Statistics of Ambiguous Rotations

Richard Arnold[1], Peter Jupp[2] and Helmut Schaeben[3]

[1] School of Mathematics, Statistics and Operations Research, Victoria University of Wellington

[2] School of Mathematics and Statistics, University of St Andrews

[3] Geophysics and Geoscience Informatics, TU Bergakademie, Freiberg

1 Introduction

Data that are rotations occur in various contexts (e.g. astronomy, geophysics and human kinematics), and methods for analysing such data are now well developed. Some rotational data are ambiguous, in that the observations can be determined only up to multiplication by a rotation in some small group. Important examples are (a) principal components of second-order physical tensors — these can be specified by orthonormal vectors which are known only up to sign, (b) orientations of crystals — these can be determined only up to the action of the symmetry group of the crystal. We outline here some methods for handling data that are ambiguous rotations.

2 Orthogonal axial frames

An *orthonormal r-frame* in $\mathbb{R}^p$ is a set $(\mathbf{u}_1, \ldots, \mathbf{u}_r)$ of r p-vectors satisfying $\mathbf{u}_i^T \mathbf{u}_j = \delta_{ij}$. The space of such orthonormal r-frames is the Stiefel manifold $V_r(\mathbb{R}^p)$. An *orthogonal axial r-frame* in $\mathbb{R}^p$ is a set $(\pm\mathbf{u}_1, \ldots, \pm\mathbf{u}_r)$ of r orthogonal axes in $\mathbb{R}^p$. The group $\mathbb{Z}_2{}^r = \{(\varepsilon_1, \ldots, \varepsilon_r) | \varepsilon_j = \pm 1\}$ acts on $V_r(\mathbb{R}^p)$ by changing the signs of the vectors $\mathbf{u}_1, \ldots, \mathbf{u}_r$, so that $(\varepsilon_1, \ldots, \varepsilon_r)$ acts by $(\mathbf{u}_1, \ldots, \mathbf{u}_r) \mapsto (\varepsilon_1\mathbf{u}_1, \ldots, \varepsilon_r\mathbf{u}_r)$. Then the collection of sets of orthogonal axial r-frames in $\mathbb{R}^p$ can be identified with the quotient manifold $V_r(\mathbb{R}^p)/\mathbb{Z}_2{}^r$. For $\mathbf{U} = (\mathbf{u}_1, \ldots, \mathbf{u}_r)$ in $V_r(\mathbb{R}^p)$ we shall denote $(\pm\mathbf{u}_1, \ldots, \pm\mathbf{u}_r)$ in $V_r(\mathbb{R}^p)/\mathbb{Z}_2{}^r$ by $[\mathbf{U}]$.

In order to perform standard algebraic operations on elements of $V_r(\mathbb{R}^p)/\mathbb{Z}_2{}^r$, we embed $V_r(\mathbb{R}^p)/\mathbb{Z}_2{}^r$ in the space of r-tuples of symmetric $p \times p$ matrices of trace 0 by

$$(\pm\mathbf{u}_1, \ldots, \pm\mathbf{u}_r) \mapsto \mathbf{t}(\pm\mathbf{u}_1, \ldots, \pm\mathbf{u}_r) = (\mathbf{u}_1\mathbf{u}_1^T - (1/p)\mathbf{I}_p, \ldots, \mathbf{u}_r\mathbf{u}_r^T - (1/p)\mathbf{I}_p).$$

2.1 Summary statistics

Observations $[\mathbf{U}_1], \ldots, [\mathbf{U}_n]$ on $V_r(\mathbb{R}^p)/\mathbb{Z}_2{}^r$ can usefully be summarised by the sample mean $\bar{\mathbf{T}}$ of their images by $\mathbf{t}$, i.e. by $\bar{\mathbf{T}} = n^{-1} \sum_{i=1}^n \mathbf{t}([\mathbf{U}_i]) = \big(\bar{\mathbf{T}}_1, \ldots, \bar{\mathbf{T}}_r\big)$, where

$$\bar{\mathbf{T}}_j = n^{-1} \sum_{i=1}^{n} \left(\mathbf{u}_{ij}\mathbf{u}_{ij}^T - p^{-1}\mathbf{I}_p\right) \qquad j = 1, \ldots, r \tag{1}$$

