Biometrika (2001), **88**, 3, *pp.* 779–791

Printed in Great Britain

A directional model for the statistical analysis of movement in three dimensions

By LOUIS-PAUL RIVEST

Department of Mathematics and Statistics, Université Laval, Ste-Foy, Québec, Canada G1K 7P4

lpr@mat.ulaval.ca

Summary

Movement of an object in three dimensions involves rotation and translation. The data for analysis are the coordinates of landmarks on the object, recorded at several time points. Statistical models for describing the rotational component of the movement of an object are proposed in this work. Under the fixed-axis model, the object rotates around an axis that does not change in time. The angular position of the object is then characterised by its rotation axis and an angle giving the extent of the rotation of the object with respect to a reference point. More complicated models occur when the rotation axis varies in time. Under the fixed-angle model the quaternions for the time-varying orientations of the object are shown to lie in a two-dimensional great circle on the surface of the unit sphere in four dimensions. Simple estimators are given of the common rotation axis and of the angles characterising the time varying orientation of the object. A score statistic for testing the fit of the fixed-axis model is constructed and methods for handling the autocorrelation of the errors at neighbouring data points are provided. The analysis of data, collected by an OPTOTRACK camera system on a rotating forearm, illustrates the methodology presented in this paper.

Some key words: Directional data; Joint kinematics; Multivariate; Quaternion; Rotation.

1. Introduction

Consider an object moving in three-dimensional space and suppose that the coordinates of J landmarks, $J > 3$, on the object are available at several time points. These landmarks are assumed to be fixed so that the position of any one relative to the others does not vary in time. This paper proposes a directional model for characterising the displacement of the object using the landmark coordinates.

The movement of an object results in a rigid body motion of its landmarks, involving a translation and a rotation; see for instance McCarthy (1990, Ch. 1). The rotation and the translation components of the motion can each be parameterised by a three-dimensional space. To study movement it is convenient to summarise the landmark data, at each time point, by a vector a and a rotation matrix $R \in \mathrm{SO}(3)$, where $\mathrm{SO}(3)$ is the set of all 3×3 rotations. If y_j $(j = 1, \ldots, J)$ are the landmark coordinates reported by the camera system, then $a = \bar{y}$ is the average coordinate, and rotation R is defined by minimising the sum, for the J landmarks on the object, of the squared distances between $y_j - \bar{y}$ and its predicted value $R(x_j - \bar{x})$, where x_j contains the true coordinates of the jth landmark in a reference position. The least squares estimator of R can be calculated by Procrustes analysis (Dryden & Mardia, 1998, Ch. 5).

The dataset for studying movement is $\{(a_i, R_i): i = 1, \ldots, n\}$, where n is the number of time points, a_i the 3×1 vector of the average landmark position at time i, and $R_i \in \mathrm{SO}(3)$ gives the orientation of the object. The movement is a pure translation if the R_i's are constant, up to experimental errors. It is pure rotation when the object is rotating with respect to a fixed point. This paper focuses on rotational movement since translations involve linear variables that are amenable to standard statistical analyses.

The simplest nontrivial rotational movement involves rotation about an axis that does not vary in time. Let R_j' denote the transpose of R_j. If we neglect errors $R_i R_j'$, the rotation needed to go from the orientation at time j to that at time i, has, under the fixed-axis model, a rotation axis μ that is independent of i and j. The angle of $R_i R_j'$ depends on i and j. In this paper we propose methods for estimating μ and the angles $\{\omega_i : i = 1, \ldots, n\}$ characterising the time varying position of the object. More complicated movements occur when the rotation axis varies in time.

This research is motivated by the analysis of human movement, the moving object being the limb of an experimental subject. The landmarks are on a marker that is attached to the limb of the subject. Their positions are recorded at a given frequency by a camera system. In some cases, see for instance Olshen et al. (1989), a marker consists of one landmark and the camera system records its time-varying coordinates. The statistical methods presented by Ramsay & Silverman (1997) are suitable for this type of data. Markers having a rigid shape and several landmarks allow a better characterisation of the body segment's motion. For instance, the data analysed in Rancourt et al. (2000) were recorded with cross-shaped markers with four landmarks, one at each extremity of the cross. The coordinates of these four landmarks permitted the calculation, at each observation time, of the pair (a, R) defined above.

Statistical models for rotations have been introduced by Downs (1972); see also Jupp & Mardia (1989) and Mardia & Jupp (2000). Prentice (1986) and Rancourt et al. (2000) used quaternions to estimate the mean rotation in a sample. Quaternions are reviewed in § 2. Under the fixed-axis model, the modal values of the quaternions in the sample are shown to lie on a great circle of S^3, the unit sphere in $\Re^4$. Each quaternion is assumed to follow an appropriate bipolar Dimroth–Watson distribution (Mardia & Jupp, 2000, p. 181). Section 2 suggests a simple least squares estimator, $\hat{\mu}$, for the fixed rotation axis. Section 3 highlights a local linear model underlying the fitting of the fixed-axis model. This model enters into the construction of a score test for the fit of the model and into the derivation in § 4 of an inference procedure that accounts for autocorrelation of the errors at neighbouring time points. The motion of a rotating forearm is analysed in § 5.

2. Descriptive statistics for rotations using quaternions

2·1. *Rotations, quaternions and skew-symmetric matrices*

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

$$R(\theta, \mu) = \exp\{\Phi(\theta\mu)\} = I + \Phi(\theta\mu) + \Phi(\theta\mu)^2/2 + \ldots$$
$$= (\cos\theta)I + (\sin\theta)\Phi(\mu) + (1 - \cos\theta)\mu\mu',$$

where $\Phi(\mu)$ is the skew symmetric matrix corresponding to $\mu=(\mu_1,\mu_2,\mu_3)'$:

$$\Phi(\mu)=\begin{pmatrix} 0 & -\mu_3 & \mu_2 \\ \mu_3 & 0 & -\mu_1 \\ -\mu_2 & \mu_1 & 0 \end{pmatrix}. \tag{1}$$

The operator $\Phi(.)$ is associated with the exterior product; if μ and ψ are vectors in $\Re^3$, then the exterior product of μ by ψ is given by $\Phi(\mu)\psi=-\Phi(\psi)\mu$. Thus $\Phi(\mu)\psi$ is a vector orthogonal to both μ and ψ whose length is equal to the product of their lengths multiplied by the sine of the angle between μ and ψ. When μ and ψ are orthogonal unit vectors, $\Phi(\mu)\psi$ is the unit vector such that the 3×3 matrix $[\mu;\,\psi;\,\Phi(\mu)\psi]$ is a right-hand rule rotation.

