Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Multivariate Analysis 97 (2006) 412–430
Journal of Multivariate Analysis
www.elsevier.com/locate/jmva

A statistical model for random rotations

Carlos A. León, Jean-Claude Massé, Louis-Paul Rivest*

Département de Mathématiques et de Statistique, Université Laval, Qué., Canada G1K 7P4

Received 10 March 2004
Available online 10 May 2005

Abstract

This paper studies the properties of the Cayley distributions, a new family of models for random $p \times p$ rotations. This class of distributions is related to the Cayley transform that maps a $p(p-1)/2 \times 1$ vector s into $SO(p)$, the space of $p \times p$ rotation matrices. First an expression for the uniform measure on $SO(p)$ is derived using the Cayley transform, then the Cayley density for random rotations is investigated. A closed-form expression is derived for its normalizing constant, a simple simulation algorithm is proposed, and moments are derived. The efficiencies of moment estimators of the parameters of the new model are also calculated. A Monte Carlo investigation of tests and of confidence regions for the parameters of the new density is briefly summarized. A numerical example is presented.

AMS 1991 subject classification: 62H11; 62H05

Keywords: Cayley transform; Multivariate t-distribution; Spherical symmetry; Statistics on manifolds

1. Introduction

The construction of distribution functions for rotation matrices has received a limited attention in the statistical literature. The matrix Fisher–von Mises distribution introduced by Downs [4] is an exponential family which has been studied by Khatri and Mardia [10], Jupp and Mardia [9], Prentice [16], and Mardia and Jupp [11]. A recent account of

* Corresponding author. Fax: +418 656 2817.
E-mail address: lpr@mat.ulaval.ca (L.-P. Rivest).

0047-259X/$ - see front matter
doi:10.1016/j.jmva.2005.03.009

the theory for this model is given by Chikuse [2]. This class of distributions is hindered by the fact that even its basic characteristics lead to complicated expressions which are difficult to evaluate. Also, the simulation of random rotations distributed according to the Fisher–von Mises distribution is not a simple task.

The present paper attempts to broaden the family of models for random rotations defined in $SO(p)$, the manifold of $p \times p$ rotations, with a proposal that leads to relatively simple statistical procedures. We parametrize rotations using the Cayley transform, that maps vectors of $\mathbb{R}^{p(p-1)/2}$ into rotations in $SO(p)$. This parametrization yields a new parametric form for the invariant measure on $SO(p)$, also called the Haar measure. In Section 3 we show that uniformly distributed rotations in $SO(p)$ are the Cayley transforms of random vectors distributed according to a generalization of the multivariate t-distribution. The new model is obtained by letting the degrees of freedom of the multivariate t vary.

This work is motivated by applications in biomechanics, where the data points are 3×3 rotation matrices. Errors are typically introduced in the modelling by assuming that the measured rotations have normal distributions in the tangent spaces to their modal values. This is the technique used by Woltring [20] in a sensitivity analysis of the Euler angle representation of a rotation, by Rivest [18] in a model for rotation matrices representing a simple flexion-extension motion about a fixed rotation axis, and by Rancourt et al. [17] to construct analysis of variance tests for rotations. An objective of this paper is to enlarge the class of densities for the error rotations in this context.

In terms of the unit Haar measure on $SO(p)$, denoted $[dP]$, the proposed Cayley family of densities is defined by

$$f_p(P|\kappa, M) = c_{\kappa,p}^{-1} \, |I_p + PM'|^{\kappa}[dP], \qquad P \in SO(p), \tag{1}$$

where $c_{\kappa,p}$ is the normalizing constant, $\kappa \geqslant 0$ is the dispersion parameter, $M \in SO(p)$, is the modal rotation for P, I_p denotes the $p \times p$ identity matrix, and $|\cdot|$ represents the determinant. Section 2 reviews the properties of the Cayley transform and derives a new parametric expression for $[dP]$. Section 3 gives a closed-form expression for $c_{\kappa,p}$, derives moments of $f_p(P|\kappa, M)$ and suggests a simple algorithm to simulate a random rotation distributed according to a Cayley density. Section 4 calculates the Fisher information matrix for (κ, M) and derives the efficiencies of simple moment estimators for M and κ. Some characteristics of the random rotations obtained when p is either 2, 3, or 4 are given in Section 5 and comparisons with the symmetric matrix Fisher–von Mises distributions are presented. Section 6 reports the results of a Monte Carlo simulation of statistical procedures presented in this paper while a numerical example is given in Section 7.

The following notation is used throughout the paper:

- $e_1^{(p)}, e_2^{(p)}, \ldots, e_p^{(p)}$ denote the columns of the identity matrix I_p,
- If s denotes a vector in $\mathbb{R}^{p(p-1)/2}$, then $\mathcal{S}(s)$ denotes a $p \times p$ skew-symmetric matrix containing the entries of s, which are denoted s_{jk} with $p \geqslant k > j \geqslant 1$ so that $\mathcal{S}(s)_{jk} = s_{jk}$.
- $\mathcal{S}^{-1}(S)$ denotes the $p(p-1)/2 \times 1$ vector of the entries above the diagonal of the $p \times p$ skew-symmetric matrix S defined in such a way that $\mathcal{S}\{\mathcal{S}^{-1}(S)\} = S$.
- $\beta(p, q)$ denotes a beta distribution with parameters p and q.

2. The Haar measure on $SO(p)$

2.1. The Cayley transform

The rotations in $SO(p)$ are the real $p \times p$ matrices P such that $P'P = I_p$ and $|P| = 1$. Endowed with the standard matrix multiplication, $SO(p)$ is an algebraic group and a Lie group. For an illuminating exposition on abstract manifolds and statistical applications see [6]; a good reference for the mathematical background on manifolds and Lie groups is Warner [19, Chapters 1–4], whereas Farrell [5, Chapters 6–8] gives a modern account from a statistical perspective. Here we will be interested mainly in the general case, but most of the applications concern $SO(3)$—a concise treatment of $SO(3)$ can be found in [12].

In what follows, an important role is played by the vector space tangent to a point P in $SO(p)$. This vector space can be represented by the set of $P\mathcal{S}(s)$ for s in $\Re^{p(p-1)/2}$. The skew-symmetric matrices $\mathcal{S}(s)$ form the Lie algebra of $SO(p)$. Given a vector $s \in \mathbb{R}^{p(p-1)/2}$ the *Cayley* transform constructs a rotation in $SO(p)$ using $S = \mathcal{S}(s)$ as

$$P = (I_p - S)^{-1}(I_p + S) = 2(I_p - S)^{-1} - I_p. \tag{2}$$

Observe that this map is always well defined since $I_p - S$ is invertible for any skew-symmetric matrix. The inverse of (2) is

$$S = I_p - 2(I_p + P)^{-1}.$$

This is well defined as long as -1 is not an eigenvalue for P. Thus the Cayley transform allows us to construct all the rotations in $SO(p)$ except for those having an eigenvalue of -1. This set has Haar measure 0 and is therefore negligible.

2.2. The invariant measure on $SO(p)$

The existence of invariant measures in $SO(p)$ has been known since the late 19th century and systematic methods to obtain them were developed by Deltheil [3] even before the general existence result of A. Haar. Since then many authors have supplied a derivation for the Haar measure on $SO(3)$ under different parametrizations (see e.g. [13,16], and the references therein). The modern approach using differential forms and their exterior algebra yields the Haar measure for general compact manifolds [6,10,14,5].

