The Annals of Statistics
1989, Vol. 17, No. 1, 307–317

SPHERICAL REGRESSION FOR CONCENTRATED FISHER–VON MISES DISTRIBUTIONS[1]

BY LOUIS-PAUL RIVEST

Université Laval

Spherical regression studies models which postulate that the unit vector v is equal to an unknown rotation P of the unit vector u "plus" an experimental error. The case where the experimental errors follow a Fisher–von Mises distribution with a large concentration parameter κ is considered in this work. Asymptotic ($\kappa \to \infty$) inferential procedures for P are proposed when n, the sample size, is fixed. Diagnostic methods for spherical regression are suggested. The key for their derivation is the fact that spherical regression is "locally" identical to ordinary least square regression. The results are presented in an arbitrary dimension. For the three-dimensional case, asymptotic tests and confidence regions for the axis and the angle of P are obtained. The data from a plate tectonic analysis of the Gulf of Aden, presented by Cochran, illustrate the proposed methodology.

1. Introduction. In spherical regression, the dependent vector v, which belongs to S_{k-1}, the unit sphere in R^k, is assumed to be equal to an unknown rotation P of a fixed S_{k-1}-vector u perturbed by an experimental error. If the experimental error follows the Fisher–von Mises distribution, the density of v is equal to

$$\frac{\kappa^{k/2-1}\exp(\kappa v'Pu)}{I_{k/2-1}(\kappa)(2\pi)^{k/2}}, \tag{1}$$

where $\kappa > 0$ is the concentration parameter and $I_{k/2-1}$ denotes a modified Bessel function [Abramowitz and Stegun (1972)]. Its distribution is labelled $F(Pu, \kappa)$; v is uniformly distributed and $\kappa = 0$, if and only if u and v are independent. When it is not, the assumption of independence cannot be written in terms of the parameters indexing (1). If rotational dependence is in doubt, it should first be ascertained with correlation measures [Jupp and Mardia (1980) and Rivest (1988)].

Spherical regression was first considered by Chang (1986); see also Chang (1987). He derived large sample inferential procedures for $\hat{P}$, the maximum likelihood estimate of P, based on observations $\{v_i, u_i\}_{i=1}^n$. Chang's results only assumed a rotationally symmetric error distribution. This work looks at spherical regression from a different angle: n is fixed and $\kappa \to \infty$. Watson (1984) and Rivest (1986) studied the distributions of directional statistics in this setting.

Received August 1986; revised June 1988.

[1]This work was supported by a Natural Sciences and Engineering Research Council of Canada grant.

AMS 1980 *subject classifications*. Primary 62H15; secondary 62J05, 62P99.

Key words and phrases. Cook's D statistic, diagnostic methods, directional data, Fisher–von Mises distribution, linear models.

Section 2 shows that, when $\kappa \to \infty$, spherical regression is identical to ordinary least squares. Section 3 adapts the least square tests to the spherical context. It provides statistics to test hypotheses of the type H_0: $P \in G$, where G is a closed subset of $SO(k)$, the set of $k \times k$ rotations. Diagnostic statistics based on residuals are proposed. Section 4 studies the special case $k = 3$. A new analysis of Chang's (1986) example is presented in Section 5.

2. The small sample asymptotic distribution of $\hat{P}$. The rotation $\hat{P}$ is the one maximizing $\sum v_i' P u_i / n$. MacKenzie (1957) and Stephens (1979) showed how to calculate $\hat{P}$. If

$$\sum \frac{u_i v_i'}{n} = S \operatorname{diag}(l_1, \ldots, l_k) T'$$

is a singular value decomposition [S and T are $SO(k)$ matrices and $l_1 \geq l_2 \geq \cdots \geq |l_k|$ are the singular values], then $\hat{P} = TS'$. For large κ, a good approximation to the maximum likelihood estimate of κ is [Watson (1983), page 163]

$$\hat{\kappa} = \frac{k-1}{2}\left(1 - \overline{R}(SO(k))\right)^{-1},$$

where $\overline{R}(SO(k)) = \sum v_i' \hat{P} u_i / n$.

