The Annals of Statistics
1986, Vol. 14, No. 3, 907–924

SPHERICAL REGRESSION[1]

By Ted Chang

Simon Fraser University and University of Kansas

Suppose $u_1, \dots, u_n$ are fixed points on the sphere, $v_1, \dots, v_n$ are random points such that the distribution of v_i depends only upon $v_i^t A u_i$ for some unknown rotation A. This paper provides asymptotic tests and confidence regions for A and for the axis of rotation of A. Results are given in arbitrary dimension.

Let S^p be the unit radius sphere in p-dimensional Euclidean space and let $\mathrm{SO}(p)$ be the $p \times p$ orthogonal matrices (that is matrices A such that $AA^t = I$) of determinant 1. We consider in this paper "spherical regression" problems on the following model: $u_1, \dots, u_n$ are fixed points in S^p (written as column vectors), $v_1, \dots, v_n$ are random points in S^p such that $v_1, \dots, v_n$ are independent and such that the density of v_i, with respect to uniform measure on S^p, is of the form $g(v_i^t A u_i)$ for some unknown A in $\mathrm{SO}(p)$. We want to develop statistical procedures for estimating and testing the unknown parameter A.

The case of the circle ($p = 2$) is essentially well known because A is counterclockwise rotation by an unknown angle θ. If θ_i is the angle from u_i to v_i, then $\theta_1, \dots, \theta_n$ are independent and identically distributed with a density of the form $g(\theta_i - \theta)$.

The case of the sphere ($p = 3$) is of considerable practical importance. The following two problems are abstractions of problems proposed to the author by workers in other fields; the first from geology and the second from petroleum exploration. It was the simultaneous and fortuitous presentation of these problems that lead to the present study.

Problem 1. A rigid body, confined to the surface of the earth, has moved in an unknown manner. For certain points (u_i) on S^3, estimates of past position at a fixed point in time (v_i) are available. What was its previous position?

In this problem the body's past position relative to its present position is determined uniquely by an element A of $\mathrm{SO}(3)$. The v_i are estimates of Au_i and the problem is to determine A.

Problem 2. The directions (v_i) of certain signals have been measured in an unknown coordinate system. The directions (u_i) of the same signals in a known coordinate system can be calculated. What is the unknown coordinate system?

Received August 1984; revised December 1985.

[1] Listings of FORTRAN programs implementing in three dimensions the procedures outlined in this paper are available from the author. They are available in single precision using the IMSL library or in double precision using the NAG library.

AMS 1980 *subject classification*. Primary 62J99.

Key words and phrases. Estimated rotations on spheres.

In this problem if the rows of A are the components of the coordinate axis of the unknown coordinate system with respect to the known one, the v_i are measurements of Au_i with error and the problem is to determine A.

Variations on Problem 1 are of especial interest in the study of plate tectonics. Geophysicists have been fitting rotations to the motion of tectonic plates for 20 years. Only some of the data they use can be modeled in the form of problem 1. The approach has been to define an error sum squares SSE(A) which depends upon the choice of a candidate rotation A, to iteratively minimize SSE(A), thus arriving at an estimate $\hat{A}$ of the unknown rotation A, and to assume an approximating distribution for

$$\frac{\mathrm{SSE}(A) - \mathrm{SSE}(\hat{A})}{\mathrm{SSE}(\hat{A})}.$$

Examples of this procedure can be found in Le Pichon (1968, 1973), Chase (1972) and Engebretson, Cox and Gordon (1984). No attempt is made to prove the correctness of the assumed asymptotic distribution. The author has found that for the choice of error sum squares studied in this paper, the asymptotic distribution is not $2n\chi^2(3)$ as one might assume. Nevertheless, if the error distribution is concentrated, as those in plate tectonics seem to be, the true asymptotic distribution is in fact extremely close to $2n\chi^2(3)$. The author hopes that this paper can be a start towards a more rigorous and mathematical understanding of these problems.

If $c_0 = E(v_i^t A u_i) > 0$, it is reasonable to estimate A by the matrix $\hat{A}$ which minimizes

$$\sum_i |v_i - Au_i|^2 = 2n - 2\sum_i \left(v_i^t A u_i\right).$$

Letting U_n and V_n be the $p \times n$ matrices whose columns are u_i and v_i, respectively, the solution for $\hat{A}$ was found by MacKenzie (1957) and Stephens (1979). It is readily computable from a modified singular value decomposition

$$U_n V_n^t = O_1 \Lambda O_2^t$$

where $O_1, O_2 \in \mathrm{SO}(p)$ and Λ is diagonal with entries $\lambda_1, \dots, \lambda_p$ satisfying $\lambda_1 \geq \lambda_2 \geq \cdots \geq |\lambda_p|$. If the rank of U_n is p, the determinant of $U_n V_n^t$ is nonzero with probability 1 and in that case, $\hat{A}$ is uniquely given by $O_2 O_1^t$. We will call $\hat{A}$ the "least squares estimate of A." In this paper we will find the asymptotic distribution of $\hat{A}$ under the assumption that $(1/n)U_n U_n^t$ converges as $n \to \infty$ to a positive definite symmetric matrix Σ (Theorem 1). We propose that asymptotic confidence regions for A be based upon Theorem 1.