An integral with respect to $[dP]$ can be viewed as an integral defined on the manifold $SO(p)$ viewed as a subsurface of $\Re^{p^2}$. The form of $[dP]$ in terms of an Euclidean product measure depends, of course, on the parametrization of this manifold. The Cayley map allows us to determine $[dP]$ using standard multivariate calculus. For any matrix X, let dX denote the matrix of differentials (dX_{ij}). If X, Y are matrices for which the matrix product XY is defined, then $d(XY) = (dX)\,Y + X\,dY$. For an invertible matrix A, the preceding result implies $0 = d(A^{-1})A + A^{-1}\,dA$ so that $d(A^{-1}) = -A^{-1}\,dA\,A^{-1}$. In particular, $I_p - S$ being always invertible, we have $d(I_p - S)^{-1} = (I_p - S)^{-1}\,dS\,(I_p - S)^{-1}$.

For the Cayley transform, the differential is $dP = 2(I_p - S)^{-1}dS(I_p - S)^{-1}$. Following James [6], an invariant measure on $SO(p)$ is given by the exterior product of the elements of $\mathcal{S}^{-1}(P'dP) = \mathcal{S}^{-1}\{2(I_p + S)^{-1}dS(I_p - S)^{-1})\}$. From Muirhead [14, Theorem 2.1.7], this exterior product is given by $(2^{p/2}|I_p + S|^{-1})^{p-1}$ times the exterior product of the elements of $s = \mathcal{S}^{-1}(S)$. This gives the following form to an invariant measure for $SO(p)$:

$$2^{p(p-1)/2}|I_p + S|^{-(p-1)}\, ds, \tag{3}$$

where the exterior product $ds = \wedge_{j>i} ds_{ij}$ is the standard product measure in $\mathbb{R}^{p(p-1)/2}$. Now integrating (3) over $\mathbb{R}^{p(p-1)/2}$ can be shown to yield $2^{p-1}\prod_{i=2}^{p}\pi^{i/2}/\Gamma(i/2)$ as the volume of $SO(p)$. This is demonstrated in Section 3.2. Normalizing (3) gives the following expression for the unit invariant measure:

$$[dP] = \prod_{i=2}^{p}\frac{\Gamma(i/2)}{\pi^{i/2}}\,\frac{2^{(p-1)(p-2)/2}}{|I_p + S|^{p-1}}\, ds. \tag{4}$$

3. The Cayley family of densities

3.1. The multivariate t-distribution

This section recalls some elementary properties of the multivariate t-distribution that are useful for investigating the properties of the new model for rotation matrices. Let Z be a $p \times 1$ vector of independent $N(0, 1)$ random variables and let W be distributed independently of Z according to a $\chi^2_{2\kappa+p}$ distribution. Then the marginal distribution of $X = Z/\sqrt{W}$ is a scaled multivariate $t_{2\kappa+p}$; its density can be shown to be given by, see [8, Chapter 37],

$$h_p(x|\kappa) = \frac{\Gamma(\kappa + p)}{\pi^{p/2}\Gamma(\kappa + p/2)}\,\frac{1}{(1 + x'x)^{\kappa+p}} \qquad x \in \mathbb{R}^p. \tag{5}$$

Moments of functions of X are given in the next proposition. All the proofs are given in the appendix.

Proposition 3.1. *If the random vector X is distributed according to $h_p(x|\kappa)$, then*

(i) $1/(1 + X'X)$ *has a* $\beta(\kappa + p/2, p/2)$ *distribution*;
(ii) $E(X_i^2) = 1/(2\kappa + p - 2)$ *provided that* $2\kappa + p > 2$;
(iii) $E\left(\dfrac{X_i^2}{1 + X'X}\right) = \dfrac{1}{2(\kappa + p)}$;
(iv) $E\left(\dfrac{X_i^4}{(1 + X'X)^2}\right) = \dfrac{3}{4(\kappa + p)(\kappa + p + 1)}$;
(v) $E\left(\dfrac{X_i^2 X_j^2}{(1 + X'X)^2}\right) = \dfrac{1}{4(\kappa + p)(\kappa + p + 1)}$ *if* $i \neq j$;
(vi) $E\left(\dfrac{X_i^2}{(1 + X'X)^2}\right) = \dfrac{2\kappa + p}{4(\kappa + p)(\kappa + p + 1)}$.

3.2. The normalizing constant

If P is distributed as $f_p(\cdot|\kappa, M)$, then PM' has the $f_p(\cdot|\kappa, I_p)$ distribution. In the sequel, $f_p(\cdot|\kappa)$ denotes the Cayley density with $M = I_p$. Thus, we evaluate the normalizing constants and some moments of the centered rotation matrix PM'. Let $s \in \Re^{p(p-1)/2}$ and

$$\mathcal{S}(s) = \begin{pmatrix} S_{11} & S_{12} \\ -S_{12}' & 0 \end{pmatrix}, \tag{6}$$

where S_{11} is $(p-1) \times (p-1)$ and $S_{12} \in \mathbb{R}^{p-1}$. Let $s' = (s_1', s_2')'$ be the corresponding partition of s, where s_1 is the vector of the $(p-1)(p-2)/2$ distinct entries of S_{11} and $s_2 = S_{12}$.

To derive a closed-form expression for the normalizing constant observe that, from (2), $|I_p + P| = 2^p/|I_p + S|$, where S is the skew-symmetric matrix associated to the rotation P. In view of the expression for the invariant probability measure for $SO(p)$ given in (4), the normalizing constant of the proposed model is given by

$$c_{\kappa,p} = 2^{\kappa p+(p-1)(p-2)/2} \prod_{i=2}^{p} \frac{\Gamma(i/2)}{\pi^{i/2}} \int_{\mathbb{R}^{p(p-1)/2}} \frac{1}{|I_p + S|^{\kappa+p-1}}\, ds, \tag{7}$$

where $S = \mathcal{S}(s)$. A standard result on the determinant of a partitioned matrix yields

$$|I_p + S| = |I_{p-1} + S_{11}|\{1 + S_{12}'(I_{p-1} + S_{11})^{-1} S_{12}\}, \tag{8}$$

where S_{11} and S_{12} are defined in (6). The integral in (7) can be split as

$$\int_{\mathbb{R}^{(p-1)(p-2)/2}} \frac{1}{|I_{p-1} + S_{11}|^{\kappa+p-1}} \int_{\mathbb{R}^{p-1}} \frac{1}{\{1 + S_{12}'(I_{p-1} + S_{11})^{-1} S_{12}\}^{\kappa+p-1}}\, ds_2\, ds_1.$$

Observe that

$$\begin{aligned} S_{12}'(I_{p-1} + S_{11})^{-1} S_{12} &= \tfrac{1}{2} S_{12}'\{(I_{p-1} + S_{11})^{-1} + (I_{p-1} - S_{11})^{-1}\} S_{12} \\ &= S_{12}'(I_{p-1} - S_{11})^{-1}(I_{p-1} + S_{11})^{-1} S_{12}. \end{aligned} \tag{9}$$

Changing variable $x = (I_{p-1} + S_{11})^{-1} S_{12}$ in the integral on $\mathbb{R}^{p-1}$ yields a jacobian equal to $|I_{p-1} + S_{11}|$. Thus,

$$\begin{aligned} &\int_{\mathbb{R}^{p(p-1)/2}} \frac{1}{|I_p + S|^{\kappa+p-1}}\, ds \\ &= \int_{\mathbb{R}^{(p-1)(p-2)/2}} \frac{1}{|I_{p-1} + S_{11}|^{\kappa+p-2}}\, ds_1 \int_{\mathbb{R}^{p-1}} \frac{1}{(1 + x'x)^{\kappa+p-1}}\, dx \\ &= \int_{\mathbb{R}^{(p-1)(p-2)/2}} \frac{1}{|I_{p-1} + S_{11}|^{\kappa+p-2}}\, ds_1\, \frac{\pi^{(p-1)/2}\Gamma(\kappa + (p-1)/2)}{\Gamma(\kappa + p - 1)}, \end{aligned}$$

since the integral on $\Re^{p-1}$ involves a function proportional to density (5). Iterating this result yields the following formula for the normalizing constant:

$$c_{\kappa,p} = \frac{2^{\kappa p+(p-1)(p-2)/2}}{\pi^{(p-1)/2}} \prod_{i=1}^{p-1} \frac{\Gamma\{(i+1)/2\}\Gamma(\kappa + i/2)}{\Gamma(\kappa + i)}. \tag{10}$$

When $\kappa = 0$, since $\Gamma(i) = 2^{i-1}\Gamma(i/2)\Gamma\{(i+1)/2\}/\sqrt{\pi}$, the normalizing constant reduces to $c_{0,p} = 1$.

