96-F-15

Theory of Cross Sectionally Contoured Distributions and Its Applications

by

Akimichi Takemura
University of Tokyo

and

Satoshi Kuriki
The Institute of Statistical Mathematics

July 1996

Theory of cross sectionally contoured distributions and its applications

Akimichi Takemura
Faculty of Economics, University of Tokyo
and
Satoshi Kuriki
The Institute of Statistical Mathematics

July, 1996

Abstract

We discuss generalization of elliptically contoured distributions to densities whose contours are arbitrary cross sections in the framework of group invariance. This generalization leads to much richer family of distributions compared to the elliptically contoured distributions. The basic property of the elliptically contoured distribution is the independence of the "length" and the "direction" of the random vector. We show that in our generalized framework this independence still holds if we define the length appropriately. Our examples include "star-shaped distributions" and their generalization to random matrices.

Key words: Elliptically contoured distribution, star-shaped distribution, group action, invariance, relatively invariant measure.

1 Introduction

Consider a continuous elliptically contoured distribution in R^p. Its density $f(x)$ is written as

$$f(x) = h(x'\Sigma^{-1}x), \tag{1}$$

where x is considered as a p dimensional column vector and Σ is a positive definite matrix. Let

$$r(x) = (x'\Sigma^{-1}x)^{1/2}$$

be the Mahalanobis distance. Then the "length" $r(x)$ and the "direction" $x/r(x)$ are independent under (1). Furthermore by changing h in (1) we can construct elliptically contoured distribution with arbitrary continuous distribution of $r(x)$. In this sense the family of elliptically contoured distributions is a generalization of the multivariate normal distribution. Furthermore because the distribution of $x/r(x)$ is common to all the elliptically contoured distributions with the same Σ, distributional results concerning $x/r(x)$ derived under the assumption of normality remain to hold for all elliptically contoured distributions. This property is often referred to as "null robustness" and has been extensively discussed in literature (see Kariya and Sinha (1989) for example). See Fang and Anderson (1990), Fang and Zhang (1990), Anderson (1993), or Gupta and Varga (1993) for comprehensive treatment of elliptically contoured distribution.

Note that elliptically contoured distribution differs from the multivariate normal distribution only in the distribution of the one dimensional length. Therefore in the framework of elliptically contoured distribution we can not consider non-normality which is exhibited in skewness or asymmetry of distributions.

For illustration consider the following density in R^2.

Example 1.1

$$f(x) = f(x_1, x_2) = h(\max(-x_1, -x_2, x_1 + x_2)). \tag{2}$$

The contours of $f(x)$ defined by

$$c = \max(-x_1, -x_2, x_1 + x_2), \qquad c > 0, \tag{3}$$

are concentric right triangles of Figure 1.

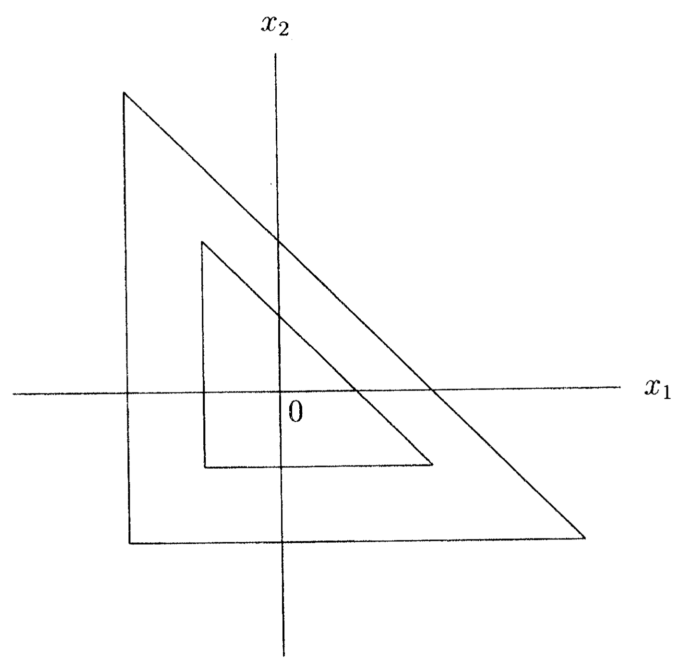

Figure 1: Triangle Contours

The contours are no longer symmetric with respect to the origin. Now define the length of the random vector by

$$r(x) = \max(-x_1, -x_2, x_1 + x_2).$$

Then our Theorem 2.1 shows that $r(x)$ and $x/r(x)$ are independent under (2). Furthermore as in the case of the elliptically contoured distribution we can assign arbitrary distribution to $r(x)$ by specifying h in (2). In this sense (2) is a generalization of the elliptically contoured distribution to the case of contours which are not ellipses. Actually we can have arbitrary star-shaped sets as contours of the density as fully discussed in Section 3.

The independence of $r(x)$ and $x/r(x)$ is a consequence of the general result on factorization of a relatively invariant measure into equivariant and invariant parts. The group action we consider in Example 1.1 is the action of the multiplicative group R^*_+ of positive reals:

$$g(x_1, x_2) = (gx_1, gx_2), \quad g \in R^*_+. \tag{4}$$

The orbits are rays starting at the origin and the triangle considered in (3) is a particular cross section. In Section 3 we will show that the independence of $r(x)$ and $x/r(x)$ in Example 1.1 is a consequence of the factorization of the two dimensional Lebesgue measure which is relatively invariant with multiplier $\chi(g) = g^2$ under the group action (4).

The group actions considered in this paper are free actions. In the case of non-free actions the equivariant has to be defined as left cosets with respect to isotropy subgroups and this considerably complicates the general theory. Generalization of the present theory to the non-free action is fully treated in our subsequent paper (Kamiya and Takemura (1996)).

Organization of the paper is as follows.

General theory of cross sectionally contoured distributions is developed in Section 2. In Section 2.2 we present a proof of the factorization of a relatively invariant measure into equivariant and invariant parts based on factorization theorem of sufficient statistic for a group family of distributions. Construction of general cross section and associated contoured distribution from a standard cross section is discussed in Section 2.3. Furthermore results on distributions of the invariants with respect to two different cross sections are given. In Section 2.4 we discuss calculation of Jacobians concerning cross sectionally contoured densities with respect to the Lebesgue measure on Euclidean sample space. In Section 3 we introduce star-shaped distributions on R^p and discuss their properties. In Section 4 we generalize star-shaped distributions to random matrices. In particular we discuss distributions of random matrices involving Gram-Schmidt orthonormalization in Section 4.1 and generalization of matrix valued beta distributions in Section 4.2.

2 General Theory

In this section we develop a general theory of cross sectionally contoured distributions. The primary example we keep in mind is the star-shaped distributions in Section 3. However it is advantageous to develop a general theory first. General theory allows us to treat more difficult examples in Section 4 and will be a basis for subsequent generalization to non-free actions (Kamiya and Takemura (1996)).

2.1 Cross section and associated equivariant function

Consider a group G acting from left on the sample space $\mathcal{X}$. We consider the case that the action of G is not transitive, so that the orbit space is not trivial. Let r be an equivariant function from $\mathcal{X}$ to G:

$$r(gx) = gr(x), \qquad g \in G, \ x \in \mathcal{X}. \tag{5}$$

We first establish the one-to-one relation between the equivariant function $r(x)$ and its associated cross section in a series of lemmas.

Lemma 2.1 *Let $r(x)$ be an equivariant function from $\mathcal{X}$ to G. Fix $g \in G$. Then the set*

$$Z_g = \{x \mid r(x) = g\} \tag{6}$$

is a cross section.

Proof. Suppose that there exist two different points x_1, x_2 on an orbit such that $g = r(x_1) = r(x_2)$. Since they are on the same orbit there exists $h \in G$ such that $x_2 = hx_1$. Therefore $r(x_1) = r(x_2) = hr(x_1)$. Canceling $r(x_1)$ we obtain $h = e$, the identity element of G, and $x_1 = x_2$ contrary to our assumption. Therefore Z_g meets each orbit at most once. Now from an arbitrary orbit pick a point x and let $h = r(x)$. Then $r(gh^{-1}x) = g$ and this show that Z_g meets each orbit. ∎

Given $r(x)$ we call

$$Z = Z_e = \{x \mid r(x) = e\}$$

the *unit cross section* associated with $r(x)$.

For $x \in \mathcal{X}$ let

$$z = z(x) = r(x)^{-1}x$$

then $z(x) \in Z$ is invariant. More precisely $z(x)$ is a maximal invariant, i.e. $z(x_1) = z(x_2)$ implies that x_1 and x_2 are on the same orbit (cf. pages 30–31 of Eaton (1989)). Consider the pair $(r, z) = (r(x), z(x))$. By this correspondence between x and (r, z), $\mathcal{X}$ is in one-to-one relation with the direct product $G \times Z$ of G and Z:

$$\mathcal{X} \leftrightarrow G \times Z.$$

We call $r = r(x)$ the *equivariant* function (or part) and $z = z(x)$ the *invariant* function (or part) of this representation, and the decomposition

$$x = rz, \qquad r \in G, \ z \in Z,$$

is referred to as *orbital decomposition* (Section 2 of Barndorff-Nielsen et al. (1989)). With star-shaped distribution in mind we sometimes refer to $r(x)$ as *length* and $z(x)$ as *direction* of x.

Remark 2.1 *For any* g, $\tilde{r}(x) = r(x)g^{-1}$ *is equivariant and the unit cross section associated with* $\tilde{r}(x)$ *coincides with* Z_g *in (6). This implies that there is nothing special about the unit cross section. However because of the convenient product representation* $\mathcal{X} \leftrightarrow G \times Z$, *we often work with the unit cross section.*

Z_{g_1} *and* Z_{g_2} *are disjoint for* $g_1 \neq g_2$ *and* $\mathcal{X} = \bigcup_{g \in G} Z_g$ *forms a partition of* $\mathcal{X}$ *into disjoint cross sections. We sometimes refer to these cross sections as family of proportional cross sections.*

We now show that the existence of r satisfying (5) implies that the group action is free.

Lemma 2.2 *If there exists an equivariant* r *from* $\mathcal{X}$ *to* G, *then the action of* G *is free, i.e. for any* x

$$gx = x \;\Rightarrow\; g = e$$

where e *is the identity element of* G.

Proof. Suppose $gx = x$. By (5) we have

$$r(x) = gr(x).$$

Since $r(x) \in G$ we can cancel $r(x)$ and obtain $g = e$. ∎

So far we have discussed how a cross section Z_g is obtained from a given equivariant function $r(x)$. Sometimes a cross section Z is given first and we need to construct the equivariant function $r(x)$. Because our action is free, this is straightforward.

Lemma 2.3 *Let* Z *be a cross section. For* x *choose* $z \in Z$ *on the same orbit and define* $r(x) \in G$ *by*

$$x = r(x)z. \tag{7}$$

Then $r(x)$ *is equivariant. Furthermore* Z *is the unit cross section for* $r(x)$.

Proof is obvious and omitted. Note that $r(x)$ in (7) is uniquely defined because of free action.

The above lemmas do not require any topological or measurability assumptions. Now we start making regularity assumptions. We assume that $\mathcal{X}$ is a measurable space and we only consider measurable functions on $\mathcal{X}$. Let π be a relatively invariant measure on $\mathcal{X}$ with multiplier χ:

$$\pi(gA) = \chi(g)\pi(A), \qquad A \subset \mathcal{X}. \tag{8}$$

We consider distributions absolutely continuous with respect to π with the density function $f(x)$. We call π the dominating measure.

Generalizing (2) of Example 1.1 we make the following definition.

Definition 2.1 *Distribution* F *on* $\mathcal{X}$ *with density* f *with respect to* π *is called cross sectionally contoured (with the associated equivariant function* $r(x)$ *) if for some* h

$$f(x) = h(r(x)). \tag{9}$$

Here we list some further regularity assumptions we make on G and π. For the terminology of these assumptions see Chapter 2 of Wijsman (1990).

Assumption 2.1

1. *π is a σ-finite measure.*
2. *G is a second countable locally compact group.*

2.2 Independence of the equivariant and the invariant

The basic result in our framework is the independence of the equivariant $r = r(x)$ and the invariant $z = r(x)^{-1}x$ for cross sectionally contoured distributions. Although this result is a straightforward consequence of the factorization theorem of relatively invariant measures (see Chapter 8 of Wijsman (1990) for example) we present our own proof based on factorization theorem for sufficient statistic.

