A Characterization of Spherical Distributions

MORRIS L. EATON*

University of Minnesota

Communicated by P. R. Krishnaiah

It is shown that when the random vector $X$ in $\mathbb{R}^n$ has a mean and when the conditional expectation $E(u'X|v'X) = 0$ for all vectors $u, v \in \mathbb{R}^n$ which satisfy $u'v = 0$, then the distribution of $X$ is orthogonally invariant. A version of this characterization is also established when $X$ does not have a mean vector.

1. INTRODUCTION

In this paper, spherical distributions on $\mathbb{R}^n$ are characterized in a couple of different ways. These characterizations arose in part from assumptions concerning error distributions in some linear model problems—see Toyooka [7], Kariya and Toyooka [5], and Eaton [3]. A discussion of this is given in Section 3.

To describe the main results here, recall that a random vector $X \in \mathbb{R}^n$ has a spherical distribution if $L(X) = L(\Gamma X)$ for all $n \times n$ orthogonal matrices $\Gamma$. Here, vectors in $\mathbb{R}^n$ are written as columns and $L(X)$ denotes the distribution of $X$. For many properties of spherical distributions, see Cambanis, Huang, and Simons [1]. One characterization of spherical distributions is the following.

**Theorem 1.** Suppose the random vector $X \in \mathbb{R}^n$ has a mean vector. Assume that for each vector $v \neq 0$ and for each vector $u$ which is perpendicular to $v$ (that is, $u'v = 0$),

$$E(u'X|v'X) = 0.$$  \hspace{1cm} (1.1)

Then $X$ is spherical and conversely.
When $X$ does not have a mean vector, an alternative characterization is possible.

**Theorem 2.** Consider a random vector $X \in \mathbb{R}^n$. Assume that for each vector $v \neq 0$ and each $u$ which is perpendicular to $v$,

$$
L(u'X | v'X) = L(-u'X | v'X).
$$

(1.2)

Then $X$ is spherical and conversely.

The notation in (1.2) means the conditional distribution of $u'X$ given $v'X$ is the same as the conditional distribution of $-u'X$ given $v'X$. A slightly different version of Theorem 2 is described in Section 3. When $X$ has a mean, then of course, (1.2) implies (1.1).

The results of Theorem 1 should be compared to characterizations of spherical distributions based on linear regression which appeared in Nimmo-Smith [6] and Hardin [4]. Also, the work of Vershik [8] is relevant.

2. Technical Details

Given a random vector $X$ in $\mathbb{R}^n$,

$$
g(t) = \mathbb{E} \exp[it'X]
$$

(2.1)

is the characteristic function of $X$. Obviously, $X$ is spherical iff

$$
g(t) = g(\Gamma t)
$$

(2.2)

for all $\Gamma$ in the orthogonal group $O_n$.

**Proof of Theorem 1.** Since $X$ has a mean vector, the gradient of $g$ exists and is given by

$$
\nabla g(t) - i\mathbb{E}X \exp[it'X].
$$

(2.3)

When (1.1) holds, an easy conditioning argument yields

$$
\mathbb{E}\{u'X \exp[iu'X]\} = 0
$$

(2.4)

for all $u$ perpendicular to $v \neq 0$. But (2.4) is equivalent to

$$
u' \nabla g(v) = 0.
$$

(2.5)

To show $X$ is spherical, (2.5) is now used to verify (2.2) when $t \neq 0$. Since $t$ and $\Gamma t$ have the same length, say $\|t\| = r$, there exists a smooth curve $c$
mapping \((0, 1)\) into \(\{ x \mid \| x \| = r \}\) such that \(c(x_1) = t\) and \(c(x_2) = \Gamma t\) for some \(x_1, x_2 \in (0, 1)\). Since \(\| c(x) \|^2 = r^2\) for \(x \in (0, 1)\)

\[(\dot{c}(x))'c(x) = 0 \quad \text{for all} \quad x \in (0, 1) \tag{2.6}\]

where \(\dot{c}(x)\) is the vector of derivatives of the curve \(c\). Using (2.5) and (2.6) we have

\[
\frac{d}{dx} g(c(x)) = (\dot{c}(x))' \nabla g(c(x)) = 0
\]

for \(x \in (0, 1)\). Thus \(g(c(x))\) is constant in \(x\) so (2.2) holds and hence \(X\) is spherical.

The converse is well known and a proof can be found in Cambanis, Huang, and Simons [1].