3.3. A simulation algorithm

Let P be a random rotation distributed according to $f_p(P|\kappa)$ and define $s = \mathcal{S}^{-1}\{I_p - 2(I_p + P)^{-1}\}$. Changing variables yields the following expression for the density of s:

$$g_p(s|\kappa) = \prod_{i=1}^{p-1}\left[\pi^{-i/2}\frac{\Gamma(\kappa+i)}{\Gamma(\kappa+i/2)}\right]\frac{1}{|I_p + S|^{\kappa+p-1}}, \qquad s \in \mathbb{R}^{p(p-1)/2}.$$

Proceeding as in Section 3.2 one can factor $g_p(s|\kappa)$ as the product of the marginal distribution of s_1 times the conditional distribution of s_2 given s_1,

$$g_p(s|\kappa) = g_{p-1}(s_1|\kappa)\,\frac{h_{p-1}\{(I_{p-1} + S_{11})^{-1}s_2|\kappa\}}{|I_{p-1} + S_{11}|}. \tag{11}$$

If s_1 and X are independent random vectors, respectively, distributed as $g_{p-1}(s_1|\kappa)$ and $h_{p-1}(x|\kappa)$, then $s = (s_1, (I_{p-1} + S_{11})X)$ is distributed according to $g_{,p}(s|\kappa)$. In addition, $g_2(s|\kappa) = h_1(s|\kappa)$. This is the key to the simulation of a random vector s with density $g_p(s|\kappa)$ and of a random rotation having a Cayley distribution.

3.4. Moments and marginal distributions

Let $S = \mathcal{S}(s)$ and define S_{11}, S_{12}, and s_1 as in (6); let $X = (I_{p-1}+S_{11})^{-1}S_{12}$, $X \in \Re^{p-1}$. When $p = 2$, set $S_{11} = 0$. Using standard results on the inverse of a partitioned matrix,

$$\begin{aligned}(I_p - S)^{-1} &= \begin{pmatrix}(I_{p-1} - S_{11})^{-1} & 0\\ 0 & 0\end{pmatrix}\\ &\quad + \frac{\begin{pmatrix}(I_{p-1} - S_{11})^{-1}S_{12}\\ 1\end{pmatrix}\begin{pmatrix}-\{(I_{p-1} - S_{11})^{-1}\}'S_{12}\\ 1\end{pmatrix}'}{1 + S_{12}'(I_{p-1} - S_{11})^{-1}S_{12}}.\end{aligned}$$

Let $P_1 = 2(I_{p-1} - S_{11})^{-1} - I_{p-1} = (I_{p-1} - S_{11})^{-1}(I_{p-1} + S_{11})$ be the Cayley transform of S_{11} and observe that $\{(I_{p-1} - S_{11})^{-1}\}' = (I_{p-1} + S_{11})^{-1}$. Some manipulations of the above expression for $(I_p - S)^{-1}$ yields the following expression for $P = 2(I_p - S)^{-1} - I_p$:

$$\begin{aligned}P &= \begin{pmatrix}P_1 & 0\\ 0 & 0\end{pmatrix} + 2\,\frac{\begin{pmatrix}P_1X\\ 1\end{pmatrix}\begin{pmatrix}-X\\ 1\end{pmatrix}'}{1 + S_{12}'(I_{p-1} + S_{11})^{-1}S_{12}} - e_p^{(p)}e_p^{(p)\prime}\\ &= \begin{pmatrix}P_1\{I_{p-1} - 2XX'/(1 + X'X)\} & 2P_1X/(1 + X'X)\\ -2X'/(1 + X'X) & (1 - X'X)/(1 + X'X)\end{pmatrix}.\end{aligned} \tag{12}$$

If P_1 and X are independent and distributed as $f_{p-1}(P_1|\kappa)$ and $h_{p-1}(x|\kappa)$, respectively, then, as shown in Section 3.3, (12) is distributed according to $f_p(P|\kappa)$. The proof of the next proposition uses extensively formula (12) for P.

Proposition 3.2. *Let P_{ij} denote the entry (i, j) of the random rotation P, for $i, j = 1, \ldots, p$. If P is distributed as $f_p(P|\kappa)$, then*

(i) *the random variable $|I_p + P|/2^p$ is distributed as a product of $p-1$ independent beta distributions, $\prod_1^{p-1} \beta(\kappa + i/2, i/2)$;*
(ii) *the marginal distribution of $(1 + P_{ii})/2$ is a $\beta\{\kappa + (p-1)/2, (p-1)/2\}$. Thus, $E(P_{ii}) = \kappa/(\kappa + p - 1)$;*
(iii) *$E\{tr(P)\} = \kappa p/(\kappa + p - 1)$ and*

$$Var\{tr(P)\} = \frac{p(p-1)(2\kappa + p - 2)(2\kappa + p - 1)}{(\kappa + p - 2)(\kappa + p - 1)^2(\kappa + p)};$$

(iv) *for $i \neq j$*

$$E\{(P_{ij} - P_{ji})^2\} = \frac{2(2\kappa + p - 2)(2\kappa + p - 1)}{(\kappa + p - 2)(\kappa + p - 1)(\kappa + p)};$$

(v) *the entries of $\mathcal{S}^{-1}[(I_p + P)^{-1} - \{(I_p + P)^{-1}\}']$ are uncorrelated and have the same marginal distribution; their variance–covariance matrix is $I_{p(p-1)/2}/(2\kappa - 1)$ when $\kappa > 1/2$.*

Let X be a $(p-1) \times 1$ random vector with density $h_{p-1}(x|\kappa)$ and define the $p \times 1$ unit vector V as

$$V = \frac{1}{1 + X'X} \begin{pmatrix} 2X \\ 1 - X'X \end{pmatrix}.$$

From (12), V and the pth column of a random rotation distributed as $f_p(P|\kappa)$, have the same density. To derive the density of V one parametrizes S^{p-1}, the unit sphere in $\Re^p$, in terms of $x \in \Re^{p-1}$ as,

$$v = \frac{1}{1 + x'x} \begin{pmatrix} 2x \\ 1 - x'x \end{pmatrix}.$$

To calculate the parametric expression for the uniform density on S^{p-1} corresponding to v, note that the $p \times (p-1)$ matrix of partial derivatives of v with respect to x is

$$\frac{\partial v}{\partial x} = \frac{2}{1 + x'x} \begin{pmatrix} I_{p-1} - 2xx'/(1 + x'x) \\ -2x'/(1 + x'x) \end{pmatrix}.$$

This is equal to $2/(1 + x'x)$ times the first $p-1$ columns of the orthogonal matrix $I_p - 2uu'/u'u$ where $u = (x', 1)'$. Thus, the exterior product of the $p-1$ columns of $\partial v/\partial x$ gives $\{2/(1 + x'x)\}^{p-1}$ as the density of the invariant measure on S^{p-1}. In view of (5), this leads to the following expression for the uniform density on S^{p-1}:

$$du = \frac{1}{\pi^{(p-1)/2}} \frac{\Gamma(p-1)}{\Gamma((p-1)/2)} \frac{1}{(1 + x'x)^{p-1}} \, dx, \quad x \in \Re^{p-1}.$$

Now v_p, the pth component of v, satisfies $(1 + v_p)/2 = 1/(1 + x'x)$, and the marginal density of V is given in the next proposition.

