The Annals of Statistics
1984, Vol. 12, No. 3, 1050–1057

SYMMETRIC DISTRIBUTIONS FOR DEPENDENT UNIT VECTORS[1]

BY LOUIS-PAUL RIVEST

Université Laval

This paper introduces several notions of symmetry for the joint distribution of two dependent unit vectors. Bivariate generalizations of $\mathscr{Q}$-symmetry (Rivest, 1984) and rotational symmetry are introduced. If the joint distribution of two unit vectors is at least $\mathscr{Q}$-symmetric the information matrix for the parameters indexing it is shown to have a simple shape.

1. Introduction. All the standard distributions for a random unit vector belonging to S_k, the unit sphere in R^k, are to different extents symmetric. Rivest (1984) expressed these symmetries in terms of the invariance of the distribution with respect to a subgroup of $O(k)$, the group of orthogonal transformations in R^k. The purpose of this work is to introduce various notions of symmetry for bivariate directional distributions and to study their statistical implications.

Section 2 gives a general definition of symmetry for the joint distribution of $(\mathbf{u}, \mathbf{v})$, two random vectors belonging to S_k. Definitions of bivariate rotational symmetry and bivariate $\mathscr{Q}$-symmetry (Rivest, 1984) are presented. When the underlying distribution is $\mathscr{Q}$-symmetric, a simple expression is derived for ρ^2, Jupp and Mardia (1980) coefficient of correlation defined as the sum of the canonical correlations:

$$\rho^2 = \operatorname{tr}(\textstyle\sum_{12} \sum_{22}^{-1} \sum_{21} \sum_{11}^{-1})$$

where $\sum_{12} = E(\mathbf{uv}') - E(\mathbf{u})E(\mathbf{v}')$, $\sum_{11} = E(\mathbf{uu}') - E(\mathbf{u})E(\mathbf{u})'$ and $\sum_{22} = E(\mathbf{vv}') - E(\mathbf{v})E(\mathbf{v})'$. The concept of cluster dependence (Rivest, 1982) is given a formal definition.

In Section 3 the information matrix for the parameters indexing a $\mathscr{Q}$-symmetric distribution is shown to have a simple form. As in Rivest (1984), it is made of two blocks: one for the generalized location and one for the shape. If the generalized location of a $\mathscr{Q}$-symmetric distribution is parametrized in terms of skew-symmetric matrices, the corresponding information matrix is shown to be made of four diagonal blocks. For rotationally symmetric models another parametrization is investigated.

2. Bivariate symmetry. This section generalizes the notion of $\mathscr{A}$-symmetry of Rivest (1984) to bivariate distributions. Let $\mathbf{u}$ and $\mathbf{v}$ be random vectors belonging to S_k; let $f(\mathbf{u}, \mathbf{v})$ denote their joint density with respect to the Lebesgue measure on $S_k \times S_k$.

Received October 1983; revised March 1984.

[1] Research supported by NSERC grant No. A5244 and FCAC grant No. EQ1023.

AMS 1980 *subject classifications.* Primary 62F10, 62F12; secondary 62H20.

Key words and phrases. Fisher information matrix, rotational symmetry, directional data.

DEFINITION 1. $\mathscr{A}$-symmetry. The density $f(\mathbf{u}, \mathbf{v})$ is said to be $\mathscr{A}$-symmetric if there exist $\mathbf{P}$ and $\mathbf{Q}$ in $O(k)$ such that $g(\mathbf{r}, \mathbf{s})$ the joint density of $\mathbf{r} = \mathbf{P}'\mathbf{u}$ and $\mathbf{s} = \mathbf{Q}'\mathbf{v}$ satisfies

$$\begin{aligned} &\text{(i) } E(r_1) \quad \text{and} \quad E(s_1) > 0 \\ (1) \quad &\text{(ii) } E(r_2 s_2) \geq E(r_3 s_3) \geq \cdots \geq E(r_k s_k) \geq 0 \\ &\text{(iii) } g(\mathbf{Hr}, \mathbf{Hs}) = g(\mathbf{r}, \mathbf{s}) \quad \text{for any } H \text{ in } \mathscr{A} \text{ a subgroup of } O(k). \end{aligned}$$

Note that since $|\det \mathbf{P}| = |\det \mathbf{Q}| = 1$, $f(\mathbf{u}, \mathbf{v}) = g(\mathbf{P}'\mathbf{u}, \mathbf{Q}'\mathbf{u})$; g can be seen as a standardization of f.