To study the distribution of $\hat{P}$, a parametrization in terms of skew-symmetric matrices is useful: If $E(l, m)$ denotes, for $l > m$, a $k \times k$ matrix of 0's except for its (l, m) and (m, l) components which are equal to 1 and -1, respectively, and $\{A_{lm}\}_{l>m}$ are real numbers, then

$$A = \sum_{l>m} A_{lm} E(l, m)$$

is a skew-symmetric matrix (it satisfies $A = -A'$) and

$$P = \exp A = \sum_0^{\infty} \frac{A^j}{j!}$$

is a rotation. Let $a = (A_{21}, A_{31}, \ldots, A_{k1}, A_{32}, \ldots, A_{k(k-1)})'$ be the component vector of A, and note that for any u_i in S_{k-1}, $Au_i = U_i a$, where U_i is a $k \times k(k-1)/2$ matrix whose (l, m) column is $E(l, m)u_i$. In differential geometry, the set of the $k \times k$ skew-symmetric matrices is called the Lie algebra of the Lie group $SO(k)$ [Warner (1983), page 84].

Concentrated Fisher–von Mises random variables can be represented in terms of normal variates [Watson (1983), page 157]: If $\varepsilon_i = (\varepsilon_{i1}, \ldots, \varepsilon_{ik-1})'$, $i = 1, \ldots, n$, are independent random vectors distributed as $N_{k-1}(0, \kappa^{-1}I)$, then

$$P'v_i =_d \left(1 - \frac{\|\varepsilon_i\|^2}{2} + o_p(\kappa^{-1})\right) u_i + u_{i(\cdot)}\left(\varepsilon_i + o_p(\kappa^{-1/2})\right), \tag{2}$$

where $u_{i(\cdot)}$ is a $k \times (k-1)$ matrix whose columns form an orthogonal basis of the vector space orthogonal to u_i and $=_d$ means equality in distribution.

One can write $\hat{P} = P \exp \hat{A}$. The distribution of $\hat{P}$ is characterized by that of $\hat{a}$, the component vector of $\hat{A}$.

THEOREM. *Assuming that n is fixed and that* $\Sigma_i^n U_i' U_i$ *is nonsingular, as* κ *goes to* ∞, *the following results hold*:

(i) $\sqrt{\kappa}\,\hat{a} \to_l N_{k(k-1)/2}(0, (\Sigma U_i' U_i)^{-1})$.

(ii) $2n\kappa(1 - \bar{R}(SO(k))) \to_l \chi^2_{(k-1)(n-k/2)}$.

(iii) $\hat{a}$ *and* $\bar{R}(SO(k))$ *are asymptotically independent.*

PROOF. One can let, without losing generality, $P = I$. For A's that are $O(\kappa^{-1/2})$,

$$R(a) = \frac{\Sigma v_i' \exp A u_i}{n} = \frac{\Sigma v_i' u_i}{n} + \frac{\Sigma v_i' A u_i}{n} + \frac{\Sigma v_i' A^2 u_i}{2n} + o_p(\kappa^{-1}).$$

Using (2), and the equality $Au_i = U_i a$, the last expression is, up to $o_p(\kappa^{-1})$, equal to

$$\frac{\Sigma v_i' u_i}{n} + \frac{\Sigma v_i' U_i a}{n} - \frac{a' \Sigma U_i' U_i a}{n}. \tag{3}$$

Straightforward calculus shows that (3) is maximized by

$$\tilde{a} = \left(\sum U \Sigma_i' U_i\right)^{-1} \sum U_i' v_i.$$

Using a Taylor series expansion and the fact that R is maximum at $\hat{a}$,

$$R(\hat{a}) - R(\tilde{a}) = (\hat{a} - \tilde{a}) \frac{\partial}{\partial a} R(a)\Big|_{a=\hat{a}} + \frac{(\hat{a} - \tilde{a})'}{2} \frac{\partial^2}{\partial a^2} R(a)\Big|_{a=a_0} (\hat{a} - \tilde{a})'$$

$$= \frac{(\hat{a} - \tilde{a})'}{2} \frac{\partial^2}{\partial a^2} R(a)\Big|_{a=a_0} (\hat{a} - \tilde{a}),$$

where a_0 belongs to the segment joining $\tilde{a}$ and $\hat{a}$. Since $R(\tilde{a}) - R(\hat{a})$ is $o_p(\kappa^{-1})$ and $\partial^2/\partial a^2\, R(a)$ is $o_p(1)$, $\hat{a} - \tilde{a}$ is $o_p(\kappa^{-1/2})$.

