RAC$\mathbb{SAM}$
Rev. R. Acad. Cien. Serie A. Mat.
VOL. **97 (**1), 2003, pp. 41–51
Estadística e Investigación Operativa / Statistics and Operations Research

A study of the tangent space model of the von Mises-Fisher distribution

A. Chakak and L. Imlahi

Abstract. For a random rotation $X = M_0 e^{\phi(\varepsilon)}$ where M_0 is a 3×3 rotation, ε is a trivariate random vector, and $\phi(\varepsilon)$ is a skew symmetric matrix, the least squares criterion consists of seeking a rotation M called the mean rotation minimizing $tr[(M - E(X))^t(M - E(X))]$. Some conditions on the distribution of ε are set so that the least squares estimator is unbiased. Of interest is when ε is normally distributed $N(0, \Sigma)$. Unbiasedness of the least squares estimator is dealt with according to eigenvalues of Σ.

Un estudio del espacio tangente modelo de la distribución de von Mises-Fisher

Resumen. Dada una rotación aleatoria $X = M_0 e^{\phi(\varepsilon)}$, donde M_0 es una rotación de 3×3, un vector aleatorio trivariante ε y $\phi(\varepsilon)$ es una matriz antisimétrica, el criterio de mínimos cuadrados consiste en hallar una rotación M denominada rotación minimizante $tr[(M - E(X))^t(M - E(X))]$. Algunas condiciones sobre la distribución de ε son dadas de manera que el estimador de mínimos cuadrados sea insesgado. Es relevante el caso en el que ε está normalmente distribuido $N(0, \Sigma)$. La carencia de sesgo del estimador de mínimos cuadrados es tratada mediante los autovalores de Σ.

1. Introduction

Downs (1972) introduced the matrix von Mises-Fisher distribution to describe a random position of a rigid object. This general exponential distribution is parametrized by an $n \times p$ $(p \leq n)$ matrix F, which is decomposed as a product of two square matrices $M_0 K$ where M_0 is an $(n \times n)$ matrix called the polar component minimizing $tr[(F - X)'(F - X)]$, and K is a $p \times p$ called the elliptical component. Large eigenvalues of K correspond to a concentrated distribution around its modal value M_0. This matrix distribution has been studied by Downs (1972), Khatri and Mardia (1977), Jupp and Mardia (1989), and Mardia and Jupp (2000). Prentice (1986) notes that almost in every practical application $n = 3$, and suggests developing statistical inference on $SO(3)$ the space of 3×3 rotations. In applications where data are close to a fixed rotation, it is better to develop statistical inference on the tangent space. The tangent space to $SO(3)$ at M_0 is the tri-dimensional space of 3×3 skew symmetric matrices. If $\phi(\varepsilon)$ is a skew symmetric matrix whose elements are components of ε, $M_0(I + \phi(\varepsilon))$ describes the rotations around M_0 when ε is close to $(0, 0, 0)^t$. However $M_0(I + \phi(\varepsilon))$ is not a rotation. Instead $M_0 e^{\phi(\varepsilon)}$ is a rotation when ε is close to $(0, 0, 0)^t$. The space of such rotations is called the tangent space approximation at M_0. Under the matrix

Presentado por Francisco Javier Girón.
Recibido: 21 de Noviembre de 2001. *Aceptado*: 8 de Mayo de 2002.
Palabras clave / Keywords: Matrix von Mises-Fiser distribution, rotation, unbiased least squares estimation, tangent approximation model.
Mathematics Subject Classifications: 62 E

von Mises-Fisher distribution, Downs (1972) shows that the tangent approximation model is such that ε is distributed as $N(0,\Sigma)$. An application of this model is provided by Rancourt *et al* (2000).

In Downs (1972) the maximum likelihood estimation of M_0, under the matrix von Mises-Fisher distribution, is the rotation M called the mean rotation closest to $E(X)$ in the least squares sense; that is M is the rotation minimizing $tr[(E(X)-M)^t(E(X)-M)]$, where $tr(A)$ designates the trace of the square matrix A. Equivalently M is the rotation maximizing $tr[M^tE(X)]$. For a random sample of rotations of size n , $\overline{X_n}=\frac{1}{n}\sum_{i=1}^{n}X_i$ is substituted to $E(X)$ and that the maximum likelihood estimate of M_0 is obtained from the singular value decomposition of $\overline{X_n}$ as $M=PQ^t$ where $\overline{X_n}=Pdiag(\gamma_1,\gamma_2,\gamma_3)Q^t$, P et Q are rotations and $\gamma_1>\gamma_2>|\gamma_3|$. Unfortunately this latter inequality is not always satisfied. In this paper we investigate the unbiased least squares estimate of the mean rotation under the tangent approximation model. Precise results are obtained when in addition $\varepsilon\sim N(0,\Sigma)$.

In section 2, we set some conditions so that the least squares estimate of the mean rotation is unbiased. These conditions are mainly obtained from singular decomposition of the mean of $e^{\phi(\varepsilon)}$ relative to the distribution of ε, in which case not all of its eigenvalues are null and the sum of pairs of its eigenvalues must be non-negative. In section 3 we determine the set of matrices Σ satisfying some sufficient conditions set in section 2 when $\varepsilon\sim N(0,\Sigma)$. It turns out that the expressions obtained are sometimes messy. A temptative fit is then provided when necessary.

2. General setup

A rotation is a matrix M satisfying $M^tM=I$ and $det(M)=1$, where M^t is the transpose of the matrix M. Every non nul vector provides a skew symmetric matrix, which provides a rotation by exponentiation. Let $\varepsilon=(\varepsilon_1,\varepsilon_2,\varepsilon_3)^t\in R^3$, and let $\phi(\varepsilon)$ be the skew symmetric matrix associated to ε as

$$\phi(\varepsilon)=\begin{pmatrix}0&-\varepsilon_3&\varepsilon_2\\ \varepsilon_3&0&-\varepsilon_1\\ -\varepsilon_2&\varepsilon_1&0\end{pmatrix}.$$

By definition of the exponential of a matrix,

$$e^{\phi(\varepsilon)}=\sum_{i=0}^{\infty}\frac{1}{n!}[\phi(\varepsilon)]^n=cos(\theta)I+\frac{sin(\theta)}{\theta}\phi(\varepsilon)+\frac{1-cos(\theta)}{\theta^2}\varepsilon\varepsilon^t.$$

This is a rotation of angle $\theta=(\varepsilon^t\varepsilon)^{1/2}$ around axis $\varepsilon/(\varepsilon^t\varepsilon)^{1/2}$.