We use the following basic lemma by Eaton (1983), Proposition 7.19. (Theorem 4.1 of Wijsman (1986) is a version of this lemma.)

Lemma 2.4 *Let $\mathcal{P} = \{P_\theta, \theta \in \Theta\}$ be a family of distributions on a sample space $\mathcal{X}$ with sufficient statistic $T : \mathcal{X} \to \mathcal{T}$. Let G be a group acting on $\mathcal{X}$ as well as on $\mathcal{T}$. Suppose that T is equivariant, the action of G on $\mathcal{T}$ is transitive, and $\mathcal{P}$ is closed under the action of G. Then T is independently distributed of any invariant statistic under $\mathcal{P}$.*

Using Lemma 2.4 we prove the following theorem.

Theorem 2.1 *For a cross sectionally contoured distribution $r = r(x)$ and $z = r(x)^{-1}x$ are independently distributed. The distribution of z does not depend on h in (9) and the distribution of r is given as $h(r)\pi_r(dr)$ where π_r is a relatively invariant measure on G with multiplier χ in (8).*

Proof. Here we give a proof for the case where h is everywhere positive on G. The case of h with smaller support requires technical arguments, which we give in Appendix.

We introduce a group family of distributions based on $f(x) = h(r(x))$ with parameter $g \in G$. Let $w = gx$, then $\pi(dw) = \chi(g)\pi(dx)$, $r(w) = gr(x)$, and the density of w with respect to π is given as

$$f(w; g) = \frac{1}{\chi(g)} h(g^{-1}r(w)).$$

Consider the group family of distributions

$$\{f(x; g) \mid g \in G\}$$

which is a dominated family with respect to π. By factorization theorem on sufficient statistic we see that $r(x)$ is sufficient for g. Since $z = r(x)^{-1}x$ is invariant, $r(x)$ and $z(x)$ are independent by Lemma 2.4.

Now we can write $h(r)\pi(dx)$ as

$$h(r)\pi(dx) = h(r)\pi_r(dr)\pi_z(dz), \tag{10}$$

where $h(r)\pi_r(dr)$ is a probability measure on G and $\pi_z(dz)$ is a probability measure on Z. Since $h(r)$ is positive everywhere, $h(r(x))\pi(dx)$ and $\pi(dx)$ are absolutely continuous with respect to each other. Therefore we can divide both sides of (10) by $h(r)$ and obtain

$$\pi(dx) = \pi_r(dr)\pi_z(dz). \tag{11}$$

Since π_z is a probability measure we have

$$\pi_r(A) = \pi_r(A)\pi_z(Z) = \pi(AZ), \qquad A \subset G, \tag{12}$$

where

$$AZ = \{gz \mid g \in A, \ z \in Z\} = \{x \mid r(x) \in A\}.$$

Therefore π_r does not depend on h. Since $\pi(dx)$ itself does not depend on h, we see from (11) that $\pi_z(dz)$ does not depend on h either.

It remains to show that π_r is relatively invariant with multiplier χ. Note that

$$\pi(d(gx)) = \pi_r(d(gr))\pi_z(dz).$$

On the other hand

$$\pi(d(gx)) = \chi(g)\pi(dx) = \chi(g)\pi_r(dr)\pi_z(dz).$$

Canceling $\pi_z(dz)$ we obtain

$$\pi_r(d(gr)) = \chi(g)\pi_r(dr).$$

∎

In the above proof we derived the factorization (11) of the relatively invariant measure π under the assumption that there exists a cross sectionally contoured distribution $h(r(x))\pi(dx)$. Because the factorization of $\pi(dx)$ is of independent interest, we state it as a corollary.

Corollary 2.1 *Let π be a relatively invariant measure on $\mathcal{X}$ with multiplier χ. Suppose that there exists a Borel subset A of G such that*

$$0 < \pi(AZ) < \infty,$$

where $AZ = \{x \mid r(x) \in A\}$. Then π can be factored as (11), where π_r is a relatively invariant measure on G with multiplier χ and π_z is a probability measure on Z.

Proof. Let $I_A(r)$ denote the indicator function of A. Then

$$f(x) = I_A(r(x))/\pi(AZ)$$

defines a cross sectionally contoured distribution. Therefore π can be factored as shown in the proof of Theorem 2.1. ∎

Remark 2.2 *As mentioned above the factorization of* π *under certain regularity conditions is a well established fact (Section 3.4 of Farrell (1985), Chapter 8 of Wijsman (1990), Section 5 of Barndorff-Nielsen et al. (1989)) and our Theorem 2.1 is a direct consequence of this factorization. We think that our proof is of some merit since it is primarily based on sufficiency rather than group theoretic argument.*

2.3 Construction of general cross sections and associated cross sectionally contoured distributions

In group invariance arguments researchers tend to look at nice cross sections. Often these cross sections are expressed in the form of convenient maximal invariants. For example consider the action of the multiplicative group R_+^* of positive reals in (4). We usually take the ordinary Euclidean distance as length and the unit circle as associated cross section. The advantage of the unit circle is that it has the additional invariance property under rotation.

Our viewpoint in this paper is that restricting our attention to these standard cross sections severely narrows the distributions we investigate. However a standard cross section is still useful as a building block of more general cross sections.

Suppose we are given a cross section $Z \subset \mathcal{X}$. An arbitrary cross section can be constructed as follows.

Lemma 2.5 *Let* s *be a function from* Z *to* G. *Then*

$$Y = \{s(z)z \mid z \in Z\} \tag{13}$$

is a cross section. Conversely every cross section can be written in this form.

Proof. $s(z)z$ is on the same orbit as z. Since Z contains exactly one point from each orbit the same holds for Y. Therefore Y is a cross section. Conversely assume that Y is a cross section. Fix a particular orbit and pick $y \in Y$ and $z \in Z$ from this orbit. Then there exists $s(z) \in G$ such that $y = s(z)z$. Therefore Y can be written as (13). ∎

We see that a general cross section is obtained by arbitrarily moving the point $z \in Z$ by a member of G for each orbit.

In Lemma 2.3 we discussed how the equivariant function can be constructed from a given cross section. Let $r(x)$ and $q(x)$ be the equivariant functions associated with Z and Y respectively. Each point $x \in \mathcal{X}$ has two different representations

$$x = r(x)z(x) = q(x)y(x).$$

Let us confirm how these equivariant functions are related to $s(z)$ in (13). For $y = s(z)z$ we have

$$e = q(y) = q(s(z)z) = s(z)q(z)$$

and hence $s(z) = q(z)^{-1}$, which means that

$$Y = \{y(z) = q(z)^{-1}z \mid z \in Z\} = \{y(x) = q(x)^{-1}x \mid x \in \mathcal{X}\}.$$

Note that

$$q(x) = q(r(x)z(x)) = r(x)q(z(x)) = r(x)s(z(x))^{-1}.$$

Similarly we can show that $s(z) = r(y)$ and

$$Z = \{z(y) = r(y)^{-1}y \mid y \in Y\} = \{z(x) = r(x)^{-1}x \mid x \in \mathcal{X}\}.$$

The next theorem gives a class of cross sectionally contoured distributions associated with the general cross sections.

Theorem 2.2 *Let* x *have cross sectionally contoured density* $h(r(x))$ *with respect to the relatively invariant dominating measure* π . *Let* $q(x)$ *be an equivariant function. Then the density function of the random variable*

$$w = r(x)q(x)^{-1}x = r(x)q(z(x))^{-1}z(x) \tag{14}$$

with respect to π *is given by*

$$h(q(w))\Delta(r(w)^{-1}q(w)).$$

Here Δ *is the right hand modulus of* π_r *:*

$$\pi_r(d(rg)) = \Delta(g)\pi_r(dr).$$

Proof. For $A \subset \mathcal{X}$ let

$$Z_A = \{r(x)^{-1}x \in Z \mid x \in A\}$$

and

$$A(z) = \{g \in G \mid gz \in A\}.$$

Then one can write

$$A = \bigcup_{z \in Z_A} A(z)\{z\},$$

which is a disjoint partition of A . Note that

$$I_A(x) = I_{Z_A}(z(x))I_{A(z(x))}(r(x)).$$

Then the distribution of w is

$$\begin{aligned}
P(w \in A) &= \int_{\mathcal{X}} I_A(r(x)q(z(x))^{-1}z(x))h(r(x))\pi(dx) \\
&= \int_Z \int_G I_A(rq(z)^{-1}z)h(r)\pi_r(dr)\pi_z(dz) \\
&= \int_Z I_{Z_A}(z)\Big(\int_G I_{A(z)}(rq(z)^{-1})h(r)\pi_r(dr)\Big)\pi_z(dz) \\
&= \int_Z I_{Z_A}(z)\Big(\int_G I_{A(z)}(rq(z)^{-1})h(rq(z)^{-1}q(z))\Delta(q(z))\pi_r(d(rq(z)^{-1}))\Big)\pi_z(dz) \\
&= \int_Z I_{Z_A}(z)\Big(\int_G I_{A(z)}(r)h(rq(z))\Delta(q(z))\pi_r(dr)\Big)\pi_z(dz) \\
&= \int_Z \int_G I_A(rz)h(q(rz))\Delta(r(z)^{-1}q(z))\pi_r(dr)\pi_z(dz) \\
&= \int_{\mathcal{X}} I_A(x)h(q(x))\Delta(r(x)^{-1}q(x))\pi(dx).
\end{aligned}$$

∎

Remark 2.3 *The transform (14) can be decomposed into the equivariant and invariant parts:*

$$\begin{cases} q(w) = r(x), \\ y(w) = q(z(x))^{-1}z(x). \end{cases}$$

x *and* w *are one-to-one. The inverse transform of (14) is easily shown to be*

$$x = q(w)r(y(w))^{-1}y(w) = q(w)r(w)^{-1}w,$$

or

$$\begin{cases} r(x) = q(w), \\ z(x) = r(y(w))^{-1}y(w). \end{cases}$$

By Theorem 2.2, when $h(r(x))$ is a density function with respect to π , we can define a family of distributions dominated by π :

$$\{f(x;q) = h(q(x))\Delta(r(x)^{-1}q(x)) \mid q \in \mathcal{C}\} \tag{15}$$

or more generally

$$\{f(x;g,q) = \frac{1}{\chi(g)}h(g^{-1}q(x))\Delta(r(x)^{-1}q(x)) \mid g \in G,\ q \in \mathcal{C}\}, \tag{16}$$

where $\mathcal{C}$ is the set of all measurable equivariant functions from $\mathcal{X}$ to G . The distribution given by (15) or (16) is a cross sectionally contoured distribution because it has the cross sectionally contoured density $h(q(x))$ or $\chi(g)^{-1}h(g^{-1}q(x))$ with respect to the relative invariant measure on $\mathcal{X}$

$$\tilde{\pi}(dx) = \Delta(r(x)^{-1}q(x))\pi(dx). \tag{17}$$

Note that $\tilde{\pi}$ has the same multiplier χ as π , and that $\tilde{\pi}$ and π are absolutely continuous with respect to each other because

$$0 < \Delta(r(x)^{-1}q(x)) < \infty.$$

Next we will focus on the distribution of $y = q(x)^{-1}x$ when the distribution of x is cross sectionally contoured with $r(x)$ as the equivariant function. Applying Theorem 2.1 to the probability measure derived in Theorem 2.2 we obtain the following.