The maximization of (3) can locally be viewed as a least squares problem. Using (2), (3) can be rewritten as

$$1 - \frac{1}{2n}\left(\sum \|\varepsilon_i\|^2 - 2\sum \varepsilon_i' u_{i(\cdot)}' U_i a + a' \sum U_i' U_i a\right).$$

Since $U_i' U_i = U_i' u_{i(\cdot)} u_{i(\cdot)}' U_i$, if $\varepsilon = (\varepsilon_1', \varepsilon_2', \ldots, \varepsilon_n')'$ and if X denotes an $n(k-1) \times k(k-1)/2$ matrix whose rows $(i-1)(k-1)+1$ to $i(k-1)$ are equal to $u_{i(\cdot)}' U_i$, then (3) is equal to

$$1 - \frac{1}{2n}(\varepsilon - Xa)'(\varepsilon - Xa) \tag{4}$$

and $\tilde{a}$ is the value maximizing this expression. By standard least squares theory, up to $o_p(\kappa^{-1/2})$,

$$\kappa^{1/2}\tilde{a} \sim N_{k(k-1)/2}\left(0, (X'X)^{-1}\right),$$

$$\kappa(y - X\tilde{a})'(y - X\tilde{a}) \sim \chi^2_{(k-1)(n-k/2)};$$

furthermore $\tilde{a}$ and $(y - X\tilde{a})'(y - X\tilde{a})$ are independent. Since $(\tilde{a} - \hat{a})$ is $o_p(\kappa^{-1/2})$ these results remain true if $\tilde{a}$ is replaced by $\hat{a}$. □

As a function of the skew-symmetric matrix A, the asymptotic density of $\hat{A}$ (or $\hat{a}$) is proportional to

$$\exp\left(\frac{\operatorname{tr}\left(A^2 \Sigma u_i u_i'\right)}{2\kappa}\right).$$

Theorem 1 of Chang (1986) presents a similar expression for the large sample density of $\hat{A}$ (or in Chang's notation H_n) when κ is fixed.

3. Hypothesis testing. Let G be a g-dimensional closed subset of $SO(k)$. Let $\hat{P}_G$ be the G-rotation maximizing $\Sigma v_i' P u_i / n$ and $\overline{R}(G) = \Sigma v_i' \hat{P}_G u_i / n$. To test H_0: $P \in G$ versus H_A: $P \notin G$, the theorem applies in most cases. Under H_0, $\hat{P}_G$ is in an infinitesimal neighborhood of P. If this neighborhood can be parametrized by $V(G, P)$, a g-dimensional vectorial subspace of $R^{k(k-1)/2}$, then

$$\hat{P}_G = P\left(I + \hat{A}_G\right) + o_p(\kappa^{-1/2}),$$

where $\hat{a}_g \in V(G, P)$ and, as in the theorem,

$$\overline{R}(G) =_d 1 - \frac{1}{2n} \min_{a \in V(G, P)} (y - Xa)'(y - Xa) + o_p(\kappa^{-1}).$$

Classical least squares theory suggests use of

$$F_{\text{obs}} = \frac{(k-1)(n-k/2)}{k(k-1)/2 - g} \frac{\overline{R}(SO(k)) - \overline{R}(G)}{1 - \overline{R}(SO(k))} \tag{5}$$

as a test statistic; its null distribution is $F_{k(k-1)/2-g,(k-1)(n-k/2)}$.

G is locally parametrizable by a vectorial subspace of $R^{k(k-1)/2}$ if it is a closed submanifold [see Warner (1983), page 22]. Most, if not all, subsets G of $SO(k)$ that are of statistical interest satisfy this requirement.

