A New Statistical Model for Random Unit Vectors

by

Karim Oualkacha & Louis-Paul Rivest

Departement of Mathematics and Statistics

Université Laval, Ste-Foy

Québec

Canada G1K 7P4

Journal of Multivariate Analysis, 2008, doi:10.1016/j.jmva.2008.03.004

Abstract

This paper proposes a new statistical model for symmetric axial directional data in dimension p. This proposal is an alternative to the Bingham distribution and to the angular central Gaussian family. The statistical properties for this model are presented. An explicit form for its normalizing constant is given and some moments and limiting distributions are derived. The proposed density is shown to apply to the modeling of 3×3 rotation matrices by representing them as quaternions, which are unit vectors in $\Re^4$. The moment estimators of the parameters of the new model are calculated; explicit expressions for their sampling variances are given. The analysis of data measuring the posture of the right arm of subjects performing a drilling task illustrates the application of the proposed model.

Keywords: Axial distribution; Directional data; Multivariate statistics;

Spherical symmetry; Quaternion; Rotation.

1 Introduction

This paper is motivated by the statistical analysis of samples of 3×3 rotation matrices. These matrices are used to characterize the orientations of the limbs of human subjects or the posture of human joints in biomechanics. Recording a 3×3 rotation matrix typically involves two reference frames. The x, y, and z axes of the laboratory reference frame depend on the camera system making the measurements while the local axes are characteristics of the object being measured. When measuring the posture of a limb the local axes typically represent the flexion axis and the direction of the limb. Statistical models for 3×3 rotation matrices are useful to characterize the variability within a sample and to compare several samples of rotation matrices.

The main statistical model for rotation matrices is the exponential family of Downs (1972); some of its properties are reviewed in Khatri and Mardia (1977), Mardia and Jupp (2000) and Chikuse (2002). It has a complicated normalizing constant so that its moments and the maximum likelihood estimator of its shape parameter are relatively difficult to evaluate. The simulation of random rotations following Downs model is not simple. Léon et al. (2006) proposed an alternative density that leads to relatively simple statistical procedures. Its high degree of symmetry makes it unsuitable for many of the samples of rotation matrices found in applications.

This paper constructs a model for 3×3 rotation matrices by proposing a new class of densities for axial unit vectors defined on S^{p-1}. The proposed

model applies to 3×3 rotation matrices since they can be represented as quaternions which are 4×1 unit vectors. Prentice (1986) and Rancourt, Rivest and Asselin (2000) use this representation.

The proposed density is an alternative to the exponential model of Bingham (1974), and to the angular central Gaussian family of Tyler (1987) which are reviewed in Section 9.4 of Mardia and Jupp (2000). Prentice (1986) noted that when a quaternion follows the Bingham distribution, the corresponding 3×3 rotation matrix has the matrix Fisher von Mises distribution. A distribution for 3×3 rotation matrices can be derived in a similar way from the angular Gaussian model.

Section 2 presents the new density in an arbitrary dimension p; it is parameterized by a vector of shape parameters $\gamma \in \Re^{p-1}$ and $M \in SO(p)$, where $SO(p)$ is the set of $p \times p$ rotation matrices. Random unit vectors distributed according to the proposed model are shown to be simple functions of independent random variables having beta distributions. Thus calculating moments and simulating vectors from the new distribution is simple. Section 3 studies the model in dimension 4 for the statistical analysis of a sample of quaternions representing 3×3 rotation matrices. Section 4 gives moment estimators for γ and M and derive their sampling distributions. Section 5 applies this methodology to the drilling data and suggests a goodness-of-fit test.

2 A General Model for Unsigned Unit Directions in S^{p-1}

The proposed density with respect to the Lebesgue measure on S^{p-1} is

$$g_{M,\gamma,p}(r) \;=\; \frac{1}{c_{\gamma,p}} \prod_{k=1}^{p-1} \Big[\sum_{l=1}^{k} (M_l^T r)^2 \Big]^{\gamma_k - \gamma_{k-1}} \qquad r \in S^{p-1},$$

where S^{p-1} is the unit sphere in $\Re^p$, $M = (M_1, \ldots, M_p) \in SO(p)$ is a $p \times p$ rotation matrix, $\gamma_0 = 0$, $\gamma = (\gamma_1, \ldots, \gamma_{p-1})^T \in \Re^{p-1}$, with $\gamma_{p-1} > \gamma_{p-2} > \ldots > \gamma_1 > 0$, $c_{\gamma,p}$ is the normalizing constant, and A^T denotes the transpose the of the matrix A. The constraint that all the γ_k's are different ensures that all the column of the matrix M are identifiable. When $\gamma_k = \gamma_{k+1}$ for $k < p-1$, one cannot distinguish M_k from M_{k+1}. Thus some elements of the parameter M are not estimable. The proposed model is axial since $g_{M,\gamma,p}(r) = g_{M,\gamma,p}(-r)$.

If r is distributed according to $g_{M,\gamma,p}$, then $u = M^T r$ is distributed according to $g_{I_p,\gamma,p}$. This is the density of the reduced model, denoted by $g_{\gamma,p}$, that is given by

$$g_{\gamma,p}(u) \;=\; [c_{\gamma,p}]^{-1} \prod_{k=1}^{p-1} \Big[\sum_{l=1}^{k} u_l^2 \Big]^{\gamma_k - \gamma_{k-1}} \quad u = (u_1, \ldots, u_p)^T \in S^{p-1}. \quad (2.1)$$

The normalizing constant of this model has an explicit form. It is given in the following proposition. All the proofs appear in the Appendix.

Proposition 1: The normalizing constant is given by

$$c_{\gamma,p} \;=\; 2(\pi)^{\frac{p-1}{2}} \prod_{k=1}^{p-1} \frac{\Gamma(\gamma_k + \frac{k}{2})}{\Gamma(\gamma_k + \frac{k+1}{2})}.$$

If the γ_j's are equal with $\gamma_1 = \gamma_2 = \ldots = \gamma_{p-1} = \gamma$, then the model parameters are the unit vector M_1 and a univariate shape parameter γ. The distribution of r is rotationally symmetric about M_1; its density can be written as $g^{rs}_{M_1,\gamma,p}(r)$. The reduced model (2.1) becomes

$$g^{rs}_{\gamma,p}(u) \quad = \quad \frac{\Gamma(\gamma + \frac{p}{2})}{2(\pi)^{\frac{p-1}{2}}\Gamma(\gamma + \frac{1}{2})} u_1^{2\gamma}, \quad u \in S^{p-1}. \tag{2.2}$$

If the common value of γ is 0, one gets the uniform distribution on S^{p-1} and $c_{0,p} = 2\pi^{p/2}/\Gamma(p/2)$ is the Lebesgue measure of S^{p-1}. Observe however that, for any $\gamma > 0$, $g^{rs}_{\gamma,p}(u) = 0$ if $u_1 = 0$. Thus as the shape vector goes to 0, $g_{M,\gamma,p}(r)$ does not converge uniformly to the uniform distribution. Following Watson (1983, p. 92) one can show that the marginal distribution of u_1, is

$$g^{rs}_{\gamma}(u_1) \quad = \quad \frac{\Gamma(\gamma + \frac{p}{2})}{\Gamma(\frac{p-1}{2})\Gamma(\gamma + \frac{1}{2})} u_1^{2\gamma}(1 - u_1^2)^{\frac{p-3}{2}}, \quad u_1 \in [-1, 1],$$

that is u_1^2 follows a $beta(\gamma+1/2, (p-1)/2)$ distribution and that $(u_2, \ldots, u_p)^T/\sqrt{1 - u_1^2}$ is uniformly distributed in S^{p-2}.