for $[\mathbf{U}_i] = (\pm\mathbf{u}_{i1}, \ldots, \pm\mathbf{u}_{ir})$. The *sample mean* $[\bar{\mathbf{U}}]$ of $[\mathbf{U}_1], \ldots, [\mathbf{U}_n]$ is defined as the element $[\mathbf{U}] = (\pm\mathbf{u}_1, \ldots, \pm\mathbf{u}_r)$ of $V_r(\mathbb{R}^p)/\mathbb{Z}_2{}^r$ that maximises $\sum_{j=1}^r \mathbf{u}_j^T \bar{\mathbf{T}}_j \mathbf{u}_j$.

The *dispersion* of $[\mathbf{U}_1], \ldots, [\mathbf{U}_n]$ is $d = \sum_{j=1}^r \left[1 - \operatorname{tr}\left\{(\bar{\mathbf{T}}_j + p^{-1}\mathbf{I}_p)^2\right\}\right]$.

2.2 A test of uniformity

The most basic hypothesis about a distribution of orthogonal axial frames is that it is uniform, i.e. that it is invariant under the action of the rotation group $SO(p)$ on $V_r(\mathbb{R}^p)$ in which $\mathbf{V}$ in $SO(p)$ maps $[\mathbf{U}] = (\pm\mathbf{u}_1, \ldots, \pm\mathbf{u}_r)$ to $[\mathbf{V}\mathbf{U}] = (\pm\mathbf{V}\mathbf{u}_1, \ldots, \pm\mathbf{V}\mathbf{u}_r)$.

It is intuitively reasonable to reject uniformity if $\bar{\mathbf{T}}_1, \ldots, \bar{\mathbf{T}}_r$ are far from $\mathbf{0}$. There is an appealing statistic, S, which measures the distances of $\bar{\mathbf{T}}_1, \ldots, \bar{\mathbf{T}}_r$ from $\mathbf{0}$. For $r = p$,

$$S \quad = \quad \{n(p-1)(p+2)/2\} \sum_{j=1}^{p} \operatorname{tr}(\bar{\mathbf{T}}_j^2). \tag{2}$$

For $r = 1$, S is the Bingham statistic for testing uniformity on the projective space $\mathbb{R}P^{p-1}$. For $r = p$, $S = \{n(p-1)(p+2)/2\}\,(p-1-d)$, where d is the sample dispersion. Uniformity is rejected if S is large. For $r = p$, under uniformity, the distribution of the statistic S is

$$S \sim \chi^2_{(p-1)^2(p+2)/2},$$

asymptotically as $n \to \infty$.

2.3 Distributions

An appealing family of distributions on $V_r(\mathbb{R}^p)/\mathbb{Z}_2{}^r$ is the family of *frame Bingham* distributions, in which the probability density functions (with respect to the uniform distribution) have the form

$$f([\mathbf{U}]; \mathbf{A}_1, \ldots, \mathbf{A}_r) = c_p(\mathbf{A}_1, \ldots, \mathbf{A}_r)^{-1} \exp\left(\sum_{j=1}^{r} \mathbf{u}_j^T \mathbf{A}_j \mathbf{u}_j \right), \tag{3}$$

where $[\mathbf{U}] = (\pm\mathbf{u}_1, \ldots, \pm\mathbf{u}_r)$, $\mathbf{A}_1, \ldots, \mathbf{A}_r$ are symmetric $p \times p$ matrices, which may be assumed to have trace zero, and $c_p(\mathbf{A}_1, \ldots, \mathbf{A}_r)$ is the normalising constant. Saddlepoint approximations to the normalising constant $c_p(\mathbf{A}_1, \ldots, \mathbf{A}_r)$ can be obtained from the general results of Kume, Preston and Wood (2013).