Letting $\mathrm{O}(p)$ denote the $p \times p$ orthogonal matrices, we define, following Stephens (1979), for a subset S of $\mathrm{O}(p)$ the vector correlation $r(S)$ by

$$r(S) = \sup_{A \in S} \frac{1}{n} \sum_i v_i^t A u_i.$$

Stephens studied the distribution of $r(\mathrm{SO}(p))$ and $r(\mathrm{O}(p))$ when the u_i and v_i are independently and uniformly distributed on the sphere S^p. Using Theorem 1, we will find for closed subgroups $G' \subseteq G$ of $O(p)$, the asymptotic distribution of $r(G')$ and of $r(G) - r(G')$ when $A \in G'$ (Theorem 2).

With $G = SO(p)$ and $G' = \{I\}$, we propose to use Theorem 2 to test whether A is some specified A_0 or not. The resulting test is based upon the test statistic $r(SO(p)) - r(A_0)$ (where, abusing the notation, we write $r(A_0)$ for $r(\{A_0\})$).

When $p = 3$, the matrix A will represent rotation of an angle θ about an axis ξ. In both of the problems cited above it is of interest to test if ξ is some predetermined ξ_0. If we let G' be the group (isomorphic to SO(2)) of rotations around ξ_0, we are testing the hypotheses $A \in G'$. We propose to base an asymptotic test on Theorem 2 and on the test statistic $r(SO(3)) - r(G')$.

Gould (1969) considered another inequivalent type of spherical regression model. For the sphere S^3, the Gould model is that the v_ι are independently Fisher distributed with model vector $u_\iota = (\cos\phi_i, \sin\phi_i \cos\theta_i, \sin\phi_i \sin\theta_i)^t$ with ϕ_ι and θ_ι known linear functions of the unknown parameters. Gould also considers a similar model on the circle S^2.

In Section 1, we state and prove Theorems 1 and 2 in arbitrary dimensions. In Section 2, we describe asymptotic hypotheses tests with special attention to three dimensions. Section 3 contains a numerical example and Section 4 discusses display of confidence regions for A.

If the underlying distribution is Fisher, $d(\kappa)\exp(\kappa v^t A u)$, the procedures in this paper are just maximum likelihood estimation and likelihood ratio testing. For other distributions, the author believes that use of least squares estimates $\hat{A}$ are justified by the relative ease of computing $\hat{A}$.

In this paper, the u_ι play the role of the predictor variables in linear regression: They are assumed fixed and v_i is assumed to have a rotationally symmetric distribution centered at Au_ι. If they are instead random but with a distribution independent of A, the results are still valid for inference conditional on the u_i. When the distribution of u and the conditional distribution of v are both Fisher, aspects of this problem were studied by Rivest (1984).

1. Statements and proofs of the main theorems. We will think of Euclidean p^2 space R^{p^2} as the collection of $p \times p$ matrices with the usual inner product $(A, B) = \operatorname{tr}(AB^t)$. Let $O(p) \subseteq R^{p^2}$ be those matrices A such that $AA^t = I$. Then $O(p)$ is closed (in the usual metric space sense), has dimension $\frac{1}{2}(p(p-1))$ and consists of two connected components. One of these is $SO(p)$, the matrices in $O(p)$ of determinant $+1$, and the other consists of the elements of $SO(p)$ followed by any reflection.

The tangent space at the identity I of $O(p)$ (and hence of $SO(p)$) is the collection of skew-symmetric $p \times p$ matrices; that is the matrices H such that $H + H^t = 0$. We denote the collection of such H by $L(SO(p))$.

The exponential map ϕ: $L(SO(p)) \to SO(p)$ is defined by

$$\phi(H) = I + H + \frac{H^2}{2!} + \frac{H^3}{3!} + \cdots .$$

If G is a closed subgroup of $O(p)$ with the metric space topology, we let $L(G)$ denote the tangent space at I of G. $L(G)$ is by definition a vector subspace of $L(\mathrm{SO}(p))$. It can be shown (see Theorem 15 and its proof, Spivak (1979), page 530) that $L(G)$ is the set of H in $L(\mathrm{SO}(p))$ such that $\exp tH$ is in G for all real t. The dimension of G is the dimension of $L(G)$. If $G = \mathrm{SO}(p)$ or $\mathrm{O}(p)$, $\dim G = \frac{1}{2}(p(p-1))$.

Let $\hat{A}_n(G)$ be the "least squares estimate of A in G," namely the element of G which maximizes

$$\frac{1}{n}\sum_{i=1}^{n} v_i^t A u_i = \operatorname{tr}\left(A\frac{U_n V_n^t}{n}\right)$$

as A varies over G. Thus $\hat{A}_n(\mathrm{O}(p))$ is the statistic defined by MacKenzie (1957) and $\hat{A}_n(\mathrm{SO}(p))$ is Stephens's (1979) modification of MacKenzie's statistic.