When κ is large, the likelihood ratio statistic for H_0, under model (1), is approximately equal to (5). One can also test H_0 with a Wald statistic which avoids the computation of $\hat{P}_G$. Examples will be given in Section 4.

The analogy with linear regression allows for local power calculations. Let A_κ be an $O(\kappa^{-1/2})$, $k \times k$ skew-symmetric matrix with component vector a_κ. Under H_A: $P = P_1 \exp A_\kappa$ for some P_1 in G, it can be shown that the limiting distribution of (5) is $F_{k(k-1)/2-g,(k-1)(n-k/2)}(\delta^2)$, where δ^2, the noncentrality parameter, is equal to

$$\delta^2 = \kappa a_\kappa' Q_\perp \left(Q_\perp' (X'X)^{-1} Q_\perp\right)^{-1} Q_\perp' a_\kappa,$$

where $Q_\perp$ is a basis of the orthogonal complement of $V(G, P_1)$. The key to the proof of this result is the fact that, under H_A, (2) holds for the distribution of $P_1' v_i$ with ε_i distributed as $N_{k-1}(u_{i(\cdot)}' U_i a_\kappa, I/\kappa)$.

3.1. *Residual analysis.* As in linear regression, residuals r_i can be defined as the estimates of the errors ε_i. In geometrical terms, r_i is the R^{k-1}-vector of the coordinates of the projection of $\hat{P}'v_i$ in the vector space orthogonal to u_i. For a given basis $u_{i(\cdot)}$ of that vector space, $r_i = u'_{i(\cdot)}\hat{P}'v_i$. The joint distribution of $\{r_i; i = 1, \ldots, n\}$ is the same as that of the residuals of the linear model defined in the theorem.

Two diagnostics from linear regression can be useful in spherical regression. To test whether datum i is an outlier, a test statistic based on r_i is given by [Cook and Weisberg (1982), page 30]

$$t_i^2 = \frac{(n - 1 - k/2) r_i \Sigma_i^{-1} r_i}{2n(1 - \bar{R}(SO(k))) - r_i \Sigma_i^{-1} r_i},$$

where Σ_i is the covariance matrix of $\kappa^{1/2} r_i$. Its null distribution is the $F_{k-1,(k-1)(n-1-k/2)}$ distribuiton. A measure of the leverage of datum i is the multiple Cook's D statistic [Cook and Weisberg (1982), page 136]

$$D_i = \left(\frac{1}{k} - \frac{1}{2n}\right) \frac{r_i'(\Sigma_i^{-2} - \Sigma_i^{-1}) r_i}{(1 - \bar{R}(SO(k)))}.$$

4. Further results for S_2-regression. Most applications of spherical regression will be to S_2-data. This section studies the special case $k = 3$. It takes advantage of the mathematical properties of R^3 to get some further results.

First the parametrization used so far is modified as follows: If A is a 3×3 skew-symmetric matrix, redefine its component vector as $a = (A_{32}, -A_{31}, A_{21})$ and U_i by

$$U_i = \begin{pmatrix} 0 & -u_{i3} & u_{i2} \\ u_{i3} & 0 & -u_{i1} \\ -u_{i2} & u_{i1} & 0 \end{pmatrix}.$$

Now $Au_i = -U_i a$ and U_i is itself skew-symmetric, its component vector is u_i. Note that for any v in R^3, $U_i v$ is the exterior or cross product of u_i by v.

Writing $u_{i(\cdot)} = (u_{i(1)}, u_{i(2)})$ and changing the sign of $u_{i(2)}$ if necessary, we can assume that $u_i, u_{i(1)}, u_{i(2)}$ form a right-hand rule oriented orthonormal basis. Then the skew-symmetric matrix U_i can be written in terms of $u_{i(\cdot)}$, defined in (2), as

$$U_i = u_{i(\cdot)} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} u'_{i(\cdot)}. \tag{6}$$

With this parametrization, the two lines of X, the design matrix of (4), for observation i, can be written as $u_{i*} = (u_{i(2)}, -u_{i(1)})'$ and $X'X = n(I - S)$, where $S = \Sigma u_i u_i'/n$.