Proposition 3.3. *The marginal density, with respect to the uniform distribution on* S^{p-1}, *of V, the last column of a rotation matrix P distributed according to* $f_p(\cdot|\kappa)$ *is*

$$f_{\kappa,p}^{(p)}(v) = \frac{\Gamma((p-1)/2)}{2^\kappa \Gamma(p-1)} \frac{\Gamma(\kappa+p-1)}{\Gamma(\kappa+(p-1)/2)} \left(1+v_p\right)^\kappa,$$

where $v = (v_1, \ldots, v_p)' \in S^{p-1}$.

Proposition 3.2(ii) applies to *V*. Thus $E(V) = e_p^{(p)}\kappa/(\kappa+p-1)$ and, properly scaled, the *p*th element of *V* has a beta distribution.

The last result shows that as κ goes to ∞ a Cayley distribution approaches a normal distribution in the tangent space to $SO(p)$ at I_p.

Proposition 3.4. *Let* $s = \mathcal{S}^{-1}\{I_p - 2(I+P)^{-1}\}$, *where P is distributed as* $f_p(P|\kappa)$. *As* κ *goes to* ∞, *the elements of* $\sqrt{2\kappa}s$ *converge in distribution to independent* $N(0, 1)$ *random variables.*

4. Parameter estimation

4.1. General

This section assumes that $P_1, P_2, \ldots, P_n$ is a sample of $p \times p$ rotation matrices distributed according to $f_p(PM'|\kappa)$, where $\kappa > 0$ and $M \in SO(p)$ are unknown parameters. First, the estimation of the parameters by maximum likelihood is investigated and a closed-form expression for the Fisher information matrix is presented. Then simple moment estimators for M and κ, both functions of $\bar{P} = \sum P_i/n$, are investigated and their efficiencies are derived. Without loss of generality, we assume that the true value of M is I_p.

4.2. Maximum likelihood estimators for κ and M

The log-likelihood for κ and M is

$$\ell(\kappa, M) = \kappa \sum \log|I_p + P_i M'| - n \log c_{\kappa,p}.$$

Rotations in an infinitesimal neighborhood of I_p, the true value of M, can be written as $I_p + A + o(A)$ where $A = \mathcal{S}(a)$ and $a \in \mathbb{R}^{p(p-1)/2}$. Let A_{jk}, $k > j$, denote the entry (j, k) of A. Observe that $A = \sum_{k>j}^p A_{jk}(e_j^{(p)} e_k^{(p)\prime} - e_k^{(p)} e_j^{(p)\prime})$. The score vector for M is the vector of the coefficients of a in the first-order expansion for $\ell(\kappa, M)$ around $M = I_p$. This first order expansion can be derived by noting that

$$\begin{aligned}\log|I_p + P_i(I_p - A)| - \log|I_p + P_i| &= \log|I_p - (I_p + P_i)^{-1} P_i A| \\ &= \log[1 - tr\{(I_p + P_i)^{-1} P_i A\}] + o_p(a)\end{aligned}$$

$$
\begin{aligned}
&= -tr\{(I_p + P_i)^{-1} P_i A\} + o_p(a) \\
&= \sum_{k>j}^{p} A_{jk}\{e_k^{(p)\prime}(I_p + P_i)^{-1} e_j^{(p)} - e_j^{(p)\prime}(I_p + P_i)^{-1} e_k^{(p)}\} + o_p(a),
\end{aligned}
$$

since $(I_p + P_i)^{-1} P_i = I_p - (I_p + P_i)^{-1}$. Thus the (j, k) element of the score vector involves the coefficient of A_{jk} in the above expansion,

$$
s(M, j, k) = \kappa \sum_{i=1}^{n} e_j^{(p)\prime}(I_p + P_i)^{-1} e_k^{(p)} - e_k^{(p)\prime}(I_p + P_i)^{-1} e_j^{(p)}.
$$

The Fisher information matrix for M is the variance covariance matrix of the vector of the $s(M, j, k)$. According to Proposition 3.2 (v), this is equal to $n\kappa^2 I_{p(p-1)/2}/(2\kappa - 1)$ provided that $\kappa > 1/2$.

The partial derivatives of $\ell(\kappa, M)$ with respect to both M and κ have null expectations so that the Fisher information matrix is block-diagonal. To determine the (κ, κ) term observe that, from expression (10) for the normalizing constant $c_{\kappa,p}$,

$$
\begin{aligned}
\frac{\partial^2 \ell(\kappa, M)}{\partial \kappa^2} &= n \left[\sum_{i=1}^{p-1} \frac{\partial^2 \log\{\Gamma(\kappa + i)\}}{\partial \kappa^2} - \frac{\partial^2 \log\{\Gamma(\kappa + i/2)\}}{\partial \kappa^2} \right] \\
&= n \sum_{i=1}^{p-1} \{\psi'(\kappa + i) - \psi'(\kappa + i/2)\},
\end{aligned}
$$

where $\psi'(z)$ denotes the trigamma function (Abramowitz and Stegun, [21]). These findings can be summarized in the following proposition:

Proposition 4.1. *The Fisher information matrix for the parameters* κ, M, *is given by* $nI(\kappa, M)$ *where*

$$
I(\kappa, M) = \begin{pmatrix} \sum_{i=1}^{p-1} \psi'(\kappa + i/2) - \psi'(\kappa + i) & 0 \\ 0 & \kappa^2 I_{p(p-1)/2}/(2\kappa - 1) \end{pmatrix}, \tag{13}
$$

provided that $\kappa > 1/2$.

When $1/2 \geqslant \kappa > 0$, the Fisher information matrix for M is infinite.

When the Fisher information matrix exists, the asymptotic distribution of the maximum likelihood estimators is normal. The asymptotic distribution of $\hat{M}$ can be expressed in terms of the vector $\hat{a}$ in $\mathbb{R}^{p(p-1)/2}$ such that $\hat{M} = I_p + \mathcal{S}(\hat{a}) + o_p(n^{-1/2})$, where I_p is the true value of M. Indeed the limiting distribution of $n^{1/2}(\hat{\kappa} - \kappa, \hat{a})$ is $N_{p(p-1)/2+1}\{0, I(\kappa, M)^{-1}\}$.

4.3. Moment estimators for M and κ

The moment estimator of M is defined as the rotation M in $SO(p)$ that maximizes $tr(\bar{P}M')$. This estimator is easily calculated using Procrustes techniques. If $\bar{P} = U\text{diag}(\lambda_1, \ldots, \lambda_p)V'$ denotes a singular value decomposition, where $\lambda_1 > \cdots > \lambda_{p-1} > |\lambda_p|$ denote the singular values and U, V are rotations in $SO(p)$, then $\hat{M}_m = UV'$ maximizes the trace and is therefore the moment estimator (index m denotes moment estimators). To estimate κ one sets the maximum value of the trace equal to $\kappa p/(\kappa + p - 1)$, the expectation of the trace according to Proposition 3.2. This yields the following moment estimator for κ:

$$\hat{\kappa}_m = (p-1)\,\frac{\bar{\lambda}}{1-\bar{\lambda}},$$

when $\hat{\lambda}$ is the average of the singular values. This section derives the large sample distribution of $(\hat{M}_m, \hat{\kappa}_m)$.

