the probability mass is concentrated in the positive orthant of $R^{n,+}$.

In dependence of the properties of $\widehat{K}$, the Minkowski functional $h_{\widehat{K}}$ of the unit ball of the chosen non-Euclidean geometry may be a norm, an antinorm, a semi-antinorm or a homogeneous functional of another type. Antinorms and semi-antinorms are introduced in Moszyńska and Richter(2012), norm contoured and antinorm contoured distributions are special cases of the distributions considered in Richter(2014).
Let $a = (a_1, ..., a_n)^T \in R^n$ with $a_i > 0, i = 1, ..., n$, $p > 0$ and $|.|_{a,p} : R^n \to [0, \infty)$ the function defined by $|x|_{a,p} = (\sum\limits_{i=1}^{n} |\frac{x_i}{a_i}|^p)^{1/p}$, $x \in R^n$, $K = K_{a,p} = \{x \in R^n : |x|_{a,p} \le 1\}$. Further, denote the boundary of K by $S = S_{a,p}$. For defining $O_S = O_{S_{a,p}}$, one can chose $\widehat{K} = K_{\frac{1}{a},q}$ in the surface measure representation(6) where $\frac{1}{a} = (\frac{1}{a_1}, ..., \frac{1}{a_n})$ and $\frac{1}{p} + \frac{1}{q} = 1$, i.e. $q = \frac{p}{p-1}, p \neq 1$. Thus

$$O_{S_{a,p}}(A) = \int\limits_{G(A)} |N(x)|_{\frac{1}{a},q} dx = a_n \int\limits_{G(A)} \frac{d(x_1, ..., x_{n-1})}{\left(1 - \sum\limits_{i=1}^{n-1} |\frac{x_i}{a_i}|^p\right)^{(p-1)/p}}.$$

If $p \ge 1, p \in (0, 1]$ or $p < 0$ then $h_K(.) = |.|_{a,p}$ is a norm, an antinorm or a semi-antinorm, respectively. Moreover, if $p \ge 1$ or $p \in (0, 1)$ then $q \ge 1$ or $q < 0$ and $h_{\widehat{K}}(.) = |.|_{\frac{1}{a},q}$ is a norm or a semi-antinorm, respectively. For the case of 'classical' elliptically contoured distributions, $p = 2$, we refer to Richter(2013).

Let Y follow the continuous p-generalized elliptically contoured distribution with parameters $g, a, p, \nu \in R^n$ and $O \in \mathcal{O}(n)$ where $g : [0, \infty) \to [0, \infty)$ is a density generating function and $\mathcal{O}(n)$ denotes the set of all orthogonal $n \times n$-matrices. Note that the random vector Y follows the density

$$f_Y(x) = C(g, a, p) g(|O^T(x - \nu)|_{a,p}), x \in R^n$$

and $X = O^T(Y - \nu)$ allows the stochastic representation $X \stackrel{d}{=} RU$ where R and U are independent, R has density

$$f_R(r) = r^{n-1} g(r) I_{[0,\infty)}(r) / \int\limits_0^\infty r^{n-1} g(r) dr,$$

and U follows the star-generalized uniform distribution

$$\omega_{a,p}(A) = O_{S_{a,b}}(A) / O_{S_{a,b}}(S_{a,b}).$$

The corresponding geometric measure representation of $P(Y \in B) = \Phi_{g,a,p,\nu,O}(B)$ is

$$\Phi_{g,a,p,\nu,O}(B) = C(g, a, p) \int\limits_0^\infty r^{n-1} g(r) O_{S_{a,p}}([\frac{1}{r} O^T(B - \nu)] \cap S_{a,p}) dr, B \in \mathcal{B}^n. \tag{7}$$

5. Normalizing constants and ball numbers

It is remarkably that the normalizing constant $C(g, a, p)$ in (7) is closely connected with a relatively new mathematical constant which is called a ball number. A ball number $\pi(K_{a,p})$ generalizes the notion of the circle number π with respect to both dimension and shape of the sphere. According to Richter(2014),

$$C(g, a, p) = \frac{1}{\int\limits_0^\infty r^{n-1} g(r) dr O_{S_{a,p}}(S_{a,p})} \tag{8}$$

where $\pi(K_{a,p}) = \frac{1}{n} O_{S_{a,p}}(S_{a,p}) = \mu(K_{a,p})$ satisfies the equations

$$\frac{O_{S_{a,p}}(S_{a,p}(r))}{n r^{n-1}} = \pi(K_{a,p}) = \frac{\mu(K_{a,p})(r)}{r^n}, r > 0.$$

For ball numbers of Platonic bodies being closely related to the cases $p = 1$ and $p = \infty$, we refer to Richter and Schicker(2014), for the general case we refer to the literature mentioned there and in Richter(2014).

Ball numbers of two-dimensional star bodies are called circle numbers of star discs. Let us write U_S instead of O_S, in this case. According to Example 1,

$$\frac{U_S(S(r))}{2r} = \pi(K^{||\cdot||}) = \frac{\mu(K^{||\cdot||}(r))}{r^2}$$

where U_S is according to (6) uniquely defined for any norm $||.||$ with the help of $h_{\widehat{K}}, \widehat{K} = O(\frac{\pi}{4})K^*$, and $h_{K^*}(.) = ||.||^*$ being the dual norm of $||.||$. Various representations of ball numbers are to be found in Richter(2014) and the references given there.

Without going into any details, we finally mention that ball numbers are closely connected with solutions to isoperimetric problems, see in Bobkov and Hondré(1997) and Gardner(2002).