Theorem 2.3 *Let* x *have cross sectionally contoured density* $h(r(x))$ *with respect to* π *. Let q(x) be any equivariant function with the associated cross section* Y *and let* $x = qy$ *,* $q \in G,\ y \in Y$ *, be the orbital decomposition.*

(i) The measure π *has the factorization*

$$\pi(dx) = \pi_r(dq)\pi_y(dy), \tag{18}$$

where π_r *is the relatively invariant measure on* G *with the multiplier* χ *such that* $h(q)\pi_r(dq)$ *is a probability measure on* G *and* $\pi_y(dy)$ *is a measure on* Y *.* π_y *does not depend on* h *.*

(ii) The distribution of $y(x) = q(x)^{-1}x$ *is written as*

$$\tilde{\pi}_y(dy) = \Delta(r(y)^{-1})\pi_y(dy). \tag{19}$$

Furthermore $y(x)$ *and* $r(x)$ *are independent.*

Proof. Let $w = r(x)q(x)^{-1}x$. By applying Theorem 2.1 to the probability measure $h(q(w))\tilde{\pi}(dw)$ of w we have the representation

$$\tilde{\pi}(dw) = \tilde{\pi}_q(d(q(w)))\tilde{\pi}_y(d(y(w))), \tag{20}$$

where the distributions of $q(w)$ and $y(w)$ are given by $h(q)\tilde{\pi}_q(dq)$ and $\tilde{\pi}_y(dy)$, respectively. Note that $\tilde{\pi}_y$ does not depend on h. Since $q(w) = q(r(x)q(x)^{-1}x) = r(x)$, the density function of $q(w)$ has to be identical to that of $r(x)$, namely

$$h(q)\tilde{\pi}_q(dq) = h(q)\pi_r(dq). \tag{21}$$

If $h(q) > 0$ a.e. on G, we have

$$\tilde{\pi}_q(dq) = \pi_r(dq) \tag{22}$$

directly. Otherwise we can derive (22) from (21) using the same argument as in Appendix. Writing x instead of w in (20), from (22) we have

$$\tilde{\pi}(dx) = \pi_r(dq)\tilde{\pi}_y(dy) \tag{23}$$

with $q = q(x),\ y = y(x)$. On the other hand from (17)

$$\begin{aligned} \tilde{\pi}(dx) &= \Delta(r(x)^{-1}q(x))\pi(dx) \\ &= \Delta(r(y)^{-1})\pi(dx). \end{aligned} \tag{24}$$

Comparing (23) and (24) we can write $\pi(dx)$ as in (18) where π_y is the measure on Y satisfying (19). π_y does not depend on h, since $\tilde{\pi}_y$ does not either.

Because of

$$y(w) = q(w)^{-1}w = r(x)^{-1}w = q(x)^{-1}x = y(x),$$

$y(x) = q(x)^{-1}x$ is independent of $q(w) = r(x)$, and $\tilde{\pi}_y(dy)$ is the probability distribution of $y(x)$. ∎

Remark 2.4 *In many applications it is easier to find the factorization (18) than (20) (or equivalently (23)). Theorem 2.3 is useful in that case.*

By (19) we see that we can construct various distributions on Y by appropriately choosing the equivariant function $r(x)$. Here we can ask the following question: "Given a density $f(y)$ on Y is there a cross sectionally contoured distribution with the density $h(r(x))$ such that the distribution of $y(x) = q(x)^{-1}x$ coincides with $f(y)$?" The following corollary gives the answer.

Corollary 2.2 *Let $f(y)\pi_y(dy)$ be a distribution on Y such that $f(y)$ is almost everywhere positive on Y with respect to π_y. Suppose that the right hand modulus Δ is not identically equal to 1 on G. Then there exists a cross sectionally contoured distribution with density $h(r(x))$ with respect to π such that the distribution of $y(x) = q(x)^{-1}x$ coincides with $f(y)\pi_y(dy)$.*

Proof. Since Δ is a continuous homomorphism from G to R_+^*, its range is either $\{1\}$ or the whole R_+^*. Therefore by the assumption Δ is a surjection. Fix y and consider the inverse image of the positive real number $f(y)$:

$$\Delta^{-1}(f(y)) = \{g \mid \Delta(g) = f(y)\}.$$

Choose $r(y) \in G$ such that $r(y)^{-1} \in \Delta^{-1}(f(y))$. Furthermore for $x = gy$ define $r(x) = gr(y)$. Then $r(x)$ is an equivariant function from $\mathcal{X}$ to G with the associated cross section

$$Z = \{z(y) = r(y)^{-1}y \mid y \in Y\}.$$

For a cross sectionally contoured distribution with density $h(r(x))$ the distribution of $y(x) = q(x)^{-1}x$ is $f(y)\pi_y(dy)$ by (19). ∎

Finally we illustrate Theorems 2.2 and 2.3 with two examples. In each example two different cross sections and associated cross sectionally contoured distributions are discussed. The first example is somewhat trivial, but it is useful for confirming the logic behind Theorem 2.2.

Example 2.1 *Bivariate gamma distribution (Beta and F distributions)*

Let $\mathcal{X} = R_+ \times R_+$ and $G = R_+^*$ whose action is

$$g(x_1, x_2) = (gx_1, gx_2), \qquad \text{for } (x_1, x_2) \in \mathcal{X},\ g \in G.$$

Let

$$\begin{aligned} z(x_1, x_2) &= \left(\frac{x_1}{x_1 + x_2}, \frac{x_2}{x_1 + x_2}\right) = (z_1, z_2), \quad z_2 = 1 - z_1, \\ y(x_1, x_2) &= \left(\frac{x_1}{x_2}, 1\right) = (y_1, 1). \end{aligned}$$

The associated equivariant functions are $r(x_1, x_2) = x_1 + x_2$ and $q(x_1, x_2) = x_2$. The unit cross sections are

$$Z = \{(z_1, 1 - z_1) \mid 0 < z_1 < 1\}, \qquad Y = \{(y_1, 1) \mid y_1 > 0\},$$

respectively. Z and Y are conveniently parameterized by z_1 and y_1.

Now let x_1, x_2 be independently distributed according to gamma distributions with shape parameters a, b, respectively. Then (x_1, x_2) has the cross sectionally contoured density function

$$h(r(x_1, x_2)) = e^{-r(x_1,x_2)} = e^{-(x_1+x_2)}$$

with respect to the dominating measure

$$\pi(d(x_1,x_2)) = \frac{1}{\Gamma(a)\Gamma(b)} x_1^{a-1} x_2^{b-1} dx_1 dx_2, \quad a,b>0, \tag{25}$$

with multiplier $\chi(g) = g^{a+b}$. The factorization (11) of π is given as

$$\begin{aligned}\pi(d(x_1,x_2)) &= \pi_r(dr) \times \pi_z(dz) \\ &= \frac{1}{\Gamma(a+b)} r^{a+b-1} dr \times \frac{1}{\mathrm{Beta}(a,b)} z_1^{a-1}(1-z_1)^{b-1} dz_1,\end{aligned}$$

and hence we see that $r(x_1,x_2) = x_1 + x_2$ is distributed as gamma distribution with shape parameter $a+b$, and that $z_1 = x_1/(x_1+x_2)$ is distributed as beta distribution with parameters a,b, independently of $r(x_1,x_2)$.

Define the random variables (w_1,w_2) by

$$\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \frac{r(x_1,x_2)}{q(x_1,x_2)} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{x_1+x_2}{x_2} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Then, by Theorem 2.2, the density of (w_1,w_2) is

$$h(q(w_1,w_2)) = \mathrm{e}^{-q(w_1,w_2)} = \mathrm{e}^{-w_2},$$

where the dominating measure is

$$\begin{aligned}\tilde{\pi}(d(w_1,w_2)) &= \Delta\Big(\frac{q(w)}{r(w)}\Big)\pi(dw) \\ &= \Big(\frac{q(w)}{r(w)}\Big)^{a+b}\pi(dw) \\ &= \frac{1}{\Gamma(a)\Gamma(b)}\Big(\frac{w_2}{w_1+w_2}\Big)^{a+b} w_1^{a-1} w_2^{b-1} dw_1 dw_2.\end{aligned}$$

Theorem 2.3 assures the factorization of π in (25) associated with the orbital decomposition $x = qy$. In this case, the factorization (18) is given by

$$\begin{aligned}\pi(d(x_1,x_2)) &= \pi_r(dq) \times \pi_y(dy) \\ &= \frac{1}{\Gamma(a+b)} q^{a+b-1} dq \times \frac{1}{\mathrm{Beta}(a,b)} y_1^{a-1} dy_1,\end{aligned}$$

and hence the distribution of $y_1 = x_1/x_2$ is expressed as

$$\Delta(r(y)^{-1})\pi_y(dy) = \frac{1}{\mathrm{Beta}(a,b)}\Big(\frac{1}{y_1+1}\Big)^{a+b} y_1^{a-1} dy_1.$$

$(b/a)y_1$ is distributed as F distribution with degrees of freedom $2a, 2b$. Here $(y_1, 1) = y(x_1,x_2) = y(w_1,w_2)$ and $q(w_1,w_2) = r(x_1,x_2) = x_1 + x_2$ are independently distributed. This example will be generalized in Section 4.2.

Example 2.2 *Multivariate normal distribution*

Let $\mathcal{X} = R^p - \{0\}$ and $G = R_+^*$, where the action is

$$g(x_1, \ldots, x_p) = (gx_1, \ldots, gx_p), \qquad \text{for } x = (x_1, \ldots, x_p)' \in \mathcal{X}, \ g \in G.$$

Let $r(x) = (x'\Sigma^{-1}x)^{1/2}$, $q(x) = (x'x)^{1/2} = \|x\|$ be two equivariant functions. Corresponding invariant functions are $z(x) = x/(x'\Sigma^{-1}x)^{1/2}$ and $y(x) = x/\|x\|$, respectively. Assume that x is distributed according to p dimensional normal distribution $N_p(0, \Sigma)$. Then x has the density function

$$h(r(x)) = \exp\Big(-\frac{1}{2}r(x)^2\Big)$$

with respect to the dominating measure

$$\pi(dx) = \frac{1}{(2\pi)^{p/2}(\det\Sigma)^{1/2}}dx \tag{26}$$

with multiplier $\chi(g) = g^p$, where dx is the Lebesgue measure on R^p.

Consider the transformation

$$w = \frac{r(x)}{q(x)}x = \Big(\frac{x'\Sigma^{-1}x}{x'x}\Big)^{1/2}x.$$

Noting that

$$\tilde{\pi}(dw) = \Delta\Big(\frac{q(w)}{r(w)}\Big)\pi(dw) = \Big(\frac{w'w}{w'\Sigma^{-1}w}\Big)^{p/2}\pi(dw),$$

we obtain the distribution of w by Theorem 2.2 as

$$\begin{aligned} h(q(w))\tilde{\pi}(dw) &= \frac{1}{(2\pi)^{p/2}(\det\Sigma)^{1/2}}\exp\Big(-\frac{1}{2}q(w)^2\Big)\Delta\Big(\frac{q(w)}{r(w)}\Big)dw \\ &= \frac{1}{(2\pi)^{p/2}(\det\Sigma)^{1/2}}\exp\Big(-\frac{1}{2}w'w\Big)\Big(\frac{w'w}{w'\Sigma^{-1}w}\Big)^{p/2}dw. \end{aligned} \tag{27}$$

By applying Theorem 2.1 to the cross sectionally contoured distribution (27) we see the independence of the length $q(w) = (w'w)^{1/2}$ and the direction $y(w) = w/(w'w)^{1/2}$ under the distribution (27).

The factorization (18) of $\pi(dx)$ in (26) associated with the orbital decomposition $x = qy$ is given by

$$\begin{aligned} \pi(dx) &= \pi_r(dq) \times \pi_y(dy) \\ &= \frac{1}{2^{(p-2)/2}\Gamma(p/2)}q^{p-1}dq \times \frac{1}{\omega_p(\det\Sigma)^{1/2}}dy \end{aligned}$$

where dy is the "volume" element of the unit sphere $Y = S^{p-1} = \{x \mid q(x) = \|x\| = 1\}$ in R^p and

$$\omega_p = \frac{2\pi^{p/2}}{\Gamma(p/2)} \tag{28}$$

is the total volume of the unit sphere. Therefore from (ii) of Theorem 2.3 the distribution of $y = y(x)$ $(= y(w))$ is

$$\Delta(r(y)^{-1})\pi_y(dy) = \frac{1}{\omega_p(\det \Sigma)^{1/2}} \frac{1}{(y'\Sigma^{-1}y)^{p/2}} dy. \tag{29}$$

This example will be discussed again in Example 3.3 in the context of star-shaped distributions.