If $v \in S^p$ has density of the form $g(v^t u)$ for some $u \in S^p$, we define constants c_0, c_1, and c_2 by

$$E(v) = c_0 u,$$

$$E\left[(v - c_0 u)(v - c_0 u)^t\right] = c_1 u u^t + c_2 I.$$

That $E(v)$ is a multiple of u is obvious by symmetry. That $E[(v - c_0 u)(v - c_0 u)^t]$ can be written in the form $c_1 u u^t + c_2 I$ is obvious when u is the "north pole" and follows for general u by rotating S^p.

Theorem 1. *Let G be a closed subgroup of $\mathrm{O}(p)$. Suppose each v_i has a density $g(v_i^t A_0 u_i)$ where A_0 is in G. Suppose furthermore $c_0 > 0$ and that $1/n\Sigma_i u_i u_i^t$ converges to a positive definite symmetric matrix Σ. Then*

(a) *$\hat{A}_n(G)$ is consistent for A_0.*

(b) *Write $A_0^t \hat{A}_n(G) = \phi(H_n)$ for $H_n \in L(G)$. Then H_n is asymptotically multivariate normal with mean 0 and density (with respect to a Lebesgue measure on $L(G)$) proportional to*

$$\exp\left[\frac{c_0^2}{2c_2} n \operatorname{tr}\left(H_n^2 \Sigma\right)\right]$$

Thus $-nc_0^2/c_2 \operatorname{tr}(H_n^2\Sigma)$ is asymptotically $\chi^2(\dim G)$.

Most of the proof of Theorem 1 is a mimic of the proofs in the asymptotic theory of the mle with the log likelihood function replaced by $\operatorname{tr}(A(U_n V_n^t)/n)$ and with nonidentically distributed variates. We will therefore omit many details.

Let $U_n = [u_1 \cdots u_n]$, $V_n = [v_1 \cdots v_n]$, and $X_n = 1/nU_n V_n^t$.

Lemma 1. $X_n \to c_0 \Sigma A_0^t$ *(strong convergence).*

Proof. Let $W_i = u_i v_i^t - c_0 u_i u_i^t A_0^t$. W_i is a $p \times p$ matrix with expected value 0. By Kolmogorov's criterion for the strong law of large numbers (see Billingsley (1979), page 250), $1/n\Sigma_{i=1}^n W_i$ converges to 0 with probability 1. The lemma follows. □

LEMMA 2. *$\hat{A}_n(G)$ is strongly consistent for A_0.*

PROOF. $\hat{A}_n(G)$ maximizes $\mathrm{tr}(AX_n)$ as A varies over G. By Lemma 1, $X_n \to c_0 \Sigma A_0^t$ with probability 1. Since Σ is positive definite and $c_0 > 0$,

$$\mathrm{tr}(Ac_0\Sigma A_0^t) = c_0 \mathrm{tr}(A_0^t A \Sigma)$$

is maximized uniquely when $A_0^t A = I$ or $A = A_0$. The lemma follows from the following observation:

Suppose f is a continuous function on $\mathscr{X} \times \mathscr{Y}$ with $\mathscr{X}$ compact and suppose furthermore that for a specific $y_0 \in \mathscr{Y}$, $f(x, y_0)$ has a unique maximum at $x = x_0$. Suppose $y_n \to y_0$ and each x_n is a choice of a maximum for $f(x, y_n)$. Then $x_n \to x_0$. □

Since $\hat{A}_n(G) \to A_0$, for large enough n we can write $A_0^t \hat{A}_n(G) = \phi(H_n)$ where $H_n \in L(G)$ is chosen to have smallest magnitude. By replacing v_i with $A_0^t v_i$, we can assume $A_0 = I$. Pick a specific $B \in L(G)$ and define a real valued function on $L(G)$

$$g_n^B(H) = \frac{d}{dt}\bigg|_{t=0} \mathrm{tr}(\phi(H + tB)X_n).$$

We have $g_n^B(H_n) = 0$. We expand g_n^B in a Taylor series around 0:

$$g_n^B(0) = \frac{d}{dt}\bigg|_{t=0} \mathrm{tr}(\phi(tB)X_n) = \mathrm{tr}(BX_n).$$

If $H \in L(G)$,

$$\begin{aligned}(g_n^B)'(0)H &= \frac{d}{ds}\bigg|_{s=0} \frac{d}{dt}\bigg|_{t=0} \mathrm{tr}(\phi(sH + tB)X_n)\\ &= \mathrm{tr}\left(\frac{HB + BH}{2} X_n\right).\end{aligned}$$

Thus

$$g_n^B(H) = \mathrm{tr}(BX_n) + \mathrm{tr}\left(\frac{HB + BH}{2} X_n\right) + R.$$

Defining for a matrix H, the ordinary Euclidean metric

$$\|H\|^2 = \mathrm{tr}(HH^t),$$

it can be easily shown that $|R| \le \|H\|^2 \|B\| e^{\|H\|}$. Since $g_n^B(H_n) = 0$, and since

$$\mathrm{tr}(H\Sigma B) = \mathrm{tr}(B^t \Sigma^t H^t) = \mathrm{tr}(B \Sigma H),$$

the following lemma is obtained:

LEMMA 3. *For $B \in L(G)$,*

$$-\mathrm{tr}(B\sqrt{n}X_n) = c_0 \mathrm{tr}(\sqrt{n}H_n \Sigma B) + R_n,$$

where

$$|R_n| \le \sqrt{n}\|H_n\|\|B\|(\|H_n\| e^{\|H_n\|} + \|X_n - c_0\Sigma\|).$$

Let $L(G)^*$ be the dual space to $L(G)$. Define $\alpha_n \in L(G)^*$ by $\alpha_n(B) = -\operatorname{tr}(B\sqrt{n}X_n)$. Each α_n is a random variable with values in $L(G)^*$.