Let M_0 be a 3×3 rotation and $X=M_0e^{\phi(\varepsilon)}$ be a random rotation around the rotation M_0, where ε is the random vector representing the experimental error. Denote by $E(X)$ the mathematical expectation of X with respect to the distribution of ε. Usually $E(X)$ is not a rotation, then it differs from M_0.

To define a measure of location for the random rotation X, we take the closest rotation to $E(X)$ in the least squares sense. In $SO(3)$, the least squares estimate M minimizes $tr[\{E(X)-M\}^t\{E(X)-M\}]$. However $tr[\{E(X)-M\}^t\{E(X)-M\}]=tr[E(X)^tE(X)]+3-2tr[M^tE(X)]$, then M is the rotation maximizing $tr[M^tE(X)]$. This criterion provides a mean rotation equivariant to changes in the system of axes in which it is recorded. If M is the mean rotation of X, then the mean rotation of M_1XM_2 is M_1MM_2, where M_1,M_2 are rotations.

It seems natural to require that the mean rotation around M_0 is M_0. The next proposition presents two conditions on the distribution of ε so that M_0 is the mean rotation of X.

Proposition 1 *The mean rotation of $X=M_0e^{\phi(\varepsilon)}$ is M_0 provided that the two conditions hold:*
1) $E(\frac{sin(\Theta)}{\Theta}\varepsilon)=(0,0,0)^t$.
2) The eigenvalues γ_1,γ_2, and γ_3 of the matrix $H=E[\cos(\Theta)I+\frac{1-\cos(\Theta)}{\Theta^2}\,\varepsilon\varepsilon^t]$, written in decreasing order are not all equal to 0 and satisfy $\gamma_1\geq\gamma_2\geq|\gamma_3|$, where $\Theta=\sqrt{\varepsilon^t\varepsilon}=\|\varepsilon\|$.

PROOF. One has $E(X) = M_0H$. To find the rotation M maximizing $tr(M^tM_0H)$, following Mackenzie (1957) and Stephens (1979), one uses the singular value decomposition of M_0H. This decomposition can be expressed in terms of the eigenvalues of H as $H = P\Gamma P^t$. Hence $M_0H = M_0P\Gamma P^t$ where Γ is the diagonal matrix of the γ_i's and $tr(M^tM_0H) = tr(M^tM_0P\Gamma P^t) = tr(P^tM^tM_0P\Gamma)$. This expression is maximum when $P^tM^tM_0P = I$, that is $M = M_0$ and the maximum value is $\sum_{i=1}^3 \gamma_i$. ■

Remark 1 The condition 2 of Proposition 1 holds when the three sums $S_{ij} = \gamma_i + \gamma_j \geq 0, 1 \leq i < j \leq 3$, and $tr(H) = \gamma_1 + \gamma_2 + \gamma_3 > 0$. ■

Although the assumption that $\gamma_1 \geq \gamma_2 \geq |\gamma_3|$ is not explicitly stated in the proof, it is necessary for the conclusion of the Proposition 1 to hold. When it fails, $tr(M_0^tE(X))$ may not be maximum anymore. For instance if ε gives the probability mass $1/2$ to both $(\pi, 0, 0)$ and $(-\pi, 0, 0)$, then $\gamma_1 = -\gamma_2 = -\gamma_3 = 1$. The mean rotation of X is not M_0 but $M_0diag(1, -1, -1)$.

If the distribution of ε is invariant to changes in the sign of any of its components then $E(\frac{\sin(\Theta)}{\Theta}\varepsilon) = (0, 0, 0)^t$, and the conclusion of Proposition 1 holds provided that $\gamma_1 \geq \gamma_2 \geq |\gamma_3|$ and $\sum_{i=1}^3 \gamma_i > 0$. A sufficient condition for this is $E[cos(\Theta)] > 0$, since $E(\frac{1-\cos(\Theta)}{\Theta}\varepsilon\varepsilon^t)$ is a diagonal matrix with elements $E(\frac{1-\cos(\Theta)}{\Theta}\varepsilon_i^2) > 0$. This is the case when most of the probability mass of ε is in the sphere $\varepsilon^t\varepsilon \leq (\pi/2)^2$.

Proposition 2 *The singular value decomposition of the matrix H is:*

$$H = \sum_{i=1}^{3} E(\cos(\Theta) + \frac{1-\cos(\Theta)}{\Theta}\lambda_i Z_i^2)e_ie_i^t$$

where $\Sigma = \sum_{i=1}^3 \lambda_i e_i e_i^t$ is the singular value decomposition of Σ, and the Z_1, Z_2, Z_3 are i.i.d. $N(0, 1)$.

PROOF. Let $\lambda_1 \geq \lambda_2 \geq \lambda_3 (\geq 0)$ be the eigenvalues of Σ, written in a decreasing order, associated to the eigenvectors $e_i, i = 1, 2, 3$ respectively. The sequence $\{e_1, e_2, e_3\}$ form an orthonormal basis. The singular value decomposition of Σ is $\Sigma = \lambda_1e_1e_1^t + \lambda_2e_2e_2^t + \lambda_3e_3e_3^t$. A decomposition of ε on the basis of the eigenvectors is $\varepsilon = \varepsilon_1e_1 + \varepsilon_2e_2 + \varepsilon_3e_3$, where for $\varepsilon_i = \varepsilon^te_i$ is normally distributed with zero mean and variance λ_i. Moreover $cov(\varepsilon_i, \varepsilon_j) = \delta_{ij}\lambda_i,\ \delta_{ij} = 1$, for $i = j$ and 0 otherwise. Normalizing the components, $\varepsilon_i = \sqrt{\lambda_i}Z_i, i = 1, 2, 3$, the Z_i 's are i.i.d. $N(0, 1)$. Thus $\varepsilon = \sum_{i=1}^3 \sqrt{\lambda_i}Z_ie_i$, and $\|\varepsilon\| = \Theta = \sqrt{\varepsilon^t\varepsilon} = \sqrt{\lambda_1Z_1^2 + \lambda_2Z_2^2 + \lambda_3Z_3^2}$. One has $\varepsilon\varepsilon^t = \sum_{i,j=1}^3 \sqrt{\lambda_i\lambda_j}Z_iZ_je_ie_j^t$. Θ is an even function of the Z_i's, which are independent with zero mean. Then $E[\frac{1-\cos(\Theta)}{\Theta^2}Z_i] = 0$, and $E[\sqrt{\lambda_i\lambda_j}\frac{1-\cos(\Theta)}{\Theta^2}Z_iZ_j] = 0, 1 \leq i \neq j \leq 3$. Therefore