The following properties are easily derived: for all $\mathbf{H}$ in $\mathscr{A}$

(2)
(i) $\mathbf{Hr}$, $\mathbf{Hs}$ and $\mathbf{r}$, $\mathbf{s}$ have the same joint distribution
(ii) the marginal distributions of $\mathbf{r}$(resp., $\mathbf{s}$) and $\mathbf{Hr}$(resp., $\mathbf{Hs}$) are the same.

Property (2) (ii) shows that if the joint density of $\mathbf{u}$, $\mathbf{v}$ is $\mathscr{A}$-symmetric, their marginal densities satisfy conditions analogous to those of univariate $\mathscr{A}$-symmetry (Rivest, 1984). Hence Definition 1 provides a bivariate extension of univariate $\mathscr{A}$-symmetry. As will be shown in Proposition 1, for the following examples the matrices $\mathbf{P}$ and $\mathbf{Q}$ can be obtained from the first two moments of $\mathbf{u}$ and $\mathbf{v}$.

EXAMPLE 1. $O(k)$-symmetry. When $\mathscr{A} = O(k)$ the marginal densities of $\mathbf{u}$ and $\mathbf{v}$ are uniform. For any $\mathscr{A}$ in $O(k)$ such that $\mathbf{Hr} = \mathbf{r}$, $g(\mathbf{r}, \mathbf{s}) = g(\mathbf{Hr}, \mathbf{Hs}) = g(\mathbf{r}, \mathbf{Hs})$; this shows that the conditional density of $\mathbf{s}$ given $\mathbf{r}$ is rotationally symmetric about $\mathbf{r}$. Thus there exists a function h for which

$$g(\mathbf{r}, \mathbf{s}) = h(\mathbf{r}'\mathbf{s}).$$

This is the general form of $O(k)$-symmetric densities. Johnson and Wehrly (1977) presented a class of $O(2)$-symmetric densities. Saw (1983) discussed the properties of an $O(k)$-symmetric density obtained by projecting a suitable $2k$-dimensional normal random vector on $S_k \times S_k$.

Note that there exist densities with uniform marginals that are not $O(k)$-symmetric. If $g(\mathbf{r}, \mathbf{s})$ is proportional to

$$1 +' r_1 s_1/2$$

the marginals of $\mathbf{r}$ and $\mathbf{s}$ are uniform but g is not $O(k)$-symmetric for any $k > 1$.

EXAMPLE 2. Bivariate rotational symmetry. If

$$\mathscr{A} = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & \mathbf{H} \end{pmatrix} : \mathbf{H} \in O(k-1) \right\}$$

one obtains a bivariate generalization of rotational symmetry. Given r_1 and s_1, the conditional density of $(1 - r_1^2)^{-1/2}(r_2, \cdots, r_k)$ and $(1 - s_1^2)^{-1/2}(s_2, \cdots, s_k)'$ is

$O(k-1)$-symmetric. Thus a rotationally symmetric density can always be written as:

$$g(\mathbf{r}, \mathbf{s}) = h(r_1, s_1, \textstyle\sum_2^k r_i s_i). \tag{3}$$

The mean direction mixture model gives a rotationally symmetric density:

$$g(\mathbf{r}, \mathbf{s}) = h(\mathbf{r}'\mathbf{s}) g_1(r_1)$$

the marginal density of $\mathbf{r}$ is rotationally symmetric about $(1, 0, \cdots, 0)'$ while given $\mathbf{r}$, $\mathbf{s}$ has a density which is rotationally symmetric about $\mathbf{r}$. Saw (1983) constructed some rotationally symmetric densities.

EXAMPLE 3. $\mathscr{D}$-symmetry. Let

$$\mathscr{D} = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & \operatorname{diag} \pm 1 \end{pmatrix} \right\}.$$

A $\mathscr{D}$-symmetric density can be written as

$$g(\mathbf{r}, \mathbf{s}) = h(r_1, s_1, r_2^2, \cdots, r_k^2, s_2^2, \cdots, s_k^2, r_2 s_2, \cdots, r_k s_k).$$

This symmetry is weaker than the ones presented in Examples 1 and 2.