Lemma 4. *α_n has a limiting multivariate normal distribution with covariance quadratic form $c_2Q(B_1, B_2) = -c_2\operatorname{tr}(B_1\Sigma B_2)$ for $B_1, B_2 \in L(G)$.*

Proof. To say that the random vector α in $L(G)^*$ has covariance quadratic form c_2Q means that if B_1, B_2 are nonrandom vectors in $L(G)$ the covariance of the real valued random variables $\alpha(B_1)$ and $\alpha(B_2)$ is $c_2Q(B_1, B_2)$.

The characteristic function of α_n is

$$F_n(B) = E\left[\exp\left(\sqrt{-1}\,\alpha_n(B)\right)\right], \qquad B \in L(G).$$

Substituting $X_n = 1/n\Sigma_i u_i v_i^t$ and noting that

$$0 = \operatorname{tr}\left[Bu_iu_i^t\right] = u_i^tBu_i,$$

since B is antisymmetric and $u_iu_i^t$ is symmetric,

$$F_n(B) = \prod_{i=1}^{n} E\left[\exp\frac{-\sqrt{-1}}{\sqrt{n}}(v_i - c_0u_i)^tBu_i\right].$$

Since $E(v_i - c_0u_i) = 0$ and

$$\begin{aligned} E\left[(Bu_i)^t(v_i - c_0u_i)(v_i - c_0u_i)^tBu_i\right] &= u_i^tB^t\left(c_1u_iu_i^t + c_2I\right)Bu_i \\ &= -c_1\left(u_i^tBu_i\right)^2 - c_2\left(u_i^tBBu_i\right) \\ &= -c_2\operatorname{tr}\left(Bu_iu_i^tB\right), \end{aligned}$$

we have

$$F_n(B) = \prod_{i=1}^{n}\left[1 + \frac{c_2}{n}\operatorname{tr}\left(Bu_iu_i^tB\right) + o\left(\frac{\|B\|^2}{n}\right)\right].$$

The remainder $o(\|B\|^2/n)$ is bounded uniformly in i by $\min[\|B\|^3/6n^{3/2}, \|B\|^2/n]$ (see Billingsley (1979), equation (26.5)), and hence as $n \to \infty$

$$F_n(B) \to \exp\left(c_2\operatorname{tr}(B\Sigma B)\right) \qquad \square$$

Now let ρ: $L(G) \to L(G)^*$ be

$$\rho(H)B = Q(H, B) = -\operatorname{tr}(H\Sigma B).$$

Since Σ is positive definite $\rho(H)H > 0$ and ρ is nonsingular. Lemmas 3 and 4 imply that

$$\alpha_n = \rho\left(-c_0\sqrt{n}H_n\right) + o_p(1),$$

and hence $\rho(-c_0/\sqrt{c_2})\sqrt{n}H_n)$ has a limiting multivariate normal distribution with covariance quadratic form Q.

Now Q defines an identification of $L(G)$ with $L(G)^*$ and this identification is ρ. It follows that H_n has an asymptotic multivariate normal distribution with a density proportional to

$$\exp\left[-\frac{1}{2}Q\left(\frac{-c_0}{\sqrt{c_2}}\sqrt{n}\,H_n, \frac{-c_0}{\sqrt{c_2}}\sqrt{n}\,H_n\right)\right]$$

$$= \exp\left[\frac{c_0^2}{2c_2} n \operatorname{tr}(H_n \Sigma H_n)\right].$$

[Let X^* be a random vector with values in a dual space $\mathscr{V}^*$ and let the quadratic form Q on $\mathscr{V}$ be the covariance of X^*. Let X be a random vector in $\mathscr{V}$ defined by $Q(X, B) = X^*(B)$ for all $B \in \mathscr{V}$. If we pick a basis $e_1, \dots, e_k$ of $\mathscr{V}$ and write $X = \sum_\iota x_i e_i$, let V be the matrix $\operatorname{cov}(x_\iota, x_j)$. Then $Q(X, X) = [x_1 \cdots x_k]V^{-1}[x_1 \cdots x_k]^t$.] This proves Theorem 1.

Theorem 2. (a) *If $A_0 \in G$, then $r(G)$ has a limiting normal distribution with mean c_0 and variance $(c_1 + c_2)/n$.*

(b) *If $A_0 \in H \subseteq G$, then $2nc_0/c_2(r(G) - r(H))$ has a limiting $\chi^2(\dim G - \dim H)$ distribution.*

(c) *If $A_0 \in K \subseteq H \subseteq G$, then*

$$\frac{\dim G - \dim H}{\dim H - \dim K}\,\frac{r(H) - r(K)}{r(G) - r(H)}$$

is asymptotically $F(\dim H - \dim K, \dim G - \dim H)$.