6. Conclusions

Exact distributions of functions of sample vectors under non-standard assumptions with respect to the distribution of the sample vector are derived applying suitable geometric measure representations, e.g., in Batún-Cutz et al. (2013), Dietrich et al. (2013), Henschel and Richter(2002), Müller and Richter (2015), Richter and Venz (2014) and Stehlik et al. (2014). The wide research area on exact distributions under non-standard model assumptions can be further developed now based upon the present results. Author's work in this field is surveyed in Richter (2012, 2014). Moreover, statistical problems are dealt with in Dietrich et al.(2013) and Stehlik et al.(2014), and those considerations should hopefully stimulate more related statistical studies in the future. A specific aspect of estimation might be to answer the following question: should one take into account the validity of equation (8) when estimating in p-generalized elliptically contoured sample distributions, or not? In other words, should estimation of the 'shape parameter' $O_S(S)$ and of the 'tail parameter' $\int_{[0,\infty)} r^{n-1}g(r)dr$ be separated, or not? Finally, the results in Section 5 may be a stimulation for further studies on ball numbers and their applications.

References

Anderson, T.W. (1993). Nonnormal multivariate distributions: inference based on elliptically contoured distributions. In: C.R. Rao(Ed.) Multivariate analysis:future directions, North-Holland Series in Statistics and Probability, vol. 5, Amsterdam, 1-24.

Arellano-Valle, R. B., Richter, W.-D. (2012). On skewed continuous $l_{n,p}$-symmetric distributions. Chilean J. Stat. **3**,2, 193-212.

Balkema, A.A., Embrechts, P. and N. Nolde (2010). Meta densities and the shape of their sample clouds, *Journal of Multivariate Analysis*, **101**, 1738-1754.

Batún-Cutz, J., González-Farías, G. and W.-D. Richter (2013). Maximum distributions for $l_{2,p}$-symmetric vectors are skewed $l_{1,p}$-symmetric distributions. Stat. Prob. Letters **83**, 2260-2268.

Bobkov, S.G., Hondré, C. (1997). Isoperimetric constants for product probability measures. Ann. Prob. **25**, 1, 184-205.

Cambanis, S., Huang, S. and G. Simons (1981). On the theory of elliptically contoured distributions. J. Multiv. Anal. **11**, 368-385.

Dietrich, T., Kalke, S. and W.-D. Richter (2013). Stochastic representations and a geometric parametrization of the two-dimensional Gaussian law. Chilean J. Stat. **4**, 2, 27-59.

Fang, K.-T., Kotz, S. and K.-W. Ng (1990). Symmetric multivariate and related distributions. Chapman and Hall, London.

Fang, K.-T., Zhang, Y. (1990). Generalized multivariate analysis. Springer.

Gardner, R.J. (2002). The Brunn-Minkowski inequality. Bull. Amer. Math. Soc. **39**, 3, 355-405(electronic).

Henschel, V., Richter, W.-D. (2002). Geometric generalization of the exponential law. J. Multiv. Anal. **81**, 189-204.

Hilbert, D. (1900). Mathematische Probleme, Mathematikerkongress, Paris.

Joenssen, D. W., Vogel, J. (2012). A power study of goodness-of-fit tests for multivariate normality implemented in R, *J. Statist. Comput. Simul.*, **75**, 93-107.

Kelker, D. (1970). Distribution theory of spherical distributions and a location-scale parameter generalization. Sankhya, Ser., A **32**, 419-430.

Mosler, K. (2013). Depth statistics, *arXiv.orgmath.arXiv:1207.4988v2*.

Moszyńska, M., Richter, W.-D. (2012). Reverse triangle inequality. Antinorms and semi-antinorms. *Studia Scientiarum Mathematicarum Hungarica*, **49** ,1 ,120-138.

Müller, K., Richter, W.-D. (2015). Exact extreme value, product and ratio distributions under non-standard assumptions. AStA Advances in Stat. Anal. **99**, 1, 1-30.

Pynnönen, S. (2012). Distribution of an arbitrary linear transformation of internally Studentized residuals of multivariate regression with elliptical errors. J. Multiv. Anal. **107**, 40-52.

Racev, S. T. (1991). Probability metrics and the stability of stochastic models. J. Wiley and Sons.

Richter, W.-D. (1985). Laplace-Gauß integrals, Gaussian measure asymptotic behaviour and probabilities of moderate deviations. Z. Anal. Anw. **4**, 3, 257-267.

Richter, W.-D. (1991). A geometric method in stochastics (in German). Rostock. Math. Kolloqu. **44**, 63-72.

Richter, W.-D. (2012). Exact distributions under non-standard model assumptions. AIP Conf. Proc. **1479**, 442-445. doi 10.1063/1.4756160.

Richter, W.-D. (2013). Geometric and stochastic representations for elliptically contoured distributions. Comm. Stat.-Theory and Methods **42**, 579-602.

Richter, W.-D. (2014). Geometric disintegration and star-shaped distributions. Journal of Statistical Distributions and Applications, **1**: 20; http://www.jsdajournal.com/content/1/1/20.

Richter, W.-D., Schicker, K. (2014). Ball numbers of Platonic bodies. J. Mathem. Anal. Appl. **416**, 783-799.

Richter, W.-D., Venz, J (2014). Geometric representations of multivariate skewed elliptically contoured distributions. Chilean J. Stat. **5**, 2, 71-90.

Schoenberg, I.J. (1938). Metric spaces and complete monotone functions. Ann. Math. **39**, 811-841.

Song, D., Gupta, A.K. (1997). L_p-norm uniform distributions. Proc. Amer. Math. Soc. **125**, 2, 595-601.

Stehlik, M., Economou, P., Kiselak, J. and W.-D. Richter (2014). Kullback-Leibler life testing. Applied Mathematics and Computation **240**, 122-139.

Szablowski, P. J. (1998). Uniform distributions on spheres in finite dimensional l_α and their generalizations. J. Multiv. Anal. **64**, 2, 103-117.