For statistical manipulations, it is convenient to represent rotation matrices as quaternions. The quaternion associated with $R(\theta,\mu)$ is a unit vector in $\Re^4$ defined by $q(\theta,\mu)=(\cos(\theta/2),\sin(\theta/2)\mu')'$ (Hamilton, 1969). One has $q(\theta,\mu)=-q(\theta+2\pi,\mu)$, so that q and $-q$ represent the same rotation. Thus, all the statistical techniques described in this paper need to be invariant to sign changes in the sample quaternions.

Quaternions are endowed with a special product corresponding to rotation multiplication. Let q_1 and q_2 be the quaternions for rotations R_1 and R_2. As mentioned in McCarthy (1990, p. 61), the quaternion for rotation R_1R_2 is the matrix product $M(q_2)q_1$, where $M(q_2)$ is a 4×4 rotation matrix defined by

$$M(q_2)=\begin{pmatrix} q_{21} & -q_{22}' \\ q_{22} & q_{21}I-\Phi(q_{22}) \end{pmatrix}, \tag{2}$$

in which q_{21} is the first entry of q_2 and q_{22} is the vector of its last three entries. Observe that $q_2^*=(q_{21},-q_{22}')'$ is the quaternion corresponding to R_2' and that $M(q_2^*)=M(q_2)'$. By construction, the function $M(.)$ corresponding to the product R_1R_2 is equal to the product of the M-matrices for R_2 and R_1. In other words,

$$M\{M(q_2)q_1\}=M(q_2)M(q_1). \tag{3}$$

This is used in § 2·2 to set up the deterministic part of the fixed-axis model.

2·2. *The geometry of the fixed-axis model*

We assume that the rotations R_i obey the fixed-axis model exactly, free of experimental errors; that is $R_i=R(\omega_i,\mu)R_0$, for $i=1,\ldots,n$, where μ is the rotation axis, ω_i is the angle giving the position of the object at the ith measurement, and R_0 is related to the orientation of the landmarks on the object. Without loss of generality we can assume that the mean direction of the angles $\{\omega_i : i=1,\ldots,n\}$ is 0; this means that

$$\sum \sin\omega_i=0, \quad \sum\cos\omega_i\geqslant 0. \tag{4}$$

Let q_i be the quaternion corresponding to rotation $R(\omega_i,\mu)R_0$, and let q_0 be the quaternion for R_0. The developments are based on the eigenvalue decomposition of the quaternion cross-product matrix, $\sum q_iq_i'/n$.

The quaternion associated with R_i is $q_i=M(q_0)q_i^0$, where $M(.)$ is defined in (2) and

$$q_i^0=\cos(\omega_i/2)(1,0,0,0)'+\sin(\omega_i/2)(0,\mu')'.$$

The quaternion cross-product matrix is equal to

$$\sum q_i q_i'/n = M(q_0)\begin{pmatrix} \sum \cos^2(\omega_i/2)/n & 0' \\ 0 & \{\sum \sin^2(\omega_i/2)/n\}\mu\mu' \end{pmatrix} M(q_0)'. \tag{5}$$

This result holds since, from (4), $\sum \cos(\omega_i/2)\sin(\omega_i/2) = \sum \sin(\omega_i)/2 = 0$. Since

$$\sum \cos^2(\omega_i/2) - \sum \sin^2(\omega_i/2) = \sum \cos(\omega_i) \geqslant 0,$$

$\sum q_i q_i'/n$ has only two nonnull eigenvalues, $\lambda_i = \sum \cos^2(\omega_i/2)/n$ and $\lambda_2 = \sum \sin^2(\omega_i/2)/n$, with corresponding eigenvectors

$$v_1 = M(q_0)(1, 0, 0, 0)', \quad v_2 = M(q_0)(0, \mu')'.$$

From (2), $q_0 = v_1$, and all the quaternions q_i belong to the great circle on S^3 spanned by v_1 and $v_2 = M(v_1)(0, \mu')'$. Vector μ corresponds to the second, third and fourth entries of $M(v_1)'v_2$. This is recorded formally in the following proposition.

PROPOSITION 1. *The quaternions q_i corresponding, for $i = 1, \ldots, n$, to rotations $R(\omega_i, \mu)R_0$ belong to a great circle of S^3. The fixed rotation axis μ is obtained from the eigenvectors v_1 and v_2 associated with the nonnull eigenvalues $\lambda_1 \geqslant \lambda_2$ of $\sum q_i q_i'/n$:*

$$\mu(v_1, v_2) = -v_{12}v_{21} + v_{11}v_{22} + \Phi(v_{12})v_{22}, \tag{6}$$

where, for $j = 1, 2$, v_{j1} and v_{j2} denote respectively the first and the last three entries of eigenvector v_j.

The function μ, defined by (6), plays a central role in this paper. When the first entries of the quaternions x_1 and x_2 are null, $\mu(x_1, x_2)$ is the exterior product. In terms of the rotations R_1 and R_2 associated with x_1 and x_2, $\mu(x_1, x_2)$ is equal to the axis of the rotation R_2R_1' times the sine of its half angle. The next proposition collects some properties of μ.

PROPOSITION 2. *The function μ defined by (6) maps a pair of vectors in $\Re^4$ into $\Re^3$; it satisfies the following properties, where x_1, x_2, x_3 and x_4 are vectors in $\Re^4$:*

(i) *$\mu(x_1, x_2) = -\mu(x_2, x_1)$ and in particular $\mu(x_1, x_1) = 0$;*
(ii) *μ is linear in its argument, with $\mu(bx_1 + cx_3, x_2) = b\mu(x_1, x_2) + c\mu(x_3, x_2)$, where $b, c \in \Re$;*
(iii) *when x_1, x_2 and x_3 are orthogonal unit vectors, $\mu(x_1, x_2)$ and $\mu(x_1, x_3)$ are orthogonal unit vectors in S^2;*
(iv) *when x_1 is a unit vector then $\mu\{M(x_1)x_2, M(x_1)x_3\} = \mu(x_2, x_3)$.*

The properties listed in Proposition 2 are proved without much difficulty; observe that (iv) comes from (3). When v_1 and v_2 are orthogonal unit vectors, items (i) and (ii) of Proposition 2 imply that

$$\mu(v_1, v_2) = \mu\{(\cos\alpha)v_1 - (\sin\alpha)v_2, (\sin\alpha)v_1 + (\cos\alpha)v_2\},$$

for any angle α in $(0, 2\pi)$. This highlights that, in Proposition 1, the axis μ is a function of the two-dimensional vector space spanned by v_1 and v_2, since it is invariant to changes in the basis of that vector space.