Proof of (a). With the notation of Theorem 1,

$$r(G) = \operatorname{tr}\left[A_0 \phi(H_n)\frac{U_n V_n^t}{n}\right].$$

As before, we can assume $A_0 = I$. Then

$$\sqrt{n}\,(r(G) - c_0) = \sqrt{n}\left(\operatorname{tr}\left[(I + H_n)\frac{U_n V_n^t}{n}\right] - c_0\right) + o_p(1)$$

$$= \sqrt{n}\operatorname{tr}\left[\frac{U_n V_n^t}{n} - c_0\frac{U_n U_n^t}{n}\right] + \sqrt{n}\,c_0 \operatorname{tr}[H_n \Sigma] + o_p(1)$$

$$= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}(v_i - c_0 u_i)^t u_i + o_p(1).$$

The summands are identically distributed with mean 0 and variance

$$E\left[u^t(v - c_0 u)(v - c_0 u)^t u\right] = u^t(c_1 u u^t + c_2 I)u$$

$$= c_1 + c_2. \square$$

Proof of (b). Let $Q(B_1, B_2) = -\operatorname{Tr}(B_1 \Sigma B_2)$. We define $H_n(G)$ by $\hat{A}_n(G) = A_0 \phi(H_n(G))$ and similarly $H_n(H)$ and $H_n(K)$, and again we set $A_0 = I$. If $B \in L(H) \subseteq L(G)$ then using Lemma 3 with both $H_n(G)$ and $H_n(H)$, we see that $Q(\sqrt{n}(H_n(G) - H_n(H)), B)$ is $o_p(\|B\|)$. Thus if β_n is the projection under Q of $\sqrt{n}\,H_n(G)$ to the perpendicular complement of $L(H)$ in $L(G)$,

$$\sqrt{n}\,(H_n(G) - H_n(H)) = \beta_n + o_p(1).$$

Thus, using Lemma 3,

$$\begin{aligned} 2n(r(G) - r(H)) &= 2n(r(G) - r(I)) - 2n(r(H) - r(I)) \\ &= c_0 Q\big(\sqrt{n}\,H_n(G), \sqrt{n}\,H_n(G)\big) \\ &\quad - c_0 Q\big(\sqrt{n}\,H_n(H), \sqrt{n}\,H_n(H)\big) + o_p(1) \\ &= c_0 Q(\beta_n, \beta_n) + o_p(1). \end{aligned}$$

Thus $2nc_0/c_2(r(G) - r(H))$ is asymptotically $\chi^2(\dim G - \dim H)$. □

Proof of (c). From part (b) we see that up to terms $o_p(1)$, $\sqrt{n}\,H_n(H)$ is the projection under Q of $\sqrt{n}\,H_n(G)$ to $L(H)$ and that $\sqrt{n}\,(H_n(H) - H_n(K))$ is its projection to the orthogonal complement of $L(K)$ in $L(H)$. Part (c) follows. □

Theorem 3. *Let $A_0^t \hat{A}_n(G) = \phi(H_n(G))$ for $H_n(G) \in L(G)$. Then $\sqrt{n}\,(r(G) - c_0)$ and $\sqrt{n}\,H_n(G)$ are asymptotically independent.*

Proof. Using the proof of Theorem 2(a),

$$\sqrt{n}\,(r(G) - c_0) = \sqrt{n}\,(\mathrm{tr}(X_n) - c_0) + o_p(1).$$

From Lemma 3, for $B \in L(G)$,

$$\alpha_n(B) = -\mathrm{tr}\big(B\sqrt{n}\,X_n\big) = c_0 \mathrm{tr}\big(\sqrt{n}\,H_n(G)\Sigma B\big) + o_p(1).$$

Let

$$F_n(t, B) = E\big[\exp\sqrt{-1}\,\big(\alpha_n(B) + t\sqrt{n}\,(\mathrm{tr}(X_n) - c_0)\big)\big]$$

be the joint characteristic function of α_n and $\sqrt{n}\,(\mathrm{tr}(X_n) - c_0)$. Using a proof similar to Lemma 4,

$$F_n(t, B) \to \exp\big(c_2 \mathrm{tr}(B\Sigma B) - t^2(c_1 + c_2)\big)$$

and the theorem follows. □

If the density g is unknown, to use Theorems 1 and 2, we need consistent estimates of c_0 and c_2. Using Theorem 2(a), we can estimate c_0 consistently by

$$\hat{c}_0 = r(G) \quad \text{if } A_0 \in G.$$

The following proposition provides a consistent estimator $\hat{c}_2$ of c_2. Using Problem 29.4 and Theorem 29.2 of Billingsley (1979) it follows that Theorems 1, 2(b) and 2(c) are still valid if $\hat{c}_0$ is replaced by $\hat{c}_0$ and c_2 is replaced by $\hat{c}_2$.