If A is a 3×3 skew-symmetric matrix, then

$$\exp A = \cos\|a\| I + \frac{\sin\|a\|}{\|a\|} A + \frac{1 - \cos\|a\|}{\|a\|^2} aa'$$

is a rotation of angle $\|a\|$ about a. Thus the natural parametrization of $SO(3)$ is (w, θ), where w, the axis of the rotation, is an S_2-vector and θ is its angle. Let $(\hat{w}, \hat{\theta})$ denote $\hat{P}$ the maximum likelihood estimate of P.

PROPOSITION. *Let P be the rotation (w, θ) where $\theta \neq 0$. If A is skew-symmetric with $O(\kappa^{-1/2})$ components, then $P\exp A$ is, up to $o(\kappa^{-1/2})$, the rotation (w_a, θ_a), where $\theta_a = \theta + w'a$,*

$$w_a = w + \frac{1}{2} w_{(\cdot)} \begin{pmatrix} \dfrac{1 + \cos\theta}{\sin\theta} & -1 \\ 1 & \dfrac{1 + \cos\theta}{\sin\theta} \end{pmatrix} w'_{(\cdot)} a$$

and $w_{(\cdot)}$ is a 3×2 matrix containing a basis of the vector space orthogonal to w such that w, $w_{(1)}$ and $w_{(2)}$ form a right-hand rule oriented orthonormal basis and $w_{(1)}$ and $w_{(2)}$ denote the first and the second column of $w_{(\cdot)}$.

Conversely if θ_a and w_a satisfies $\theta_a - a$ and $w_a - w$ are $O(\kappa^{-1/2})$, then, up to $o(\kappa^{-1/2})$, (w_a, θ_a) is the rotation $P\exp A$, where the component vector of A is equal to

$$a = (\theta_a - \theta) w + w_{(\cdot)} \begin{pmatrix} \sin\theta & 1 - \cos\theta \\ \cos\theta - 1 & \sin\theta \end{pmatrix} w'_{(\cdot)} w_a.$$

PROOF. One can write $w_a = w + dw$ where dw is an $O(\kappa^{-1/2})$ vector orthogonal to w. It satisfies

$$(7) \qquad \begin{aligned} (P + PA)(w + dw) &= w + dw \\ \Leftrightarrow (I - P)\, dw &= PAw + o(\kappa^{-1/2}). \end{aligned}$$

One can express $I - P$ as

$$w_{(\cdot)} \begin{pmatrix} 1 - \cos\theta & \sin\theta \\ -\sin\theta & 1 - \cos\theta \end{pmatrix} w'_{(\cdot)}.$$

The only solution of (7) which is orthogonal to w is

$$dw = \frac{1}{2} w_{(\cdot)} \begin{pmatrix} \dfrac{\sin\theta}{1 - \cos\theta} & -1 \\ 1 & \dfrac{\sin\theta}{1 - \cos\theta} \end{pmatrix} w'_{(\cdot)} a.$$

Note that $\operatorname{tr}(P) = 1 + 2\cos\theta$. Thus, using (6) one can write

$$\begin{aligned} \operatorname{tr}(P\exp A) &= \operatorname{tr} P + \operatorname{tr}(\sin\theta WA) + o(\kappa^{-1/2}) \\ &= \operatorname{tr} P - 2w'a\sin\theta + o(\kappa^{-1/2}), \\ &= 1 + 2\cos(\theta + w'a) + o(\kappa^{-1/2}). \end{aligned}$$

The converse is proved in a similar way. □

4.1. *Inference on the axis.* To test that the axis of P is w is testing

$$H_0: \quad P \in G_w + \{\cos\theta I + \sin\theta W + (1-\cos\theta)ww' : \theta \in (0, 2\theta)\}.$$

Chang (1986) derived a formula for the maximum $\overline{R}(w)$, on G_w, of $\Sigma v_i' P u_i / n$.

$$F_{\text{obs}} = \left(n - \frac{3}{2}\right) \frac{\overline{R}(SO(3)) - \overline{R}(w)}{1 - \overline{R}(SO(3))}$$

has an $F_{2,2n-3}$ asymptotic null distribution.