When $p = 2$, (2.1) becomes

$$g_{\gamma,2}(u_1, u_2) = \frac{\Gamma(\gamma + 1)}{2\sqrt{\pi}\Gamma(\gamma + \frac{1}{2})} u_1^{2\gamma}, \qquad (u_1, u_2)^T \in S^1. \tag{2.3}$$

This is related to the circular beta density with parameters $(\gamma + 1/2, 1/2)$, see Jammalamadaka & SenGupta (2001, p. 51), whose density is given by

$$g_{\gamma,2}(\theta) = \frac{\Gamma(\gamma + 1)}{2^{\gamma+1}\sqrt{\pi}\Gamma(\gamma + \frac{1}{2})} \left[1 + \cos(\theta)\right]^{\gamma}, \qquad -\pi \leq \theta \leq \pi.$$

If θ has this circular beta density, then $u = (\cos(\theta/2), \epsilon \sin(\theta/2))^T$ is distributed according to (2.3) where ϵ is uniformly distributed on $\{-1, 1\}$.

The distribution of u_p, the last component of u, in (2.1) can be determined using Watson's (1983, p. 44) parametrization of S^{p-1},

$$u = t\begin{pmatrix} 0 \\ 1 \end{pmatrix} + \sqrt{1-t^2}\begin{pmatrix} v \\ 0 \end{pmatrix}, \quad t \in [-1,1], \quad v \in S^{p-2},$$

whose Jacobian is $du = (1-t^2)^{\frac{p-3}{2}} dt dv$. Thus the joint density of (t, v) is

$$g_{\gamma,p}(t,v) = [c_{\gamma,p-1}]^{-1} \prod_{k=1}^{p-2} \left[\sum_{l=1}^{k} v_l^2\right]^{\gamma_k - \gamma_{k-1}} \frac{\Gamma(\gamma_{p-1} + \frac{p}{2})}{\sqrt{\pi}\Gamma(\gamma_{p-1} + \frac{p-1}{2})}(1-t^2)^{\gamma_{p-1} + \frac{p-3}{2}},$$

where $v \in S^{p-2}$ and $t \in [-1,1]$. Thus t and v are independent, the marginal density of v is $g_{\gamma,p-1}(v)$, with $\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_{p-2})^T$, and the marginal distribution of t is given by

$$f_T(t) = \frac{\Gamma(\gamma_{p-1} + \frac{p}{2})}{\sqrt{\pi}\Gamma(\gamma_{p-1} + \frac{p-1}{2})}(1-t^2)^{(\frac{p-3}{2} + \gamma_{p-1})}, \qquad t \in [-1,1].$$

This is the density function of $(2\beta_{p-1} - 1)$, where β_{p-1} is distributed according to a $beta(\gamma_{p-1} + (p-1)/2, \gamma_{p-1} + (p-1)/2)$. Hence, u satisfies

$$u \stackrel{d}{=} \begin{pmatrix} 2\sqrt{\beta_{p-1}(1-\beta_{p-1})}v \\ (2\beta_{p-1} - 1) \end{pmatrix},$$

where $\stackrel{d}{=}$ means equality in distribution. In a similar way, one can write the distribution of the last entry of v in terms of a beta random variable. Iterating this procedure proves the following proposition.

Proposition 2: Let β_j be independent random variables distributed according to $beta(\gamma_j + j/2, \gamma_j + j/2)$ distributions, for $j = 1, \ldots, p-1$ and let ϵ be distributed according to the discrete uniform distribution on {-1,1}, then

the unit vector

$$u = \begin{pmatrix} 2^{p-1}\prod_1^{p-1}\sqrt{\beta_j(1-\beta_j)}\epsilon \\ . \\ . \\ 2^{p-k}\prod_k^{p-1}\sqrt{\beta_j(1-\beta_j)}(2\beta_{k-1}-1) \\ . \\ . \\ (2\beta_{p-1}-1) \end{pmatrix}_{p\times 1}, \tag{2.4}$$

is distributed according to $g_{\gamma,p}$.

Proposition 2 shows that, starting from independent beta random variables, a random vector distributed according to the proposed distribution is easily constructed. If we let $u^{(k)} = (u_1, \ldots, u_k)^T$, for $k = 1, \ldots, p$, then from (2.4), we can write $u^{(k)}$ as

$$u^{(k)} = c_k v^{(k)}, \tag{2.5}$$

where $c_k = \sqrt{u_1^2 + \ldots + u_k^2} = 2^{p-k}\prod_k^{p-1}\sqrt{\beta_j(1-\beta_j)}$ and $v^{(k)} \in S^{k-1}$. Since c_k is a function of $\beta_k, \ldots, \beta_{p-1}$ and $v^{(k)}$ depends only on $\beta_{k-1}, \ldots, \beta_1$, the random variable c_k is independent of the unit vector $v^{(k)}$, which is distributed according to $g_{\gamma,k}$, with $\gamma = (\gamma_1, \ldots, \gamma_{k-1})^T$.

If in (2.4) we let $y_j = 4\beta_j(1-\beta_j)$, then one can show that y_j is distributed according to a $beta(\gamma_j + j/2, 1/2)$. Thus an alternative form for (2.4) is

$$u_k = \left(\prod_{j=k}^{p-1}\sqrt{y_j}\right)\sqrt{1-y_{k-1}}\epsilon_k, \quad k = 1, 2, \ldots, p,$$

where ϵ_k's are random variables distributed according to the discrete uniform distribution in {-1,1}, $y_0 = 0$, and the product is equal to one when $k = p$.

2.1 Limiting Cases

This section derives limiting distributions obtained when some elements of the shape parameter vector γ go to infinity. The derivations rely on the following result. If $\gamma_j = \alpha_j \tau$, then as τ goes to infinity,

$$\begin{aligned} \sqrt{\tau}(2\beta_j - 1) &\xrightarrow{d} N\left(0, \frac{1}{2\alpha_j}\right), \\ \sqrt{\beta_j(1-\beta_j)} &\xrightarrow{prob} \frac{1}{2}, \end{aligned}$$

where β_j is distributed according to a $beta(\gamma_j + j/2, \gamma_j + j/2)$. Together with (2.4), these results can be used to derive the following limiting distribution.

Proposition 3: Suppose that γ_j is fixed, for $j = 1, \ldots, k-1$ and $\gamma_j = \alpha_j \tau$, for $j = k, \ldots, p-1$, for some $1 \leq k \leq p$. If u is distributed as $g_{\gamma,p}$ then, as $\tau \to \infty$,

1. The limiting density of $(u_1, \ldots, u_k)^T$ is $g_{\gamma,k}(.)$, with $\gamma = (\gamma_1, \ldots, \gamma_{k-1})$;
2. The vector $\sqrt{\tau}(u_{k+1}, \ldots, u_p)^T$ converges in distribution to a $N_{p-k}\left(0, diag\left(\frac{1}{2\alpha_j}\right)\right)$.