$$\begin{aligned} E\left\{[1-\cos(\Theta)]\varepsilon\varepsilon^t/\Theta^2\right\} &= \sum_{i,j=1}^{3} \sqrt{\lambda_i\lambda_j}E[Z_iZ_j(1-\cos(\Theta))/\Theta^2)]e_ie_j^t \\ &= \sum_{i=1}^{3} \lambda_iE[Z_i^2(1-\cos(\Theta))/\Theta^2]e_ie_i^t. \end{aligned}$$

Writing $I = \sum_{i=1}^3 e_ie_i^t$, the singular value decomposition of H is

$$H = \sum_{i=1}^{3} E[cos(\Theta) + \frac{1-\cos(\Theta)}{\Theta^2}\ \lambda_iZ_i^2\ ]e_ie_i^t,$$

with eigenvalues $\gamma_i = E[cos(\Theta) + \frac{1-\cos(\ \Theta)}{\Theta^2}\ \lambda_iZ_i^2\ ], i = 1, 2, 3$ associated to the eigenvectors $e_i,\ i = 1, 2, 3$ respectively. ■

3. Application to the tangent approximation model

Recall that under the von Mises-Fisher distribution, the tangent approximation model is $M_0 e^{\phi(\varepsilon)}$ with $\varepsilon \sim N(0, \Sigma)$. Then $E(\frac{sin(\Theta)}{\Theta}\, \varepsilon) = (0,0,0)^t$. The condition (1) of Proposition 1 is always satisfied in this case.

Proposition 3 *For $\lambda_3 = \lambda_2 = 0 < \lambda_1$, the condition 2) of Proposition 1 holds.*

PROOF. When $\lambda_3 = \lambda_2 = 0 < \lambda_1$, $\Theta = \sqrt{\lambda_1}\,|Z_1|$. The eigenvalues of H are $\gamma_1 = 1,\ \gamma_2 = \gamma_3 = E[cos(\sqrt{\lambda_1}\,|Z_1|)] = E[cos(\sqrt{\lambda_1} Z_1)] = E(e^{i\sqrt{\lambda_1} Z_1}) = e^{-\lambda_1/2}$, as the characteristic function of a standardized $N(0,1)$ evaluated at $\sqrt{\lambda_1}$. Therefore $tr(H) = 1 + 2e^{-\lambda_1/2} > 0$, and $S_{ij} > 0,\ 1 \le i \ne j \le 3$.

When $\lambda_3 = 0 < \lambda_2 \le \lambda_1$, $\Theta = \sqrt{\lambda_1 Z_1^2 + \lambda_2 Z_2^2}$. The eigenvalues of H are

$$\begin{aligned} \gamma_1 &= E[cos(\Theta) + \frac{1 - cos(\Theta)}{\Theta^2}\lambda_1 Z_1^2], \\ \gamma_2 &= E[cos(\Theta) + \frac{1 - cos(\Theta)}{\Theta^2}\lambda_2 Z_2^2], \\ \gamma_3 &= E[cos(\Theta)]. \end{aligned}$$

Then the sums of pairs of eigenvalues are

$$\begin{aligned} S_{12} &= \gamma_1 + \gamma_2 = 1 + E[cos(\Theta)], \\ S_{13} &= \gamma_1 + \gamma_3 = E[2cos(\Theta) + \frac{1 - cos(\Theta)}{\Theta^2}\lambda_1 Z_1^2], \\ S_{23} &= \gamma_2 + \gamma_3 = E[2cos(\Theta) + \frac{1 - cos(\Theta)}{\Theta^2}\lambda_2 Z_2^2]. \end{aligned}$$

Note that (Z_1, Z_2) and (Z_2, Z_1) have the same distribution, S_{13} is obtained from S_{23} by interchanging λ_1 and λ_2. Particularly, $S_{13} = S_{23}$ when $\lambda_1 = \lambda_2$.

First $S_{12} = 1 + E[cos(\Theta)] \ge 0$. To derive expressions of S_{13} and S_{23}, we evaluate $E[cos(\Theta)]$. The distribution of $\Theta^2 = \lambda_1 Z_1^2 + \lambda_2 Z_2^2$ is a linear combination of two independent chi-square χ_1^2 distributions. Put $\Theta = \sqrt{W\Psi}$, with $W = Z_1^2 + Z_2^2$ and $\Psi = (\lambda_1 Z_1^2 + \lambda_2 Z_2^2)/(Z_1^2 + Z_2^2)$. Clearly W is distributed as a χ_2^2 (with density $\frac{e^{-w/2}}{2}$), independently distributed from $\beta_1 = \frac{Z_1^2}{Z_1^2 + Z_2^2} \sim B(1/2, 1/2)$ the beta distribution. Then W is independent of $\Psi = (\lambda_1 - \lambda_2)\beta_1 + \lambda_2$. Given $[\Psi = \psi]$, and expanding the cosine in series,

$$E^{[\Psi=\psi]}[cos(\Theta)] = E[cos(\sqrt{W\psi})] = E[\Sigma_{i=0}^{\infty} \frac{(-\psi W)^i}{(2i)!}].$$

The moments of W are $E(W^i) = 2^i i!$. Given $[\Psi = \psi]$,

$$E^{[\Psi=\psi]}[cos(\Theta)] = G(\psi) = 1 - \psi + \frac{\psi^2}{3.1} - \frac{\psi^3}{5.3.1} + \ldots$$