NOTATION. Let $\{\mathbf{p}_i\}_{i=1}^k$ and $\{\mathbf{q}_i\}_{i=1}^k$ denote the columns of $\mathbf{P}$ and $\mathbf{Q}$.

A straightforward generalization of Proposition 1 of Rivest (1984) is:

PROPOSITION 1. *If* $(\mathbf{u}, \mathbf{v})$ *has a* $\mathscr{D}$*-symmetric density,*

$$E(\mathbf{u}) = \mathbf{p}_1 E(r_1), \quad E(\mathbf{v}) = \mathbf{q}_1 E(s_1)$$

$$E(\mathbf{u}\mathbf{u}') = \mathbf{P} \operatorname{diag}\{E(r_i^2)\}\mathbf{P}'$$

$$E(\mathbf{v}\mathbf{v}') = \mathbf{Q} \operatorname{diag}\{\mathrm{E}(s_i^2)\}\mathbf{Q}'$$

$$E(\mathbf{u}\mathbf{v}') = \mathbf{P} \operatorname{diag}\{\mathrm{E}(r_i s_i)\}\mathbf{Q}'.$$

This implies that the matrices $\mathbf{P}$ and $\mathbf{Q}$ of Definition 1 can be defined in terms of the moments of $\mathbf{u}$ and $\mathbf{v}$: when $E(r_1)$ and $E(s_1)$ are positive, $\mathbf{p}_1$ and $\mathbf{q}_1$ are the mean directions of $\mathbf{u}$ and $\mathbf{v}$ respectively while $\mathbf{P}$ and $\mathbf{Q}$ are eigenvector matrices of $E(\mathbf{u}\mathbf{u}')$ and $E(\mathbf{v}\mathbf{v}')$ and left and right singular vector matrices of $E(\mathbf{u}\mathbf{v}')$.

Using Proposition 1, ρ^2, Jupp and Mardia coefficient of correlation can be written as

$$\rho^2 = \textstyle\sum_{i=1}^k \operatorname{corr}^2(\mathbf{p}_i'\mathbf{u}, \mathbf{q}_i'\mathbf{v})$$

where corr denotes Pearson's correlation coefficient. When $\rho^2 \neq 0$ Jupp and Mardia following Mackenzie's (1957) proposal suggested predicting $\mathbf{v}$ for a given $\mathbf{u}$ by $\mathbf{QP}'\,\mathbf{u}$. If the distribution of $(\mathbf{u}, \mathbf{v})$ is $O(k)$-symmetric, $\mathbf{QP}'\,\mathbf{u}$ is the mean direction of the conditional distribution of $\mathbf{v}$ given $\mathbf{u}$; it is a good predictor of $\mathbf{v}$

(Rivest, 1982). However there are situations where it is not appropriate:

DEFINITION 2. Cluster Dependence (Rivest, 1982). Two unit vectors $\mathbf{u} = \mathbf{Pr}$ and $\mathbf{v} = \mathbf{Qs}$ having a $\mathscr{Q}$-symmetric density are cluster dependent if given r_1 and s_1, $(1 - r_1^2)^{-1/2}(r_2, \cdots, r_k)$ and $(1 - s_1^2)^{-1/2}(s_2, \cdots, s_k)$ are independent.

A cluster dependent $\mathscr{Q}$-symmetric density can be written as:

$$g(r, s) = h_1(r_1, s_1, r_2^2, \cdots, r_k^2)h_2(r_1, s_1, s_2^2, \cdots, s_k^2).$$

If (u, v) has a cluster dependent $\mathscr{Q}$-symmetric density, given r_1 and s_1, then r_i, s_i and r_is_i have a null expectation for $i > 1$; therefore

$$\rho^2 = \mathrm{corr}^2(r_1, s_1) \quad E(\mathbf{v} \mid \mathbf{u}) = \mathbf{q}_1 E(s_1 \mid r_1).$$

For a cluster dependent distribution $\rho^2 \neq 0$ and the mean direction of $\mathbf{v}$ given $\mathbf{u}$ is $\mathbf{q}_1$ if $E(s_1 \mid r_1) > 0$ and $-\mathbf{q}_1$ if not. In this case $\mathbf{Q}'\mathbf{Pu}$ is a poor predictor of $\mathbf{v}$ unless $\mathbf{q}_1 = \mathbf{p}_1$ and $\mathbf{u}$ is highly concentrated about $\mathbf{p}_1$.