When $k = 1$, $|u_1|$ tends to 1 in probability and u is distributed in one of the two hyperplanes tangent to S^{p-1} at $(\pm 1, 0, \ldots, 0)^T$. When $k > 1$, the unit vector u is distributed close to the subspace of S^{p-1} of dimension $k-1$ defined by the equation $u_1^2 + \ldots + u_k^2 = 1$. The distance between u and this subspace is characterized by $(u_{k+1}, \ldots, u_p)^T$ that has a limiting normal distribution.

2.2 A Closure Property

Suppose that given $x \in S^{p-1}$, the random vector r has a rotationally symmetric density about x, $g^{rs}_{\gamma,p}(r^T x)$ which is given in (2.2). Now suppose that x

is uniformly distributed in a q dimensional subspace of S^{p-1}. Then $x = Uv$, where $U = (U_1, \ldots, U_q)_{p\times q}$, $p > q$, $U^T U = I_q$ and v is uniform in S^{q-1}. The marginal distribution of r is given by

$$\begin{aligned} g(r) &= \int_{S^{q-1}} \frac{\Gamma(\gamma + p/2)\Gamma(q/2)}{4\pi^{(p+q-1)/2}\Gamma(\gamma + 1/2)} (r^T U v)^{2\gamma} dv \\ &= \left(\sqrt{r^T U U^T r}\right)^{2\gamma} \frac{\Gamma(\gamma + p/2)\Gamma(q/2)}{4\pi^{(p+q-1)/2}\Gamma(\gamma + 1/2)} \int_{S^{q-1}} \left(\frac{v^T U^T r}{\sqrt{r^T U U^T r}}\right)^{2\gamma} dv \\ &= \frac{\Gamma(\gamma + p/2)\Gamma(q/2)}{2\pi^{p/2}\Gamma(\gamma + q/2)} \left\{\sum_{i=1}^{q} (U_i^T r)^2\right\}^{\gamma}. \end{aligned}$$

This is the reduced model $g_{\gamma^*,p}(r)$ where the first $q-1$ components of the shape parameters γ^* are equal to γ while its last $p-q$ components are 0. Such models are considered in Chapter 5 of Watson (1983). The competing models of Bingham and Tyler do not satisfy such a closure property.

2.3 Moment Calculations

The moments of the unit vector u distributed as $g_{\gamma,p}$ are given next. As shown in the Appendix, they are derived from (2.4), by evaluating moments of beta random variables.

Proposition 4: Let u be distributed according to $g_{\gamma,p}(u)$, where the $p-1$ entries of γ satisfy $\gamma_{p-1} > \gamma_{p-2} > \ldots > \gamma_1 > 0$; the matrix of second order moments of u is given by $E(uu^T) = diag(\lambda_k)$, where $\lambda_k = E(u_k^2)$ is given by

$$\lambda_k = \frac{1}{2(\gamma_{k-1} + \frac{k}{2})} \prod_{j=k}^{p-1} \frac{(\gamma_j + \frac{j}{2})}{(\gamma_j + \frac{j+1}{2})} \quad \text{and} \quad \lambda_1 > \lambda_2 > \ldots > \lambda_p,$$

and $\gamma_0 = 0$. Moreover,

$$E(u_k^4) = \frac{3}{4(\gamma_{k-1} + \frac{k}{2})(\gamma_{k-1} + \frac{k+2}{2})} \prod_{j=k}^{p-1} \frac{(\gamma_j + \frac{j}{2})(\gamma_j + \frac{j+2}{2})}{(\gamma_j + \frac{j+1}{2})(\gamma_j + \frac{j+3}{2})},$$

$$
\begin{aligned}
E(u_k^2 u_l^2) &= \frac{1}{4(\gamma_{k-1}+\frac{k}{2})(\gamma_{l-1}+\frac{l+2}{2})} \prod_{j=k}^{p-1} \frac{(\gamma_j+\frac{j}{2})}{(\gamma_j+\frac{j+1}{2})} \prod_{j=l}^{p-1} \frac{(\gamma_j+\frac{j+2}{2})}{(\gamma_j+\frac{j+3}{2})}, \quad k<l, \\
&= \frac{\lambda_k}{2(\gamma_{l-1}+\frac{l+2}{2})} \prod_{j=l}^{p-1} \frac{(\gamma_j+\frac{j+2}{2})}{(\gamma_j+\frac{j+3}{2})}, \quad k<l, \qquad (2.6) \\
E(u_k) &= E(u_k u_l) = E(u_k^3 u_l) = 0, \quad k \neq l,
\end{aligned}
$$

where the product is equal to 1 when $k = p$.

Let $r = Mu$, then the matrix of second order moments of r is given by

$$
E(rr^T) = M diag(\lambda_1, \ldots, \lambda_p) M^T, \qquad (2.7)
$$

where $\lambda_1 > \ldots > \lambda_p > 0$ are the eigenvalues of $E(rr^T)$. Furthermore the jth column of M, M_j, is the eigenvector associated with λ_j.

3 The Model in the Special Case $p = 4$

When $p = 4$, $g_{M,\gamma,p}$ gives a model for quaternions, a representation of 3×3 rotation matrices. This section investigates the application of the proposed model to 3×3 rotation matrices. First, the correspondence between 3×3 rotation matrices and quaternions is reviewed in Section 3.1. To our knowledge $p = 4$ is the only instance of such a correspondance between unit vectors and rotation matrices.

3.1 3×3 Rotation Matrices and Quaternions

Let $R(\theta, \mu)$ denote a rotation of angle θ, $\theta \in (-\pi, \pi]$, around the unit vector μ in $\Re^3$. We have

$$\begin{aligned} R(\theta, \mu) &= \exp \mathcal{S}(\theta\mu) = I_3 + \mathcal{S}(\theta\mu) + \mathcal{S}(\theta\mu)^2/2 + ... \\ &= \cos\theta I_3 + \sin\theta \mathcal{S}(\mu) + (1 - \cos\theta)\mu\mu^t, \end{aligned}$$

where $\mathcal{S}(\mu)$ is the skew-symmetric matrix corresponding to $\mu = (\mu_1, \mu_2, \mu_3)^T$, given by

$$\mathcal{S}(\mu) = \begin{pmatrix} 0 & -\mu_3 & \mu_2 \\ \mu_3 & 0 & -\mu_1 \\ -\mu_2 & \mu_1 & 0 \end{pmatrix}.$$

The quaternion associated with $R(\theta, \mu)$ is a unit vector in $\Re^4$ defined by $q(\theta, \mu) = (\cos(\theta/2), \sin(\theta/2)\mu^T)^T$ (Hamilton, 1969). Note that, $q(\theta, \mu) = -q(\theta + 2\pi, \mu)$, so that q and $-q$ represent the same rotation. The rotation matrix R can be expressed in terms of its quaternion q as (Prentice, 1986),

$$R = \Phi(q) = \begin{pmatrix} q_1^2 + q_2^2 - q_3^2 - q_4^2 & 2(q_2q_3 - q_1q_4) & 2(q_1q_3 + q_2q_4) \\ 2(q_1q_4 + q_2q_3) & q_1^2 + q_3^2 - q_2^2 - q_4^2 & 2(q_3q_4 - q_1q_2) \\ 2(q_2q_4 - q_3q_1) & 2(q_3q_4 + q_1q_2) & q_1^2 + q_4^2 - q_2^2 - q_3^2 \end{pmatrix} \quad (3.1)$$