Put $\psi = \phi^2$, leads to $G(\phi^2) = 1 - \phi D(\phi)$, where

$$D(\phi) = \phi - \frac{\phi^3}{3.1} + \frac{\phi^5}{5.3.1} - \frac{\phi^7}{7.5.3.1} + ...$$

$D(\phi)$ satisfies a simple differential equation: $D'(\phi) = 1 - \phi D(\phi)$, whose solution is

$$D(\phi) = e^{-\phi^2/2} \int_0^{\phi} e^{u^2/2} du.$$

Therefore

$$G(\phi^2) = 1 - \phi D(\phi) = 1 - \phi e^{-\phi^2/2} \int_0^{\phi} e^{u^2/2} du.$$

Transforming back to ψ, we have

$$E^{[\Psi=\psi]}[cos(\Theta)] = 1 - \sqrt{\psi}e^{-\psi/2}\int_0^{\sqrt{\psi}} e^{u^2/2}du = 1 - \sqrt{\psi}F(\psi),$$

where $F(x) = e^{-x/2}\int_0^{\sqrt{x}} e^{u^2/2}du,\ x > 0$. Recall that $\lambda_2 Z_2^2/(\lambda_1 Z_1^2 + \lambda_2 Z_2^2) = \lambda_2(1-\beta_1)/\Psi$. Then

$$\begin{aligned}
S_{23}^{[\Psi=\psi]} &= E^{[\Psi=\psi]}[2cos(\Theta) + (1-cos(\Theta))\frac{\lambda_2 Z_2^2}{\lambda_1 Z_1^2+\lambda_2 Z_2^2}] \\
&= 2E^{[\Psi=\psi]}(cos(\Theta)) + (1 - E^{[\Psi=\psi]}(cos(\Theta)))\frac{\lambda_2(1-\beta_1)}{\psi} \\
&= 2 + \sqrt{\psi}e^{-\psi/2}(\int_0^{\sqrt{\psi}} e^{\frac{u^2}{2}}du)[\frac{\lambda_2(1-\beta_1)}{\psi} - 2] \\
&= 2 + \sqrt{\psi}F(\psi)[\frac{\lambda_2(1-\beta_1)}{\psi} - 2].
\end{aligned}$$

We consider the two situations: $\lambda_1 > \lambda_2$, and $\lambda_1 = \lambda_2$ separately. ■

Proposition 4 *For* $\lambda_3 = 0 < \lambda_2 < \lambda_1$, *the condition 2) of Proposition 1 holds on the set* $\{(\lambda_1, \lambda_2, 0) : \lambda_1 > \lambda_2 > 0;\ \frac{44}{\lambda_2 log(1+\lambda_2)} \le \lambda_1 \le 6e^{-1.8\lambda_2} + \frac{1}{2}e^{2.3(\lambda_2-3)} + 4.8\}$.

PROOF. When $\lambda_1 > \lambda_2,\ \beta_1 = \frac{\psi-\lambda_2}{\lambda_1-\lambda_2}$. We have

$$S_{23}^{[\Psi=\psi]} = 2 + F(\psi)[\frac{\psi(\lambda_2-2\lambda_1)+\lambda_1\lambda_2}{\sqrt{\psi}(\lambda_1-\lambda_2)}].$$

Ψ is an affine transformation of $B(1/2, 1/2)$ with density $\frac{1}{\pi\sqrt{(\psi-\lambda_2)(\lambda_1-\psi)}}$. The explicit expression is then

$$\begin{aligned}
S_{23} &= 2 + E\left\{e^{-\psi/2}(\int_0^{\sqrt{\psi}} e^{\frac{u^2}{2}}du)[\frac{\psi(\lambda_2-2\lambda_1)+\lambda_1\lambda_2}{\sqrt{\psi}(\lambda_1-\lambda_2)}]\right\} \\
&= 2 + \int_{\lambda_2}^{\lambda_1}\frac{1}{\pi\sqrt{(\psi-\lambda_2)(\lambda_1-\psi)}}e^{-\psi/2}(\int_0^{\sqrt{\psi}} e^{\frac{u^2}{2}}du)[\frac{\psi(\lambda_2-2\lambda_1)+\lambda_1\lambda_2}{\sqrt{\psi}(\lambda_1-\lambda_2)}]d\psi.
\end{aligned}$$

S_{23} has no explicit expression, and the equation $S_{23}(\lambda_1, \lambda_2) = 0$ has no explicit solution. Numerical methods are used to determine the approximate contour of solutions. Then the subset $\{(\lambda_1, \lambda_2, 0) : \lambda_1 > \lambda_2 > 0;\ \frac{44}{\lambda_2 log(1+\lambda_2)} \le \lambda_1 \le 6e^{-1.8\lambda_2} + \frac{1}{2}e^{2.3(\lambda_2-3)} + 4.8\}$ provides a convenient fit to the region where S_{23} is positive.

Now interchanging λ_1 and λ_2, in S_{23},

$$\begin{aligned}
S_{13} &= 2 + E\left\{e^{-\psi/2}(\int_0^{\sqrt{\psi}} e^{\frac{u^2}{2}}du)[\frac{\psi(\lambda_1-2\lambda_2)+\lambda_1\lambda_2}{\sqrt{\psi}(\lambda_2-\lambda_1)}]\right\} \\
&= 2 + \int_{\lambda_2}^{\lambda_1}\frac{1}{\pi\sqrt{(\psi-\lambda_2)(\lambda_1-\psi)}} \times \frac{e^{-\psi/2}(\int_0^{\sqrt{\psi}} e^{\frac{u^2}{2}}du)}{\sqrt{\psi}} \times \frac{\psi(\lambda_1-2\lambda_2)+\lambda_1\lambda_2}{(\lambda_1-\lambda_2)}d\psi.
\end{aligned}$$

For $\psi \in [\lambda_2, \lambda_1]$:

$$min(\lambda_1, 2\lambda_2) \le \frac{\psi(\lambda_1-2\lambda_2)+\lambda_1\lambda_2}{\lambda_1-\lambda_2} \le max(\lambda_1, 2\lambda_2).$$