By the proposition

$$\hat{w} = w + \frac{w_{(\cdot)}}{2} \begin{pmatrix} \dfrac{\cos\theta_0 + 1}{\sin\theta_0} & -1 \\ 1 & \dfrac{1+\cos\theta_0}{\sin\theta_0} \end{pmatrix} w_{(\cdot)}' \hat{a},$$

where θ_0 is the true angle. Hence, under H_0,

$$\kappa^{1/2} \begin{pmatrix} \sin\theta_0 & 1-\cos\theta_0 \\ \cos\theta_0 - 1 & \sin\theta_0 \end{pmatrix} w_{(\cdot)}' \hat{w} \sim N_2\left(0, \frac{w_{(\cdot)}(I-S)^{-1} w_{(\cdot)}}{n}\right)$$

and the Wald statistic,

$$\text{(8)} \qquad \begin{aligned} &\frac{2n-3}{4} \hat{w}' w_{(\cdot)} \begin{pmatrix} \sin\hat{\theta} & \cos\hat{\theta} - 1 \\ 1 - \cos\hat{\theta} & \sin\hat{\theta} \end{pmatrix} \left(w_{(\cdot)}'(I-S)^{-1} w_{(\cdot)} \right)^{-1} \\ &\times \begin{pmatrix} \sin\hat{\theta} & 1-\cos\hat{\theta} \\ \cos\hat{\theta} - 1 & \sin\hat{\theta} \end{pmatrix} w_{(\cdot)}' \hat{w} / \left(1 - \overline{R}(SO(3))\right), \end{aligned}$$

follows a $F_{2,2n-3}$ distribution. It can be used to construct a confidence region for w.

4.2. *Inference on the angle.* Let $G_\theta = \{\cos\theta I + \sin\theta W + (1-\cos\theta)ww',\ w \in S_2\}$ be the two-dimensional closed set of rotations of angle θ. Here

$$\begin{aligned} \overline{R}(\theta) &= \max_w \cos\theta \frac{\Sigma v_i' u_i}{n} - \sin\theta \frac{\Sigma v_i' U_i w}{n} + (1-\cos\theta) w' \frac{\Sigma v_i u_i'}{n} w \\ &= \cos\theta \frac{\Sigma v_i' u_i}{n} + \max_w (1-\cos\theta) w' \frac{\Sigma v_i u_i' + u_i v_i'}{2n} w - \sin\theta \frac{\Sigma v_i' U_i w}{n}. \end{aligned}$$

This maximization problem is studied by Forsythe and Golub (1965) and Bingham and Mardia (1978). To test H_0: $P \in G_\theta$, the test statistic is

$$F_{\text{obs}} = (2n-3) \frac{\overline{R}(SO(3)) - \overline{R}(\theta)}{1 - \overline{R}(SO(3))}.$$

It has an $F_{1,2n-3}$ distribution. The proposition provides a simple way to derive the Wald statistic. Under H_0, $\hat{\theta} = \theta + w_0' a$, where w_0 is the true axis. Thus, under H_0,

$$t_{\text{obs}} = \hat{\theta} - \theta \Bigg/ \sqrt{\frac{2(1 - \bar{R}(SO(3)))}{2n - 3} \hat{w}'(I - S)^{-1}\hat{w}}$$

follows a t_{2n-3} distribution. A confidence interval for θ can be constructed using this statistic. To construct confidence region for both θ and w, one proceeds as in Section 4 of Chang (1986) with χ^2 critical values replaced by F's.

In the calculation of the statistics presented in this section, one can take $w_{(\cdot)}$ as the last two columns of $(e_1 + w)(e_1 + w)'/(1 + e_1'w) - I$, where $e_1 = (1, 0, 0)'$; $u_{i(\cdot)}$ can be defined in an analogous way. When $k = 3$, the calculations of t_i^2 and D_i defined in Section 3, are simplified by noting that

$$\Sigma_i = u_{i*}\left(I - \frac{(I - S)^{-1}}{n}\right)u_{i*}'.$$

5. Numerical example. Chang (1986) undertook to fit model (1) to Cochran (1981) data which are presented in Chang's paper. He calculated $\hat{P}$ as a rotation of 2.38° about the axis 25.31°N and 24.29°E and $\hat{\kappa} = 1.78 \times 10^6$.