Quaternions are endowed with a special product corresponding to rotation multiplication. Let p and q be the quaternions for the rotation matrices R_1 and R_2 respectively. As mentioned in McCarthy (1990, p. 61), the quaternion for the product R_1R_2 is the vector $P_+q = Q_-p$, where P_+ and Q_- are 4×4 rotation matrices defined by

$$P_+ = p_1 I_4 + S_+(p_2, p_3, p_4), \qquad Q_- = q_1 I_4 + S_-(q_2, q_3, q_4), \qquad (3.2)$$

and

$$\mathcal{S}_+(x) = \begin{pmatrix} 0 & -x^T \\ x & \mathcal{S}(x) \end{pmatrix}, \quad \mathcal{S}_-(x) = \begin{pmatrix} 0 & -x^T \\ x & -\mathcal{S}(x) \end{pmatrix}, \quad x \in \Re^3.$$

Observe that ${}^t p = (p_1, -p_2, -p_3, -p_4)^T$ is the quaternion for the rotation matrix inverse of R_1,R_1^T. Thus, $P_+^T q$ is the quaternion for $R_1^T R_2$, moreover, $P_+^T q = Q_-({}^t p)$.

Moran (1976) and Kim (1991) observed that if the rotation matrix R is distributed according to the uniform distribution in $SO(3)$ then its quaternion r is such that ϵr is uniformly distributed on the unit sphere S^3 where ϵ takes the values -1 and $+1$ with a probability of $1/2$. Thus the jacobian of the transformation that maps the upper half sphere of S^3 into $SO(3)$ is 1.

Any 4×4 rotation matrix $M = (M_{ij})_{1 \le i,j \le 4}$, can be written as the matrix product $P_+ Q_-$, where P_+ and Q_- are derived from the quaternions p and q as in (3.2). Given M, we can find p and q as follows

$$\begin{aligned} p_1 &= \frac{1}{4}\sqrt{A_1^2 + A_2^2 + A_3^2 + [\mathrm{tr}(M)]^2}, \\ q_1 &= \frac{\mathrm{sign}\{\mathrm{tr}(M)\}}{4}\sqrt{B_1^2 + B_2^2 + B_3^2 + [\mathrm{tr}(M)]^2}, \end{aligned}$$

$$\begin{pmatrix} p_2 \\ p_3 \\ p_4 \end{pmatrix} = -\frac{1}{4q_1}\begin{pmatrix} B_1 \\ B_2 \\ B_3 \end{pmatrix}, \qquad \begin{pmatrix} q_2 \\ q_3 \\ q_4 \end{pmatrix} = -\frac{1}{4p_1}\begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix},$$

where sign(x) is -1 if x is negative and 1 otherwise and

$$\begin{aligned}
A_1 &= M_{12} - M_{21} - M_{34} + M_{43}, \\
A_2 &= M_{13} - M_{31} + M_{24} - M_{42}, \\
A_3 &= M_{14} - M_{41} - M_{23} + M_{32}, \\
B_1 &= M_{12} - M_{21} + M_{34} - M_{43}, \\
B_2 &= M_{13} - M_{31} - M_{24} + M_{42}, \\
B_3 &= M_{14} - M_{41} + M_{23} - M_{32}.
\end{aligned}$$

These results are derived by noting that $\mathrm{tr} P_+ Q_- = 4p_1q_1$ and that $q_1 S_+(p_2, p_3, p_4) + p_1 S_-(q_2, q_3, q_4)$ is the skew-symmetric part of P_+Q_-.

3.2 Moment Calculations

Let r be a quaternion distributed according to $g_{M,\gamma,4}$ and let R be the rotation matrix associated to r. We have $r = Mu$, $M \in SO(4)$. From Section 3.1, there exist two quaternions p and q such as $r = P_+Q_-u = P_+U_+q$, where U_+ is a 4×4 rotation matrix, associated to u by (3.2). In terms of 3×3 rotation matrices, this relationship can be written as $R = PUQ$, where $P = \Phi(p)$, $U = \Phi(u)$ and $Q = \Phi(q)$, are the 3×3 rotation matrices associated to quaternions p, u and q respectively and $\Phi(.)$ is given in (3.1). Since u is distributed as $g_{\gamma,p}$, equation (3.1) and Proposition 2 imply that $E(U)$ is a diagonal matrix whose elements can be expressed in term of the second moments λ_k of Proposition 4. Consequently, we can write

$$E(R) = PE(U)Q \quad = \quad P\ diag \begin{pmatrix} \lambda_1 + \lambda_2 - \lambda_3 - \lambda_4 \\ \lambda_1 + \lambda_3 - \lambda_2 - \lambda_4 \\ \lambda_1 + \lambda_4 - \lambda_2 - \lambda_3 \end{pmatrix} Q,$$

see also Section 4 of Prentice (1986). This is the singular value decomposition for $E(R)$. The fact that $\lambda_1 \geq \ldots \geq \lambda_4 \geq 0$ implies that its singular values satisfy $E(U_{11}) > E(U_{22}) > |E(U_{33})|$. We conclude that the mean rotation is PQ see (Rivest, Rancourt, and Asselin 2000). The corresponding quaternion is $P_+q = M_1$, where M_1 is the first column of the 4×4 rotation matrix M. This is the eigenvector associated to the largest eigenvalue λ_1 of $E(rr^T)$.

When $\gamma_1 = \gamma_2 = \gamma_3 = \gamma$, the reduced model in (2.1) becomes $g^{sr}_{\gamma,4}(u)$ given in (2.2). Using the transformation $U = \Phi(u)$ given in (3.1), that has the Jacobian $[dU] = du/(2\pi^2)$ where $[dU]$ is the unit invariant measure on $SO(3)$. One can write (2.1) in terms of 3×3 rotation matrices as

$$g_\gamma(U) \quad = \quad \frac{\sqrt{\pi}\Gamma(\gamma+2)}{2^{2\gamma}\Gamma(\gamma+\frac{1}{2})}\left[1+tr(U)\right]^\gamma .$$

This is equal to the model of León, Rivest and Massé (2006) when $p = 3$.