Finally in this section, an orthogonal basis for $\Re^4$ is constructed in terms of arbitrary unit vectors, $v_1 \in S^3$ and $\mu \in S^2$, satisfying $\Phi(\mu)v_{12} \neq 0$. Observe that μ,

$$\mu_1 = \frac{\Phi(\mu)v_{12}}{\{1 - v_{11}^2 - (\mu'v_{12})^2\}^{\frac{1}{2}}}, \quad \mu_2 = \frac{v_{12}'\mu\mu - v_{12}}{\{1 - v_{11}^2 - (\mu'v_{12})^2\}^{\frac{1}{2}}} \tag{7}$$

are orthogonal unit vectors in $\Re^3$ such that the 3×3 matrix (μ, μ_1, μ_2) is a rotation. Thus the 4×4 block diagonal matrix of 1 and (μ, μ_1, μ_2) is a rotation in so(4). Premultiplying each column of the above diagonal matrix by $M(v_1)$, as defined in (2), gives the following orthonormal basis for $\Re^4$:

$$v_1, v_2 = \begin{pmatrix} -\mu' v_{12} \\ v_{11}\mu - \Phi(v_{12})\mu \end{pmatrix}, \quad v_3 = \begin{pmatrix} 0 \\ \gamma(v, \mu) \end{pmatrix}, \quad v_4 = \begin{pmatrix} \gamma(v, \mu)'\mu \\ -\Phi(\mu)\gamma(v, \mu) \end{pmatrix}, \tag{8}$$

where $\gamma(v, \mu) = \{v_{11}\Phi(v_{12})\mu - \Phi(v_{12})^2\mu\}/\{1 - v_{11}^2 - (\mu' v_{12})^2\}^{\frac{1}{2}}$ is a unit vector in $\Re^3$. From (iv) of Proposition 2, $\mu(v_1, v_2) = \mu$.

Observe that the fixed axis μ can be recovered from v_3 and v_4: $\mu(v_1, v_2) = \mu(v_4, v_3)$. Since four parameters, two for each of μ and γ, determine the vector space spanned by v_3 and v_4, a great circle in S^3 is determined by four parameters.

2·3. *Properties of the quaternion cross-product matrix*

Suppose now that the rotations R_i $(i = 1, \ldots, n)$ and their associated quaternions q_i are observed with error. The quaternion cross-product matrix is $\sum q_i q_i'/n$; its eigenvalues are $\hat{\lambda}_1 > \hat{\lambda}_2 > \hat{\lambda}_3 > \hat{\lambda}_4$, with corresponding eigenvectors p_1, p_2, p_3 and p_4. Eigenvector p_1 is the quaternion corresponding to the mean rotation $\bar{R}$, a measure of location for the sample of rotations; see Prentice (1986) and Rancourt et al. (2000). The residual, for rotation i, with respect to the mean rotation can be expressed as $R_i \bar{R}'$. The residual quaternion is $M(p_1)' q_i$, where $M(.)$ is defined in (2).

Let $\{\cos(\omega_i/2), \sin(\omega_i/2)\phi_i'\}'$ denote the loadings of q_i on p_1, p_2, p_3 and p_4, where $\phi_i = (\phi_{i1}, \phi_{i2}, \phi_{i3})'$ is a unit vector in $\Re^3$. Note that $\sin(\omega_i/2)\phi_i = \mu(p_1, q_i)$. One has

$$q_i = \cos(\omega_i/2) p_1 + \sin(\omega_i/2)[p_2 \;\; p_3 \;\; p_4]\phi_i,$$

and the quaternion associated with the residual rotation matrix $M(p_1)' q_i$ is

$$M(p_1)' q_i = \begin{pmatrix} \cos(\omega_i/2) \\ \sin(\omega_i/2)\{\phi_{i1}\mu(p_1, p_2) + \phi_{i2}\mu(p_1, p_3) + \phi_{i3}\mu(p_1, p_3)\} \end{pmatrix},$$

where μ is defined in (6). Thus the angle of the residual rotation is ω_i, while its axis is $\phi_{i1}\mu(p_1, p_2) + \phi_{i2}\mu(p_1, p_3) + \phi_{i3}\mu(p_1, p_3)$. When $\hat{\lambda}_3$ and $\hat{\lambda}_4$ are small, $\phi_i \simeq (1, 0, 0)'$ and the axis of the residual rotation does not change much with i. In this case the fixed-axis model fits well; one can estimate the fixed axis by $\hat{\mu} = \mu(p_1, p_2)$ or $\hat{\mu} = \mu(p_3, p_4)$. When only $\hat{\lambda}_4$ is small, the axis of the residual rotation belongs to the great circle of S^2 spanned by $\mu(p_1, p_2)$ and $\mu(p_1, p_3)$, which are two orthogonal unit vectors. Inferential procedures for $\hat{\mu}$ under the fixed rotation model are derived in § 4·1.

3. The fixed-axis model for samples of rotations

3·1. *Model construction*

In this section $\{q_i : i = 1, \ldots, n\}$ is a sample of quaternions and v_1, v_2, v_3 and v_4 are defined in (8). Under the fixed axis model, the modal value for q_i is

$$u(\omega_i) = v_1 \cos(\omega_i/2) + v_2 \sin(\omega_i/2),$$

where the unknown angles ω_i are assumed to satisfy (4). It is convenient to let $\theta_i = \omega_i/2$, for $i = 1, \ldots, n$.