Remark 2.5 *Figure 2 is a surface plot of the density function (27) in two dimensional case ($p = 2$) with its contours plotted on the bottom plane. Here we set* $\Sigma = \mathrm{diag}(\sigma_{11}, \sigma_{22}) = \mathrm{diag}(\sqrt{2}, 1/\sqrt{2})$. *Note that the density function is not continuous at the origin. In fact, if we let* $w = (\rho\cos\phi, \rho\sin\phi)$ *and* $\rho \to +0$ *with* ϕ *fixed, it holds that*

$$\exp\Big(-\frac{1}{2}w'w\Big)\Big(\frac{w'w}{w'\Sigma^{-1}w}\Big)^{2/2} \to \frac{1}{\sigma_{11}^{-1}\cos^2\phi + \sigma_{22}^{-1}\sin^2\phi},$$

which depends on ϕ *unless* $\sigma_{11} = \sigma_{22}$. *Figure 2 is drawn by connecting finite mesh points of the density (27) and its behavior around the origin is not perfectly rendered in Figure 2.*

Remark 2.6 *The distribution (29) has been noted in several literatures. Section 3.6 of Watson (1983) referred to this distribution as angular Gaussian distribution, and discussed some properties. Several arguments on statistical inferences based on this model are given in Tyler (1987). The special case where* $p = 2$ *is treated in Section 3.4.7 of Mardia (1972). See also Kent and Tyler (1988) for the case of* $p = 2$.

The distribution (29) of $y = x/\|x\|$, *as well as that of* $y'\Sigma^{-1}y = x'\Sigma^{-1}x/x'x$, *plays an important role in null robust testing problems. See, e.g., Kariya and Eaton (1977) and King (1980).*

2.4 Density of the equivariant and the invariant for Euclidean sample space

In this section we consider the case that the sample space $\mathcal{X}$ is Euclidean R^n endowed with the standard inner product $\langle\cdot,\cdot\rangle$. We assume that $G \subset GL(p)$ is a Lie group of dimension m ($m \le p^2$) and the action of G on $\mathcal{X}$ is of class C^1 .

Furthermore we assume that the Lebesgue measure dx is relatively invariant under the action of G with multiplier χ .

Given a cross section Z each orbit can be identified with G . Since the action of G is of class C^1 , each orbit is a manifold of dimension m of class C^1 . Regarding cross sections, we restrict our attention to cross sections Z which are manifolds of class C^1 of dimension $n - m$. We can allow Z to be piecewise of class C^1 as in Example 1.1, where Z consists of 3 straight line segments. In terms of the equivariant function we equivalently assume that $r(x)$ is piecewise of class C^1 . In the following we will ignore singularities of Z because for piecewise smooth Z , the singularities form a set of

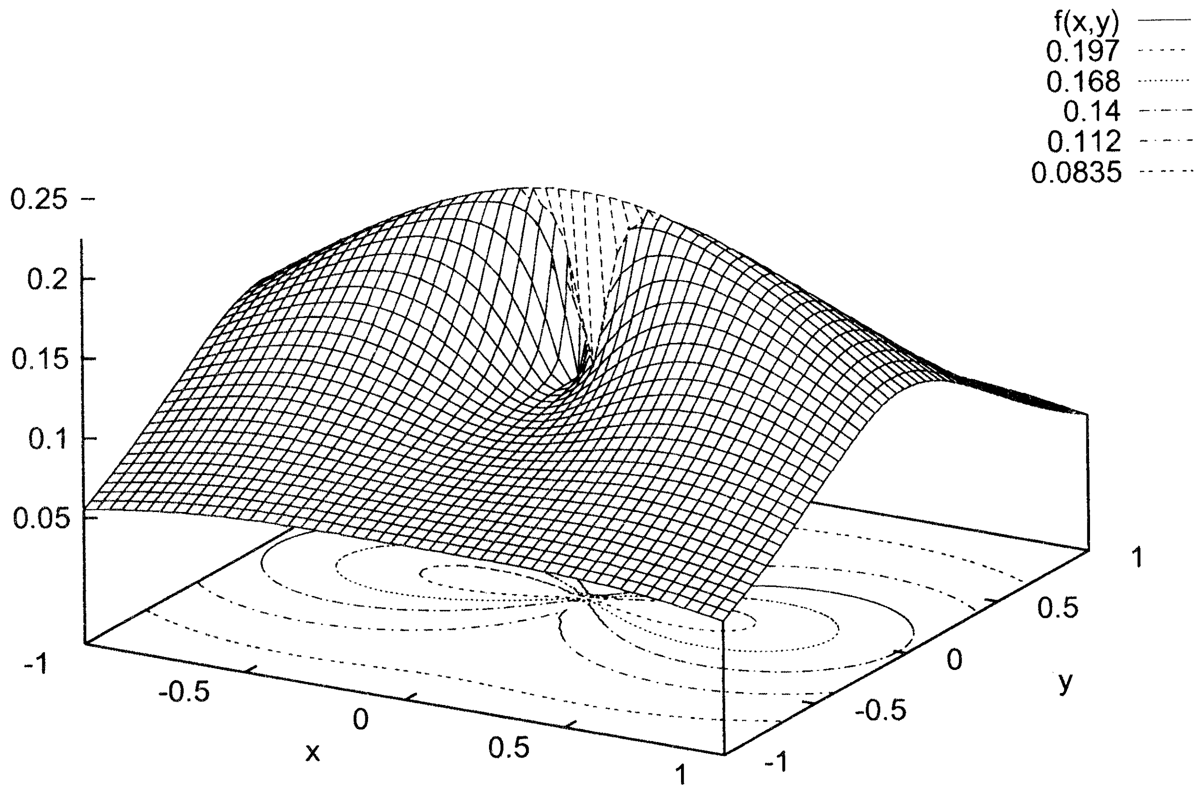

Figure 2: Surface plot and contour of the density (27)

Lebesgue measure 0. Note that Z or $r(x)$ may even be discontinuous at the points of singularity.

Fix a point $x_0 \in \mathcal{X}$ and write $x_0 = r_0 z_0$. Let $u_1, \ldots, u_m$ be local coordinates of G in a neighborhood of the identity element e of G and let $u_{m+1}, \ldots, u_n$ be local coordinates of Z in a neighborhood of z_0. Then x in a neighborhood of x_0 can be expressed as

$$x = r_0 r(u_1, \ldots, u_m) z(u_{m+1}, \ldots, u_n). \tag{30}$$

Here we are using the fact that a neighborhood of e translated by r_0 is a neighborhood of r_0. For convenience we assume $r(0, \ldots, 0) = e$ and $z(0, \ldots, 0) = z_0$.

In the examples considered in Sections 3 and 4 the groups are R_+^* and the lower triangular group $LT(p)$. These groups have a natural global coordinate system. However for the general consideration, it is advantageous to use local coordinates at the unit cross section as (30).

The usual Jacobian calculation at $x = x_0$ gives us

$$dx = J(r_0, z_0) du_1 \cdots du_m \times du_{m+1} \cdots du_n$$

where J is the Jacobian. Now under the regularity conditions of Theorem 2.1, $J(r_0, z_0)$ can be factored as

$$J(r_0, z_0) = J_r(r_0)J_z(z_0).$$

Now dx can be written as

$$dx = J_r(r_0)du_1 \cdots du_m \times J_z(z_0)du_{m+1} \cdots du_n. \tag{31}$$

$J_r(r_0)du_1 \cdots du_m$ is a relatively invariant measure on G with multiplier χ.

Consider translating x by g. Then

$$d(gx) = \chi(g)dx = \chi(g)J_r(r_0)du_1 \cdots du_m \times J_z(z_0)du_{m+1} \cdots du_n. \tag{32}$$

On the other hand $gx = gr_0z_0$ and by (31)

$$d(gx) = J_r(gr_0)du_1 \cdots du_m \times J_z(z_0)du_{m+1} \cdots du_n. \tag{33}$$

(32) and (33) imply

$$J_r(r_0) = \chi(r_0)J_r(e)$$

and (31) can be written as

$$dx = \chi(r_0)J_r(e)du_1 \cdots du_m \times J_z(z_0)du_{m+1} \cdots du_n. \tag{34}$$

It might be mathematically more pleasing to express the above relation in terms of exterior differential forms. Taking the differential of $x = rz$ and forming the wedge product (31) can be written as

$$dx_1 \wedge \cdots \wedge dx_n = \omega_r \wedge \omega_z,$$

where ω_r is a relatively invariant m-form on G with multiplier χ and ω_z is an $(n-m)$-form on Z. In the following we employ both the usual Jacobian calculation and calculation using differential forms.

For the rest of this section we consider orbits and cross sections as submanifolds of R^n with volume elements induced from the usual inner product of R^n. Unfortunately we could not establish general results relating the volume elements of the orbits or the cross sections to the group action. However in the case of star-shaped distributions in Section 3, the arguments below give very clear interpretation of the behavior of the group action. Also for the generalization of star-shaped distribution in Section 4.1 the arguments below are useful.

Define the *orbit manifold* $M_O(x_0)$ and *cross section manifold* $M_C(x_0)$ through x_0 by

$$M_O(x_0) = \{rz_0 \mid r \in G\}, \qquad M_C(x_0) = \{r_0z \mid z \in Z\}.$$

Denote the tangent space of $M_O(x_0)$ at x_0 by $T_{x_0}(M_O)$ and the tangent space of $M_C(x_0)$ at x_0 by $T_{x_0}(M_C)$. Similarly denote the tangent space of $\mathcal{X}$ itself at x_0 by $T_{x_0}(\mathcal{X})$.

dx on the left hand side of (34) is the (Lebesgue) volume element at x_0. Now both $M_O(x_0)$ and $M_C(x_0)$ as submanifolds of R^n have volume elements induced by the usual inner product of R^n. We can investigate these volume elements using the first fundamental form.

First consider the orbit and cross section manifolds at z_0. Regard $x = r(u_1, \ldots, u_m) z(u_{m+1}, \ldots, u_n)$ as an n dimensional vector and let

$$v_i = \frac{\partial x}{\partial u_i}\Big|_{u_1=\cdots=u_n=0}, \qquad i = 1, \ldots, n.$$

Then $\{v_1, \ldots, v_m\}$ forms a basis of $T_{z_0}(M_O)$ and $\{v_{m+1}, \ldots, v_n\}$ forms a basis of $T_{z_0}(M_C)$. Let

$$\Gamma = \Gamma(z_0) = (\gamma_{ij}) = (\langle v_i, v_j \rangle), \qquad i, j = 1, \ldots, n,$$

be the matrix of inner products of v_i 's and let $\Gamma(z_0)$ be partitioned as

$$\Gamma(z_0) = \begin{pmatrix} \Gamma_{11}(z_0) & \Gamma_{12}(z_0) \\ \Gamma_{21}(z_0) & \Gamma_{22}(z_0) \end{pmatrix} \tag{35}$$

where $\Gamma_{11}(z_0)$ is $m \times m$. Then around z_0

$$\begin{aligned} dx &= \sqrt{\det \Gamma(z_0)} du_1 \cdots du_n \\ &= \sqrt{\det \Gamma_{11}(z_0)} du_1 \cdots du_m \times \sqrt{\det \Gamma_{22}(z_0)} du_{m+1} \cdots du_n \\ &\qquad \times \sqrt{\det(I_m - \Gamma_{12}(z_0)\Gamma_{22}^{-1}(z_0)\Gamma_{21}(z_0)\Gamma_{11}^{-1}(z_0))}. \end{aligned} \tag{36}$$

On the right hand side the first term is the volume element of the orbit manifold $M_O(z_0)$ at z_0 and the second term is the volume element of the cross section manifold $M_C(z_0)$ at z_0. Furthermore the third term can be written as

$$\prod_{i=1}^{\min(m,n-m)} \sin \theta_i,$$

where $\theta_1, \ldots, \theta_{\min(m,n-m)}$ are the canonical angles between the tangent spaces $T_{z_0}(M_O)$ and $T_{z_0}(M_C)$.