3.3 A Great Circle Model

When modeling rotational data it may happen that λ_3 and λ_4 are very close to 0. For these models, γ_2 and γ_3 are large and the unit vector r takes its value in a great circle of S^3. In this case, the standardized quaternion u satisfies $u \approx (u_1, u_2, 0, 0)^T$, where $(u_1, u_2)^T \sim g_{\gamma,2}(u_1, u_2)$, see (2.3). Thus $r = Mu$ can be written as

$$\begin{aligned} r &\approx \cos(\theta/2)M_1 + \sin(\theta/2)M_2 \\ &= [M_1]_+ \left\{\cos(\theta/2)(1,0,0,0)^T + \sin(\theta/2)[M_1]_+^T M_2\right\}, \end{aligned}$$

where θ has a circular beta distribution with parameters $(\gamma + 1/2, 1/2)$. One has $[M_1]_+^T M_2 = (0, \mu^T)^T$, where μ is a S^2 vector since $[M_1]_+^T M_2$ is a unit vector

in $\Re^4$ whose first component is null. In terms of 3×3 rotation matrices, the above expression for r is $R = R_0 R(\theta, \mu)$, where R_0 is the rotation matrix corresponding to M_1 and $R(\theta, \mu)$ is the rotation matrix corresponding to the quaternion $(\cos\theta/2, \sin\theta/2\mu^T)^T$. This is a situation where the variability in R can be expressed as rotations around a fixed axis μ; the rotation angles have a circular beta distribution when r is distributed according to $g_{M,\gamma,4}$.

From a geometrical point of view, μ is the rotation axis in the so called local reference frame. An alternative expression for the fixed axis model, with respect to the rotation axis $R_0\mu$ in the laboratory reference frame, is $R = R(\theta, R_0\mu)R_0$. Fixed axis models for rotation matrices are investigated in Rivest (2001).

4 Parameter Estimation

Consider $\{r_1, r_2, \ldots, r_n\}$, a sample of unit vectors in $\Re^p$ distributed according to $g_{M,\gamma,p}(r)$, where $\gamma \in \Re^{p-1}$ and $M \in SO(p)$ are unknown parameters. This section discusses the estimation of γ and M. Moment estimators for γ and M which are functions of the sample cross-product matrix $\sum r_i r_i^T/n$ are derived; their asymptotic distributions are calculated.

This section emphasizes the method of moments to estimate parameters because it is simple and it has a large efficiency. The information matrix for the parameters of γ and M when $p = 4$, is given in Oualkacha (2004, Section 4.3). It shows that the efficiency of the moment estimators of γ and M is greater than 90% when the components of γ are relatively large, i. e. $(\gamma_1 > 2, \gamma_2 > 4)$. For the rotationally symmetric models, the efficiency of

the moment estimators is calculated in section 5.2 of León Rivest and Massé (2006), it is greater than 90% when $\gamma > 4$. This suggests that the lost of information associated with the moment estimators is small, especially when the data is clustered around its first principal direction.

4.1 Moment Estimators

The estimating equation for (M, γ) is $\hat{B} = E(rr^T)$, where $E(rr^T)$ is given in (2.7) and $\hat{B} = \sum_i^n r_i r_i^T/n$. The matrix $\hat{B}$ is positive definite; its spectral decomposition is

$$\hat{B} \;=\; \frac{1}{n}\sum_i^n r_i r_i^T = \hat{M}\big[diag(\hat{\lambda}_j)\big]_{1\le j\le p}\hat{M}^T, \tag{4.1}$$

where $\hat{M} = (\hat{M}_1, \hat{M}_2, \ldots, \hat{M}_p)$ is a matrix of eigenvectors associated to the eigenvalues $\hat{\lambda}_1 > \hat{\lambda}_2 > \ldots > \hat{\lambda}_p$. Consequently, the moment estimator of M is $\hat{M}$ and the moment estimator of γ, $\hat{\gamma}$, is defined implicitly by the equations $\hat{\lambda}_j = \lambda_j$, for $j = 1, \ldots, p$, when λ_j is defined in Proposition 4. The solution to these equations is

$$\hat{\gamma}_k = \frac{1}{2}\Big(\frac{\sum_{j=1}^k \hat{\lambda}_j}{\hat{\lambda}_{k+1}} - k\Big), \quad k = 1, 2, \ldots, p-1.$$

These moment estimates satisfy $\hat{\gamma}_{k+1} = \hat{\gamma}_k\hat{\lambda}_{k+1}/\hat{\lambda}_{k+2} + (k+1)(\hat{\lambda}_{k+1} - \hat{\lambda}_{k+2})/(2\hat{\lambda}_{k+2})$. This implies that $\hat{\gamma}_1 < \hat{\gamma}_2 < \ldots < \hat{\gamma}_{p-1}$.

The asymptotic distributions of $\hat{\gamma}$ and $\hat{M}$ are now derived. For this, let $m = vect(m_{jk})_{1\le j<k\le p}$ a vector in $\Re^{(p-1)p/2}$ close to zero, so

$$\begin{aligned} M\exp\big(\mathcal{S}(m)\big) &= M(I_p + \mathcal{S}(m) + \frac{\mathcal{S}(m)^2}{2!} + \cdots) \\ &= M(I_p + \mathcal{S}(m) + o(m)) \\ &\approx M(I_p + \mathcal{S}(m)), \end{aligned}$$

describes the rotation about M, where $\mathcal{S}(m)$ is a $p \times p$ skew-symmetric matrix containing the entries of m, such that $\mathcal{S}(m)_{jk} = m_{jk}$ for $1 \leq j < k \leq p$. Thus

$$M^T \hat{M} = I_p + \mathcal{S}(\hat{m}), \tag{4.2}$$

where $\hat{m} = vect(\hat{m}_{jk})_{1 \leq j < k \leq p}$ measures the discrepancy between M and $\hat{M}$. The asymptotic distributions of $\hat{\gamma}$ and $\hat{M}$ are given in the next proposition which is proved in Appendix.

Proposition 5: As the sample size n becomes large, we have

$$\textbf{i)} \;\; n^{1/2}\big(\hat{\gamma} - \gamma\big) \;\; \rightarrow \;\; N_{p-1}(0_{p-1}, \Sigma_\gamma),$$

where Σ_γ is a $(p-1) \times (p-1)$ diagonal matrix whose diagonal entries are given by

$$\Sigma_\gamma(k,k) \;\; = \;\; \frac{(\gamma_k + \frac{k}{2})(\gamma_k + \frac{k+1}{2})}{\lambda_{k+1}(\gamma_k + \frac{k+3}{2})} \prod_{j=k+1}^{p-1} \frac{(\gamma_j + \frac{j+2}{2})}{(\gamma_j + \frac{j+3}{2})},$$

where the product is equal to 1 when $k = p-1$.

$$\textbf{ii)} \;\; n^{1/2}\hat{m} \rightarrow N_{\frac{(p-1)p}{2}}(0_{\frac{(p-1)p}{2}}, \Sigma_m),$$

where

$$\Sigma_m = diag\left\{\Sigma_{m_{kl}}\right\}_{\frac{(p-1)p}{2} \times \frac{(p-1)p}{2}}, \quad 1 \leq k < l \leq p,$$

where $\Sigma_{m_{kl}}$ is the variance of the component $\hat{m}_{kl}$ of $\hat{m}$ that is given by

$$\Sigma_{m_{kl}} \;\; = \;\; \frac{\lambda_k}{2(\lambda_l - \lambda_k)^2(\gamma_{l-1} + \frac{l+2}{2})} \prod_{j=l}^{p-1} \frac{(\gamma_j + \frac{j+2}{2})}{(\gamma_j + \frac{j+3}{2})}, \quad 1 \leq k < l \leq p.$$

iii) $\hat{\gamma}$ and $\hat{m}$ are asymptotically independent.