$[\psi(\lambda_1 - 2\lambda_2) + \lambda_1\lambda_2]/(\lambda_1 - \lambda_2)$ is an increasing function of ψ for $\lambda_1 \geq 2\lambda_2$ with minimum value $2\lambda_2$ at $\psi = \lambda_2$, and maximum value λ_1 at $\psi = \lambda_1$. Therefore $S_{13} > 2 + min\{\lambda_1, 2\lambda_2\}\frac{F(\lambda_1)}{\sqrt{\lambda_1}} > 2$, which shows that S_{13} is positive. ■

Proposition 5 *For* $0 = \lambda_3 < \lambda_2 = \lambda_1$, *the condition 2) of Proposition 1 holds.*

PROOF. When $\lambda_1 = \lambda_2$, $\Theta = \sqrt{\lambda_1(Z_1^2 + Z_2^2)}$ and Ψ is a constant equal to λ_1. One has $E[cos(\Theta)] = 1 - \sqrt{\lambda_1}F(\lambda_1)$. This provides

$$\begin{aligned} S_{23} &= 2E[cos(\Theta)] + E[(1 - cos(\Theta))\beta_1] \\ &= 2 + \sqrt{\lambda_1}e^{-\lambda_1/2}\int_0^{\sqrt{\lambda_1}} e^{\frac{u^2}{2}}du[E(\beta_1) - 2]. \\ &= 2 - \frac{3}{2}\sqrt{\lambda_1}e^{-\lambda_1/2}\int_0^{\sqrt{\lambda_1}} e^{\frac{u^2}{2}}du. \end{aligned}$$

Noting that $0 \leq \sqrt{\lambda_1}F(\lambda_1) \leq 1.28495$, then $S_{23} = S_{13} \geq 0.07257 > 0$, which implies that all the γ_i's are not null.

The last three cases deal with situations when λ_3 is positive. We have $tr(H) = 1 + 2E[\cos(\Theta)]$. The distribution of Θ^2 is a linear combination of i.i.d. χ_1^2 distributions. Let's write $\Theta = \sqrt{W\Psi}$, with $W = Z_1^2 + Z_2^2 + Z_3^2$, and $\Psi = \lambda_1\beta_1 + \lambda_2\beta_2 + \lambda_3\beta_3$, where $\beta_i = Z_i^2/(Z_1^2 + Z_2^2 + Z_3^2) \sim B(1/2, 1)$. $W \sim \chi_3^2$, independently distributed from β_i, $i = 1, 2, 3$; and hence independent of Ψ. Given $[\Psi = \psi]$ in $[\lambda_3, \lambda_1]$, one has

$$E[\cos(\sqrt{W\psi})] = \sum_{i=0}^{\infty} \frac{(-\psi)^i E(W^i)}{(2i)!}.$$

The moments of W are

$$\begin{aligned} E(W^i) &= \int_0^{\infty} \frac{x^{i+1/2}e^{-x/2}}{\Gamma(3/2)2^{3/2}}dx \\ &= \frac{2^i}{\Gamma(3/2)}\int_0^{\infty} y^{i+1/2}e^{-y}dy \\ &= \frac{2^i}{\Gamma(3/2)}\Gamma(i + 3/2) \\ &= \frac{(2i+1)!}{2^i i!}. \end{aligned}$$

Therefore

$$\begin{aligned} E[\cos(\sqrt{W\psi})] &= \sum_{i=0}^{\infty}(-1)^i \frac{\psi^i}{(2i)!}\frac{(2i+1)!}{2^i i!} \\ &= \sum_{i=0}^{\infty} \frac{(-1)^i}{i!}(\frac{\psi}{2})^i(2i+1) \\ &= e^{-\psi/2}(1 - \psi). \end{aligned}$$

We get

$$tr(H) = 1 + 2E[\cos(\Theta)] = E[1 + 2(1 - \Psi)\exp(-\frac{\Psi}{2})].$$

However $1 + 2\exp(-\frac{\Psi}{2})(1 - \Psi) \geq 1 - 4\exp(-\frac{3}{2}) \geq 0.1074$. Thus $tr(H) > 0$ implying that not all eigenvalues of H are null when $\lambda_3 > 0$. ■

Proposition 6 *For* $0 < \lambda_3 = \lambda_2 = \lambda_1$, *the condition 2) of Proposition 1 holds* .

PROOF. When $\lambda_1 = \lambda_2 = \lambda_3 > 0$, the eigenvalues of H are all equal and $\psi = \lambda_1$. The three sums of pairs of eigenvalues are

$$S_{12} = S_{13} = S_{23} = \frac{2}{3}[1 + 2(1 - \lambda_1)\exp(-\frac{\lambda_1}{2})] \geq 0.0716$$

which ends the proof. ∎

For the remaining cases we show that S_{12} is positive first. Using the same argument as for the trace, one has

$$\begin{aligned} \gamma_i &= E[cos(\Theta) + \frac{1 - cos(\Theta)}{\Theta^2}\lambda_i Z_i^2] \\ &= E[e^{-\Psi/2}(1 - \Psi)(1 - \frac{\lambda_i\beta_i}{\Psi}) + \frac{\lambda_i\beta_i}{\Psi}], \ \ i = 1, 2, 3. \end{aligned}$$

makes

$$S_{ij} = E[e^{-\Psi/2}(1 - \Psi)(1 + \frac{\lambda_k\beta_k}{\Psi}) + 1 - \frac{\lambda_k\beta_k}{\Psi}], \ 1 \leq i \neq j \neq k \leq 3.$$

That is,

$$\begin{aligned} S_{12} &= E[e^{-\Psi/2}(1 - \Psi)(1 + \frac{\lambda_3\beta_3}{\Psi}) + 1 - \frac{\lambda_3\beta_3}{\Psi}], \\ S_{13} &= E[e^{-\Psi/2}(1 - \Psi)(1 + \frac{\lambda_2\beta_2}{\Psi}) + 1 - \frac{\lambda_2\beta_2}{\Psi}], \\ S_{23} &= E[e^{-\Psi/2}(1 - \Psi)(1 + \frac{\lambda_1\beta_1}{\Psi}) + 1 - \frac{\lambda_1\beta_1}{\Psi}]. \end{aligned}$$

Expression S_{13} is obtained from S_{23} by interchanging λ_1 and λ_2, since the joint distribution of (Z_1^2, Z_2^2, Z_3^2) is the same for all permutations of the Z_i^2's.