**Proof of Theorem 2.** It is an easy argument which shows that (1.2) is equivalent to

\[
g(au + bv) = g(-au + bv) \tag{2.7}
\]

for all \(a, b \in \mathbb{R}^1\) and \(u\) which are perpendicular to \(v \neq 0\). Continuity shows (2.7) holds when \(v = 0\). To verify (2.2), consider \(t\) and \(\Gamma t\), and set \(v = \frac{1}{2}(\Gamma t + t)\). With \(u = \frac{1}{2}(\Gamma t - t)\), \(u'v = 0\) and

\[
u + v = \Gamma t, \quad -u + v = t.
\]

Thus, (2.7) yields (2.2) so \(X\) is spherical. Again, the converse is trivial.

### 3. A Linear Model Application

The linear model problem which gave rise to Theorems 1 and 2 is the following: Consider a linear model on \(\mathbb{R}^n\), \(Y = \mu + \varepsilon\), where \(\mu\) is in a known linear subspace \(M\) and the error vector \(\varepsilon\) has a mean of zero and a covariance \(\Sigma\) which belongs to a known set \(\gamma\) of positive definite matrices. When \(\Sigma\) is known, the Gauss–Markov estimator of \(\mu\) is

\[
\hat{\mu} = P_X Y \tag{3.1}
\]

where \(P_X\) is the projection onto \(M\) whose null space is \(\Sigma(M^\perp)\) (see Eaton [2, 3] for a discussion). It is often the case that \(\Sigma\) is not known and must be estimated from the data \(Y\). Typically such estimators \(\tilde{\Sigma}\) satisfy

(i) \(\tilde{\Sigma}(y) = \tilde{\Sigma}(y + x)\) for \(y \in \mathbb{R}^n, x \in M\),

(ii) \(\tilde{\Sigma}(-y) = \tilde{\Sigma}(y)\) for \(y \in \mathbb{R}^n\),
and such estimators are called *residual type estimators* in Eaton [3]. A common statistical method is to use

$$\tilde{\mu} = \bar{P} Y$$  \hspace{1cm} (3.2)

as an estimator of \( \mu \), where

$$\bar{P} = P_\Sigma$$ \hspace{1cm} (3.3)

and \( \Sigma \) is a residual-type estimator. When \( L(\varepsilon) = L(-\varepsilon) \), it is easy to show that

$$L(\tilde{\mu} - \mu) - L(\mu)$$

so when \( \tilde{\mu} \) has an expectation (which it may not), then \( E\tilde{\mu} = \mu \). When \( \tilde{\mu} \) has a covariance, the usual Gauss–Markov theorem suggests that

$$\text{Cov}(\tilde{\mu}) \leq \text{Cov}(\mu)$$ \hspace{1cm} (3.4)

where \( \text{Cov}(\cdot) \) denotes covariance matrix; that is, \( \text{Cov}(\tilde{\mu}) - \text{Cov}(\mu) \) is non-negative definite.

It is shown in Eaton [3] that a sufficient condition for (3.4) to hold is that

$$E(P_\Sigma \varepsilon | Q_\Sigma \varepsilon) = 0,$$ \hspace{1cm} (3.5)

where \( Q_\Sigma = I - P_\Sigma \). Set \( X = \Sigma^{-1/2} \varepsilon, P = \Sigma^{-1/2} P_\Sigma \Sigma^{1/2} \), and \( Q = I - P \). Since \( \Sigma \) is non-singular by assumption, (3.5) is equivalent to

$$E(PX | QX) = 0.$$ \hspace{1cm} (3.6)

It is easily verified that \( P \) is an orthogonal projection of rank \( k \) equal to the dimension of the regression subspace \( M \). That (3.6) holds for spherical \( X \)'s which have a mean is well known.

Now, assume that (3.6) holds for all rank \( k \) orthogonal projections \( P \). The claim is that Theorem 1.1 is applicable so that \( X \) is spherical. To see this, consider \( v \neq 0 \) and \( u \) which is perpendicular to \( v \). Pick a projection \( P \) of rank \( k \) such that \( Pu = u \) and \( Qu = v \). Since \( v'X = v'QX \), we can write

$$E(u'X | v'X) = E\{ (u'PX | v'QX) \} = E(u'E(PX | v'QX) = u'E[ E(PX | QX) | v'QX]$$

which is zero by (3.6). Thus \( X \) is spherical so \( \varepsilon = \Sigma^{1/2}X \) is elliptical by definition (\( \varepsilon \) is *elliptical* if it is a linear transformation of a spherical random vector).
The implication of the above argument is that a natural sufficient condition for a non-linear version of the Gauss–Markov theorem to hold is (3.5). But if (3.5) holds for all regression subspaces of a fixed dimension, then the error vector must be elliptical.

ACKNOWLEDGMENT

I would like thank one of the referees whose comments lead to a better presentation of this material.

REFERENCES