The small sample biases of the asymptotic variances given in the above proposition have been investigated in a Monte-Carlo study that is not reported here. When $n \geq 50$ $\hat{\Sigma}_j(k,k)/\hat{\lambda}_k^2$ provides reliable variance estimates for $\log \hat{\lambda}_k$, where $\hat{\Sigma}_j(k,k)$ is the plug-in variance estimate. The variance estimates obtained from Proposition 5 *ii*) also have small biases when $n \geq 50$. For small sample sizes, the parametric bootstrap can be used to estimate the variances.

4.2 Estimation of the fixed axis model when $p = 4$

When $p = 4$ and when γ_2 and γ_3 are large, one has a fixed-axis model for the 3×3 rotation matrices as discussed in Section 3.3. This axis is estimated by $\hat{\mu}$, the vector of the second, the third and the fourth entries of $[\hat{M}_1]_+^T \hat{M}_2$. The asymptotic distribution of $\hat{\mu}$ is given next.

Proposition 6: As the sample size n becomes large, we have

$$n^{1/2}\big(\hat{\mu} - \mu\big) \;\; \rightarrow \;\; N_3(0, \Sigma_\mu),$$

where Σ_μ is given by

$$\Sigma_\mu = \left[\Sigma_{m_{23}} + \Sigma_{m_{14}}\right] \mu_1 \mu_1^T + \left[\Sigma_{m_{13}} + \Sigma_{m_{24}}\right] \mu_2 \mu_2^T,$$

where $(0, \mu_1^T)^T = [M_1]_+^T M_3$ and $(0, \mu_2^T)^T = [M_1]_+^T M_4$.

When γ_2 and γ_3 are large a convenient expression for this covariance matrix is

$$\Sigma_\mu = \left\{\frac{\lambda_3}{\lambda_2} + \frac{\lambda_4}{\lambda_1}\right\} \mu_1 \mu_1^T + \left\{\frac{\lambda_4}{\lambda_2} + \frac{\lambda_3}{\lambda_1}\right\} \mu_2 \mu_2^T + o(\frac{1}{\gamma_2}).$$

When $\gamma_2 = \gamma_3$, $\lambda_3 = \lambda_4$ and this expression coincides with the variance estimate given in Section 4.1 of Rivest (2001).

5 Data analysis

To illustrate the methodology presented in this paper, we fit the proposed model to the data collected from the experiment given in Rancourt et al. (2000). The sample consists of $n = 30$ observations that measure the orientations of the upper right arm of a subject performing drilling tasks. The arm pose is defined via one marker attached in the arm. The marker orientation is characterized by a 3×3 rotation matrix $R = [\mu_x, \mu_y, \mu_z]$, where μ_x, μ_y and μ_z are the orientations of the local's x, y and z axes of the marker in the laboratory coordinate system. When resting, the arm is in a vertical position, the local x axis then points backward, the local y axis goes upward and the local z axis points left. Thus the local y-axis is the direction of the upper arm, and the local z-axis is the rotation axis of the elbow. The subject is asked to point a drill at various targets 30 times. The rotation matrices in the sample record the orientations of the local coordinate system at each repetition. The $n = 30$ quaternions for the sample 3×3 rotation matrices are given in Table 1.

The moment estimators of $\log \gamma_j$'s and their parametric bootstrap standard errors are $\log \hat{\gamma}_1 = 2.60 \quad s.e. = 0.28$, $\log \hat{\gamma}_2 = 5.35 \ s.e. = 0.28$, and $\log \hat{\gamma}_3 = 8.10 \ s.e. = 0.31$. The large sample standard errors derived from Proposition 5 are 10% to 20% smaller than those obtained with the parametric bootstrap. Since the $\hat{\gamma}_2$ and $\hat{\gamma}_3$ are large we have a fixed axis model. Thus $R_i = \hat{R}_0 R(\theta_i, \hat{\mu})$ and the variability of R_i in the local coordinate system is characterized by θ_i that has a circular beta distribution with parameters $(\hat{\gamma}_1 + 1/2, 1/2)$ around the fixed axis $\hat{\mu}$. Since $\hat{\gamma}_1 = 13.42$, the range of possible values for θ_i ± 40 degrees, with a probability of 95%.

r_i	r_{i1}	r_{i2}	r_{i3}	r_{i4}
r_1	0.664	0.193	0.390	-0.608
r_2	-0.623	-0.167	-0.416	0.640
r_3	-0.605	-0.195	-0.425	0.644
r_4	-0.602	-0.178	-0.416	0.657
r_5	-0.562	-0.276	-0.480	0.614
r_6	0.791	0.098	0.369	-0.477
r_7	0.802	0.056	0.391	-0.448
r_8	0.755	0.098	0.381	-0.525
r_9	0.789	0.079	0.371	-0.483
r_{10}	0.732	0.109	0.393	-0.545
r_{11}	0.859	0.067	0.395	-0.318
r_{12}	0.853	0.042	0.372	-0.364
r_{13}	0.866	0.023	0.364	-0.341
r_{14}	0.829	0.033	0.366	-0.421
r_{15}	0.852	0.054	0.361	-0.374

r_i	r_{i1}	r_{i2}	r_{i3}	r_{i4}
r_{16}	0.920	0.059	0.368	-0.123
r_{17}	0.895	0.034	0.360	-0.260
r_{18}	0.910	0.050	0.378	-0.170
r_{19}	0.916	0.043	0.355	-0.181
r_{20}	0.926	0.008	0.333	-0.178
r_{21}	0.795	0.053	0.386	-0.464
r_{22}	0.780	0.042	0.344	-0.521
r_{23}	0.772	0.064	0.355	-0.523
r_{24}	0.791	0.016	0.352	-0.500
r_{25}	0.701	0.104	0.383	-0.593
r_{26}	0.876	0.009	0.349	-0.332
r_{27}	0.850	0.045	0.358	-0.383
r_{28}	0.837	0.039	0.380	-0.391
r_{29}	0.898	0.005	0.334	-0.285
r_{30}	0.874	0.055	0.351	-0.330

Table 1: Sample of $n = 30$ quaternions for the right arm pose in a drilling task.

The moment estimator of M_1 is $\hat{M}_{11} = 0.813$ $s.e. = 0.017$, $\hat{M}_{12} = 0.077$ $s.e. = 0.011$, $\hat{M}_{13} = 0.383$ $s.e. = 0.006$ and $\hat{M}_{14} - 0.431$ $s.e. = 0.027$, while the moment estimator of the fixed axis is $\hat{\mu}_1 = -0.524$ $s.e. = .019$, $\hat{\mu}_2 = -0.365$ $s.e. = .043$, and $\hat{\mu}_3 = 0.773$ $s.e. = .029$. These standard errors were evaluated using the parametric bootstrap. Since the largest entry of $\hat{\mu}$ is the third one, the arm changes its posture by moving about an axis closed

to the z-axis. From Proposition 2, the angle of the residual rotation not explained by the fixed axis model has an $N\{0, (2\hat{\gamma}_2)^{-1}\}$ distribution. The standard deviations is 3.9 degrees; this highlights that the residual rotation is small.