Section 5 of Saw (1983) provides an example of a cluster dependent rotationally symmetric distribution.

3. The information matrix. The information matrix for the parameters indexing a Q-symmetric density is shown to have a simple form. The model is

$$f(\mathbf{u}, \mathbf{v}) = g(\mathbf{P}'\mathbf{u}, \mathbf{Q}'\mathbf{v}; \phi) \tag{4}$$

where ϕ is a shape parameter and $\mathbf{P}$ and $\mathbf{Q}$ belong to $O(k)$.

Let $\mathbf{P}_0$ and $\mathbf{Q}_0$ be the true values of $\mathbf{P}$and $\mathbf{Q}$ respectively. The parameter spaces for $\mathbf{P}$ and $\mathbf{Q}$ are made of two disjoint isomorphic subsets: $O^+(k)$, the set of rotations and $O^-(k)$, the orthogonal transformations whose determinant is -1. Thus $\mathbf{P}$ and $\mathbf{Q}$ can each be parametrized by a rotation and a $0 - 1$ parameter indicating to which subset of $O(k)$ $\mathbf{P}_0'\mathbf{P}$ (resp., $\mathbf{Q}_0'\mathbf{Q}$) belongs. In calculating the information matrix one can ignore the $0 - 1$ parameter and consider only matrices $\mathbf{P}$ and $\mathbf{Q}$ for which $\mathbf{P}_0'\mathbf{P}$ and $\mathbf{Q}_0'\mathbf{Q}$ are rotations.

Rotations can be parametrized in terms of skew-symmetric matrices. If $\mathbf{R}$ is a rotation, there exists a skew-symmetric matrix $\mathbf{A}$ ($\mathbf{A}' = -\mathbf{A}$) such that

$$\mathbf{R} = \exp(\mathbf{A}) = \textstyle\sum_{i=0}^{\infty} \mathbf{A}^i/i!.$$

Straightforward manipulations show that if $\mathbf{A}$ is 2×2

$$\exp(\mathbf{A}) = \begin{pmatrix} \cos A_{21} & -\sin A_{21} \\ \sin A_{21} & \cos A_{21} \end{pmatrix}$$

while if it is 3×3

$$\exp(\mathbf{A}) = \cos \| \mathbf{a} \| \mathbf{I} + \frac{1 - \cos \| \mathbf{a} \|}{\| \mathbf{a} \|^2} \mathbf{aa}' + \frac{\sin \| \mathbf{a} \|}{\| \mathbf{a} \|} \mathbf{A}$$

is a rotation of angle $\| \mathbf{a} \|$ about $\mathbf{a} = (A_{32}, -A_{31}, A_{21})'$. There exist other parametrizations of $\mathbf{R}$ in terms of skew-symmetric matrices; for instance

$$\mathbf{R} = \mathbf{A} + (\mathbf{I} + \mathbf{A}^2)^{1/2} = \mathbf{I} + \mathbf{A} + \textstyle\sum_1^\infty \binom{1/2}{i}\mathbf{A}^{2i}$$

where the last equality holds if the eigenvalues of A^2 are in absolute value less than one. Thus any $(\mathbf{P}, \mathbf{Q})$ in the neighborhood of $(\mathbf{P}_0, \mathbf{Q}_0)$ can be written as:

$$\mathbf{P} = \mathbf{P}_0\{\mathbf{I} + \Theta + O(\Theta^2)\}, \quad \mathbf{Q} = \mathbf{Q}_0\{\mathbf{I} + \Psi + O(\Psi^2)\}$$

where Θ and Ψ are skew-symmetric matrices. Let $\boldsymbol{\theta} = (\Theta_{21}, \Theta_{31}, \cdots, \Theta_{k1}, \Theta_{32}, \cdots, \Theta_{kk-1})$ and $\boldsymbol{\psi} = (\Psi_{21}, \cdots, \Psi_{k1}, \Psi_{32}, \cdots, \Psi_{kk-1})$; note that $(\boldsymbol{\phi}, \boldsymbol{\theta}, \boldsymbol{\psi})$ parametrizes f. Let $\mathbf{U} = \mathbf{U}(\mathbf{P}_0, \mathbf{Q}_0, \boldsymbol{\phi}, \mathbf{u}, \mathbf{v}) = (\mathbf{U}'_\phi, \mathbf{U}'_\theta, \mathbf{U}'_\psi)'$ be the score vector evaluated at $\boldsymbol{\phi}$, $\boldsymbol{\theta} = \boldsymbol{\psi} = \mathbf{0}$:

$$\mathbf{U} = \frac{\partial}{\partial(\boldsymbol{\phi}, \boldsymbol{\theta}, \boldsymbol{\psi})} \ln f(\mathbf{u}, \mathbf{v}) \,|_{\theta=\psi=\mathbf{0}}$$

and $\mathbf{i} = E(\mathbf{U}\mathbf{U}')$ be the Fisher information matrix evaluated at $\boldsymbol{\phi}$, $\boldsymbol{\theta} = \boldsymbol{\psi} = \mathbf{0}$:

$$\mathbf{i} = \begin{pmatrix} \mathbf{i}_{\phi\phi} & i_{\phi\theta} & i_{\phi\psi} \\ \mathbf{i}_{\theta\phi} & \mathbf{i}_{\theta\theta} & \mathbf{i}_{\theta\psi} \\ \mathbf{i}_{\psi\phi} & i_{\psi\theta} & \mathbf{i}_{\psi\psi} \end{pmatrix}$$

where $\mathbf{i}_{\phi\phi} = E(\mathbf{U}_\phi \mathbf{U}'_\phi)$ etc $\cdots$.

The following propositions summarize the properties of $\mathbf{i}$ when f is $\mathscr{Q}$-symmetric. The first one is an easy generalization of Proposition 2 of Rivest (1984).

PROPOSITION 2. *If f is $\mathscr{Q}$-symmetric and if $g(\mathbf{r}, \mathbf{s}; \boldsymbol{\phi})$ is differentiable,*

$$\mathbf{i}_{\phi\theta} = i_{\phi\psi} = \mathbf{0}.$$

PROPOSITION 3. *If f is $\mathscr{Q}$-symmetric and if $g(\mathbf{r}, \mathbf{s}; \boldsymbol{\phi})$ is a differentiable function of $\mathbf{r}$ and $\mathbf{s}$, $\mathbf{i}_{\theta\psi}$, $\mathbf{i}_{\theta\theta}$ and $\mathbf{i}_{\psi\psi}$ are diagonal matrices.*

PROOF. For any $m > 1$ let $\mathbf{H}_m$ be a diagonal matrix of 1 except for the (m, m) entry which is -1. Let

$$\mathbf{g}^{(1)}(\mathbf{r}, \mathbf{s}) = \frac{\partial}{\partial \mathbf{r}} g(\mathbf{r}, \mathbf{s}; \boldsymbol{\phi}), \quad \mathbf{g}^{(2)}(\mathbf{r}, \mathbf{s}) = \frac{\partial}{\partial \mathbf{s}} g(\mathbf{r}, \mathbf{s}; \boldsymbol{\phi}).$$

The assumptions imply that

$$\mathbf{g}^{(i)}(\mathbf{H}_m\mathbf{r}, \mathbf{H}_m\mathbf{s}) = \mathbf{H}_m\mathbf{g}^{(i)}(\mathbf{r}, \mathbf{s}), \quad i = 1, 2. \tag{5}$$

Define $\mathbf{J}_{ij}$ as a $k \times k$ matrix of 0 except for the (i, j) and (j, i) entries which are equal to 1 and -1 respectively. For $i > j$ one can write

$$\begin{aligned} U_{\Theta_{ij}} &= \frac{\partial}{\partial \Theta_{ij}} \ln f(\mathbf{u}, \mathbf{v}) \,|_{\theta=\psi=\mathbf{0}} \\ &= -\{\mathbf{g}^{(1)}(\mathbf{r}, \mathbf{s})\}'\mathbf{J}_{ij}\mathbf{r}/g(\mathbf{r}, \mathbf{s}; \boldsymbol{\phi}) \qquad (6) \\ &= -\{g_i^{(1)}(\mathbf{r}, \mathbf{s})r_j - g_j^{(1)}(\mathbf{r}, \mathbf{s})r_i\}/g(\mathbf{r}, \mathbf{s}; \boldsymbol{\phi}) \end{aligned}$$

where $\mathbf{r} = \mathbf{P}_0'\mathbf{u}$ and $\mathbf{s} = \mathbf{Q}_0'\mathbf{v}$. In a similar way it can be shown that for $i' > j'$

$$U_{\Psi_{i'j'}} = -\{\mathbf{g}^{(2)}(\mathbf{r}, \mathbf{s})\}'\mathbf{J}_{i'j'}\mathbf{s}/g(\mathbf{r}, \mathbf{s}; \phi).$$