A useful subfamily of the frame Bingham family (3) is the *commuting frame Bingham* distributions, obtained by imposing the condition that $\mathbf{A}_1, \ldots, \mathbf{A}_r$ commute. In this case they can be diagonalised simultaneously, so that

$$\mathbf{A}_j = \mathbf{V}\,\mathbf{D}_j \mathbf{V}^T \qquad j = 1, \ldots, r, \tag{4}$$

for some $\mathbf{V}$ in $O(p)$, where $\mathbf{D}_1, \ldots \mathbf{D}_r$ are diagonal. The parameter $[\mathbf{V}]$ in $O(p)/\mathbb{Z}_2^p$ is an orthogonal axial frame that describes location and the parameters $\mathbf{D}_1, \ldots \mathbf{D}_r$ measure concentration.

The family of *frame Watson* distributions on $V_r(\mathbb{R}^p)/\mathbb{Z}_2{}^r$ is the subfamily of the frame Bingham distributions for which the matrices $\mathbf{A}_1, \ldots, \mathbf{A}_r$ have rank 1 and are orthogonal. Thus $\mathbf{A}_j = \kappa_j \boldsymbol{\mu}_j \boldsymbol{\mu}_j^T$ for $j = 1, \ldots, r$, where $\kappa_1, \ldots, \kappa_r$ are scalars and $[\mathbf{M}] = (\pm\boldsymbol{\mu}_1, \ldots, \pm\boldsymbol{\mu}_r)$ is an orthogonal axial frame. Then the probability density functions have the form

$$f([\mathbf{U}]; \kappa_1, \ldots, \kappa_r, [\mathbf{M}]) = c_p(\kappa_1, \ldots, \kappa_r)^{-1} \exp\left\{ \sum_{j=1}^{r} \kappa_j (\mathbf{u}_j^T \boldsymbol{\mu}_j)^2 \right\}, \tag{5}$$

where $\mathbf{U} = (\pm\mathbf{u}_1, \ldots, \pm\mathbf{u}_r)$ and $c_p(\kappa_1, \ldots, \kappa_r)$ is the normalising constant.

Orthogonal axial frames arise as sets of principal axes of variance matrices of multivariate distributions. For both concentrated and diffuse multivariate distributions, the distribution of the principal axes is approximately commuting frame Bingham.

3 Symmetrical objects in $\mathbb{R}^3$

The orientation of a rigid object in $\mathbb{R}^3$ can be described by a rotation that transforms it into some standard position. If the object is asymmetrical then this rotation is unique, so that the orientations of the object correspond to elements of the rotation group $SO(3)$. If the object is symmetrical then the set of rotations that have no visible effect on the object forms a subgroup K of $SO(3)$. Thus the orientations of the object correspond to elements of the quotient group $SO(3)/K$. In the cases that we consider, K is finite. The orthogonal axial frames of Section 2 (with $p = r = 3$) fit into this setting; the symmetry group K is that of orthorhombic crystals. For $\mathbf{U}$ in $SO(3)$ we shall denote the equivalence class of $\mathbf{U}$ in $SO(3)/K$ by $[\mathbf{U}]$.

The cases of greatest practical interest are crystals and regular tetrahedra. The orientation of such an object can be specified by a suitable frame, i.e. ordered or unordered set of unit vectors or axes. For example, the orientation of an orthorhombic crystal is given by an orthogonal axial frame $(\pm\mathbf{u}, \pm\mathbf{v} \pm \mathbf{w})$, the orientation of a cube is given by an unordered set $\{\pm\mathbf{u}, \pm\mathbf{v}, \pm\mathbf{w}\}$ of orthogonal axes, and the orientation of a tetrahedron is given by an unordered set $\{\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{x}\}$ of unit vectors (unique up to cyclic order) that are at angle $\cos^{-1}(-1/3)$ to one another.