To interpret this analysis one must bear in mind that a change of the orientation of the upper arm is the composition of a rotation of the back plus a motion of the shoulder. For the subject considered here, the back did not move much since the analysis of the rotation data obtained from the back marker gives $\hat{\gamma}_1 = 113, s.e. = 29$. Most of the changes in orientation take place at the shoulder joint. The changes in the posture of this joint occur mostly through rotations about $\hat{\mu}$ which is relatively close to the z axis. During the experiment the upper arm stays in a plane close to the $z = 0$ plane that is spanned by the x (backward direction) and the y (upward direction) axis.

We now investigate the fit of the model. Since $\hat{\gamma}_2$ and $\hat{\gamma}_3$ are large, the centered quaternions satisfy $u_i \approx (\cos(\theta_i/2), \sin(\theta_i/2), 0, 0)^T$, where $\cos^2(\theta_i/2)$ is distributed according to a $beta(\hat{\gamma}_1 + 1/2, 1/2)$. A goodness of fit test for the proposed distribution amounts to testing whether $\{\cos^2(\theta_i/2)\}$ has a $beta(\hat{\gamma}_1 + 1/2, 1/2)$ distribution. First note that $\cos^2(\theta_i/2)$ is estimated by $(\hat{M}_1 r_i)^2$; the $beta(\hat{\gamma}_1 + 1/2, 1/2)$ Q-Q plot is given in Figure 1.

The beta distribution fits reasonably well. To carry out formal goodness of fit tests, we use the correlation coefficient in the Q-Q plot and the Kolmogorov-Smirnov statistic. The observed value for these two statistics are 0.967 and 0.157 respectively. To calculate p-values, we use the parametric bootstrap. The sampling distributions of these statistics are approximated by

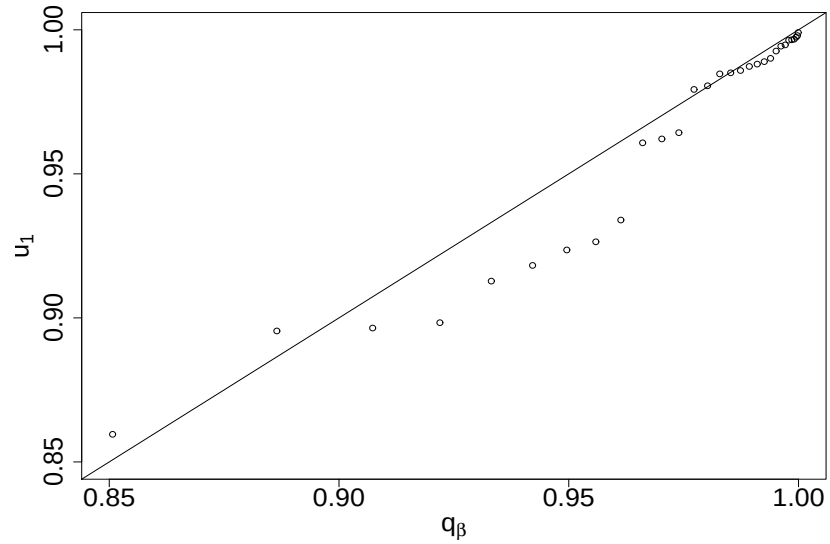

Figure 1: Q-Q plot for the fit of the $beta(\hat{\gamma}_1 + 1/2, 1/2)$ distribution to the sample $\{(\hat{M}_1^t r_i)^2\}$.

evaluating them repeatedly on data simulated from the proposed distribution with parameters equal to their moment estimates. The bootstrap p-values are 0.244 for the correlation test and 0.09 for the Kolmogorov-Smirnov test. The proposed model provides a reasonable fit.

6 Discussion

This paper has proposed a flexible model for axial data of arbitrary dimension. The proposed density is well suited to analyze samples of 3×3 rotation matrices. Simple moments estimators of the parameters are available and

the simulation of data from the proposed distribution is simple making the parametric bootstrap an appealing strategy to determine the sampling distributions of interest.

Appendix A

A.1. Proof of Proposition 1. We prove this proposition by induction. We can verify easily that for $p = 2$, $c_{\gamma,2}$ is given by (2.3), now suppose that proposition 1 true for $p-1$. Using Watson's (1983, p. 44) parametrization of S^{p-1} given in Section 2, we have

$$\begin{aligned} c_{\gamma,p} &= \int_{v\in S^{p-2}} \prod_{k=1}^{p-2} \left[\sum_{l=1}^{k} v_l^2\right]^{\gamma_k-\gamma_{k-1}} dv \int_{-1}^{1} (1-t^2)^{\gamma_{p-1}+\frac{p-3}{2}} dt \\ &= c_{\gamma(p-1)} \frac{\sqrt{\pi}\Gamma(\gamma_{p-1}+\frac{p-1}{2})}{\Gamma(\gamma_{p-1}+\frac{p}{2})} \\ &= 2(\pi)^{\frac{p-2}{2}} \prod_{k=1}^{p-2} \frac{\Gamma(\gamma_k+\frac{k}{2})}{\Gamma(\gamma_k+\frac{k+1}{2})} \frac{\sqrt{\pi}\Gamma(\gamma_{p-1}+\frac{p-1}{2})}{\Gamma(\gamma_{p-1}+\frac{p}{2})}. \end{aligned}$$

This completes the proof of Proposition 1.

A.2. Proof of Proposition 4. The expressions for λ_k, $E(u_k^4)$ and $E(u_k^2 u_l^2), 1 \leq k < l \leq p$ come from the decomposition of u as a product of beta random variables given in Proposition 2. They are derived by noting that if X is distributed as a $\beta(\gamma + k/2, \gamma + k/2)$ random variable, then

$$4E\{X(1-X)\} = \frac{\gamma + k/2}{\gamma + (k+1)/2},\ E\{(2X-1)^2\} = \frac{1}{2\{\gamma + (k+1)/2\}},$$

$$16E\{X^2(1-X)^2\} = \frac{(\gamma + k/2)(\gamma + 1 + k/2)}{\{\gamma + (k+1)/2\}\{\gamma + (k+3)/2\}},$$

$$E\{(2X-1)^4\} = \frac{3}{4(\gamma + k/2)(\gamma + 1 + k/2)}.$$

A.3. Proof of Proposition 5. Following Bellman (1970, chapter 4), one can write

$$\begin{aligned}\hat{\lambda}_j - \lambda_j &= \frac{1}{n}\sum_{i=1}^{n}\left[(M_j^T r_i)^2 - \lambda_j\right] + O_p(\frac{1}{n}) \\ &= \frac{1}{n}\sum_{i=1}^{n} u_{ji}^2 \lambda_j + O_p(\frac{1}{n}),\end{aligned}$$

and

$$\begin{aligned}\hat{M}_j - M_j &= \sum_{k\neq j}^{p}\left[\frac{1}{n}\sum_{i=1}^{n}\frac{M_j^T r_i r_i^T M_k}{\lambda_j - \lambda_k}\right] M_k + O_p(\frac{1}{n}) \\ &= \sum_{k\neq j}^{p}\left[\frac{1}{n}\sum_{i=1}^{n}\frac{u_{ji}u_{ki}}{\lambda_j - \lambda_k}\right] M_k + O_p(\frac{1}{n}),\end{aligned}$$

where u_{ji} is the j th component of the i th centered observation u_i. Now let $\left(\frac{\partial}{\partial\hat{\lambda}}\hat{\gamma}\right)_{\big|\hat{\lambda}=\lambda}$ the partial derivative $(p-1)\times p$ matrix of $\hat{\gamma}$ with respect to $\hat{\lambda}$ at point $\lambda = (\lambda_1, \ldots, \lambda_p)^t$. The kth row of the matrix is

$$\left(\underbrace{\frac{1}{\lambda_{k+1}}, \ldots, \frac{1}{\lambda_{k+1}}}_{k \text{ times}}, -\frac{\sum_{j=1}^{k}\lambda_j}{\lambda_{k+1}^2}, 0, \ldots, 0\right).$$