Quaternion q_i is assumed to be distributed according to a bipolar Dimroth–Watson density, proportional to $\exp[\kappa\{q_i'u(\omega_i)\}^2]$, with shape parameter $\kappa > 0$. Large-concentration inferential procedures, valid when κ goes to infinity, are proposed in this paper. When κ is large, q_i takes its value in the tangent space to S^3 at $u(\omega_i)$. This can be written formally as

$$q_i = u(\omega_i) + \varepsilon_{0i} + o_p(\kappa^{-\frac{1}{2}}) \quad (i = 1, \ldots, n), \tag{9}$$

where ε_{0i} is a 4×1 noise vector satisfying $u(\omega_i)'\varepsilon_{0i} = 0$. From Mardia & Jupp (2000, p. 236) this noise vector has a four-dimensional normal distribution with mean 0 and covariance matrix $\{I - u(\omega_i)u(\omega_i)'\}/(2\kappa)$, denoted by $N_4[0, \{I - u(\omega_i)u(\omega_i)'\}/(2\kappa)]$. Large-concentration inference is common for directional models; see Chang (1993).

The parameter space for the fixed-axis model has $n + 5$ dimensions corresponding to the n angles ω_i, the four degrees of freedom for the direction of the S^3 great circle and parameter κ for error variance. Rivest (1999) investigates a similar model for 3×1 unit vectors scattered about a circle on the surface of S^2.

3·2. *Estimation of the parameters of the fixed-axis model*

The maximum likelihood estimators of the $n + 4$ location parameters maximise the least squares criterion

$$\frac{1}{n}\sum\left\{\cos\left(\frac{\omega_i}{2}\right) q_i'v_1 + \sin\left(\frac{\omega_i}{2}\right) q_i'v_2\right\}^2.$$

For fixed v_1 and v_2, the maximising ω_i's are given by $\hat{\omega}_i = 2\hat{\theta}_i$, where

$$\hat{\theta}_i = \arctan(v_2'q_i, v_1'q_i) \quad (i = 1, \ldots, n), \tag{10}$$

and $\arctan(x, y)$ is the angle θ such that $\sin\theta = x/(x^2 + y^2)$ and $\cos\theta = y/(x^2 + y^2)$. Observe that going from q_i to $-q_i$ changes $\hat{\theta}_i$ to $\hat{\theta}_i + \pi$. This property is used in § 4 to derive regression tests for the fit of the fixed-axis model that are invariant to changes in the signs of the sample quaternions.

The maximum likelihood estimators of the unit vectors v_1 and v_2 for the direction of the great circle maximise

$$\frac{1}{n}\sum\{(q_i'v_1)^2 + (q_i'v_2)^2\}.$$

This is a standard maximisation problem (Rao, 1973, Ch. 1). The maximum value is $\hat{\lambda}_1 + \hat{\lambda}_2$. Any pair of orthogonal unit vectors that spans the same vector space as the two eigenvectors, p_1 and p_2, of $\sum q_iq_i'/n$ associated with eigenvalues $\hat{\lambda}_1$ and $\hat{\lambda}_2$ is a possible solution. Define $\hat{v}_1$ and $\hat{v}_2$ as the pair for which the estimated angles $\hat{\omega}_i$ defined in (10) satisfy (4). It is convenient to let $\hat{v}_j$ $(j = 1, \ldots, 4)$ be the estimators of the vectors in (8) obtained with $\hat{v}_1$ and $\hat{v}_2$.

It is also possible to characterise the vector space spanned by $\hat{v}_1$ and $\hat{v}_2$ through its orthogonal complement which is spanned by orthogonal unit vectors v_3 and v_4 minimising

$$\frac{1}{n}\sum\{(q_i'v_3)^2 + (q_i'v_4)^2\}. \tag{11}$$

The unit vectors $\hat{v}_3$ and $\hat{v}_4$ defined above are a solution to this problem. The minimum value is $\hat{\lambda}_3 + \hat{\lambda}_4$. Finally note that the maximum likelihood estimator for κ, derived by solving an equation similar to (10.3.31) of Mardia & Jupp (2000), is $3/\{2(\hat{\lambda}_3 + \hat{\lambda}_4)\}$. The large number of nuisance parameters makes this estimator heavily biased, and a better one is derived after Proposition 3.

3·3. *Fitting the model when the axis of rotation is known*

Let μ_0 denote the known axis of rotation. From (8), fitting the fixed-axis model involves estimating the unit vector γ. We have

$$v_3 = \begin{pmatrix} 0 \\ I \end{pmatrix} \gamma = A_1 \gamma, \quad v_4 = \begin{pmatrix} \mu_0' \\ -\Phi(\mu_0) \end{pmatrix} \gamma = A_2 \gamma.$$

The optimal γ minimises $\sum\{(q_i' A_1 \gamma)^2 + (q_i' A_2 \gamma)^2\}/n$. The minimising vector is the eigenvector corresponding to the smallest eigenvalue of $\sum(A_1' q_i q_i' A_1 + A_2' q_i q_i' A_2)/n$, while the smallest eigenvalue, $\hat{\lambda}_3^0$ say, is the minimal value for (11) when $\mu = \mu_0$.

4. Statistical inference for the fixed-axis model

4·1. *The sampling distribution of the estimator of the rotation axis*

The statistical analysis is concerned with the estimation of the two-dimensional subspace of $\Re^4$ spanned by the unit vectors v_3 and v_4. Its least squares estimator, obtained by minimising (11), should be a subspace in the neighbourhood of the true value. An orthogonal basis for the subspace spanned by the unit vectors minimising (11) is

$$\tilde{v}_3 = v_3 + \hat{\beta}_1 v_1 + \hat{\beta}_2 v_2 + o_p(\kappa^{-\frac{1}{2}}), \quad \tilde{v}_4 = v_4 + \hat{\beta}_3 v_1 + \hat{\beta}_4 v_2 + o_p(\kappa^{-\frac{1}{2}}), \tag{12}$$

where the $O_p(\kappa^{-\frac{1}{2}})$ vector $\hat{\beta} = (\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3, \hat{\beta}_4)'$ determines the position of the fitted subspace with respect to that spanned by v_3 and v_4. Observe that $\tilde{v}_3 \neq \hat{v}_3$ and $\tilde{v}_4 \neq \hat{v}_4$; for instance, from (8), the first entry of $\hat{v}_3$ is null while that of $\tilde{v}_3$ is possibly nonnull. Indeed $\tilde{v}_3$ and $\tilde{v}_4$ are not estimable individually. They are a convenient mathematical representation for the basis of the vector space spanned by $\hat{v}_3$ and $\hat{v}_4$ since they are linear in v_1 and v_2. By contrast, $\hat{v}_3$ and $\hat{v}_4$ parameterise the fitted vector space using two S^2 unit vectors, μ and γ. This nonlinear parameterisation, used in a previous version of this work, resulted in laborious derivations.