To study the fit of the model, the *spatial residual plot* of Figure 1 is useful. Each datum is represented by an arrow starting at the predicted value ($\hat{P}u_i$) and pointing toward the observed value (v_i). The length of the arrow is proportional to $\|r_i\|$. Up to a multiplicative constant the arrow joins the estimated mean direction of a Fisher random vector to its realization, if the model fits well. The longitude–latitude coordinate system was used to match Figure 8 of Cochran (1981); also, since all the latitudes were less than 20°, it did not bring in significant distortions (which is not true for large latitudes). Each observation is labelled by its case number.

The graph actually contains nine arrows and two points corresponding to observations 4 and 10 that have very small residuals. Observations 3 and 11 have the largest residuals and, as shown in Table 1, are the most influential.

The t_i^2 have to be compared with critical values from an F distribution with 2 and 17 degrees of freedom. Observations 3 and 11 are significant at the 0.025 and 0.05 levels, respectively. Even if observation 3 has the largest residual, observation 11 has the largest Cook's statistic. This is so because the latter is located on the boundary on the design space, and has therefore more leverage than the former. It is tempting to declare observations 3 and 11 outliers; however the removal of these two points does not clean up the data completely, since observation 10 then has a t_i^2 of 2.13 and a D_i equal to 0.73.

If the model fits well, $\{r_i\}$ estimates a sample from $N_2(0, I/\kappa)$. Therefore $\{\|r_i\|^2\}$ is approximately an exponential sample. This was checked with a Q–Q plot. It did not display any strong departure from the exponential model. The conclusion of this study is a cautious acceptance of the proposed model.

Chang (1986) tested

$$H_0^{(w)}\colon w \text{ is } 26.5°N, 21.5°E.$$

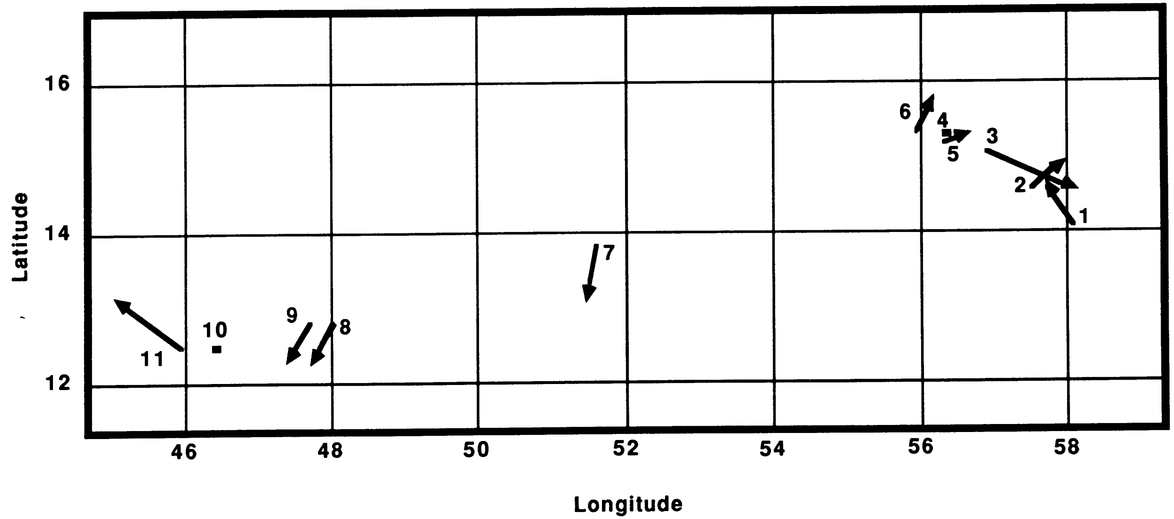

FIG. 1. *Spatial residual plot. Each observation is represented by an arrow starting at the predicted value, pointing towards the observed value and with length equal to ten times the distance between the observed and the predicted values.*