To evaluate S_{12}, we write $\Psi = \lambda_3\beta_3 + (1 - \beta_3)\lambda_3 C_{12}$, where $C_{12} = (\lambda_1 Z_1^2 + \lambda_2 Z_2^2)/[\lambda_3(Z_1^2 + Z_2^2)]$, which takes values in $(\lambda_2/\lambda_3, \lambda_1/\lambda_3)$. The random variables β_3 and C_{12} are independent, because C_{12} is independent of $(Z_1^2 + Z_2^2, Z_3^2)$. Then it is also independent of β_3 as a continuous function of $(Z_1^2 + Z_2^2, Z_3^2)$. Writing

$$\begin{aligned} S_{12} &= E\left\{[e^{-\Psi/2}(1 - \Psi) + \frac{1}{2}][1 + \frac{\beta_3}{\beta_3 + C_{12}(1 - \beta_3)}]\right\} \\ &\quad + E[\frac{1}{2} - (\frac{3}{2})\frac{\beta_3}{\beta_3 + C_{12}(1 - \beta_3)})] \end{aligned}$$

The first term $[e^{-\Psi/2}(1-\Psi)+(1/2)][1+\beta_3/(\beta_3+C_{12}(1-\beta_3)] \geq e^{-\Psi/2}(1-\Psi)+1/2 > 0$. As to the second term, $C_{12} \geq \lambda_2/\lambda_3 \geq 1$ implies $\beta_3 + C_{12}(1 - \beta_3) \geq 1$ making $E[1/2 - (3/2)(\beta_3/(\beta_3 + C_{12}(1 - \beta_3))] \geq 1/2 - (3/2)E(\beta_3) = 0$. This shows that $S_{12} > 0$.

Proposition 7 *For* $0 < \lambda_3 < \lambda_2 \leq \lambda_1$, *the condition 2) of Proposition 1 holds in the following situations:*

i) $0.7\lambda_1 \leq \lambda_3 < \lambda_2 \leq \lambda_1$.

ii) $\sqrt{\frac{1-C}{C}}((2 - C)e^{-\lambda_1/2} - C) + arctg(\sqrt{\frac{1-C}{C}}) - \int_0^{arctg(\sqrt{\frac{(1-C)}{C}})} e^{-\frac{\lambda_1 C}{2cos^2 z}}(1 + \frac{\lambda_1 C(1-C)}{cos^2 z})dz \geq 0,$

where $\frac{\lambda_3}{\lambda_1} \leq C \ \leq \min(\frac{\lambda_2}{\lambda_1}, 0.7)$.

PROOF. For $\lambda_1 \geq \lambda_2 > \lambda_3 > 0$, we study $S_{23} = \gamma_2 + \gamma_3$ first. One has

$$S_{23} = E[e^{-\Psi/2}(1 - \Psi)(1 + \frac{\lambda_1\beta}{\Psi}) + 1 - \frac{\lambda_1\beta}{\Psi}],$$

where $\Psi = (\lambda_1 Z_1^2 + \lambda_2 Z_2^2 + \lambda_3 Z_3^2)/(Z_1^2 + Z_2^2 + Z_3^2)$. To evaluate S_{23}, we write $\Psi = \lambda_1\beta + (1-\beta)\lambda_1 C$, where $C = (\lambda_2 Z_2^2 + \lambda_3 Z_3^2)/(\lambda_1(Z_2^2 + Z_3^2))$, and $\beta = Z_1^2/(Z_1^2 + Z_2^2 + Z_3^2)$. Note that β and C are statistically independent, and that $\lambda_3/\lambda_1 \leq C \leq \lambda_2/\lambda_1 \leq 1$. Furthermore, $\beta \sim B(1/2, 1)$. Thus, one has $P(\beta < t) = t^{1/2}$. In the sequel let $E(\cdot)$ denote the expectation with respect to β when C is fixed. One has,

$$S_{23} = E(e^{-(\lambda_1\beta + (1-\beta)\lambda_1 C)/2}(1-\Psi)(1 + \frac{\lambda_1\beta}{\Psi}) + 1 - \frac{\lambda_1\beta}{\Psi})$$

One can write

$$\begin{aligned} S_{23} &= E\left\{[(e^{-\Psi/2}(1-\Psi) + \frac{1}{2}](1 + \frac{\beta}{\beta + C(1-\beta)})\right\} \\ &\quad + E(\frac{1}{2} - (\frac{3}{2})\frac{\beta}{\beta + C(1-\beta)}). \end{aligned}$$

Noting that $\forall\psi$: $e^{-\psi/2}(1-\psi) + 1/2 \geq 1/2 - 2e^{-3/2} = .0537$ and that $C \leq 1$. One has

$$E(\frac{\beta}{\beta + (1-\beta)C}) = \frac{\sqrt{C}}{(1-C)^{3/2}}[\sqrt{\frac{1-C}{C}} - arctg(\sqrt{\frac{1-C}{C}})]$$

A lower bound for S_{23} is then given by

$$1 - 2e^{-3/2} + (-1 - 2e^{-3/2})\frac{\sqrt{C}}{(1-C)^{3/2}}[\sqrt{\frac{1-C}{C}} - arctg(\sqrt{\frac{1-C}{C}})]$$

This is positive as long as C is bigger than 0.7. We now consider the case where $0 \leq C \leq 0.7$. Direct calculation gives

$$\begin{aligned} &E(\exp(-\Psi/2)(1-\Psi)(1 + \frac{\lambda_1\beta}{\Psi})) = \\ &\quad \exp(-\lambda_1 C/2)\int_0^1 \exp(-\lambda_1(1-C)x^2/2)[1 - \lambda_1 C - \lambda_1 x^2(2-C) \\ &\quad + \frac{x^2}{C + (1-C)x^2}]dx \end{aligned}$$

One has

$$\int_0^1 e^{-\lambda_1(1-C)x^2/2}[\lambda_1 x^2(2-C)]dx = \frac{2-C}{1-C}\int_0^1 e^{-\lambda_1(1-C)x^2/2}dx - \frac{2-C}{1-C}e^{-\lambda_1(1-C)/2}$$