According to the weak law of large numbers, as n goes to ∞, $\bar{P}M'$ converges weakly to $E(P_1)M' = \kappa M'/(\kappa + p - 1)$. Thus $M = I_p$ maximizes $tr\{E(P_1)M'\}$ and, by Slutsky's Theorem, $\hat{M}_m$ is convergent. The underlying density $f_p(P|\kappa)$ and the moment estimator $\hat{M}_m$ satisfy the symmetry conditions of Chang and Rivest [1]. Thus, applying their Proposition 4, if $\hat{a}_m$ is such that $\hat{M}_m = I_p + \mathcal{S}(\hat{a}_m) + o_p(n^{-1/2})$, then the limiting distribution of $n^{1/2}\hat{a}_m$ is $N_{p(p-1)/2}\{0, c_1 I_{p(p-1)/2}/(2d_1^2)\}$, where c_1 and d_1 can be evaluated using Chang and Rivest [1] Eqs. (36) and (37), with $\rho_0(t) = t$. This yields

$$c_1 = \frac{E\{(P_{12} - P_{21})^2\}}{2} = \frac{(2\kappa + p - 2)(2\kappa + p - 1)}{(\kappa + p - 2)(\kappa + p - 1)(\kappa + p)}$$

and $d_1 = E(P_{11}) = \kappa/(\kappa + p - 1)$. Thus the asymptotic covariance matrix of $n^{1/2}\hat{a}_m$ is

$$\frac{(\kappa + p - 1)(2\kappa + p - 2)(2\kappa + p - 1)}{2\kappa^2(\kappa + p - 2)(\kappa + p)}\, I_{p(p-1)/2}.$$

Considering Proposition 4.1, the efficiency of $\hat{M}_m$ is given by

$$\frac{(\kappa - 1/2)(\kappa + p - 2)(\kappa + p)}{(\kappa + p - 1)(\kappa + p/2 - 1)(\kappa + p/2 - 1/2)} = 1 - \frac{p^2 - p + 4}{4\kappa^2} + O(\kappa^{-3})$$

when $\kappa > 1/2$. Efficiency curves are presented in Fig. 1; they show that the moment estimator is quite good when the clustering is relatively large ($\kappa > 4$).

To derive the large sample distribution of $\hat{\kappa}_m$ observe that

$$\bar{\lambda} = \frac{1}{np}\sum_{i=1}^{n} tr(P_i) + \frac{tr\,\bar{P}(\hat{M}_m - I_p)}{p}.$$

The second term on the right-hand side is $o_p(n^{-1/2})$, thus applying the Central limit theorem to the first term on the right-hand side gives the asymptotic distribution of $\bar{\lambda}$; from Proposition 3.2, $n^{1/2}\{\bar{\lambda} - \kappa/(\kappa + p - 1)\}$ is approximately distributed as $N[0, (p-1)(2\kappa + p - 2)(2\kappa +$

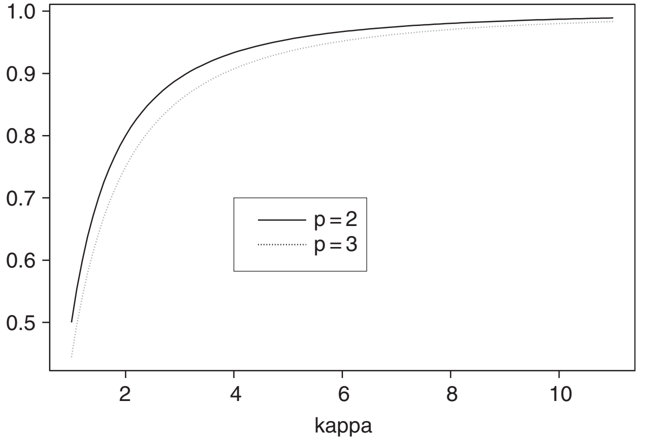

Fig. 1. Efficiency of the moment estimator for M.

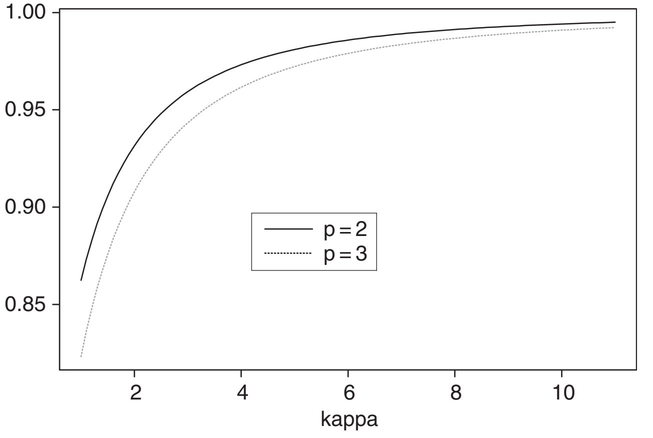

Fig. 2. Efficiency of the moment estimator for κ.

$p-1)/\{p(\kappa+p-2)(\kappa+p-1)^2(\kappa+p)\}]$. Using a standard linearization argument leads to the following limiting distribution:

$$n^{1/2}(\hat{\kappa}_m - \kappa) \xrightarrow{d} N\left(0, \frac{(\kappa+p-1)^2(2\kappa+p-2)(2\kappa+p-1)}{p(p-1)(\kappa+p-2)(\kappa+p)}\right).$$

The efficiency curves given in Fig. 2 reveal that the moment estimator for κ is quite good when $\kappa > 2$.

5. The model in a few special cases

5.1. The case $p = 2$

When $p = 2$, it is convenient to express the density of P in terms of θ and μ, respectively, equal to the angles of P and M. The density of θ is then given by

$$f_2(\theta|\kappa, \mu) = \frac{\Gamma(\kappa+1)}{2^{\kappa+1}\pi^{1/2}\Gamma(\kappa+1/2)} \{1+\cos(\theta-\mu)\}^{\kappa}\, d\theta, \quad 0 \leqslant \theta \leqslant 2\pi.$$

This is a circular beta distribution as presented in [7, p. 51].

Let a and b be two angles such that $-\pi < a - \mu < b - \mu < \pi$, then $Pr\{a < \theta < b\}$ can be expressed in terms of $F_{2\kappa+1}(x)$, the t-distribution with $2\kappa + 1$ degrees of freedom. Indeed

$$Pr\{a < \theta < b\} = \frac{\Gamma(\kappa+1)}{2^{\kappa+1}\pi^{1/2}\Gamma(\kappa+1/2)} \int_{a-\mu}^{b-\mu} (1+\cos\theta)^{\kappa}\, d\theta.$$

Let $x = \sin\theta/(1+\cos\theta)$ so that $dx = d\theta/(1+\cos\theta)$. Using $1 + x^2 = 2/(1+\cos\theta)$, after some manipulations, the above integral can be expressed as

$$Pr\{a < \theta < b\} = F_{2\kappa+1}\left(\sqrt{\kappa}\, \frac{\sin(b-\mu)}{1+\cos(b-\mu)}\right) - F_{2\kappa+1}\left(\sqrt{\kappa}\, \frac{\sin(a-\mu)}{1+\cos(a-\mu)}\right).$$

The equation for the maximum likelihood estimation of μ based on a sample $\theta_1, \ldots, \theta_n$ is

$$\sum \frac{\sin(\theta_i - \mu)}{1+\cos(\theta_i-\mu)} = 0.$$

Observe that $\sin(\theta-\mu)/\{1+\cos(\theta-\mu)\}$ has a singularity at $\theta - \mu = -\pi$. It is $-\infty$ at $-\pi^+$ and $+\infty$ at $-\pi^-$. This makes the maximum likelihood estimate of μ sensitive to small changes in the data. For this reason, we prefer to use the moment estimator of Section 4.2.

It is interesting to compare the Cayley density with the von Mises density which is proportional to $\exp(\kappa\cos\theta)$. As κ goes to 0, both the von Mises and the Cayley models tend to the uniform distribution. For any positive value of κ, $f_{\kappa,2,\mu}(\mu+\pi) = 0$, so that the convergence of the Cayley density to the uniform density is not uniform. This convergence is uniform for the von Mises distribution so that the two models differ near the uniform distribution. For concentrated samples the two models are equivalent since they are approximately normal. As their concentration parameters increase the two models become more and more similar. This is illustrated in Fig. 3.