Applying (5) and using the fact that

$$\mathbf{H}_m\mathbf{J}_{ij}\mathbf{H}_m = (-1)^{\delta_{i,m}+\delta_{j,m}}\mathbf{J}_{ij}$$

where $\delta_{a,b} = 1$ if $a = b$ and 0 if not:

$$\begin{aligned} -\{\mathbf{g}^{(1)}(\mathbf{H}_m\mathbf{r}, \mathbf{H}_m\mathbf{s})\}'\mathbf{J}_{ij}\mathbf{H}_m\mathbf{r}/g(\mathbf{H}_m\mathbf{r}, \mathbf{H}_m\mathbf{s}; \phi) \\ = -\mathbf{g}^{(1)}(\mathbf{r}, \mathbf{s})'\mathbf{H}_m\mathbf{J}_{ij}\mathbf{H}_m\mathbf{r}/g(\mathbf{r}, \mathbf{s}; \phi) \\ = (-1)^{\delta_{i,m}+\delta_{j,m}}U_{\Theta_{ij}}. \end{aligned}$$

A similar result holds for $U_{\Psi_{i'j'}}$.

Consider

$$i_{\Theta_{ij}\Psi_{i'j'}} = E(U_{\Theta_{ij}}U_{\Psi_{i'j'}})$$

if $(i, j) \neq (i', j')$ let m be an integer, $m > 1$, that is equal to one and only one of $\{i, j, i', j'\}$. Changing variables $\mathbf{r}_1 = \mathbf{H}_m\mathbf{r}$ and $\mathbf{s}_1 = \mathbf{H}_m\mathbf{s}$ in the previous expectation shows that

$$i_{\Theta_{ij}\Psi_{i'j'}} = -i_{\Theta_{ij}\Psi_{i'j'}} = 0.$$

Thus $\mathbf{i}_{\theta\psi}$ is diagonal. The proofs for $\mathbf{i}_{\theta\theta}$ and $\mathbf{i}_{\psi\psi}$ are similar. □

Special cases

(i) If in addition f is *cluster dependent* then

$$\mathbf{i}_{\theta\psi} = 0.$$

PROOF. Cluster dependence implies that for $m > 1$, $\mathbf{H}_m\mathbf{r}$, $\mathbf{s}$ and $\mathbf{r}$, $\mathbf{s}$ have the same distribution; therefore

$$\mathbf{g}^{(1)}(\mathbf{H}_m\mathbf{r}, \mathbf{s}) = \mathbf{H}_m\mathbf{g}^{(1)}(\mathbf{r}, \mathbf{s}).$$

Changing variables $\mathbf{r}_1 = \mathbf{H}_i\mathbf{r}$ and $\mathbf{s}_1 = \mathbf{s}$ in $E(U_{\Theta_{ij}}U_{\Psi_{ij}})$ shows that it is null.

(ii) If in addition f has *rotational symmetry*, for $k \geq i > j > 1$

$$-i_{\Theta_{ij}\Psi_{ij}} = i_{\Theta_{ij}\Theta_{ij}} = i_{\Psi_{ij}\Psi_{ij}}.$$

PROOF. By (3), $g(\mathbf{r}, \mathbf{s}) = h(r_1, s_1, \sum_2^k r_i s_i)$, let $h^{(3)}(x, y, z) = \partial/\partial z\, h(x, y, z)$. By (6), for $i > j > 1$

$$-U_{\Theta_{ij}} = U_{\Psi_{ij}} = \frac{(s_i r_j - s_j r_i)h^{(3)}(r_1, s_1, \sum_2^k r_i s_i)}{h(r_1, s_1, \sum_2^k r_i s_i)}.$$

A univariate version of Proposition 3 is easily proved: if the generalized location of a univariate $\mathscr{D}$-symmetric model is parametrized with skew-symmetric matrices, its information matrix is diagonal.