TABLE 1
Influence statistics for the example

Case	1	2	3	4	5	6	7	8	9	10	11
t_i^2	0.57	0.38	4.90	0.00	0.14	0.48	1.05	0.90	1.11	0.04	3.87
D_i	0.09	0.04	0.31	0.00	0.01	0.04	0.08	0.10	0.12	0.01	0.63

The likelihood ratio statistic for H_0 is $F_{\text{obs}} = 1.25$ while the Wald statistic is $F_{\text{obs}} = 1.46$. Both have 2 and 19 degrees of freedom; they are not significant at the 0.05 level. This agrees with the conclusion that Chang reached using a large sample test. The large discrepancy between the two F statistics is caused by the small angle of $\hat{P}$. Indeed, for small $\hat{\theta}$ (8) can be factorized as $\hat{\theta}^2$ times a term independent of $\hat{\theta}$. Thus a Wald statistic of 1.25 would have obtained with $\hat{\theta} = 2.20°$ which is in the range of possible values for the unknown rotation angle.

Chang (1986) considered also $H_0^{(\theta)}$: $\theta = 2.04°$. The Wald statistic is $t_{\text{obs}} = 2.24$ with 19 degrees of freedom which is significant at the 0.05 level. Thus the angle of the rotation is significantly larger than 2.04°, which is similar to the conclusion reached by Chang (1986).

To check the stability of the result it is interesting to redo the analysis without observations 3 and 11. The angle of rotation becomes 2.62° while the axis is 24.27°N and 27.48°E. The likelihood ratio statistic for $H_0^{(w)}$ becomes $F_{\text{obs}} = 8.49$ with 2 and 15 degrees of freedom which is highly significant. Thus, using this data set, it is not possible to reach an unambiguous conclusion concerning $H_0^{(w)}$.

Acknowledgment. I am grateful to Nathalie Plante for her computational assistance.

REFERENCES

ABRAMOWITZ, M. and STEGUN, I. (1972). *Handbook of Mathematical Functions*. Dover, New York.

BINGHAM, C. and MARDIA, K. V. (1978). A small circle distribution on the sphere. *Biometrika* **65** 379–398.

CHANG, T. (1986). Spherical regression. *Ann. Statist.* **14** 907–924.

CHANG, T. (1987). On the statistical properties of estimated rotations. *J. Geophys. Res.* **92B** 6319–6329.

COCHRAN, J. (1981). The Gulf of Aden: Structure and evolution of a young ocean basin and continental margin. *J. Geophys. Res.* **86** 263–287.

COOK, R. D. and WEISBERG, S. (1982). *Residuals and Influence in Regression*. Chapman and Hall, London.

FORSYTHE, G. E. and GOLUB, G. H. (1965). On the stationary values of a second-degree polynomial on the unit sphere. *SIAM J.* **13** 1050–1068.

JUPP, P. E. and MARDIA, K. V. (1980). A general correlation coefficient for directional data and related regression problems. *Biometrika* **67** 163–173.

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

RIVEST, L.-P. (1986). Modified Kent's statistics for testing goodness of fit for the Fisher distribution in small concentrated samples. *Statist. Probab. Lett.* **4** 1–4.

RIVEST, L.-P. (1988). A distribution for dependent unit vectors. *Comm. Statist. A—Theory Methods* **17** 461–483.

STEPHENS, M. A. (1979). Vector correlation. *Biometrika* **66** 41–48.

WARNER, F. W. (1983). *Foundations of Differentiable Manifolds and Lie Groups*. Springer, New York.

WATSON, G. S. (1983). *Statistics on Spheres*. Wiley, New York.

WATSON, G. S. (1984). The theory of concentrated Langevin distributions. *J. Multivariate Anal.* **14** 74–82.

DÉPARTEMENT DE MATHÉMATIQUES
ET STATISTIQUE
UNIVERSITÉ LAVAL
CITÉ UNIVERSITAIRE QUÉBEC
CANADA G1K 7P4