Data and distributions on $SO(3)/K$ can be handled by embedding $SO(3)/K$ into a suitable vector space V using $\mathbf{t} : SO(3)/K \to V$, where $\mathbf{t}$ is a symmetric tensor of order 1, 2, 3, 4 or 6 (which are the numbers of axes of symmetry in the crystal point-symmetry groups; see, e.g. Chaps. 1 and 2 of Powell, 2010). In the orthorhombic case, $\mathbf{t}((\pm\mathbf{u}, \pm\mathbf{v}, \pm\mathbf{w})) = (\mathbf{u}\mathbf{u}^T - (1/3)\mathbf{I}_3, \mathbf{v}\mathbf{v}^T - (1/3)\mathbf{I}_3, \mathbf{w}\mathbf{w}^T - (1/3)\mathbf{I}_3)$. In the cubic case $\mathbf{t}(\{\pm\mathbf{u}, \pm\mathbf{v}, \pm\mathbf{w}\}) = \otimes^4\mathbf{u} + \otimes^4\mathbf{v} + \otimes^4\mathbf{w} - (1/5)\mathbf{I}_3 \otimes \mathbf{I}_3[3]$, where the 4-tensor $\mathbf{I}_3 \otimes \mathbf{I}_3[3]$ has (i, j, k, ℓ)th component $\delta_{ij}\delta_{k\ell} + \delta_{ik}\delta_{j\ell} + \delta_{i\ell}\delta_{jk}$, with δ_{ij} denoting the Kronecker delta. Thus the (i, j, k, ℓ)th component of $\mathbf{t}(\{\pm\mathbf{u}, \pm\mathbf{v}, \pm\mathbf{w}\})$ is

$$t_{i,j,k,\ell}(\{\pm\mathbf{u}, \pm\mathbf{v}, \pm\mathbf{w}\}) = u_i u_j u_k u_\ell + v_i v_j v_k v_\ell + w_i w_j w_k w_\ell - (1/5)\{\delta_{ij}\delta_{k\ell} + \delta_{ik}\delta_{j\ell} + \delta_{i\ell}\delta_{jk}\}.$$

Similarly, in the tetrahedral case $\mathbf{t}(\{\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{x}\}) = \otimes^3\mathbf{u} + \otimes^3\mathbf{v} + \otimes^3\mathbf{w} + \otimes^3\mathbf{x}$.

3.1 Inference on $SO(3)/K$

Most of the constructions and techniques that are available for orthogonal axial frames can be extended readily to $SO(3)/K$.

The uniform distribution on $SO(3)/K$ is the unique distribution that is invariant under the action of $SO(3)$ on $SO(3)/K$ in which $\mathbf{V}$ in $SO(3)$ maps $[\mathbf{U}]$ to $[\mathbf{VU}]$. Since the embeddings $\mathbf{t}$ are chosen so that $E\{\mathbf{t}([\mathbf{U}])\} = \mathbf{0}$ for $\mathbf{U}$ uniformly distributed on $SO(3)/K$, it is intuitively reasonable to reject uniformity if $\bar{\mathbf{t}} = n^{-1}\sum_{i=1}^n \mathbf{t}([\mathbf{U}_i])$ is far from $\mathbf{0}$. An appropriate test statistic is

$$S = cn\,\|\bar{\mathbf{t}}\|^2, \tag{6}$$

where c is a constant that depends on the symmetry group K. Uniformity is rejected if S is large. Under uniformity, the distribution of the statistic S is

$$S \sim \chi^2_\nu,$$

asymptotically as $n \to \infty$ where ν depends on K.

A useful family of distributions on $SO(3)/K$ has densities of the form

$$f([\mathbf{U}]; [\mathbf{V}], \kappa) = c(\kappa)\exp\{\kappa\langle \mathbf{t}([\mathbf{U}]), \mathbf{t}([\mathbf{V}])\rangle\}, \tag{7}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product on the vector space V. For $\kappa > 0$, the mode is $[\mathbf{V}]$ and the maximum likelihood estimate of $[\mathbf{V}]$ is the sample mean. In the orthorhombic case, (7) is the density of the frame Watson distribution (5) for orthogonal axial frames with $r = p = 3$ and $\kappa_1 = \kappa_2 = \kappa_3 = \kappa$.

References

Arnold, R. and Jupp, P.E. (2013). Statistics of orthogonal axial frames. *Biometrika*, to appear.

Kume, A. Preston, S.P. and Wood, A.T.A. (2013). Saddlepoint approximations for the normalising constant of Fisher–Bingham distributions on products of spheres and Stiefel manifolds. *Biometrika*, **100**, to appear.

Powell, R.C. (2010). *Symmetry, Group Theory and the Physical Properties of Crystals*. New York, Springer.