The aim of this section is to determine the first-order contributions of the errors to the vector $\hat{\beta}$. From (11) and (9), $\hat{\beta}$ is given by the values that minimise

$$\frac{1}{n} \sum_{i=1}^{n} [\{v_3' \varepsilon_{0i} + (\cos\theta_i)\beta_1 + (\sin\theta_i)\beta_2\}^2 + \{v_4' \varepsilon_{0i} + (\cos\theta_i)\beta_3 + (\sin\theta_i)\beta_4\}^2] + o_p(\kappa^{-1}).$$

Finding the vector that minimises the leading, $O_p(\kappa^{-1})$, term of this expression gives the first-order contribution of the errors to $\hat{\beta}$. This term can be expressed as $(\varepsilon - X\beta)'(\varepsilon - X\beta)$, where ε is the $2n \times 1$ vector of the errors. The two error components for the ith data point are $\varepsilon_{2i-1} = v_3' \varepsilon_{0i}$ and $\varepsilon_{2i} = v_4' \varepsilon_{0i}$, where ε_{0i} is defined in (9). The design matrix X is given by

$$X = -\begin{pmatrix} \cos\theta_1 & \sin\theta_1 & 0 & 0 \\ 0 & 0 & \cos\theta_1 & \sin\theta_1 \\ \cos\theta_2 & \sin\theta_2 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \cos\theta_n & \sin\theta_n & 0 & 0 \\ 0 & 0 & \cos\theta_n & \sin\theta_n \end{pmatrix}. \tag{13}$$

This yields an approximation for $\hat{\beta}$,

$$\hat{\beta} = (X'X)^{-1}X'\varepsilon + o_p(\kappa^{-\frac{1}{2}}),$$

and for the sum of the squared residuals,

$$n(\hat{\lambda}_3 + \hat{\lambda}_4) = \varepsilon'\{I - X(X'X)^{-1}X'\}\varepsilon + o_p(\kappa^{-1}).$$

The following proposition applies standard results on linear models to the fixed-axis model.

PROPOSITION 3. *When the errors have independent Dimroth–Watson distributions, as the shape parameter κ goes to ∞,*

(i) $(2\kappa)^{\frac{1}{2}}\varepsilon$ *converges in distribution to a $N_{2n}(0, I)$ distribution;*
(ii) $2n\kappa(\hat{\lambda}_3 + \hat{\lambda}_4)$ *converges to a chi-squared distribution with $2n-4$ degrees of freedom;*
(iii) $(2\kappa)^{\frac{1}{2}}\hat{\beta}$ *converges to a $N_4\{0, (X'X)^{-1}\}$ distribution;*
(iv) $(\hat{\lambda}_3 + \hat{\lambda}_4)$ *and $\hat{\beta}$ are asymptotically independent.*

The first part of Proposition 3 is an easy consequence of (9). Part (ii) suggests estimating κ by $\hat{\kappa} = (n-2)/\{n(\hat{\lambda}_3 + \hat{\lambda}_4)\}$. The normality of the error vector ε is the key result of Proposition 3. It can be ascertained by applying regression diagnostics to the residuals defined in § 4·2.

Consider the problem of testing whether or not the rotation axis is equal to a known unit vector μ_0. The null hypothesis is $H_0: \mu = \mu_0$. An F-test is easily constructed by comparing the two values of (11) under the null and the alternative hypotheses, giving

$$F_{\text{obs}} = \frac{(\hat{\lambda}_3^0 - \hat{\lambda}_3 - \hat{\lambda}_4)/2}{(\hat{\lambda}_3 + \hat{\lambda}_4)/(2n-4)},$$

where $\hat{\lambda}_3^0$ is defined in § 3·3. If the assumptions of Proposition 3 are met, the asymptotic distribution of this statistic is, under the null hypothesis, that of $F(2, 2n-4)$. This result holds true since one can show that the column space of the $2n \times 2$ design matrix for the local linear model for a known axis is spanned by the column space of the matrix X.

The estimated rotation axis is $\hat{\mu} = \mu(\tilde{v}_4, \tilde{v}_3)$. In view of (12) and of the linearity of μ, repeated application of part (ii) in Proposition 2 yields, if we neglect $o_p(\kappa^{-\frac{1}{2}})$ terms,

$$\begin{aligned}
\mu(\tilde{v}_4, \tilde{v}_3) &\doteq \mu(v_4, v_3) + \mu(\tilde{v}_4 - v_4, v_3) + \mu(v_4, \tilde{v}_3 - v_3) \\
&\doteq \mu(v_4, v_3) + \hat{\beta}_1\mu(v_4, v_1) + \hat{\beta}_2\mu(v_4, v_2) + \hat{\beta}_3\mu(v_1, v_3) + \hat{\beta}_4\mu(v_2, v_3) \\
&\doteq \mu(v_4, v_3) + (\hat{\beta}_2 + \hat{\beta}_3)\mu_1 + (\hat{\beta}_4 - \hat{\beta}_1)\mu_2, \qquad (14)
\end{aligned}$$

where μ_1 and μ_2 are defined in (7). Under constraints (4) on the ω_i's, $(X'X)^{-1}$ is a diagonal matrix with diagonal entries given by $1/\sum(\cos\theta_i)^2$, $1/\sum(\sin\theta_i)^2$, $1/\sum(\cos\theta_i)^2$ and $1/\sum(\sin\theta_i)^2$. Thus $\hat{\beta}_2 + \hat{\beta}_3$ and $\hat{\beta}_4 - \hat{\beta}_1$ are, for large κ's, independent normally distributed with variance $n/\{(2\kappa)\sum\cos(\omega_i/2)^2\sum\sin(\omega_i/2)^2\}$. From (5) and Proposition 3, an estimator of this variance is given by $(\hat{\lambda}_3 + \hat{\lambda}_4)/\{(2n-4)\hat{\lambda}_1\hat{\lambda}_2\}$. Therefore, an estimator of the covariance matrix of $\hat{\mu}$ is given by

$$v(\hat{\mu}) = \frac{\hat{\lambda}_3 + \hat{\lambda}_4}{(2n-4)\hat{\lambda}_1\hat{\lambda}_2}(I - \hat{\mu}\hat{\mu}').$$

4·2. *Investigating the fit of the model*

In many instances the hypotheses of Proposition 3 are not fulfilled. The rotation axis varies and the errors may exhibit autocorrelation. The linear model of § 4·1 is a convenient tool for investigating these questions.