Now consider translation by g from $x = z_0$ to $x = gz_0$. Let g_* denote the differential of the map $x \mapsto gx$. g_* is a linear map from $T_{z_0}(\mathcal{X})$ to $T_{gz_0}(\mathcal{X})$. Clearly

$$\chi(g) = |\det g_*|.$$

Furthermore we have

$$g_* T_{z_0}(M_O) = T_{gz_0}(M_O), \qquad g_* T_{z_0}(M_C) = T_{gz_0}(M_C),$$

because

$$gM_O(z_0) = M_O(gz_0) = M_O(z_0), \qquad gM_C(z_0) = M_C(gz_0).$$

Let $v_j(gz_0) = g_* v_j$, $j = 1, \ldots, n.$ Then $\{v_1(gz_0), \ldots, v_m(gz_0)\}$ forms a basis of $T_{gz_0}(M_O)$ and $\{v_{m+1}(gz_0), \ldots, v_n(gz_0)\}$ forms a basis of $T_{gz_0}(M_C)$. Let $\Gamma(gz_0) = (\gamma_{ij}(gz_0)) = (\langle v_i(gz_0), v_j(gz_0) \rangle)$, $i, j = 1, \ldots, n,$ and partition $\Gamma(gz_0)$ as in (35). Then around $x_0 =$

$r_0 z_0$ we have the same decomposition of volume element as in (36) with $\Gamma_{ij}(z_0)$ replaced by $\Gamma_{ij}(r_0 z_0)$. Now (34) implies that

$$\chi(r_0) = \frac{\sqrt{\det \Gamma_{11}(r_0 z_0)}}{\sqrt{\det \Gamma_{11}(z_0)}} \cdot \frac{\sqrt{\det \Gamma_{22}(r_0 z_0)}}{\sqrt{\det \Gamma_{22}(z_0)}} \times \frac{\sqrt{\det(I_m - \Gamma_{12}(r_0 z_0)\Gamma_{22}^{-1}(r_0 z_0)\Gamma_{21}(r_0 z_0)\Gamma_{11}^{-1}(r_0 z_0))}}{\sqrt{\det(I_m - \Gamma_{12}(z_0)\Gamma_{22}^{-1}(z_0)\Gamma_{21}(z_0)\Gamma_{11}^{-1}(z_0))}}. \quad (37)$$

The first term and the second term on the right hand side show the changes of volume elements of the orbit manifold and the cross section manifold due to translation by g respectively. Write

$$\chi_O(r_0) = \frac{\sqrt{\det \Gamma_{11}(r_0 z_0)}}{\sqrt{\det \Gamma_{11}(z_0)}}, \quad \chi_C(r_0) = \frac{\sqrt{\det \Gamma_{22}(r_0 z_0)}}{\sqrt{\det \Gamma_{22}(z_0)}}. \quad (38)$$

It is of interest to investigate whether χ_O and χ_C are homomorphisms from G to R_+^* . If χ_O is a homomorphism, we say that the volume element of the orbit manifold is relatively invariant with multiplier χ_O . We also say that the volume element on the cross section manifold is relatively invariant with multiplier χ_C if χ_C is a homomorphism from G to R_+^* . For the case of star-shaped distributions in Section 3 both χ_O and χ_C are homomorphisms. In Section 4.1 we will see an example where $\chi_O(r)$ is a homomorphism although $\chi_C(r)$ is not.

3 Star-shaped distributions in R^p

In this section we define star-shaped distributions in R^p generalizing Example 1.1 and investigate their properties. All relevant results are easy consequences of the general theory given in the previous section. We summarize the results in Theorem 3.1 below.

Let $G = R_+^*$ and define its action on R^p by

$$g(x_1, \ldots, x_p) = (gx_1, \ldots, gx_p).$$

The Lebesgue measure is relatively invariant with multiplier $\chi(g) = g^p$ and we take the Lebesgue measure as the dominating measure.

Let $r(x)$ be an equivariant function from R^p to R_+^* . We call distributions with the densities of the form $f(x) = h(r(x))$ *star-shaped* distributions. The orbits are rays starting from the origin. Note however that we omit the origin from the sample space R^p . We can do this because the Lebesgue measure of the origin is zero.

The associated cross section

$$Z = \{x \mid r(x) = 1\}$$

is a set which meets each ray exactly once. Hence

$$\bigcup_{0 \le c \le 1} cZ$$

contains every line segment connecting the origin and a point on Z, namely it is a star-shaped set with respect to the origin. This is the reason why we call the densities of the form $h(r(x))$ star-shaped. (For the term "star-shaped" see also Definition 3.1 of Naiman and Wynn (1992).)

On R_+^* we take $r^{p-1}dr$ as the standard relatively invariant measure with multiplier $\chi(r) = r^p$. By Theorem 2.1 $r = r(x)$ and $z = x/r(x)$ are independent and the joint distribution of r and z can be written as

$$(1/c_0)\, h(r) r^{p-1} dr\, \pi_z(dz), \qquad c_0 = \int_0^\infty h(r) r^{p-1} dr, \tag{39}$$

where π_z is a probability measure on Z.

For star-shaped distributions the standard cross section is obviously the unit sphere $Y = S^{p-1} = \{x \mid q(x) = 1\}$ where $q(x) = \|x\| = (x'x)^{1/2}$ is the usual Euclidean length of x. Now dx obviously factors as

$$dx = q^{p-1} dq\, dy = (1/c_0)\, q^{p-1} dq \cdot c_0 dy,$$

where dy is the volume element on the unit sphere. In this case "area element" might be a better expression, but we just use the word "volume element" regardless of the dimensionality.

Since $G = R_+^*$ is commutative the right hand modulus coincides with the multiplier

$$\Delta(g) = g^p.$$

From Theorem 2.3 the distribution of $y = x/\|x\|$ is given as

$$c_0\, r(y)^{-p} dy, \qquad 1/c_0 = \int_Y r(y)^{-p} dy.$$

c_0 is also given by (39).

We now investigate star-shaped densities more closely using the techniques of Section 2.4. We make the additional assumption that $r(x)$ is piecewise of class C^1. Let dr be the volume (length) element of R_+^* around r and du_1 be the volume element of R_+^* around $r = 1$. Then $dr = r\, du_1$. Let

$$v_1 = \frac{\partial x}{\partial u_1}\Big|_{z_0} = \frac{\partial}{\partial u_1}(u_1 z_0) = z_0.$$

Now choose local coordinates $u_2, \ldots, u_p$ of $M_C(z_0)$ such that

$$v_j = \frac{\partial}{\partial u_j} z(0, \ldots, 0, u_j, 0, \ldots, 0)\Big|_{u_j=0}, \quad j = 2, \ldots, p,$$

are orthonormal vectors. Then $du_2 \cdots du_p$ is the volume element of $M_C(z_0)$ at z_0. Writing $x = u_1 z(u_2, \ldots, u_p)$ we see that

$$dx = |\det(v_1, \ldots, v_p)| \times du_1 \times du_2 \cdots du_p, \tag{40}$$

where $(v_1, \ldots, v_p)$ denotes the matrix consisting of columns $v_1, \ldots, v_n$. Let n_{z_0} be the unit normal vector of Z at z_0 pointing outward of the star-shaped set $\bigcup_{0 \le c \le 1} cZ$. Write $v_1 = z_0$ as a linear combination of orthonormal vectors $n_{z_0}, v_2, \ldots, v_p$, as

$$z_0 = a_1 n_{z_0} + a_2 v_2 + \cdots + a_p v_p, \qquad a_1 = \langle z_0, n_{z_0} \rangle, \quad a_i = \langle z_0, v_i \rangle, \quad i = 2, \ldots, p.$$

Then

$$|\det(v_1, \ldots, v_p)| = a_1 = \langle z_0, n_{z_0} \rangle$$

and (40) is written as

$$dx = du_1 \times du_2 \cdots du_p \times \langle z_0, n_{z_0} \rangle.$$

Rewrite this further as

$$dx = \|z_0\| du_1 \times du_2 \cdots du_p \times \langle z_0 / \|z_0\|, n_{z_0} \rangle. \tag{41}$$

Note that $\|z_0\| du_1 = \sqrt{\langle v_1, v_1 \rangle} du_1$ is the volume element of $M_O(z_0)$. As mentioned above $du_2 \cdots du_p$ is the volume element of $M_C(z_0)$. Let θ denote the angle between z_0 (or $T_{z_0}(M_O)$) and the tangent space $T_{z_0}(M_C)$. Then $\pi/2 - \theta$ is the angle between z_0 and n_{z_0}. Therefore

$$\langle z_0 / \|z_0\|, n_{z_0} \rangle = \sin \theta.$$

We now see that (41) corresponds to (36) in the previous section.

We also note that the unit normal vector n_{z_0} coincides with the normalized gradient of $r(x)$, i.e.,

$$n_{z_0} = \frac{\nabla r(z_0)}{\|\nabla r(z_0)\|}.$$

Furthermore sometimes it helps to use the following fact. Let $H_{z_0} = z_0 + T_{z_0}(M_C)$ be the tangent hyperplane of Z at z_0. Then

$$\langle z_0, n_{z_0} \rangle = \text{Euclidean distance from the origin to } H_{z_0}. \tag{42}$$

For star-shaped distributions the effect of translation by g is straightforward. Since the translation by g is just the scale change, we see that the volume element of the orbit manifold $M_O(z_0)$ is multiplied by g and the volume element of the cross section manifold $M_C(z_0)$ is multiplied by g^{p-1}, $p-1$ being the dimensionality of $M_C(z_0)$. Furthermore the angle between these manifolds does not change by g. Therefore in the present setup (37) just reads

$$r \times r^{p-1} \times 1.$$

We now summarize the above results in the following theorem.

Theorem 3.1 *Let $x \in R^p$ have star-shaped density of the form $h(r(x))$. Then $r = r(x)$ and $z = x/r(x)$ are independent and the joint distribution of r and z is written as*

$$(1/c_0)\, h(r) r^{p-1}\, dr\, \pi_z(dz),$$

where $c_0 = \int_0^\infty h(r) r^{p-1} dr$ and π_z is a probability measure on $Z = \{x \mid r(x) = 1\}$.

Let dy denote the volume element of the unit sphere $Y = S^{p-1}$ in R^p. Then $1/c_0 = \int_Y r(y)^{-p} dy$ and $y = x/\|x\|$ is distributed as

$$c_0 \, r(y)^{-p} dy.$$

Under the additional assumption that $r(x)$ is piecewise of class C^1, $(1/c_0)\,\pi_z(dz)$ is written as

$$(1/c_0)\,\pi_z(dz) = \langle z, n_z \rangle \, dz,$$

where n_z is the outward unit normal vector of Z and dz on the right hand side is the volume element of Z. Furthermore the joint distribution of r and z is written as

$$h(r) \times \|z\| dr \times r^{p-1} dz \times \langle z/\|z\|, n_z \rangle,$$

where $\|z\|dr$ is the volume element of $M_O(x)$, $r^{p-1}dz$ is the volume element of $M_C(x)$, and $\langle z/\|z\|, n_z\rangle$ equals $\sin\theta$ where θ is the angle between $T_x(M_O)$ and $T_x(M_C)$.

Now consider Corollary 2.2. For the case of star-shaped distributions in R^p the right hand modulus $\Delta(g) = g^p$ is one-to-one. Therefore for a given positive probability density $f(y)$ on the unit sphere $Y = S^{p-1}$ we can uniquely determine an equivariant function $r(x)$ such that for the cross sectionally contoured density $h(r(x))$ the distribution of $y = x/\|x\|$ is given by $f(y)dy$. In fact the unique equivariant function is

$$r(x) = \|x\| (f(x/\|x\|))^{-1/p}$$

with the associated cross section

$$Z = \{(f(y))^{1/p} y \mid y \in Y\}.$$

For the rest of this section we consider some examples of star-shaped distributions.

Example 3.1 *(Example 1.1 continued.)*

First we work out Example 1.1 in detail. Just for concreteness let us specify h as $h(r) \propto \exp(-r^2)$. Therefore

$$f(x) \propto \exp(-r(x)^2), \qquad r(x) = \max(-x_1, -x_2, x_1 + x_2).$$

We need to evaluate the integration over $Z = \{x \mid \max(-x_1, -x_2, x_1 + x_2) = 1\}$ to determine the overall normalizing constant for $f(x)$.