5.2. The case $p = 3$

When $p = 3$ the three eigenvalues of a rotation P are $1, e^{i\theta}$, and $e^{-i\theta}$, where $i = \sqrt{-1}$ and θ is the angle of P. Thus $|I_3 + P| = 2\{1 + tr(P)\} = 4(1+\cos\theta)$, and the

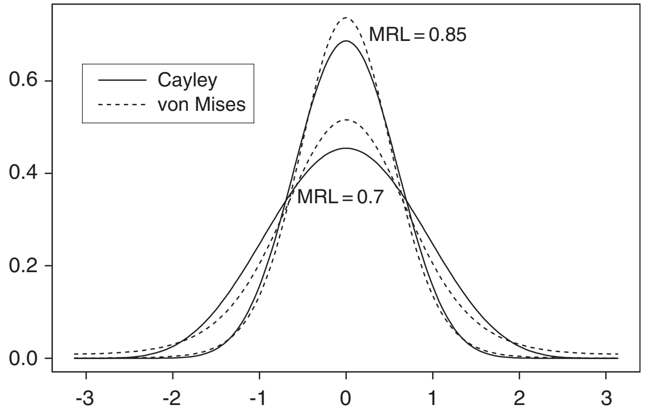

Fig. 3. Comparison of the Cayley and of the von Mises density when the mean resultant length (MRL), $E(\cos\theta)$, is equal to 0.7 and 0.85.

density function of the model takes the form

$$f_3(P|\kappa, M) = \frac{\sqrt{\pi}\Gamma(\kappa+2)}{2^{2\kappa}\Gamma(\kappa+1/2)}(1+tr(PM'))^{\kappa}.$$

It is convenient to redefine the operator $\mathcal{S}$ by

$$\mathcal{S}(s) = \begin{pmatrix} 0 & -s_3 & s_2 \\ s_3 & 0 & -s_1 \\ -s_2 & s_1 & 0 \end{pmatrix},$$

where $s = (s_1, s_2, s_3)'$ belongs to $\Re^3$. For this special case,

$$g_3(s|\kappa) = h_{\kappa-1,3}(s) = \frac{1}{\pi^{3/2}}\frac{\Gamma(\kappa+2)}{\Gamma(\kappa+1/2)}\frac{1}{(1+s's)^{\kappa+2}}, \quad s \in \Re^3.$$

The Cayley transform satisfies

$$2\{I_3 - \mathcal{S}(s)\}^{-1} - I_3 = I_3 + \frac{2}{1+s's}\{\mathcal{S}(s) + \mathcal{S}^2(s)\}.$$

This is a rotation of $\arccos\{(1-s's)/(1+s's)\}$ radians about axis s.

Following [13], one can write the uniform measure for $[dP]$ for $SO(3)$ in terms of the rotation angle $\theta \in (0, \pi)$ and of the rotation axis $u \in S^2$ as

$$[dP] = \frac{1-\cos\theta}{\pi}\, d\theta\, du,$$

where du is the uniform unit measure on S^2. Thus when P is distributed as $f_3(\cdot|\kappa)$, its axis is uniformly distributed on S^2 and the density of its angle is

$$\frac{\Gamma(\kappa+2)}{\sqrt{\pi}2^{\kappa}\Gamma(\kappa+1/2)}\,(1+\cos\theta)^{\kappa}(1-\cos\theta) \quad \theta \in (0, \pi).$$

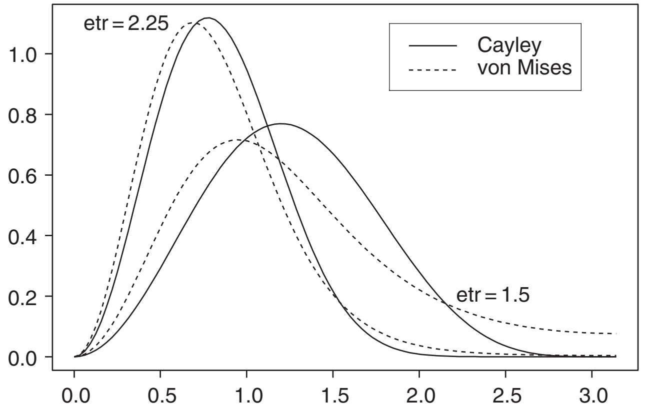

Fig. 4. Comparison of the densities for the rotation angle when the expected trace (etr), $E\{tr(P)\}$, is equal to 1.5 and 2.25.

Elementary manipulations show that $(1+\cos\theta)/2$ has a $\beta(\kappa+1/2, 3/2)$ distribution. Thus the distribution of θ can be expressed in terms of the $F_{3,2\kappa+1}$ distribution as

$$Pr\{\theta < a\} = Pr\left\{F_{3,2\kappa+1} < \frac{(2\kappa+1)(1-\cos a)}{3(1+\cos a)}\right\}.$$

The value a for which the probability that θ belongs to $(0, a)$ is $1-\alpha$ is

$$a = \arccos\left(\frac{1-3F_{3,2\kappa+1,1-\alpha}/(2\kappa+1)}{1+3F_{3,2\kappa+1,1-\alpha}/(2\kappa+1)}\right). \tag{14}$$

For instance, the κ values for which the upper bounds, in degrees, for the 95% prediction intervals for the rotation angle are 10°, 20° and 30° are equal to 510, 127 and 56, respectively. This relationship is helpful to relate the value of κ and the magnitude of the errors.

It is interesting to compare the Cayley distribution to the symmetric matrix Fisher–von Mises distribution whose density is proportional to $\exp\{\kappa\, tr(P)\}$. Since $tr(P) = 2\cos\theta + 1$, the corresponding density for θ is

$$\frac{(1-\cos\theta)\exp\{2\kappa\cos(\theta)\}}{\pi\{\mathcal{I}_0(2\kappa) - \mathcal{I}_1(2\kappa)\}}, \quad \theta \in (0, \pi),$$

where $\mathcal{I}_k(\cdot)$ denotes the modified Bessel of the first kind, see Appendix 1 of [11]. For this distribution, elementary manipulations of Bessel functions show that $E\{tr(P)\} = \mathcal{I}_0(2\kappa)/[\kappa\{\mathcal{I}_0(2\kappa) - \mathcal{I}_1(2\kappa)\}] - 1$. The situation is similar to the circular case. As κ goes to 0, the Cayley density does not converge uniformly to the uniform density because of the singularity at $\theta = \pi$ and the two models differ. They both converge to the same local normal model when their concentration parameters become large. The is exemplified in Fig. 4 which shows a very close agreement between the two models when $E\{tr(P)\} = 2.25$.

When $p = 3$, model $f_p(\cdot|\kappa)$ assumes that the errors have no preferred orientation. Statistic R, proposed at the end of Section 5 of Prentice, is available to test this hypothesis.

Considering formula (6.1) in Prentice [15], one can correct a typographical error in the formula for R which leads to the following expression for the test statistic:

$$R = n\,\frac{\sum_{i=1}^{3}(\lambda_i - \bar{\lambda})^2}{8\hat{c}_{23}},$$

where c_{23} is defined in [15] and the λ_i are the singular values defined in Section 4.3. In the notation of this paper, $c_{23} = E[\{tr(P) - 3\}^2]/(15 \times 16)$. Using Proposition 3.2, c_{23} is easily evaluated,

$$c_{23} = \frac{1}{15 \times 16}[Var\{tr(P)\} + \{E\,tr(P) - 3\}^2] = \frac{1}{4(\kappa + 2)(\kappa + 3)}.$$

A simple model based test for spherical symmetry rejects the null hypothesis if $R = n(\hat{\kappa} + 2)(\hat{\kappa} + 3)\sum(\lambda_i - \bar{\lambda})^2/2$ is large when compared to critical values of a χ^2_5 distribution.