PROPOSITION 1. *If $A_0 \in G$ and*

$$\hat{c}_2 = \frac{1}{p-1}\left[1 - \frac{1}{n}\sum_i \left(v_i^t \hat{A}_n(G) u_i\right)^2\right],$$

then $\hat{c}_2 \to c_2$ in probability.

LEMMA 5. $1 = c_0^2 + c_1 + pc_2$.

PROOF.

$$\begin{aligned} Evv^t &= E\left[(v - c_0 Au)(v - c_0 Au)^t\right] + c_0^2 Auu^t A^t \\ &= \left(c_0^2 + c_1\right) Auu^t A^t + c_2 I. \end{aligned}$$

Taking the trace of both sides, we get the lemma. □

PROOF OF THE PROPOSITION. Setting as usual $A_0 = I$, we get $\hat{A} = \hat{A}_n(G) = I + o_p(1)$. Therefore

$$\frac{1}{n}\sum_i \left(v_i^t \hat{A} u_i\right)^2 = \frac{1}{n}\sum_i \left(v_i^t u_i\right)^2 + o_p(1).$$

The right-hand side converges in distribution to $c_0^2 + c_1 + c_2$. Using the lemma, the proposition follows. □

We now consider models of the form $d(\kappa)g(\kappa v^t A u)$ where the concentration parameter κ is unknown. If $c_0(\kappa) = E(v^t A u)$ is monotonic we can estimate κ from the sample statistic $r(G)$ by solving

$$c_0(\hat{\kappa}) = r(G).$$

Theorem 2(a) can then be used for inferences on κ.

REMARK. One might wonder about the necessity of the requirement that G be closed. Since the closure $\overline{G}$ of a subgroup G is still a subgroup, and since $r(G) = r(\overline{G})$, Theorem 2 remains valid if $\dim G$ is always replaced by $\dim \overline{G}$.

Theorem 1, however, cannot be generalized to nonclosed subgroups. If G is not closed, the dimension of G will always be strictly less than the dimension of $\overline{G}$. Since $\hat{A}_n(G)$, if it exists, will also be $\hat{A}_n(\overline{G})$, we expect $\hat{A}_n(G)$ to exist with probability 0.

An example of this pathology is the infamous "real line embedded in the torus" which occurs in SO(4). If r is a fixed irrational number and

$$G = \left\{ \begin{bmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & \cos r\theta & -\sin r\theta \\ 0 & 0 & \sin r\theta & \cos r\theta \end{bmatrix} \middle| \; \theta \text{ a real number} \right\},$$

then

$$\bar{G} = \left\{ \begin{bmatrix} \cos\theta_1 & -\sin\theta_1 & 0 & 0 \\ \sin\theta_1 & \cos\theta_1 & 0 & 0 \\ 0 & 0 & \cos\theta_2 & -\sin\theta_2 \\ 0 & 0 & \sin\theta_2 & \cos\theta_2 \end{bmatrix} \right\}.$$

The author believes that the generic pathological nature of the nonclosed subgroups indicates that they have no practical statistical interest.

REMARK. If $c_0 = 0$, $\hat{A}_n$ might very well be inconsistent.

For example, if each v_i is uniformly distributed on S^p then for each $A \in \mathrm{O}(p)$, $v_1, \ldots, v_n$ and $Av_1, \ldots, Av_n$ are equally likely. If $\hat{A}_n(G)$ is the least squares fit for $v_1, \ldots, v_n$, then $A\hat{A}_n(G)$ will be the least squares fit for $Av_1, \ldots, Av_n$. It follows that the distribution of $\hat{A}_n(G)$ will be the unique left invariant Haar measure for each n and hence $\hat{A}_n(G)$ is inconsistent.

For this example, Stephens (1979) has studied the limiting distributions of $r(\mathrm{SO}(p))$ and $r(\mathrm{O}(p))$ and we observe that Theorem 2 is also false.

If $c_0 < 0$ and if $(-I) \in G$, then $\hat{A}_n(G) \to -IA_0$, and Theorems 1 and 2 could be modified to handle this case. However, if $c_0 < 0$, it is intrinsically unreasonable to study the $\hat{A}$ which maximizes $\sum v_i^t A u_i$. A more reasonable approach would be to maximize $\sum v_i^t A(-u_i)$ and if this were done, Theorems 1, 2, and 3 could still be applied with minor changes.

2. Hypothesis tests. Suppose H is a closed subgroup of $\mathrm{O}(p)$. If c_0 and c_2 are known, we can use Theorem 2(b) to asymptotically test if the true orthogonal matrix A is in H.

EXAMPLE. Suppose we wish to test H_0: $A = A_0$. Then using Theorem 2(b) with $H = \{I\}$ and each u_i replaced by $A_0 u_i$, we have $2nc_0/c_2(r(\mathrm{O}(p)) - r(A_0))$ is asymptotically $\chi^2(p(p-1)/2)$ if H_0 is true.

EXAMPLE. Suppose $p = 3$ and we wish to test if A is a rotation about a specified unit vector ξ_0. Let H be the subgroup of all rotations around ξ_0. If ξ_0 is the correct axis, we have $2nc_0/c_2(r(\mathrm{O}(p)) - r(H))$ is asymptotically $\chi^2(2)$.