If f is rotationally symmetric the rank of $\mathbf{i}_{(\theta\psi)(\theta\psi)}$, the information matrix for $\mathbf{P}$, $\mathbf{Q}$, is $(k + 2)(k - 1)/2$. In general if f is $\mathscr{A}$-symmetric for any subgroup $\mathscr{A}$ of

$O(k)$ containing $\mathscr{D}$ the rank of $\mathbf{i}_{(\theta\psi)(\theta\psi)}$ is $k(k-1)$ minus the dimension of the parameter space needed to parametrize $\mathscr{A}$. In such cases, other parametrizations for $\mathbf{P}$ and $\mathbf{Q}$ have to be considered.

If for instance f is rotationally symmetric then ϕ, $\mathbf{P}$ and $\mathbf{q}_1$ parametrize f. Without loss of generality one can assume that $\mathbf{Q}$ belongs to $O^-(k)$ and take

$$\mathbf{Q} = \mathbf{I} - \frac{(\mathbf{e}_1 - \mathbf{q}_1)(\mathbf{e}_1 - \mathbf{q}_1)'}{1 - q_{11}}$$

where $\mathbf{e}_1 = (1, 0, \cdots, 0)'$ (in what follows, $\mathbf{e}_j$ will denote a vector of 0 except for its jth component, which is 1) and q_{11} is the first component of $\mathbf{q}_1$. Given $\mathbf{q}_{01}$, the true value of $\mathbf{q}_1$ that is assumed not to be equal to $\mathbf{e}_1$ and

$$\mathbf{Q}_0 = (\mathbf{q}_{01}, \cdots, \mathbf{q}_{0k}) = \mathbf{I} - \frac{(\mathbf{e}_1 - \mathbf{q}_{01})(\mathbf{e}_1 - \mathbf{q}_{01})'}{1 - q_{011}}$$

the corresponding matrix, define ψ_i as $\mathbf{q}_1'\mathbf{q}_{0i+1}$ for $i = 1, 2, \cdots, k-1$ and $\mathbf{x} = (0, \psi_1, \cdots, \psi_{k-1})'$. Then

$$\begin{aligned}\mathbf{q}_1 &= (1 - \|\mathbf{x}\|^2)^{1/2}\mathbf{q}_{01} + \mathbf{Q}_0\mathbf{x} \\ &= \mathbf{q}_{01} + \mathbf{x} + \frac{\mathbf{q}_{01}'\mathbf{x}}{1 - q_{011}}(\mathbf{e}_1 - \mathbf{q}_{01}) + O(\|\mathbf{x}\|^2).\end{aligned}$$

Elementary manipulations show that

$$\mathbf{Q} = \mathbf{Q}_0\mathbf{Q}_0'\mathbf{Q} = \mathbf{Q}_0\left(\mathbf{I} + \frac{\mathbf{x}(\mathbf{e}_1 - \mathbf{q}_{01})' - (\mathbf{e}_1 - \mathbf{q}_{01})\mathbf{x}'}{1 - q_{011}} + O(\|\mathbf{x}\|^2)\right)$$

and

$$\begin{aligned}\frac{\partial}{\partial\psi_i}\mathbf{Q}\,|_{\psi=0} &= \mathbf{Q}_0\left(\frac{\mathbf{e}_{i+1}(\mathbf{e}_1 - \mathbf{q}_{01})' - (\mathbf{e}_1 - \mathbf{q}_{01})\mathbf{e}_{i+1}'}{1 - q_{011}}\right) \\ &= \mathbf{Q}_0\mathbf{L}_{i+1}.\end{aligned}$$

If $\mathbf{P}$ is parametrized as before by θ and $\mathbf{Q}$ by ψ, one can show that

$$U_{\psi_i} = -\{g^{(2)}(\mathbf{r}, \mathbf{s})\}'\mathbf{L}_{i+1}\mathbf{s}/g(\mathbf{r}, \mathbf{s}; \phi)$$

and
$$\mathbf{H}_{i+1}\mathbf{L}_{i+1}\mathbf{H}_{i+1} = -\mathbf{L}_{i+1}.$$

Thus with this parametrization a partial extension of Proposition 3 holds:

PROPOSITION 4. *If $g(\mathbf{r}, \mathbf{s}; \phi) = h(r_1, s_1, \sum_2^k r_js_j; \phi)$ where h is a differentiable function and if $\mathbf{P}$ and $\mathbf{q}_1$ are parametrized by θ and ψ, $\mathbf{i}_{\theta\theta}$ is diagonal and $\mathbf{i}_{\theta\psi}$ satisfies $i_{\Theta_{ij}\psi_{i'}} = 0$ if $i' + 1 \neq i$ and $i' + 1 \neq j$.*

Given a sample $\{\mathbf{u}_i, \mathbf{v}_i\}$ from f, Proposition 2 and 3 suggest an algorithm for the efficient estimation of the parameters:

i) maximize $\pi_i g(\hat{\hat{\mathbf{P}}}'\mathbf{u}_i, \hat{\hat{\mathbf{Q}}}'\mathbf{v}_i; \phi)$ to get an efficient estimator ϕ where $\hat{\hat{\mathbf{P}}}$ and

$\hat{\hat{\mathbf{Q}}}$ are $O_p(n^{-1/2})$ consistent estimators of $\mathbf{P}$ and $\mathbf{Q}$. For instance, one can take $\hat{\hat{\mathbf{P}}}$ and $\hat{\hat{\mathbf{Q}}}$ as the right and left singular vector matrices of $\sum \mathbf{u}_i \mathbf{v}_i'$.

ii) To estimate $\mathbf{P}$ and $\mathbf{Q}$ let

$$\begin{pmatrix} \theta \\ \psi \end{pmatrix} = \frac{1}{n} \hat{\mathbf{i}}^{-1}_{(\theta\psi)(\theta\psi)} \sum_{i=1}^{n} \begin{pmatrix} \mathbf{U}_\theta(\hat{\hat{\mathbf{P}}}, \hat{\hat{\mathbf{Q}}}, \hat{\phi}, \mathbf{u}_i, \mathbf{v}_i) \\ \mathbf{U}_\psi(\hat{\hat{\mathbf{P}}}, \hat{\hat{\mathbf{Q}}}, \hat{\phi}, \mathbf{u}_i, \mathbf{v}_i) \end{pmatrix}$$

and $\hat{\Theta}$ and $\hat{\Psi}$ be the corresponding skew-symmetric matrices. Then $\hat{\mathbf{P}} = \hat{\hat{\mathbf{P}}} \exp \hat{\Theta}$ and $\hat{\mathbf{Q}} = \hat{\hat{\mathbf{Q}}} \exp \hat{\Psi}$ are obtained after one iteration of Fisher's scoring method for the maximization of $\pi\, g(\mathbf{P}'\mathbf{u}_i, \mathbf{Q}'\mathbf{v}_i; \hat{\phi})$ with $\hat{\hat{\mathbf{P}}}$ and $\hat{\hat{\mathbf{Q}}}$ as starting values; they are therefore efficient estimators (Cox and Hinkley, 1974, Chapter 9). For this method to be valid $i_{(\theta\psi)(\theta\psi)}$ has to be of full rank; f has to be strictly $\mathscr{D}$-symmetric. For rotationally symmetric models, a similar algorithm can be constructed with the parametrization of Proposition 4.

Acknowledgement. I am grateful to a referee for suggesting the parametrization of the rotations used in Section 3 and for many stimulating comments.

REFERENCES

COX, D. R. and HINKLEY, D. V. (1974). *Theoretical Statistics.* Chapman, London.

JOHNSON, R. A. and WEHRLY, T. (1977). Measures and models for angular correlation and angular-linear correlation. *J. Roy. Statist. Soc. B.* **39** 602–606.

JUPP, P. E. and MARDIA, K. V. (1980). A general correlation coefficient for directional data and related regression problems. *Biometrika* **67** 163–173.

MACKENZIE, J. K. (1957). The estimation of an orientation relationship. *Acta Cryst.* **10** 61–62.

RIVEST, L. P. (1982). Some statistical methods for bivariate directional data. *J. Roy. Statist. Soc. B.* **44** 81–90.

RIVEST, L. P. (1984). On the information matrix for symmetric distributions on the hypersphere. *Ann. Statist.* **12** 1085–1089.

SAW, J. C. (1983). Dependent unit vectors. *Biometrika* **70** 665–671.

DÉPARTEMENT DE MATHÉMATIQUES
UNIVERSITÉ LAVAL
CITÉ UNIVERSITAIRE
QUÉBEC G1K 7P4
CANADA