Thus

$$\begin{aligned} &E(\exp(-\Psi/2)(1-\Psi)(1 + \frac{\lambda_1\beta}{\Psi})) = \\ &\quad \exp(-\lambda_1 C/2)\int_0^1 \exp(-\lambda_1(1-C)x^2/2)[-\frac{1}{1-C} - \lambda_1 C \\ &\quad + \frac{x^2}{C + (1-C)x^2}]dx + \frac{2-C}{1-C}\exp(-\lambda_1/2) \end{aligned}$$

Writing $\frac{x^2}{C+(1-C)x^2} = \frac{1}{1-C}(1 - \frac{C}{C+(1-C)x^2})$ and rearranging the terms, we get

$$
\begin{aligned}
&E(\exp(-\Psi/2)(1-\Psi)(1+\frac{\lambda_1\beta}{\Psi})) = \\
&\quad -\frac{\sqrt{C}}{(1-C)^{3/2}}\exp(-\lambda_1 C/2)\int_0^1 \exp(-\lambda_1 C(1-C)x^2/(2C)) \\
&\quad \left[\frac{1+\lambda_1 C(1-C)(1+(1-C)x^2/C)}{1+(1-C)x^2/C}\right] d(x\sqrt{(1-C)/C}) + \frac{2-C}{1-C}\exp(-\lambda_1/2)
\end{aligned}
$$

Changing variable $y = (1-C)^{1/2}x/C^{1/2}$ leads to

$$
\begin{aligned}
&E(\exp(-\Psi/2)(1-\Psi)(1+\frac{\lambda_1\beta}{\Psi})) = \\
&\quad -\frac{\sqrt{C}}{(1-C)^{3/2}}\int_0^{\sqrt{\frac{1-C}{C}}} \exp(-\lambda_1 C(1+y^2)/2) \\
&\quad [\frac{1+\lambda_1 C(1-C)(1+y^2)}{1+y^2}]dy + \frac{2-C}{1-C}\exp(-\lambda_1/2)
\end{aligned}
$$

Using the change of variable $z = arctg(y)$ and using $1+\tan(z)^2 = 1/\cos(z)^2$, the last part of the expectation gives

$$
\begin{aligned}
&E(\exp(-\Psi/2)(1-\Psi)(1+\frac{\lambda_1\beta}{\Psi})) = \\
&\quad -\frac{\sqrt{C}}{(1-C)^{3/2}}\int_0^{\sqrt{\frac{1-C}{C}}} \exp(-\lambda_1 C/(2cos^2 z)) \\
&\quad [1+\lambda_1\frac{C(1-C)}{\cos^2 z}]dz + \frac{2-C}{1-C}\exp(-\lambda_1/2)
\end{aligned}
$$

Adding terms, we have

$$
\begin{aligned}
S_{23} &= E[\frac{\sqrt{C}}{(1-C)^{3/2}}\{\sqrt{\frac{1-C}{C}}((2-C)e^{-\lambda_1/2}-C) + \arctan(\sqrt{\frac{1-C}{C}}) \\
&\quad -\int_0^{\arctan(\sqrt{\frac{(1-C)}{C}})} e^{-\frac{\lambda_1 C}{2\cos^2 z}}(1+\frac{\lambda_1 C(1-C)}{\cos^2 z})dz\}] \\
&= E[A(\lambda_1, C)].
\end{aligned}
$$

Since $\lambda_2 > \lambda_3$, $C = \frac{\lambda_2-\lambda_3}{\lambda_1}\beta + \frac{\lambda_3}{\lambda_1}$ has a density $\frac{\lambda_1}{\pi\sqrt{(\lambda_1 c-\lambda_3)(\lambda_2-\lambda_1 c)}}$ in $(\frac{\lambda_3}{\lambda_1}, \frac{\lambda_2}{\lambda_1})$. Then $S_{23} \geq 0$ if $A(\lambda_1, C) \geq 0$.

To evaluate S_{13} we write $\Psi = \lambda_2\beta_2 + (1-\beta_2)\lambda_2 C$, where $C = (\lambda_1 Z_1^2 + \lambda_3 Z_3^2)/(\lambda_2(Z_1^2+Z_3^2))$, and $\beta_2 = Z_2^2/(Z_1^2+Z_2^2+Z_3^2)$. C is in the interval $(\lambda_3/\lambda_2, \lambda_1/\lambda_2)$ which contains 1. Recall that

$$
S_{13} = E(e^{-\Psi/2}(1-\Psi)(1+\frac{\lambda_2\beta_2}{\Psi}) + 1 - \frac{\lambda_2\beta_2}{\Psi}).
$$

Calculation of the expectation relative to β_2 changes according to $C > 1$ or $C < 1$. When $C = c \in [\lambda_3/\lambda_2, 1]$, interchanging λ_1 and λ_2, in S_{23},

$$\begin{aligned}
&E(e^{-\Psi/2}(1-\Psi)(1+\frac{\lambda_2\beta_2}{\Psi})+1-\frac{\lambda_2\beta_3}{\Psi}) \\
&= \{\frac{\sqrt{c}}{(1-c)^{3/2}}\sqrt{\frac{1-c}{c}}[(2-c)e^{-\lambda_2/2}-c]+\arctan(\sqrt{\frac{1-c}{c}}) \\
&\quad -\int_0^{\arctan(\sqrt{\frac{(1-c)}{c}})} e^{-\frac{\lambda_2 c}{2\cos^2 z}}[1+\frac{\lambda_2 c(1-c)}{\cos^2 z}]dz\}
\end{aligned}$$