Note that in Figure 3 the closest point on L_1 from the origin is $(1/2, 1/2)$ and the closest point on L_2 from the origin is $(0, -1)$. These are interior points of the line segments. Therefore from (42) we have

$$\langle z, n_z \rangle = \begin{cases} 1/\sqrt{2}, & \text{if } z \in L_1, \\ 1, & \text{if } z \in L_2, L_3. \end{cases}$$

Noting that on Z the lengths of the line segments L_1, L_2, L_3 are $3\sqrt{2}, 3, 3$, respectively, we have

$$1/c_0 = \int_Z \langle z, n_z \rangle dz = 3 \times 3 = 9. \tag{43}$$

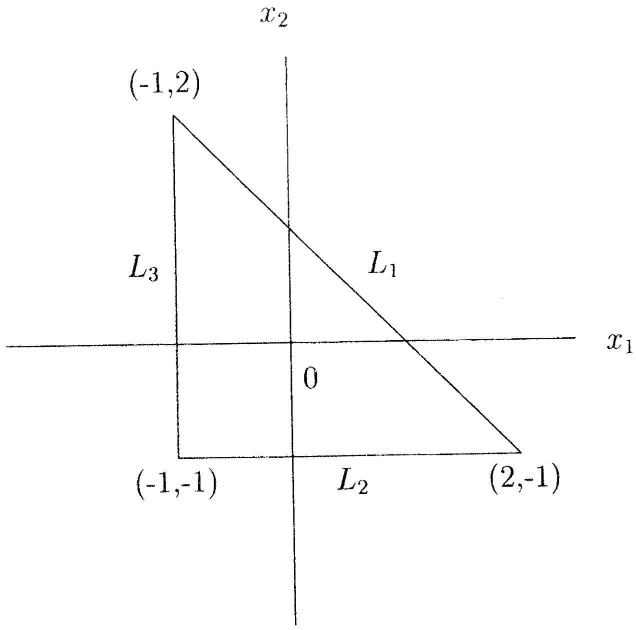

Figure 3: Line Segments of Triangle Contour

We see that under $f(x)$, each of three line segments gets equal probability 1/3, and on each of these line segments z is distributed uniformly (i.e., on each of the line segments, the probability is proportional to the length). Now

$$(1/c_0) \int_0^\infty r \exp(-r^2) dr = \frac{9}{2}.$$

Therefore after normalization

$$f(x) = \frac{2}{9} \exp(-r(x)^2), \qquad r(x) = \max(-x_1, -x_2, x_1 + x_2).$$

It is interesting to note that each of L_1, L_2, L_3 gets probability $1/3$, although L_1 is longer than L_2, L_3 . Therefore the probability density (with respect to the volume element of the line segments) on L_1 is actually lower than on L_2 or L_3 and the density of z is not "uniform" on the whole Z . This is because the distance between neighboring contours of $f(x) = h(r(x))$ is narrower (by the factor $1/\sqrt{2}$) around L_1 compared to L_2 and L_3 .

Let θ denote the angle of the vector x . Then the density of θ is proportional to $r(\cos\theta, \sin\theta)^{-2}$ with respect to $d\theta$. Let $\theta_0 = \arctan(1/2) \doteq 0.4636$. Then

$$r(\cos\theta, \sin\theta)^{-2} = \begin{cases} (\cos\theta + \sin\theta)^{-2}, & \text{if } -\theta_0 < \theta < \pi/2 + \theta_0, \\ \sin^{-2}\theta, & \text{if } -3/4\pi < \theta < -\theta_0, \\ \cos^{-2}\theta, & \text{if } \pi/2 + \theta_0 < \theta < 5/4\pi\ . \end{cases}$$

Using $(d/d\theta)\tan\theta = 1/\cos^2\theta$ we can directly evaluate the definite integrals. The results are

$$\int_{-\theta_0}^{\pi/2+\theta_0} (\cos\theta + \sin\theta)^{-2} d\theta = \int_{-3/4\pi}^{-\theta_0} \sin^{-2}\theta d\theta = \int_{\pi/2+\theta_0}^{5/4\pi} \cos^{-2}\theta d\theta = 3. \tag{44}$$

Therefore the normalizing constant is again $3 \times 3 = 9$ and the distribution of θ is given as

$$\frac{1}{9} r(\cos\theta, \sin\theta)^{-2} d\theta. \tag{45}$$

However there is no need to carry out integrals in (44). c_0 has been already calculated as $c_0 = 1/9$ in (43). This immediately leads to (45) as summarized in Theorem 3.1.

Also note that actually we already know the values of (44) because on Z each line segment gets probability $1/3$ and therefore on the unit circle each of the corresponding parts has to get probability $1/3$ under the induced probability measure.

Example 3.2 *Hypercube distribution and crosspolytope distribution*

As a next example let us consider "hypercube" distribution and "crosspolytope" distribution on R^p. Hypercube and crosspolytope are basic regular polytopes in R^p polar to each other (Chapter 0 of Ziegler (1995)).

Let

$$r(x) = \max(|x_1|, \ldots, |x_p|),$$

then $Z = \{x \mid r(x) = 1\}$ is the surface of the hypercube C_p in R^p. As in the previous example the density of z is constant on each face of Z. Furthermore by symmetry each face of Z gets equal probability. Therefore in this case the distribution of z is uniform on Z. Since $\langle z, n_z \rangle = 1$, $\int_Z \langle z, n_z \rangle dz$ coincides with the total volume of the surface of the hypercube. Therefore

$$1/c_0 = \int_Z \langle z, n_z \rangle dz = 2p \times 2^{p-1}.$$

Now Theorem 3.1 implies that

$$\int_{y_1^2 \geq y_2^2, \ldots, y_p^2} \frac{1}{|y_1|^p} dy = 2^p, \tag{46}$$

where $y = (y_1, \ldots, y_p) \in Y = S^{p-1}$.

Now let

$$r(x) = |x_1| + \cdots + |x_p|,$$

then $Z = \{x \mid r(x) = 1\}$ is the surface of the crosspolytope C_p^Δ in R^p.
Fang and Fang (1987) and their subsequent papers in Fang and Anderson (1990) studied properties of this distribution under the name "ℓ_1-norm symmetric distributions".

C_p^Δ has 2^p faces and each face is just the sign change of the standard $p-1$ dimensional simplex Δ_{p-1}:

$$\Delta_{p-1} = \{x = (x_1, \ldots, x_p) \in R^p \mid x_i \geq 0,\ i = 1, \ldots, p,\ x_1 + \cdots + x_p = 1\}.$$

It is easy to check that the $p-1$ dimensional volume of Δ_{p-1} is given by

$$V_{p-1}(\Delta_{p-1}) = \frac{\sqrt{p}}{(p-1)!}.$$

The Euclidean distance from the origin to Δ_{p-1} is $1/\sqrt{p}$. Dividing the total volume of the surface of the crosspolytope by $\sqrt{p}$ we obtain

$$\int_Z \langle z, n_z \rangle dz = 2^p \times \frac{1}{(p-1)!}.$$

This leads to the following definite integral.

$$\int_{y_1 \geq 0, \ldots, y_p \geq 0} \frac{1}{(|y_1| + \cdots + |y_p|)^p} \, dy = \frac{1}{(p-1)!}. \tag{47}$$

The definite integrals (46) and (47) seem to be rather difficult to evaluate directly.

Example 3.3 *Elliptically contoured distribution*

Finally we consider the elliptically contoured distribution in R^p. Let Σ be a $p \times p$ positive definite matrix and let $f(x) = h(r(x))$, $r(x) = (x'\Sigma^{-1}x)^{1/2}$. The gradient of $r(x)$ is given as $\Sigma^{-1}x/r(x)$.

The density of $z = x/r(x)$ on Z with respect to the volume element dz of Z is proportional to

$$\langle z, \Sigma^{-1}z/\|\Sigma^{-1}z\| \rangle = \frac{1}{\sqrt{z'\Sigma^{-2}z}}.$$

We can determine the constant by considering the particular case of the normal distribution. Note that

$$\begin{aligned} 1 &= \int_{\mathcal{X}} \frac{1}{(2\pi)^{p/2}(\det \Sigma)^{1/2}} \exp\Big(-\frac{1}{2} r(x)^2 \Big) dx \\ &= \frac{1}{(2\pi)^{p/2}(\det \Sigma)^{1/2}} \int_0^\infty \exp\Big(-\frac{1}{2} r^2 \Big) r^{p-1} dr \int_Z \langle z, n_z \rangle dz. \end{aligned}$$

Hence

$$1/c_0 = \omega_p (\det \Sigma)^{1/2} = \int_Z \langle z, n_z \rangle dz,$$

where ω_p is given in (28).

The density of z on Z with respect to the volume element dz of Z is given as

$$\frac{1}{\omega_p (\det \Sigma)^{1/2}} \frac{1}{\sqrt{z'\Sigma^{-2}z}}.$$

The distribution of the direction vector $y = x/\|x\|$ was already given in (29), because (29) does not depend on the normality assumption in Example 2.2.

4 Matrix star-shaped distributions

In this section we generalize the star-shaped distributions of the previous section to random matrices. The group $\mathcal{G}$ we consider is the group $\mathcal{LT}(p)$ of lower triangular matrices with positive diagonal elements. In Section 4.1 we consider $\mathcal{LT}(p)$ acting on $n \times p$ matrix X as $X \mapsto XT'$ where $T \in \mathcal{LT}(p)$. In Section 4.2 we consider two sample Wishart problem and consider two $p \times p$ positive definite matrices W_1, W_2. The action of $\mathcal{LT}(p)$ is given as $(W_1, W_2) \mapsto (TW_1T', TW_2T')$.

Although matrix generalizations of star-shaped distributions are not as useful as the star-shaped distributions in the previous section for application, it is of considerable theoretical interest, because $\mathcal{LT}(p)$ is not commutative and the results of the general theory of Section 2 can be fully appreciated by considering these generalizations.

Throughout this section we will write groups, cross sections, and other manifolds in script letters to distinguish them from matrices written in capital letters.

4.1 Distribution related to Gram-Schmidt orthonormalization

Let $\mathcal{X}$ be the set of $n \times p$ matrices $X = (x_{ij})$ of rank p. $\mathcal{X}$ is identified with R^{np} with the standard inner product

$$\langle X, Y \rangle = \operatorname{tr} X'Y, \qquad X,\ Y \in \mathcal{X}.$$

Let $\mathcal{LT}(p)$ be the group of lower triangular matrices with positive diagonal elements and let $\mathcal{V}_{n,p} = \{H : n \times p \mid H'H = I_p\}$ denote the Stiefel manifold. The Gram-Schmidt orthonormalization of the columns of X leads to

$$X = HT',$$

where $H \in \mathcal{V}_{n,p}$ and $T = (t_{ij}) \in \mathcal{LT}(p)$.

In this section we shall call this decomposition "HT decomposition" for convenience. Define the action of $\mathcal{LT}(p)$ on $\mathcal{X}$ by

$$X \mapsto XT', \qquad T \in \mathcal{LT}(p). \tag{48}$$

In order to make the action of $\mathcal{LT}(p)$ from the left, we are writing the lower triangular matrix T transposed.

In (48) each row of X is multiplied from the right by T'. Therefore $d(XT') = (\det T)^n dX$, $dX = \bigwedge_{i,j} dx_{ij}$, and the multiplier of the Lebesgue measure is

$$\chi(T) = (\det T)^n = \prod_{i=1}^{p} t_{ii}^n.$$

We can use elements of T themselves as global coordinates of $\mathcal{LT}(p)$. For $S = (s_{ij}) \in \mathcal{LT}(p)$, $d(ST) = \prod_{i=1}^{p} s_{ii}^i dT$ where $dT = \bigwedge_{i \geq j} dt_{ij}$. Therefore

$$\prod_{i=1}^{p} t_{ii}^{n-i} dT \tag{49}$$

is a relatively invariant measure on $\mathcal{LT}(p)$ with multiplier χ (see Muirhead (1982) for example) and we use this as the standard relatively invariant measure. Furthermore $d(TS) = \prod_{i=1}^{p} s_{ii}^{p-i+1} dT$ and the right hand modulus of the relatively invariant measure (49) is

$$\Delta(T) = \prod_{i=1}^{p} t_{ii}^{n+p-2i+1}$$

(Section 1.4 of Eaton (1989)).

Let the HT decomposition of X be $X = HT'$. Since the Gram-Schmidt orthonormalization process is uniquely defined, H and T are functions of X : $X = H(X)T(X)'$. It is obvious that $T(X)$ is an equivariant function and $H(X)$ is an invariant function. Also $\mathcal{V}_{n,p}$ is a cross section with respect to the action of $\mathcal{LT}(p)$. We consider $\mathcal{V}_{n,p}$ as the standard cross section.