To calculate $r(H)$ we note that if $A(\theta)$ is right-hand rule rotation of θ radians around ξ_0, then

$$A(\theta) = I + \sin\theta L + (1 - \cos\theta)L^2,$$

where

$$L = \begin{pmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{pmatrix}$$

and $\xi_0 = (t_1, t_2, t_3)^t$. Thus

$$r(H) = \max_{0 \le \theta \le 2\pi} \frac{1}{n} \sum v_\iota^t A(\theta) u_\iota$$
$$= a_0 + a_2 + \left(a_1^2 + a_2^2\right)^{1/2},$$

where $a_r = 1/n\Sigma_i v_i^t L^r u_\iota$. The fitted angle $\hat{\theta}$ is specified by $\sin\hat{\theta} = a_1/(a_1^2 + a_2^2)^{1/2}$ and $\cos\hat{\theta} = -a_2/(a_1^2 + a_2^2)^{1/2}$.

The critical region for both tests takes the form that the test statistic is too big, indicating, as in linear regression, that the improvement in the fit is better than can be attributed to overfitting of the model.

All these tests are still asymptotically true if c_0 and c_2 are replaced by $\hat{c}_0$ and $\hat{c}_2$ where

$$\hat{c}_0 = r(\mathrm{O}(p)),$$
$$\hat{c}_2 = \frac{1}{p-1}\left[1 - \frac{1}{n}\sum_\iota \left(v_\iota^t \hat{A}(\mathrm{O}(p))\right)u_\iota\right]^2.$$

For the convenience of the reader we note the following values of c_0, c_1, and c_2 for the Fisher distribution $d(\kappa)e^{\kappa v^t u}$ where $d(\kappa) = \kappa/\sinh\kappa$ and $p = 3$:

$$c_0 = \coth\kappa - \frac{1}{\kappa},$$
$$c_1 = \frac{2}{\kappa^2} - \frac{\coth\kappa}{\kappa} - \operatorname{csch}^2\kappa,$$
$$c_2 = \frac{1}{\kappa}\left(\coth\kappa - \frac{1}{\kappa}\right).$$

3. A numerical example. Geophysicists believe that the Gulf of Aden formed as Arabia began to separate from Africa about 20 million years ago. Table 1 gives the latitudes and longitudes of fracture zone intersections with 3′S and 3′N magnetic anomalies, digitizing from Figure 8 of Cochran (1981). Geophysical theory indicates the Arabian and Somalian plates have been moving so that the points u_i and v_i (for each i) were once coincident. The problem is to fit the relative motion of the Arabian plate from the Somalian plate, thinking of the Somalian plate as fixed in its present location.

The choice of the Somalian plate as fixed and the u_ι points as those on the South intersections is arbitrary. If, however, the roles of the two plates were reversed, the analysis would change as one would expect: For example, the fitted rotation $\hat{A}$ would be replaced by $\hat{A}^t$.

When the points u_ι and v_ι are converted to Euclidean coordinates, the matrix $U_n V_n^t / n$ is

$$\frac{U_n V_n^t}{n} = \begin{bmatrix} 0.3509 & 0.4547 & 0.1425 \\ 0.4454 & 0.5942 & 0.1867 \\ 0.1302 & 0.1738 & 0.0547 \end{bmatrix}$$

TABLE 1

u_i (Somalia)		v_i (Arabia)	
Latitude	Longitude	Latitude	Longitude
13.05	57.56	14.28	58.12
13.34	57.07	14.54	57.67
13.89	56.50	15.00	57.16
14.19	55.97	15.33	56.51
14.10	55.92	15.25	56.48
14.21	55.38	15.37	55.93
12.68	50.95	13.59	51.51
11.97	47.56	12.78	48.11
12.06	47.35	12.86	47.89
11.63	45.80	12.44	46.39
11.73	45.36	12.58	45.87
$n = 11$ points			

and $\hat{A}(\mathrm{SO}(3))$

$$\hat{A}(\mathrm{SO}(3)) = \begin{bmatrix} 0.9997 & -0.0175 & 0.0157 \\ 0.0180 & 0.9993 & -0.0341 \\ -0.0151 & 0.0343 & 0.9993 \end{bmatrix}.$$

$\hat{A}$ represents a rotation of 2.38° around an axis through 25.31°N latitude and 24.29°E longitude.

Since the true rotation A is known to be in SO(3) and since SO(3) is a connected component of O(3), Lemma 2 implies that

$$\mathrm{pr}[\hat{A}_n(\mathrm{SO}(3)) = \hat{A}_n(\mathrm{O}(3))] \to 1 \quad \text{as } n \to \infty.$$

Since $\det(U_n V_n^t/n)$ is positive, $\hat{A}(\mathrm{SO}(3)) = \hat{A}(\mathrm{O}(3))$ in this case.