The entries of the X-matrix can be estimated as the sines and the cosines of the angles $\hat{\theta}_i$ defined in (10); let $\hat{X}$ denote the estimated design matrix. Since the difference between $\hat{X}$ and X is $O(\kappa^{-\frac{1}{2}})$, replacing X by $\hat{X}$ in the calculations does not change the first-order properties of the estimators. The residuals for this linear model can be defined as a $2n \times 1$ vector r whose components for the ith data point are given by $r_{2i-1} = \hat{v}_3' q_i$ and $r_{2i} = \hat{v}_4' q_i$. Observe that replacing the quaternion q_i for the ith data point by $-q_i$ changes the signs of the two residuals and of the two rows of $\hat{X}$ for that data point.

To derive a score test for the fit of the model, suppose that the rotation axis varies with the rotation angle and that, at time i, it is equal to

$$\mu + \beta_5(\cos\theta_i^*)\mu_1 + \beta_6(\sin\theta_i^*)\mu_1 + \beta_7(\cos\theta_i^*)\mu_2 + \beta_8(\sin\theta_i^*)\mu_2 + O(\kappa^{-1}), \tag{15}$$

where $\theta_i^* = \theta_i$ modulo π and θ_i^* belongs to $(-\pi/2, \pi/2)$, μ_1 and μ_2 are defined in (7) and β_5, β_6, β_7 and β_8 are $O(\kappa^{-\frac{1}{2}})$ unknown parameters. Adding to (15) terms in sines and cosines of $2\theta_i^*, 3\theta_i^*, \ldots$ would permit us to model complex relationships between the rotation axis and the rotation angle. When the rotation axis changes according to (15), the modal value for q_i becomes

$$u(\omega_i) + v_3(\sin\theta_i)(\beta_5\cos\theta_i^* + \beta_6\sin\theta_i^*) + v_4(\sin\theta_i)(\beta_7\cos\theta_i^* + \beta_8\sin\theta_i^*) + O(\kappa^{-1}).$$

Thus, the errors ε_{2i-1} and ε_{2i} have systematic components respectively given by

$$\beta_5\sin\theta_i\cos\theta_i^* + \beta_6\sin\theta_i\sin\theta_i^*, \quad \beta_7\sin\theta_i\cos\theta_i^* + \beta_8\sin\theta_i\sin\theta_i^*,$$

and therefore the model with a varying rotation axis is related to a local linear model with a $2n \times 8$ design matrix given by $[X; X_c]$, where

$$X_c = -\begin{pmatrix} \sin\theta_1\cos\theta_1^* & \sin\theta_1\sin\theta_1^* & 0 & 0 \\ 0 & 0 & \sin\theta_1\cos\theta_1^* & \sin\theta_1\sin\theta_1^* \\ \sin\theta_2\cos\theta_2^* & \sin\theta_2\sin\theta_2^* & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \sin\theta_n\cos\theta_n^* & \sin\theta_n\sin\theta_n^* & 0 & 0 \\ 0 & 0 & \sin\theta_n\cos\theta_n^* & \sin\theta_n\sin\theta_n^* \end{pmatrix}.$$

When the fixed-axis model fits a dataset well, the hypothesis $H_0: \beta_5 = \beta_6 = \beta_7 = \beta_8 = 0$ should be accepted. Let $n\hat{\lambda}_c$ be the sum of the squared residuals for the regression of r on $[\hat{X}; \hat{X}_c]$, where $\hat{X}_c$ is an estimate for X_c. Note that $n\hat{\lambda}_c$ is also the sum of the squared residuals of the regression of the unobserved ε on $[\hat{X}; \hat{X}_c]$ since the vector $\varepsilon - r$ is in the vector space spanned by the columns of $\hat{X}$. The F-test with 4 and $2n-8$ degrees of freedom rejects H_0 at level α if

$$\frac{(\hat{\lambda}_3 + \hat{\lambda}_4 - \hat{\lambda}_c)/4}{\hat{\lambda}_c/(2n-8)} > F_{4,2n-8,\alpha}.$$

Note also that this test is invariant to changes of any quaternion from q_i to $-q_i$.

When movement is studied, the data points are ordered in time and autocorrelation is likely. This feature can be brought to the model by assuming that the odd and even entries of the error vector ε are two independent realisations of an AR(1) process with autocorrelation ρ and stationary variance $1/(2\kappa)$. An estimator of the first-order residual autocorrelation is given by

$$\hat{\rho} = \frac{\sum_{i=1}^{n-1} (r_{2i+1} r_{2i-1} + r_{2i+2} r_{2i}) \operatorname{sgn}\{\cos(\hat{\theta}_i - \hat{\theta}_{i+1})\}}{\sum_{i=1}^{n} (r_{2i-1}^2 + r_{2i}^2)}.$$

This statistic fulfils the variance condition since changing q_i into $-q_i$ changes the signs of r_{2i} and r_{2i-1} and transforms $\hat{\theta}_i$ into $\hat{\theta}_i + \pi$. In many instances the quaternions do not change much in time and one can define the sign of q_i, for given q_{i-1}, such that $q_i' q_{i-1} \geqslant 0$. In such situations, one can omit $\operatorname{sgn}\{\cos(\hat{\theta}_i - \hat{\theta}_{i+1})\}$ from the definition of $\hat{\rho}$.