When $p = 4$, the model is more complex. One can write $f_4(P|\kappa)$ in terms of $tr(P)$ and $tr(P^2)$. Also $|I_4 + S|$ appearing in the density $g_4(s|\kappa)$ is a polynomial of order 4 featuring cross-product terms. Unlike the case $p = 3$, the entries of s are not permutation-symmetric.

6. A Monte Carlo study

This section presents the results of a Monte Carlo experiment investigating the confidence levels of confidence regions for κ and M and the true level of Prentice's R test for samples coming from a Cayley density when $p = 3$. Since the log-transform stabilizes the variance of $\hat{\kappa}_m$, the Monte Carlo simulations investigated the coverage of the following 95% confidence interval for $\log\kappa$

$$\log(\hat{\kappa}_m) \pm 1.96\sqrt{\frac{(2\hat{\kappa}_m + 1)(\hat{\kappa}_m + 2)^2}{3n\hat{\kappa}_m^2(\hat{\kappa}_m + 3)}}.$$

From Section 4, $\hat{a}_m = \mathcal{S}^{-1}\{(\hat{M}_m M' - M\hat{M}_m')/2\}$ has a $N_3[0,\, (\kappa+2)(2\kappa+1)I_3/\{\kappa^2(\kappa+3)\}]$ distribution. Since $tr(\hat{M}_m' M) \approx 3 - \hat{a}_m'\hat{a}_m$, the asymptotic distribution of $3 - tr(\hat{M}_m' M)$ is proportional to a χ^2_3. A 95% confidence region for M given by

$$\left\{ M : n\,\frac{\{3 - tr(\hat{M}_m M')\}\hat{\kappa}_m^2(\hat{\kappa}_m + 3)}{n(\hat{\kappa}_m + 2)(2\hat{\kappa}_m + 1)} < 7.81 \right\},$$

was considered in the study (where $\chi^2_{3,0.95} = 7.81$). The true level of Prentice's R test, at a nominal 5% level was also investigated.

Table 1 reveals that the true levels of the confidence regions for κ and R are slightly less than 95%. Their performance improves as either n or κ increases. The level of Prentice's R test holds for all the cases considered.

Table 1
Simulation results of statistical methods for κ and M based on asymptotic approximations, for three sample sizes and four concentration parameters

n	$\kappa = 1/2$	$\kappa = 1$	$\kappa = 20$	$\kappa = 1000$
5	0.75, 0.73, 0.95	0.74, 0.82, 0.95	0.83, 0.85, 0.96	0.84, 0.84, 0.97
10	0.75, 0.80, 0.95	0.84, 0.88, 0.95	0.89, 0.90, 0.96	0.89, 0.91, 0.95
30	0.87, 0.86, 0.95	0.91, 0.92, 0.95	0.93, 0.93, 0.96	0.93, 0.94, 0.95
100	0.93, 0.94, 0.95	0.94, 0.95, 0.95	0.95, 0.95, 0.95	0.94, 0.94, 0.95

The true levels, or confidence levels, of procedures constructed at a nominal level of 95% are reported for four sample sizes. In each cell, the three entries correspond to (i) a confidence interval for $\log \kappa$, (ii) a confidence region for M, and (iii), Prentice's R test.

7. A numerical example

Rancourt et al. [17] measured the orientations of the back, the upper arm, the forearm and the hand of 8 subjects performing drilling tasks. For each segment$\times$subject treatment there are $n = 30$ rotations. The average $\bar{P}$ for the hand of subject 2 is

$$\bar{P} = \begin{pmatrix} -0.0055 & 0.9620 & 0.2641 \\ -0.9902 & 0.0248 & -0.1145 \\ -0.1173 & -0.2638 & 0.9543 \end{pmatrix}.$$

The singular values of $\bar{P}$ are 0.9982, 0.9969, and 0.9966. Thus $\hat{\kappa}_m = 721$ $(s.e. = 108)$. Furthermore, the mean rotation is

$$\hat{M}_m = \begin{pmatrix} -0.006 & 0.964 & 0.265 \\ -0.993 & 0.025 & -0.115 \\ -0.117 & -0.264 & 0.957 \end{pmatrix}.$$

This rotation matrix is interpreted in terms of the local coordinate system for the wrist marker measured when the arm is parallel to the body: the y-axis goes up while the x-axis goes back. Now $\hat{M}_m$ can roughly be regarded as a rotation of $-90°$ in the x–y plane. This is the rotation needed to bring the upper arm from its original vertical position up to the horizontal position used for drilling. To test the fit of the proposed model, one can compare Prentice's R statistics to χ^2_5 critical values. One has $R = 10.77,\ \ p = 0.056$. The hypothesis of spherical symmetry is tenable. Using (14) a 95% prediction interval for the angle of a rotation distributed as $f_3(\cdot|\hat{\kappa}_m, \hat{M}_m)$ is 8.43°. This characterizes the variability of the orientation of the wrist marker of the second subject in the experiment.

The spherical symmetry of the Cayley density is a restrictive constraint when analyzing biomechanical data. It appears to be appropriate for segments whose orientation does not vary much during the experiment. In [17], the orientations of the forearm and of the upper arm varied by more than 20°. These variations typically have preferred orientations that cannot be modeled with a Cayley density. Thus it would be useful to extend $f_3(\cdot|\kappa, M)$ to non-spherically symmetric errors; this is under investigation.

8. Discussion

This paper has introduced the Cayley densities as a class of symmetric distributions for random rotations. This model has appealing properties. As κ increases from 0 to ∞, the Cayley density goes from the uniform distribution to a locally normal distribution on the tangent space. Its moments can be expressed in terms of simple functions of κ and random Cayley rotations are easily generated. Many of the distribution functions associated to the Cayley model can be expressed in terms of standard F and t distribution functions. Thus it is an appealing alternative to the symmetric matrix Fisher–von Mises model for rotation matrices.

Appendix

Proof of Proposition 3.1. According to the construction of X, $E(X_i^2) = E(1/\chi^2_{2\kappa+p}) = 1/(2\kappa+p-2)$. In terms of the random vectors W and Z defined in Section 3 $(1+X'X)^{-1} = W/(W+Z'Z)$. This follows a $\beta(\kappa+p/2, p/2)$ distribution. Since the X_i's are exchangeable, $E\{X_i^2/(1+X'X)\} = E\{p^{-1}X'X/(1+X'X)\} = p^{-1}[1-E\{1/(1+X'X)\}] = 1/\{2(\kappa+p)\}$. To prove (iv), write

$$E\left(\frac{X_i^4}{(1+X'X)^2}\right) = \frac{(\kappa+p/2)(\kappa+p/2+1)}{(\kappa+p)(\kappa+p+1)}\, E_{\kappa+2}(Y_i^4),$$

where the index $\kappa+2$ refers to an expectation with respect to the density $h_{\kappa+2,p}(y)$. Now Y_i is distributed as the ratio of a $N(0, 1)$ random variable over $\sqrt{W}$ where W has a $\chi^2_{2\kappa+4+p}$ distribution. Thus $E(Y_i^4) = 3E(1/W^2) = 3/\{4(\kappa+p/2)(\kappa+p/2+1)\}$. This completes the proof of (iv); (v) and (vi) are proved using the same technique. □

Proof of Proposition 3.2. If P and s are, respectively, distributed as $f_{\kappa,p}(P)$ and $g_p(s|\kappa)$, then for any $p \times p$ orthogonal matrix Q, P and QPQ', resp. $\mathcal{S}(s)$ and $Q\mathcal{S}(s)Q'$, have the same distribution. Taking a Q obtained by interchanging two columns of I_p or by changing the sign, from $+$ to $-$, of one of its diagonal entries proves the following results that are used in the proof:

(i) the diagonal elements of P have a permutation-symmetric distribution,
(ii) the marginal distribution of $P_{ij} - P_{ji}$ does not depend on i and j, and
(iii) the elements of s have the same marginal distribution.