If X is distributed according to the normal distribution $N_{np}(0, I_n \otimes \Sigma)$, i.e. if the rows of X are i.i.d. multivariate normal $N_p(0, \Sigma)$, then the density of X is written as

$$\begin{aligned} f(X) &= \frac{1}{(2\pi)^{np/2}(\det\Sigma)^{n/2}} \exp\Big(-\frac{1}{2}\operatorname{tr}\Sigma^{-1}X'X\Big) \\ &= \frac{1}{(2\pi)^{np/2}(\det\Sigma)^{n/2}} \exp\Big(-\frac{1}{2}\operatorname{tr}\Sigma^{-1}TT'\Big). \end{aligned}$$

Under normality $T(X)$ and $H(X)$ are independent and $H(X)$ has the orthogonally invariant probability measure on the Stiefel manifold $\mathcal{V}_{n,p}$. Note that the distribution of $H(X)$ does not depend on Σ. The multivariate normal distribution can be generalized by considering density of the form

$$f(X) = h(T(X)).$$

This is a matrix version of elliptically contoured distribution. See Lemma 5.2.4 of Farrell (1985), Example 4.4 of Eaton (1989), Section 3.3 of Kariya and Sinha (1989), and Takemura (1993). For this class of distributions the distribution of $H(X)$ is the same as the multivariate normal distribution. Actually this is a special case of our Theorem 2.1.

Now consider general cross sections. Let $S(H)$ be a function from $\mathcal{V}_{n,p}$ to $\mathcal{LT}(p)$ and assume that $S(H)$ is piecewise of class C^1. Define

$$\mathcal{Z} = \{HS(H)' \mid H \in \mathcal{V}_{n,p}\}, \qquad Z(X) = H(X)S(H(X))'.$$

Then from (13) of Section 2.3 we see that $\mathcal{Z}$ is a cross section. The associated equivariant function is

$$R(X) = T(X)S(H(X))^{-1}.$$

Let X have the density $f(X)$ of the form

$$f(X) = h(R(X)). \tag{50}$$

Then $R(X) = (r_{ij})$ and $Z(X)$ are independently distributed and the joint distribution can be written as

$$(1/c_0)\, h(R) \prod_{i=1}^{p} r_{ii}^{n-i}\, dR\, \pi_Z(dZ),$$

where

$$c_0 = \int_{\mathcal{LT}(n)} h(R) \prod_{i=1}^{p} r_{ii}^{n-i} dR$$

and π_Z is a probability measure on $\mathcal{Z}$.

For the HT decomposition the factorization of the Lebesgue measure is given as follows. Let $h_1, \ldots, h_p$ denote the columns of H . Choose $h_{p+1}, \ldots, h_n$ so that $\{h_1, \ldots, h_n\}$ is an orthonormal basis of R^n . Define an orthogonally invariant $np - p(p+1)/2$ -form on $\mathcal{V}_{n,p}$ by

$$(H'dH) = \bigwedge_{i=1}^{p} \bigwedge_{j=i+1}^{n} h_j' dh_i.$$

Then dX is decomposed as

$$dX = \prod_{i=1}^{p} t_{ii}^{n-i} dT \cdot (H'dH) \tag{51}$$

(Theorem 2.1.13 of Muirhead (1982)). Noting $R(H) = S(H)^{-1}$, Theorem 2.3 shows that the distribution of $H(X)$ under the density (50) is given as

$$c_0 \prod_{i=1}^{p} s(H)_{ii}^{n+p-2i+1} (H'dH).$$

We now investigate the factorization of the Lebesgue measure dX from the viewpoint of Section 2.4. It turns out that the Lebesgue measure of the orbit manifold $\mathcal{M}_O(X)$ is relatively invariant with multiplier $\chi_O(T) = \prod_{i=1}^{p} t_{ii}^{i}$. However the Lebesgue measure of the cross section manifold $\mathcal{M}_C(X)$ is not relatively invariant. We shall show this by investigating the factorization of the Lebesgue measure for the case of the standard HT decomposition. We also make comments on some ambiguous statements in existing literature concerning the volume element of the Stiefel manifold.

Consider a point $X = H_0$ where H_0 is a particular point in $\mathcal{V}_{n,p}$. Let $h_1, \ldots, h_p$ be the columns of H_0 and choose $h_{p+1}, \ldots, h_n$ to form an orthonormal basis of R^n . Let δ_{ij} denote the Kronecker's delta and E_{ij} denote the matrix with 1 in (i,j) position and 0 everywhere else. Let $u_{ij},\ i \geq j,$ be local coordinates of $\mathcal{LT}(p)$ around the identity element $I \in \mathcal{LT}(p)$, i.e.,

$$I + U = (\delta_{ij} + u_{ij}) \in \mathcal{LT}(p)$$

is a lower triangular matrix in the neighborhood of I . Then $X = H_0(I+U)'$ is in the neighborhood of H_0 on the orbit manifold $\mathcal{M}_O(H_0)$. Now take the partial derivative of X with respect to u_{ij} , then

$$\frac{\partial}{\partial u_{ij}} (H_0(I+U)')\Big|_{U=0} = H_0 E_{ji} = (0, \ldots, 0, \underset{i-th}{h_j}, 0, \ldots, 0),$$

which is a matrix with h_j as the i -th column. We denote matrix of this form by $H(i;j),\ 1 \leq i \leq p,\ 1 \leq j \leq n$. It is easily seen that

$$\langle H(i;j), H(i';j') \rangle = \operatorname{tr} H(i;j)' H(i';j') = \delta_{ii'} \delta_{jj'}, \qquad 1 \leq i, i' \leq p,\ 1 \leq j, j' \leq n.$$

Hence $p(p+1)/2$ "vectors"

$$H(i;j), \qquad 1 \le j \le i \le p, \tag{52}$$

form a set of orthonormal vectors of $T_{H_0}(\mathcal{M}_O)$. Similarly $np - p(p+1)/2$ vectors

$$H(i;j), \qquad i < j,\ 1 \le i \le p,\ 2 \le j \le n,$$

form a set of orthonormal vectors of the orthogonal complement $T_{H_0}(\mathcal{M}_O)^{\perp}$ of $T_{H_0}(\mathcal{M}_O)$.

Now take the differential of $H'H = I_p$ at $H = H_0$. Then $H_0'dH + dH'H_0 = 0$, i.e. $H_0'dH$ is skew symmetric. Using

$$H_0'H(i;j) = \begin{cases} E_{ji}, & \text{if } j \le p\,, \\ 0, & \text{if } j > p\,, \end{cases}$$

we see that

$$\sum_{\substack{1\le i\le p\\ 1\le j\le n}} a_{ij}H(i;j)$$

belongs to $T_{H_0}(\mathcal{M}_C)$ iff

$$H_0' \sum_{\substack{1\le i\le p\\ 1\le j\le n}} a_{ij}H(i;j) = \sum_{i,j=1}^{p} a_{ij}E_{ji}$$

is skew symmetric, i.e.,

$$a_{ij} = -a_{ji}, \quad j \le p.$$

Other a_{ij}'s, i.e., a_{ij} for $j > p$, are arbitrary. Therefore as orthonormal basis of $T_{H_0}(\mathcal{M}_C)$ we can take the following $np - p(p+1)/2$ vectors:

$$\begin{aligned} &\frac{1}{\sqrt{2}}(H(i;j) - H(j;i)), \qquad 1 \le j < i \le p, \\ &H(i;j), \qquad 1 \le i \le p,\ p < j \le n. \end{aligned} \tag{53}$$

Note that (52) and (53) are mutually orthonormal except for the pairs $(H(i;j), (H(i;j) - H(j;i))/\sqrt{2})$, $1 \le j < i \le p$, and that the angle between $H(i;j)$ and $(H(i;j) - H(j;i))/\sqrt{2}$ is $\pi/4$. We see that among the canonical angles between $T_{H_0}(\mathcal{M}_O)$ and $T_{H_0}(\mathcal{M}_C)$, $p(p-1)/2$ angles are $\pi/4$ and the other $\min(p, (n-p)p)$ angles are $\pi/2$, i.e., orthogonal. Note that $dU = \bigwedge_{i\ge j} du_{ij}$ is the volume element of $T_{H_0}(\mathcal{M}_O)$ because of the orthonormality in (52). Therefore comparing (51) and (36) dX around $X = H_0$ can be written as

$$dX = dU \cdot (\sqrt{2})^{p(p-1)/2}(H'dH) \cdot \Big(\frac{1}{\sqrt{2}}\Big)^{p(p-1)/2}$$

where $(\sqrt{2})^{p(p-1)/2}(H'dH) = 2^{p(p-1)/4}(H'dH)$ is the volume element of the Stiefel manifold $\mathcal{V}_{n,p}$.

We now investigate the effect of translation by $T = T_0 \in \mathcal{LT}(p)$ on the volume elements. For the orbit manifold we can use the fact that orbit manifold is a relatively open subset of $p(p+1)/2$ dimensional linear subspace in R^{np} and the fact that the elements of T can be used as global coordinates of $\mathcal{LT}(p)$. Consider the elements of T as local coordinates around T_0. Differentiating $X = H_0 T'$ with respect to the t_{ij} we obtain the same orthonormal vectors $H(i;j)$ as tangent vectors of $T_{H_0 T_0}(\mathcal{M}_O)$. Therefore the volume element of $\mathcal{M}_O(H_0 T_0)$ is equal to $dT = \bigwedge_{i \geq j} dt_{ij}$. Let u_{ij} be local coordinates around $T = I$ as above. Then

$$dT = d(T_0 U) = \prod_{i=1}^{p} t_{ii}^{i} \cdot dU, \qquad T_0 = (t_{ij}),$$

and χ_O in (38) is given as

$$\chi_O(T) = \prod_{i=1}^{p} t_{ii}^{i}. \tag{54}$$

In this case we see that the volume element of the orbit manifold is relatively invariant with multiplier χ_O. Also note that dT is a relatively invariant measure on $\mathcal{LT}(p)$ itself with the multiplier (54). Therefore dT can be considered both as the relatively invariant measure on $\mathcal{LT}(p)$ and as the volume element on the orbit manifold with the same multiplier for the case of HT decomposition.

On the other hand the volume element of the cross section manifold is not relatively invariant. We shall show this by investigating the canonical angles between the orbit and the cross section manifolds. Suppose that χ_C in (38) is a homomorphism from $\mathcal{G}$ to R_+^* in addition to χ_O. Then the third term on the right hand side of (37) has to be a homomorphism as well. However the third term corresponds to the canonical angles between $\mathcal{M}_O$ and $\mathcal{M}_C$ and it is bounded. Note that bounded homomorphism from $\mathcal{G}$ to R_+^* has to be identically equal to 1. Therefore if χ_C is a homomorphism in addition to χ_O, then the third term has to be 1. Now for the special case of $p = 2$, it is easily checked that the only non-orthogonal canonical angle θ at $X = HT'$ satisfies

$$\sin^2 \theta = \frac{t_{11}^2}{t_{11}^2 + t_{21}^2 + t_{22}^2},$$

which is not constant and the third term on the right hand side of (37) is not identically equal to 1.

We have investigated the standard cross section $\mathcal{V}_{n,p}$ in detail. However the argument can be equally applied to other cross sections $\mathcal{Z}$. Note that the action of $\mathcal{G}$ on the orbit manifold $\mathcal{M}_O(z_0)$ and the volume element on $\mathcal{M}_O(z_0)$ do not depend on the choice of cross section. Therefore for any cross section $\mathcal{Z}$, the volume element of $\mathcal{M}_O$ is relatively invariant with the multiplier χ_O. We just have to take into account that the volume element of $\mathcal{M}_O(Z)$ at the unit cross section $\mathcal{Z}$ has to be multiplied by the multiplier $\chi_O(S) = \prod_{i=1}^{p} s_{ii}^{i}$ where $S = (s_{ij}) \in \mathcal{LT}(p)$ is defined by $Z = HS'$.

Now we summarize results of this section in the following theorem.

Theorem 4.1 *Let a random $n \times p$ matrix X have the density (with respect to the Lebesgue measure) of the form $h(R(X))$, where $R(X)$ is an equivariant function from $\mathcal{X}$ to $\mathcal{LT}(p)$. Let $Z(X) = XR(X)'^{-1}$. Let $X = HT$ be the HT decomposition of X*

and define $S(Z) \in \mathcal{LT}(p)$ *by* $Z = HS(Z)'$. *Then* $R(X)$ *and* $Z(X)$ *are independent and their joint distribution is given as*

$$(1/c_0)\, h(R) \prod_{i=1}^{p} r_{ii}^{n-i}\, dR\, \pi_Z(dZ) = (1/c_0)\, h(R) \prod_{i=1}^{p} s(Z)_{ii}^{i}\, dR \cdot \prod_{i=1}^{p} r_{ii}^{n-i} s(Z)_{ii}^{-i} \pi_Z(dZ),$$

where

$$c_0 = \int_{\mathcal{LT}(p)} h(R) \prod_{i=1}^{p} r_{ii}^{n-i} dR,$$

π_Z *is a probability measure on* $\mathcal{Z}$, *and* $\prod_{i=1}^{p} s(Z)_{ii}^{i}\, dR$ *is the volume element of* $\mathcal{M}_O(X)$ *at* $X = RZ$. *Furthermore the distribution of* H *is given as*

$$c_0 \prod_{i=1}^{p} s(H)_{ii}^{n+p-2i+1} (H'dH).$$

Concerning Corollary 2.2 we see that its conditions are satisfied and any positive density $f(H)$ on $\mathcal{V}_{n,p}$ with respect to $(H'dH)$ can be obtained from a density of the form $h(R(X))$. However in this case Δ is not one-to-one and $R(X)$ is not uniquely determined from $f(H)$.