Estimators and tests with the Cochrane & Orcutt (1949) correction for autocorelation are easily implemented. To obtain an estimator for μ and its covariance matrix when autocorrelation is present it suffices to regress $r_a = r_{(-1)} - \hat{\rho} r_{(-n)}$ on $X_a = X_{(-1)} - \hat{\rho} X_{(-n)}$, where the subscript $(-i)$ means that the two rows for data point i have been deleted. Let $\hat{\beta}_{1a}$, $\hat{\beta}_{2a}$, $\hat{\beta}_{3a}$ and $\hat{\beta}_{4a}$ be the least squares estimators of the regression parameters. An estimator for μ adjusted for autocorrelation is easily derived from (14):

$$\hat{\mu}_a = \frac{\hat{\mu} + \hat{\mu}_1(\hat{\beta}_{2a} + \hat{\beta}_{3a}) + \hat{\mu}_2(\hat{\beta}_{4a} - \hat{\beta}_{1a})}{\{1 + (\hat{\beta}_{2a} + \hat{\beta}_{3a})^2 + (\hat{\beta}_{4a} - \hat{\beta}_{1a})^2\}^{\frac{1}{2}}},$$

where $\hat{\mu}_1$ and $\hat{\mu}_2$ are the least squares estimators for (7). A covariance estimator corrected for autocorrelation is given by

$$v(\hat{\mu}_a) = \hat{\sigma}_a^2 [\hat{\mu}_1;\, \hat{\mu}_2] \begin{pmatrix} 0 & -1 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}' (X_a' X_a)^{-1} \begin{pmatrix} 0 & -1 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} [\hat{\mu}_1;\, \hat{\mu}_2]',$$

where $\hat{\sigma}_a^2$ is an estimator of the residual variance. The score test for a fixed rotation axis can also be corrected along the lines of Cochrane & Orcutt (1949).

The residual vector r and the estimated design matrices $\hat{X}$ and $\hat{X}_c$ can be entered into any standard package for mixed linear models to fine-tune the analysis. An alternative to the Cochrane–Orcutt analysis is to regress r on $\hat{X}$ and $\hat{X}_c$ in a mixed linear model package such as SAS proc mixed or S-Plus lme, and to model the odd and even entries of r as repeated measures, with an AR(1) covariance matrix. Estimators of μ and of its covariance matrix can then be derived from (14); tests on the parameters for $\hat{X}_c$ investigate the fit of the model.

5. DATA ANALYSIS

5·1. *The experimental set-up*

To illustrate the methodology presented in this paper, a simple experiment was conducted in the laboratory of Brad McFadyen of the Rehabilitation Research Center at Université Laval. While standing with his right elbow resting on a stool and his upper arm still, a subject extended his right forearm while holding a planar marker, with six landmarks, in his right hand. The movement recorded should therefore correspond to a rotation of the forearm at the elbow joint. Since the subject was standing, with his upper arm in a vertical position, the total extension of the forearm, from the vertical upper arm to the horizontal stool, is less than 90°. The X–Y–Z coordinates, in metres, of the six landmarks were recorded by a OPTOTRACK camera system during one second at a frequency of 100 hertz. This led to $n=99$ records for each marker. The dataset for analysis is $\{(a_i, R_i): i=1, \ldots, 99\}$, where a_i and R_i are respectively the average landmark position and the orientation at time i.

5·2. *The fixed-axis model*

The eigenvalues for the quaternion cross-product matrix are $\hat{\lambda}_1=0{\cdot}970282$, $\hat{\lambda}_2=0{\cdot}029591$, $\hat{\lambda}_3=0{\cdot}000119$ and $\hat{\lambda}_4=0{\cdot}000008$. Since the first eigenvalue is related to the mean rotation, the proportion of variability explained by the fixed-axis model is $\hat{\lambda}_2/(\hat{\lambda}_2+\hat{\lambda}_3+\hat{\lambda}_4)$. This is equal to 99·57%, suggesting that the fixed-axis model fits well. The estimate of the rotation axis is given by $\hat{\mu}=(0{\cdot}858, -0{\cdot}146, -0{\cdot}492)'$.

The score test for changing rotation axis, introduced in § 4·2, allow us to evaluate the fit of the fixed-axis model. The autocorrelation of the residuals of the regression of r on $[\hat{X}; \hat{X}_c]$ is $\hat{\rho}=0{\cdot}88$, and a Cochrane–Orcutt corrected F-statistic for $H_0: \beta_5=\beta_6=\beta_7=\beta_8=0$ is 33·12 on 4 and 188 degrees of freedom. The test therefore suggests that the rotation axis does vary in time. Significant results were also obtained using the MIXED procedure of SAS, using either the MINQUE or the restricted maximum likelihood estimator for the AR(1) variance covariance matrix for the odd and even residuals.

To investigate whether or not the rotation axis varies on a great circle, one can compare the fit of the eight-dimensional model with that of a six-dimensional model with the axis on the great circle spanned by $\hat{\mu}$ and $\mu(p_1, p_3)$. Under this model, the space of axes spanned by (15) is two-dimensional; its design matrix involves X and two linear combinations of the columns of X_c. The F-statistic for comparing these two models is 2·94 on 2 and 188 degrees of freedom giving a p-value of 5·5%, which suggests that the axis varies on a great circle.

To investigate variations in the rotation axis, the residual rotations $R_i\bar{R}'$ are available, where $\bar{R}$ is the mean rotation as defined in § 2·3. The quaternion for $R_i\bar{R}'$ is $M(p_1)'q_i$. The rotation axis of $R_i\bar{R}'$ is proportional to $t_i=\mu(p_1, q_i)$. To avoid difficulties caused by rotations through small angles, the axis of $R_i\bar{R}'$ can be estimated by the eigenvector corresponding to the largest eigenvalue of $\sum_{i=-6}^{+6} t_{i+j}t'_{i+j}$. The projection of these unit vectors in the plane spanned by $\mu(p_1, p_3)$ and $\mu(p_1, p_4)$ is given in Fig. 1, where the p_j's are the quaternion cross-product matrix eigenvectors. Sequential identification numbers in Fig. 1 show that the largest deviations in the rotation axis occur for data points 45 to 60, corresponding to residual rotations of angles less than 0·1 radians, that is 6°. If we omit these points, the variations in the rotation axis are of the order of $\pm 6°$ in the direction of the third eigenvector. This agrees with the findings in the orthopaedic literature (An et al., 1984) who reported changes of a similar magnitude in the carrying angle during flexion extensions at the elbow joint. Note also that the small changes in the direction of the fourth principal component in Fig. 1 confirm that the rotation axis varies on a great circle.