When $C = c \in [1, \lambda_1/\lambda_2]$, using the same argument as for S_{12}, we first note that $E(e^{-\Psi/2}(1-\Psi)(1+\frac{\lambda_2\beta_2}{\Psi})+1-\frac{\lambda_2\beta_2}{\Psi}) > 0$. Moreover, one has

$$\begin{aligned}
&E(e^{-\Psi/2}(1-\Psi)(1+\frac{\lambda_2\beta_2}{\Psi})+1-\frac{\lambda_2\beta_2}{\Psi}) \\
&= 1-E(\frac{\lambda_2\beta_2}{\Psi})+E[e^{-\Psi/2}(1+\frac{\beta_2}{c-(c-1)\beta_2} \\
&\quad -\lambda_2 c+\lambda_2\beta_2(c-2))],
\end{aligned}$$

and

$$E(\frac{\beta_2}{c-(c-1)\beta_2}) = \int_0^1 \frac{t}{c-(c-1)t}\frac{dt}{2\sqrt{t}}.$$

Changing the variable $y = \sqrt{t}\sqrt{\frac{c-1}{c}}$, and writing $\frac{y^2}{1-y^2} = \frac{1}{2(1-y)} + \frac{1}{2(1+y)} - 1$, direct calculation gives

$$E(\frac{\beta_2}{c-(c-1)\beta_2}) = \frac{\sqrt{c}}{2(c-1)^{3/2}}log(\frac{1+\sqrt{1-1/c}}{1-\sqrt{1-1/c}}) - \frac{1}{c-1}.$$

Also, using a change of variable $x = \sqrt{t}$,

$$\begin{aligned}
&E(e^{-\Psi/2}(1+\frac{\beta_2}{c-(c-1)\beta_2}-\lambda_2 c+\lambda_2\beta_2(c-2)) \\
&= \int_0^1 e^{[-\lambda_2 c+(c-1)\lambda_2 x^2]/2}[1-\lambda_2 c+\frac{x^2}{c-(c-1)x^2}+\lambda_2 x^2(c-2)]dx.
\end{aligned}$$

The density of C is $\frac{\lambda_2}{\pi\sqrt{(\lambda_2 c-\lambda_3)(\lambda_1-\lambda_2 c)}}$. Finally,

$$\begin{aligned}
S_{13} = &\int_{\frac{\lambda_3}{\lambda_2}}^1 \frac{\lambda_2}{\pi\sqrt{(\lambda_2 c-\lambda_3)(\lambda_1-\lambda_2 c)}}\frac{\sqrt{c}}{(1-c)^{3/2}}\{\sqrt{\frac{1-c}{c}}[(2-c)e^{-\lambda_2/2}-c] \\
&+\arctan(\sqrt{\frac{1-c}{c}}) - \int_0^{\arctan(\sqrt{\frac{(1-c)}{c}})} e^{-\frac{\lambda_2 c}{2\cos^2 z}}[1+\frac{\lambda_2 c(1-c)}{\cos^2 z}]dz\}dc \\
&+\int_1^{\frac{\lambda_1}{\lambda_2}} \frac{\lambda_2}{\pi\sqrt{(\lambda_2 c-\lambda_3)(\lambda_1-\lambda_2 c)}}\{\frac{c}{(c-1)} - \frac{\sqrt{c}}{2(c-1)^{3/2}}\log(\frac{1+\sqrt{1-1/c}}{1-\sqrt{1-1/c}}) \\
&+\int_0^1 e^{[-\lambda_2 c+(c-1)\lambda_2 x^2]/2}[1-\lambda_2 c+\frac{x^2}{c-(c-1)x^2}+\lambda_2 x^2(c-2)]dx\}dc.
\end{aligned}$$

which is always positive by numerical methods. ■

Proposition 8 *For* $0 < \lambda_3 = \lambda_2 < \lambda_1$, *the condition 2) of Proposition 1 holds on* $\{(\lambda_1, \lambda_2, \lambda_2) : 0 < \lambda_2 < \lambda_1;\ \frac{25}{\lambda_2 log(1+\lambda_2)} \leq \lambda_1 \leq 7e^{-2.3\lambda_2} + 3.35 + \frac{1}{10}e^{3(\lambda_2-2)}\}$.

PROOF. When $\lambda_1 > \lambda_2 = \lambda_3 > 0$, $C = \lambda_2/\lambda_1$. S_{23} reduces to

$$\begin{aligned} S_{23} &= \frac{\lambda_1\sqrt{\lambda_2}}{(\lambda_1-\lambda_2)^{3/2}}\{\sqrt{\lambda_1/\lambda_2 - 1}[(2-\lambda_2/\lambda_1)e^{-\lambda_1/2} - \lambda_2/\lambda_1] \\ &\quad + \arctan(\sqrt{\lambda_1/\lambda_2 - 1}) - \int_0^{\arctan(\sqrt{\lambda_1/\lambda_2 - 1})} e^{\frac{-\lambda_2}{2\cos^2 z}}[1 + \frac{\lambda_2(\lambda_1-\lambda_2)}{\lambda_1\cos^2 z}]dz\}, \end{aligned}$$

A convenient fit for the region where S_{23} is positive in the set

$$\{(\lambda_1, \lambda_2, \lambda_2) : \frac{25}{\lambda_2\log(1+\lambda_2)} \leq \lambda_1 \leq 7e^{-2.3\lambda_2} + 3.35 + \frac{1}{10}e^{3(\lambda_2-2)}\}. \quad \blacksquare$$

References

[1] Downs, T. D. (1972). Orientations Statistics. *Biometrika*, **59**, 665–676.

[2] Jupp, P. E., and Mardia, K. V. (1989). A unified view of the theory of directional statistics,1975-1988. *Int. Statist. Rev.*, **57**, 261–294.

[3] Khatri, C. G., and Mardia, K. V. (1977). The von Mises-Fisher distribution in orientation statistics. *J. R. Statist. Soc. Ser. B*, **39**, 95–106.

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

[5] Mardia, K. V., and Jupp, P. E. (2000). *Directional Statistics.* John Wiley. New York.

[6] Prentice, M. J. (1986). Orientation statistics without parametic assumptions. *J. Roy. Statist. Soc. Ser. B*, **48**, 214–222.

[7] Rancourt, D., Rivest, L-P., and Asselin, J. (2000). Using orientation statistics to investigate variations in human kinematics. *Appl. Statistics*, **49**, 81–94.

[8] Stephens, M. A. (1979). Vector correlations. *Biometrika*, **66**, 41–48.

A. Chakak and L. Imlahi
Faculty of Sciences, P.B. 2121
University Abdelmalek Essaadi
Tetouan. Morocco