Remark 4.1 *As shown above the volume element of the Stiefel manifold* $\mathcal{V}_{n,p}$ *is given by* $2^{p(p-1)/4}(H'dH)$. *The statements in Section 2.1.4 of Muirhead (1982) and Section 7.8 of Farrell (1985) are ambiguous about the factor* $2^{p(p-1)/4}$ *and seem to claim that* $(H'dH)$ *itself is the volume element. For example consider the simplest case of the group* $\mathcal{G} = \mathcal{O}^+(2)$ *of rotations in* R^2. *Let* $H \in \mathcal{O}^+(2)$ *be written as*

$$H = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Regarded as a vector in R^4, *the elements of* $\mathcal{G}$ *form a circle of radius* $\sqrt{2}$ *and the total length of the circle is* $2\sqrt{2}\pi$. *On the other hand* $\int_{\mathcal{O}^+(2)} (H'dH) = \int_0^{2\pi} d\theta = 2\pi$ *gives a wrong answer.*

4.2 Generalization of matrix beta distribution

In this section we briefly discuss how our general theory can be applied for generalization of matrix beta distributions. The group we consider is again the lower triangular group $\mathcal{G} = \mathcal{LT}(p)$. Let W_1, W_2 be $p \times p$ positive definite matrices and define the action of $\mathcal{G}$ by

$$(W_1, W_2) \mapsto (TW_1T', TW_2T').$$

As a dominating measure we consider a measure of the following form

$$(\det W_1)^{a-(p+1)/2} (\det W_2)^{b-(p+1)/2} dW_1 dW_2$$

where $a, b > (p-1)/2$ and $W_1 = (w_{1,ij})$, $W_2 = (w_{2,ij})$, $dW_1 = \bigwedge_{i \ge j} dw_{1,ij}$, $dW_2 = \bigwedge_{i \ge j} dw_{2,ij}$. This measure is relatively invariant with multiplier

$$\chi(T) = (\det T)^{2(a+b)} = \prod_{i=1}^{p} t_{ii}^{2(a+b)}.$$

The relatively invariant measure on $\mathcal{LT}(p)$ with the same multiplier is

$$\prod_{i=1}^{p} t_{ii}^{2(a+b)-i} dT$$

and its right hand modulus is

$$\Delta(T) = \prod_{i=1}^{p} t_{ii}^{2(a+b)+p-2i+1}.$$

It is interesting to note that there are two common cross sections used in literature. Let $TT' = W_1 + W_2$ be the Cholesky decomposition of $W_1 + W_2$. Then T itself is the equivariant function and $U = T^{-1}W_1T'^{-1}$ is the invariant function. If W_1 and W_2 are independent Wishart matrices, then U has the matrix beta distribution. On the other hand let $TT' = W_2$ be the Cholesky decomposition of W_2. Then the invariant $F = T^{-1}W_1T'^{-1}$ has the matrix F distribution (Dawid (1981), Section 5 of Farrell (1985)). This is a matrix version of Example 2.1.

Here we prefer to use the beta type cross section and consider the Cholesky decomposition of $W_1 + W_2$. Write

$$\mathcal{U} = \{(U, I - U) \mid U : p \times p,\ 0 < U < I\}$$

where $<$ means the Löwner order. For convenience write $W = (W_1, W_2)$ and

$$W = (TUT', T(I-U)T'), \qquad T = T(W),\ U = U(W).$$

Let $S(U)$ be a function from $\mathcal{U}$ to $\mathcal{LT}(p)$. Let

$$\mathcal{Z} = \{\ (S(U)US(U)', S(U)(I-U)S(U)') \mid U \in \mathcal{U}\ \},$$
$$Z(W) = (S(U(W))U(W)S(U(W))', S(U(W))(I-U(W))S(U(W))')$$

be a cross section with the associated equivariant function

$$R(W) = T(W)S(U(W))^{-1}.$$

Now the application of Theorems 2.1 and 2.3 gives the following result.

Theorem 4.2 *Suppose that the distribution of $W = (W_1, W_2)$ is given as*

$$h(R(W))(\det W_1)^{a-(p+1)/2}(\det W_2)^{b-(p+1)/2} dW_1 dW_2.$$

Then $R(W)$ and $Z(W)$ are independent and their joint distribution is given as

$$(1/c_0)\, h(R) \prod_{i=1}^{p} r_{ii}^{2(a+b)-i}\, dR\, \pi_Z(dZ), \tag{55}$$

where

$$c_0 = \int_{\mathcal{LT}(n)} h(R) \prod_{i=1}^{p} r_{ii}^{2(a+b)-i} dR,$$

and π_Z is a probability measure on $\mathcal{Z}$. Furthermore the distribution of $U = (u_{ij})$ is given as

$$c_0 \prod_{i=1}^{p} s(U)_{ii}^{2(a+b)+p-2i+1} (\det U)^{a-(p+1)/2} (\det(I-U))^{b-(p+1)/2} dU$$

where $dU = \bigwedge_{i \geq j} du_{ij}$.

Unfortunately the volume elements of both orbit manifold and cross section manifold do not seem to be relatively invariant with respect to the group action and interpretation of the terms of (55) from the viewpoint of Section 2.4 does not seem feasible.

5 Appendix

Here we complete the proof of Theorem 2.1 for the case of h with possibly smaller support.

Even when the support of h is not the whole G, the factorization (10) holds. Our difficulty is that outside the support of h we can not divide by $h(r)$ and π_r is not uniquely determined.

Based on h we will construct an alternative $\tilde{h}$ such that $\tilde{h}$ is everywhere positive on G and $\tilde{h}(r(x))$ defines a cross sectionally contoured distribution. If we construct such a $\tilde{h}$, the proof in Section 2.2 applies with $h = \tilde{h}$ and the proof of Theorem 2.1 is complete.

Let π_r be defined by (12). We argue that there exists an open subset A of G such that

$$0 < \pi_r(A) < \infty. \tag{56}$$

This can be shown as follows. Note that

$$1 = \int_{\mathcal{X}} h(r(x))\pi(dx) = \int_G h(r)\pi_r(dr)$$

implies $\pi_r(\{r \mid h(r) > 0\}) > 0$. For $b > 0$ write

$$\{r \mid h(r) \geq b\} = A_b.$$

Then by the continuity of measures there exists some $b_0 > 0$ such that $\pi_r(A_{b_0}) > 0$. On the other hand

$$1 \geq \int_{A_{b_0}} h(r)\pi_r(dr) \geq b_0 \pi_r(A_{b_0}).$$

It follows that $\pi_r(A_{b_0})$ is finite. Since A_{b_0} is a Borel subset of G there exists a sequence of open subsets $A_n \supset A_{b_0}$ of G such that

$$\pi_r(A_n) \downarrow \pi_r(A_{b_0}),$$

as $n \to \infty$. We see that there exists some n such that $A_n = A$ satisfies (56). Let $c = \pi_r(A)$ for this A.

Since G is second countable there exists a countable dense set $\{g_1, g_2, \ldots\}$ of G. Let

$$\tilde{h}(r) = \frac{1}{c} \sum_{k=1}^{\infty} \frac{1}{\chi(g_k)2^k} I_{g_k A}(r)$$

where $I_{g_k A}$ is the indicator function of $g_k A$. Fix arbitrary $g \in G$. Then the set $\{g_1^{-1}g, g_2^{-1}g, \ldots\}$ is dense in G. Therefore for some k, $g_k^{-1}g \in A$ or $g \in g_k A$. It follows that $\tilde{h}(r)$ is everywhere positive. Furthermore by construction

$$\int_{\mathcal{X}} \tilde{h}(r(x))\pi(dx) = \int_G \tilde{h}(r)\pi_r(dr) = 1.$$

Therefore $\tilde{h}(r(x))$ defines a cross sectionally contoured distribution. This completes the proof of Theorem 2.1.

∎

References

[1] Anderson, T.W. (1993). Nonnormal multivariate distributions: inference based on elliptically contoured distributions. in *Multivariate Analysis: Future Directions* (Editor C.R. Rao), 1–24, North-Holland, Amsterdam.

[2] Barndorff-Nielsen, O.E., Blæsild, P., and Eriksen, P.S. (1989). *Decomposition and Invariance of Measures, and Statistical Transformation Models.* Lecture Notes in Statistics, Vol. 58, Springer, New York.

[3] Dawid, A.P. (1981). Some matrix-variate distribution theory: Notational considerations and a Bayesian application. *Biometrika*, **68**, 265–274.

[4] Eaton, M.L. (1983). *Multivariate Statistics: A Vector Space Approach.* Wiley, New York.

[5] Eaton, M.L. (1989). *Group Invariance Applications in Statistics.* Regional Conference Series in Probability and Statistics, Vol. 1, Institute of Mathematical Statistics and American Statistical Association, Hayward and Alexandria.

[6] Fang, K.T. and Anderson, T.W. (1990). *Statistical Inference in Elliptically Contoured and Related Distributions* (Editors K.T. Fang and T.W. Anderson), Allerton Press, New York.

[7] Fang, K.T. and Fang, B.Q. (1987). Some families of multivariate symmetric distributions related to exponential distribution. *J. Multivariate Anal.*, **24**, 109–122.

[8] Fang, K.T. and Zhang, Y.T. (1990). *Generalized Multivariate Analysis.* Springer and Science Press, Berlin and Beijing.

[9] Farrell, R.H. (1985). *Multivariate Calculation: Use of the Continuous Groups.* Springer, New York.

[10] Gupta, A.K. and Varga, T. (1993). *Elliptically Contoured Models in Statistics.* Kluwer Academic, Dordrecht.

[11] Kamiya, H. and Takemura, A. (1996). Cross sectionally contoured distributions for non-free actions. In Preparation.

[12] Kariya, T. and Eaton, M.L. (1977). Robust tests for spherical symmetry, *Ann. Statist.*, **5**, 206–215.

[13] Kariya, T. and Sinha, B.K. (1989). *Robustness of Statistical Tests*, Academic Press, Boston.

[14] Kent, J.T. and Tyler, D.E. (1988). Maximum likelihood estimation for the wrapped Cauchy distribution. *J. Appl. Statist.*, **15**, 247–254.

[15] King, M.L. (1980). Robust tests for spherical symmetry and their application to least squares regression. *Ann. Statist.*, **8**, 1265–1271.

[16] Mardia, K.V. (1972). *Statistics of Directional Data.* Academic Press, London.

[17] Muirhead, R.J. (1982). *Aspects of Multivariate Statistical Theory.* Wiley, New York.

[18] Naiman, D.Q. and Wynn, H.P. (1992). Inclusion-exclusion-Bonferroni identities and inequalities for discrete tube-like problems via Euler characteristics, *Ann. Statist.*, **20**, 43–76.

[19] Takemura, A. (1993). Maximally orthogonally invariant higher order moments and their application to testing elliptically-contouredness. in *Statistical Science & Data Analysis* (Editors K. Matsushita et al.), 225–235, VSP, Utrecht.

[20] Tyler, D.E. (1987). Statistical analysis for the angular central Gaussian distribution on the sphere. *Biometrika*, **74**, 579–589.

[21] Watson, G.S. (1983). *Statistics on Spheres.* Wiley, New York.

[22] Wijsman, R.A. (1986). Global cross sections as a tool for factorization of measures and distribution of maximal invariants. *Sankhyā, Ser. A*, **48**, 1–42.

[23] Wijsman, R.A. (1990). *Invariant Measures on Groups and Their Use in Statistics.* Lecture Notes–Monograph Series, Vol. 14, Institute of Mathematical Statistics, Hayward.

[24] Ziegler, G.M. (1995). *Lectures on Polytopes.* Springer, New York.