Throughout the proof let X be a random vector with density $h_{p-1}(x|\kappa)$. Considering (8), (9), and (11), we see that $|I_p + P|/2^p$ is distributed as $|I_p + S_{11}|^{-1}(1 + X'X)^{-1}$, the product of two independent random variables. From Proposition 3.1(i), $(1 + X'X)^{-1}$ has a $\beta\{\kappa+(p-1)/2, (p-1)/2\}$ distribution. Since S_{11} is a skew-symmetric matrix associated to a $(p-1)\times(p-1)$ rotation, (i) is proved by iterating (8) and (9). According to (12), P_{pp} is distributed as $2/(1+X'X)-1$. The proof of (ii) is completed by using (i) of Proposition 3.1. To prove (iii) observe that, since the P_{ii}'s are permutation-symmetric the variance of the trace is equal to p times the variance of P_{11}, $(p-1)(2\kappa+p-1)/\{(\kappa+p-1)^2(\kappa+p)\}$, plus

$p(p-1)$ times the pairwise covariance. We calculate $E(P_{11}P_{pp})$. Using (12) this expectation can be written as a function of the $(p-1)\times(p-1)$ rotation P_1 and of a random vector X. It can be evaluated by noting that X and P_1 are independent, that $E(P_1)=\kappa I_{p-1}/(\kappa+p-2)$, and using Proposition 3.1. This yields

$$\begin{aligned} E(P_{11}P_{pp}) &= E\left(e_1^{(p-1)\prime}P_1 e_1^{(p-1)}\frac{1-X'X}{1+X'X}\right)\\ &\quad -2E\left(\frac{e_1^{(p-1)\prime}P_1XX'e_1^{(p-1)}(1-X'X)}{(1+X'X)^2}\right)\\ &= \frac{\kappa^2}{(\kappa+p-1)(\kappa+p-2)} - 2\,\frac{\kappa}{(\kappa+p-2)}\,E\left(\frac{X_1^2(1-X'X)}{(1+X'X)^2}\right)\\ &= \frac{\kappa^2}{(\kappa+p-1)(\kappa+p-2)} - \frac{\kappa(\kappa-1)}{(\kappa+p)(\kappa+p-1)(\kappa+p-2)}\\ &= \frac{\kappa\{\kappa^2+(p-1)\kappa+1\}}{(\kappa+p)(\kappa+p-1)(\kappa+p-2)}. \end{aligned}$$

Subtracting the product of the expectations yields $\kappa(2\kappa+p-1)/\{(\kappa+p-2)(\kappa+p-1)^2(\kappa+p)\}$ as the pairwise covariance. To prove (iv) we take $i=1$ and $j=p$. According to (12), $P_{1p}-P_{p1}=2e_1^{(p-1)\prime}(I+P_1)X/(1+X'X)$. This is the dot product of two independent random vectors, with one, namely $X/(1+X'X)$, distributed according to a rotationally symmetric density. Thus $P_{1p}-P_{p1}$ is distributed as the product of the length of the first vector, $2\sqrt{e_1^{(p-1)\prime}(2I+P_1+P_1')e_1^{(p-1)}}=\sqrt{8(1+e_1^{(p-1)\prime}P_1e_1^{(p-1)}})$, times the first component of the second, $X_1/(1+X'X)$. Now, using results of Proposition 3.1,

$$\begin{aligned} E\{(P_{1p}-P_{p1})^2\} &= 8\{1+E(e_1^{(p-1)\prime}P_1e_1^{(p-1)})\}E\left(\frac{X_1^2}{(1+X'X)^2}\right)\\ &= 8\left(1+\frac{\kappa}{\kappa+p-2}\right)\frac{2\kappa+p-1}{4(\kappa+p-1)(\kappa+p)}\\ &= 2\,\frac{(2\kappa+p-2)(2\kappa+p-1))}{(\kappa+p-2)(\kappa+p-1)(\kappa+p)}. \end{aligned}$$

To prove (v) observe that $\mathcal{S}^{-1}\{(I_p+P)^{-1}-(I_p+P)^{-1\prime}\}=-\mathcal{S}^{-1}(S)=-s$. By (11), the marginal density of one entry of s is $h_{\kappa,1}(x)$ whose second moment is given by Proposition 3.1(ii). □

Proof of Proposition 3.4. Let $t=\sqrt{2\kappa}s$ and $T=\mathcal{S}(t)$. The density of t is proportional to

$$\begin{aligned} \left|I_p+\frac{T}{\sqrt{2\kappa}}\right|^{-(\kappa+p-1)} &= \left|I_p+T'T/2\kappa\right|^{-(\kappa+p-1)/2}\\ &\approx \exp\{-tr(T'T)/4\}\\ &= \prod_{j>i}\exp(-t_{ij}^2/2). \qquad \square \end{aligned}$$

References

[1] T. Chang, L.-P. Rivest, *M*-estimation for location and regression parameters in group models: a case study using Stiefel manifolds, Ann. Statist. 29 (3) (2001) 784–814.
[2] Y. Chikuse, Statistics on Special Manifolds, Lecture Notes in Statistics, vol. 174, Springer, New York , 2003.
[3] R. Deltheil, Probabilités Géométriques. Traité du Calcul des Probabilités et de ses Applications, tome II, fasc. 2, Gauthier-Villars, Paris, 1926.
[4] T.D. Downs, Orientation statistics, Biometrika 59 (1972) 665–676.
[5] R.H. Farrell, Multivariate Calculation, Springer Series in Statistics, Springer, New York, 1985.
[6] A.T. James, Normal multivariate analysis and the orthogonal group, Ann. Math. Statist. 25 (1954) 40–75.
[7] S.R. Jammalamadaka, A. SenGupta, Topics in Circular Statistics, World Scientific Publishing Co., London, 2001.
[8] N.L. Johnson, S. Kotz, Distributions in Statistics. Continuous Multivariate Distributions, Wiley, New York, 1972.
[9] P.E. Jupp, K.V. Mardia, Maximum likelihood estimation for the matrix von Mises–Fisher matrix and Bingham distributions, Ann. Statist. 7 (1979) 599–606.
[10] C.G. Khatri, K.V. Mardia, The von Mises–Fisher matrix distribution in orientation statistics, J. Roy. Statist. Soc. Ser. B 39 (1) (1977) 95–106.
[11] K.V. Mardia, P.E. Jupp, Directional Statistics, Wiley Series in Probability and Statistics, Wiley, Chichester, 2000.
[12] J.M. McCarthy, Introduction to Theorical Kinematics, The MIT Press, Cambridge, 1990.
[13] R.E. Miles, On random rotations in R^3, Biometrika 52 (1965) 636–639.
[14] R.J. Muirhead, Aspects of Multivariate Statistical Theory, Wiley, New York, 1982.
[15] M.J. Prentice, A distribution-free method of interval estimation for unsigned directional data, Biometrika 71 (1984) 147–154.
[16] M.J. Prentice, Orientation statistics without parametric assumptions, J. Roy. Statist. Soc. Ser. B 48 (2) (1986) 214–222.
[17] D. Rancourt, L.-P. Rivest, J. Asselin, Using orientation statistics to investigate variations in human kinematics, J. Roy. Statist. Soc. Ser. C 49 (1) (2000) 81–94.
[18] L.-P. Rivest, A directional model for the statistical analysis of movement in three dimensions, Biometrika 88 (3) (2001) 779–791.
[19] F.W. Warner, Foundations of Differentiable Manifolds and Lie Groups, Graduate Texts in Mathematics, vol. 94, Springer, New York, 1983.
[20] H. Woltring, 3-d attitude representation of human joints: a standardization proposal, J. Biomech. 27 (1994) 1399–1414.
[21] M. Abramowitz, I.-A. Stegun, Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, National Bureau of Standards Applied Mathematics Series, Vol. 55, 1972.