5·3. *The kinematics of the elbow joint*

Since the fixed-axis model does not fit well, are the angles $\hat{\omega}_i$ representative of the movement? Can they be used to calculate the speed and the acceleration of the upper arm during a flexion of the elbow joint? The answer to these questions is a tentative 'yes'. The fixed-axis model forces the $R_i\bar{R}'$ rotation axis to be $\hat{\mu}$. This should have little impact on

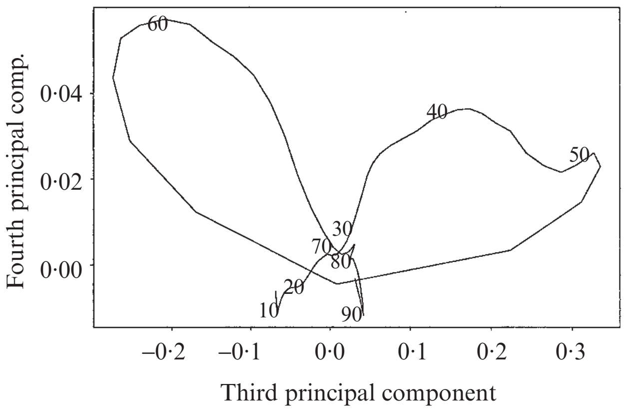

Fig. 1. Variations (radians) of the rotation axis, with identification numbers provided for nine data points.

its angle since, as shown in Rivest (1989), the angle and the axis of a rotation are locally orthogonal. To investigate this numerically one can compare the fitted rotation angle between time i and $i+1$, $\hat{\omega}_i - \hat{\omega}_{i+1}$, with its empirical counterpart $2 \arccos(q'_{i+1} q_i)$. The correlation between fitted and empirical angles is 0·984, indicating close agreement.

To investigate the kinematics of the elbow joint we can fit the 99 $\hat{\omega}_i$ values with 16 order-4 B-spline functions with equally spaced knots using S-Plus functions available with Ramsey & Silverman (1997). Figure 2 gives a time plot of the fitted ω_i's and of the acceleration. The largest accelerations are at the beginning and end of the movement. Noteworthy are the eight acceleration peaks in the graph; this figure agrees with Ramsay's (2000) acceleration pattern of 8 peaks per second.

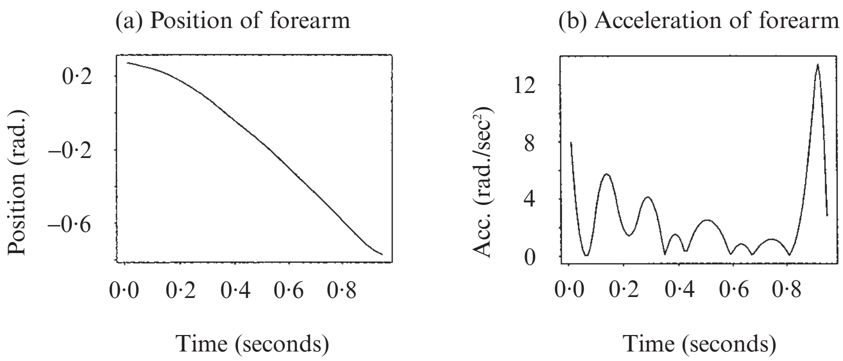

Fig. 2. (a) Angular position (radians) and (b) acceleration (radians/sec^2), of the rotating forearm.

ACKNOWLEDGEMENT

I am grateful to Denis Rancourt and Brad McFadyen for making the dataset available, to Mireille Guay and Charles Dumont for their assistance with data analysis, to a referee for his careful reading of the manuscript and to the Natural Sciences and Engineering Research Council of Canada and the Fonds pour la Formation des Chercheurs et l'aide à la Recherche du Québec for their financial support.

References

An, K. N., Morrey, B. F. & Chao, E. Y. S. (1984). Carrying angle of the human elbow joint. *J. Orthoped. Res.* **1**, 369–78.

Chang, T. D. (1993). Spherical regression and the statistics of tectonic plate reconstruction. *Int. Statist. Rev.* **61**, 299–316.

Cochrane, D. & Orcutt, G. H. (1949). Application of least squares regression to relationships containing autocorrelated error terms. *J. Am. Statist. Assoc.* **44**, 32–61.

Downs, T. D. (1972). Orientation statistics. *Biometrika* **59**, 665–76.

Dryden, I. L. & Mardia, K. V. (1998). *Statistical Analysis of Shape*. New York: John Wiley.

Hamilton, W. R. (1969). *Elements of Quaternions*. Chelsea: Chelsea Press.

Jupp, P. E. & Mardia, K. V. (1989). A unified view of the theory of directional statistics, 1975–1988. *Int. Statist. Rev.* **57**, 261–94.

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

McCarthy, J. M. (1990). *Introduction to Theoretical Kinematics*. Cambridge, MA: MIT Press.

Olshen, R. A., Biden, E. N., Wyatt, M. P. & Sutherland, D. H. (1989). Gait analysis and the bootstrap. *Ann. Statist.* **17**, 1419–40.

Prentice, M. J. (1986). Orientation statistics without parametric assumptions. *J. R. Statist. Soc.* B **48**, 214–22.

Ramsay, J. O. (2000). Functional component of variation in hand-writing. *J. Am. Statist. Assoc.* **95**, 9–15.

Ramsay, J. O. & Silverman, B. W. (1997). *Functional Data Analysis*. New York: Springer Verlag.

Rancourt, D., Rivest, L. P. & Asselin, J. (2000). Using orientation statistics to investigate variations in human kinematics. *Appl. Statist.* **15**, 39–50.

Rao, C. R. (1973). *Linear Statistical Inference and its Applications*, 2nd ed. New York: John Wiley.

Rivest, L.-P. (1989). Spherical regression for concentrated Fisher–von Mises distributions. *Ann. Statist.* **17**, 307–17.

Rivest, L.-P. (1999). Some linear model techniques for analyzing small circle spherical data. *Can. J. Statist.* **27**, 623–38.

[*Received April* 2000. *Revised December* 2000]