According to Slutzky's theorem and to the central limit theorem, as n goes to infinity, $\hat{\gamma}$ and $\hat{M}$ have asymptotic normal distributions. Now we prove that the off diagonal terms of Σ_γ are zero (i.e: $\Sigma_\gamma(k,l) = 0, k < l$). To do so, we can verify that

$$\begin{aligned}\Sigma_{\gamma(p)}(k,l) &= \frac{1}{4}E\left[\left(\frac{\partial}{\partial\hat{\lambda}}\hat{\gamma}\right)_{\big|\hat{\lambda}=\lambda} uu^T \left(\frac{\partial}{\partial\hat{\lambda}}\hat{\gamma}\right)^T_{\big|\hat{\lambda}=\lambda}\right]_{(k,l)} \\ &= \frac{1}{4}E\left[\left(\frac{u_1^2 + \ldots + u_k^2}{\lambda_{k+1}} - \frac{\sum_{j=1}^{k}\lambda_j u_{k+1}^2}{\lambda_{k+1}^2}\right)\left(\frac{u_1^2 + \ldots + u_l^2}{\lambda_{l+1}} - \frac{\sum_{j=1}^{l}\lambda_j u_{l+1}^2}{\lambda_{l+1}^2}\right)\right].\end{aligned}$$

Using (2.5), the vector $u^{(k+1)}$ of the first $k+1$ entries of u can be expressed as $u^{(k+1)} = (u_1^2 + \ldots + u_{k+1}^2)v^{(k+1)}$, where $v^{(k+1)}$ is a random S^p vector. Thus $\Sigma_\gamma(k,l)$ becomes

$$\begin{aligned}\Sigma_{\gamma(p)}(k,l) &= \frac{1}{4}E\left[\left\{\frac{v_1^2+\ldots+v_k^2}{\lambda_{k+1}} - \frac{\sum_{j=1}^k \lambda_j v_{k+1}^2}{\lambda_{k+1}^2}\right\}\right. \\ &\times \left.\left\{(u_1^2+\ldots+u_{k+1}^2)\left(\frac{\sum_{j=1}^l u_j^2}{\lambda_{l+1}} - \frac{\sum_{j=1}^l \lambda_j u_{l+1}^2}{\lambda_{l+1}^2}\right)\right\}\right].\end{aligned}$$

The expectation on the right hand side involves the product of two random variables. The first one is a function of the $(k+1)\times 1$ unit vector v with distribution $g_{\gamma,k+1}$. Considering Proposition 4, this first term has a null expectation. In terms of the *beta* random variables defined in Proposition 2, the second term depends on $\beta_{k+1},\ldots,\beta_{p-1}$; it is therefore independent of the first term. The diagonal terms of this matrix are evaluated using the following expression,

$$\Sigma_{\gamma(p)}(k,k) = \frac{1}{4\lambda_{k+1}^4}E\{(u_1^2+\ldots+u_{k+1}^2)^2\}E\{(\lambda_{k+1} - v_{k+1}^2\sum_1^{k+1}\lambda_j)^2\}.$$

The variance covariance matrix for $\hat{M}$ comes from (2.6). To prove *iii*) and that Σ_m is diagonal, observe that $E(u_{ji}u_{ki}^3) = E(u_{ji}u_{ki}u_{li}^2) = 0$, for all $j \neq k \neq l$.

A.4. Proof of Proposition 6. It is derived immediately from (4.2), since $[\hat{M}_1]_+$ and $\hat{M}_2$ can be written as

$$\begin{aligned}[\hat{M}_1]_+^T &= [M_1]_+^T - \hat{m}_{12}[M_2]_+^T - \hat{m}_{13}[M_3]_+^T - \hat{m}_{14}[M_4]_+^T, \\ \hat{M}_2 &= M_2 + \hat{m}_{12}M_1 - \hat{m}_{23}M_3 - \hat{m}_{24}M_4.\end{aligned}$$

Proposition 2 in Rivest (2001) shows that $[M_1]_+^T M_3 = [M_4]_+^T M_2$ and $[M_3]_+^T M_2 = [M_1]_+^T M_4$. A first order expansion of $[\hat{M}_1]_+^T \hat{M}_2$ yields

$$\begin{pmatrix} 0 \\ \hat{\mu} - \mu \end{pmatrix} = -\left(\hat{m}_{23} + \hat{m}_{14}\right)[M_1]_+^T M_3 - \left(\hat{m}_{13} + \hat{m}_{24}\right)[M_1]_+^T M_4 + o_p(\hat{m}'\hat{m}).$$

References

Bellman, R. (1970). *Introduction to Matrix Analysis*, 2nd ed. McGraw-Hill, New York

Bingham, C. (1974). An antipodally symmetric distribution on the sphere. *Ann. Statist.*, **2**, 1201-1225.

Chikuse, Y. (2002). *Statistics on Special Manifolds.* Springer-Verlag, New York.

Downs, T. D. (1972). Orientation statistics. *Biometrika.* **59,** 665-676.

Jammalamadaka, S. R. and SenGupta, A. (2001). *Topics in Circular Statistics.* World Scientific Publishing Co. Pte. Ltd. London.

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

Kim, P.T. (1991). Decision theoretic analysis of spherical regression. *Journal of Multivariate Analysis.* **38,** 233-240

León, C. A. , Massé, J. C. and Rivest, L. P. (2006). A statistical model for random rotations in $SO(p)$. *Journal of Multivariate Analysis.* **97,** 412-430.

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

Moran, P.A.P. (1976). Quaternions, Haar measure and the estimation of a paleomagnetic rotation. In *Perspsectives in Probability and Statistics* (J. Gani, Ed.), pp. 295-301. Applied Probability Trust and Academic Press, London.

Oualkacha, K. (2004) *Étude d'un modèle statistique pour les rotations*. Mémoire pour l'obtention du grade de maître és sciences. Québec: Université Laval.

Prentice, M. J. (1986). Orientation statistics without parametric assumptions. *J. R. Statist. Soc.* **B 48**, 214-222

Rancourt, D., Rivest, L. P. and Asselin, J. (2000). Using orientation statistics to investigate variations in human kinematics. *Journal of the Royal Statistical Society Series C-Applied Statistics*, 49, 81-94 .

Rivest, L. P. (2001). A directional model for the statistical analysis of movement in three dimensions. *Biometrika*. **88,** 3, pp. 779-791.

Watson, G. S. (1983). *Statistics on Spheres*. A Wiley-Interscience Publication, John Wiley and Sons. New York.