For this data set

$$\hat{c}_0 = r(\hat{A}(\mathrm{SO}(3))) = 1 - 0.5812 \times 10^{-6},$$

$$\hat{c}_2 = 0.5812 \times 10^{-6},$$

and

$$\hat{c}_1 + \hat{c}_2 = 0.3867 \times 10^{-12}.$$

McKenzie et al. (1970) found, by fitting the 500 fathom contours on each side of the Gulf, that the pole of rotation of the Arabian plate relative to the Somalian plate is located at 26.5°N and 21.5°E. If H is the subgroup of rotations around that axis, $\hat{A}(H)$ is a rotation of 2.20° and $r(H) = 1 - 0.6579 \times 10^{-6}$. Then

$$\frac{2n\hat{c}_0}{\hat{c}_2}[r(\mathrm{SO}(3)) - r(H)] = 2.902. \tag{1}$$

Comparing this to a $\chi^2(2)$ distribution, we see no contradiction between the data of Table 1 and the McKenzie axis.

McKenzie also fits a rotation angle of 7.6°. If, following Cochran (1981), we take 20 million years as the age of the rift and, following La Brecque et al. (1977), 5.37 million years before present as the time that the points u_ι and v_ι were coincident, this prorates to an angle of 2.04° over the 5.37 million years. If A_0 is a rotation of 2.04° around the axis 26.5°N and 21.5°E,

$$r(A_0) = 1 - 0.1691 \times 10^{-5}$$

and hence

$$\frac{2n\hat{c}_0}{\hat{c}_2}\left[r(\mathrm{SO}(3)) - r(A_0)\right] = 42.02, \tag{2}$$

which needs to be compared to a $\chi^2(3)$ distribution to test $A = A_0$.

This spectacularly high level of χ^2 should not cause any excitement. If the angle of rotation in A_0 had been between 2.14 and 2.25° the null hypotheses would have been accepted at an approximate 0.05 significance level. The imprecisions in the dating used above make 2.04° in fact indistinguishable from angles in that range.

With $r(G) = 1 - 0.5812 \times 10^{-6}$ and $\hat{c}_1 + \hat{c}_2 = 0.3867 \times 10^{-12}$ we get, using Theorem 2(a), that with 95% confidence, $1 - c_0 = (0.5812 \pm 0.3675) \times 10^{-6}$. Assuming a Fisher error distribution, $c_0(\kappa) = 1 - 1/\kappa + o(1/\kappa)$ and hence $1.1 \times 10^6 < \kappa < 4.7 \times 10^6$ with an estimate $\hat{\kappa} = (0.5812 \times 10^{-6})^{-1} = 1.72 \times 10^6$.

A computer simulation was run using IMSL generator GGUBS with 10 000 runs, the given 11 points u_ι, and a true rotation of 2.04° around 26.5°N, 21.5°E. Three runs were made with a Fisher error distribution and $\kappa = 1.0 \times 10^6$, 1.72×10^6, and 5.0×10^6. In each run, the test statistic (1) exceeded 2.902 approximately 30% of the time. This compares with a $\chi^2(2)$ distribution p-value of 23%. The test statistic (2) exceeded 42.02 0.01% of the time.

For this problem, the author has found that the programming of the formulas in this paper in single precision led to no significant figures in the computed values of χ^2. If instead the mathematically equivalent formulas

$$\begin{aligned}
1 - r(G) &= \frac{1}{2n}\sum_\iota |v_\iota - \hat{A}u_\iota|^2 = 1 - \hat{c}_0,\\
r(G) - r(H) &= (1 - r(H)) - (1 - r(G)),\\
\hat{c}_2 &= \frac{1}{2n}\sum_\iota |v_\iota - \hat{A}u_\iota|^2 - \frac{1}{8n}\sum_i |v_i - \hat{A}u_\iota|^4,\\
\hat{c}_1 + \hat{c}_2 &= \frac{1}{4n}\sum_\iota |v_i - \hat{A}u_\iota|^4 - (1 - \hat{c}_0)^2
\end{aligned} \tag{3}$$

are used, the author has found that single and double precision programming yield results agreeing in at least four significant figures. One can conclude that, at least for this data set, the formulas (3) above work satisfactorily in single precision.

In this example, the u_i suffer from a spherical regression analogue of multicollinearity: They are very close to lying on a small arc of a great circle. This can be detected using the matrix $\hat{\Sigma} = 1/n\Sigma_\iota u_i u_\iota^t$. It is easily proven that the rank of $\hat{\Sigma}$ is the dimension of the smallest vector subspace of R^p containing all the u_ι. In the instant case the eigenvalues of $\hat{\Sigma}$ are 0.99332, 0.00663, and 0.00005. From Theorem 1, we see that multicollinearity causes large variances in $\hat{A}$ and hence small changes in the data will cause unexpectedly large changes in $\hat{A}$. Furthermore, when the estimated $\hat{A}$ is close to the identity, rotations which are in fact quite close in SO(3) might have seemingly disparate axes of rotation.

The analysis assumes that the u_i are known without error or at the very least that the conditional distribution of v_i given u_i is symmetric around Au_i. A preferable model would be: u_i has a distribution of the form $g(u_i^t\xi_i)$; v_i has a distribution of the form $g(v_i^tA\xi_i)$; u_i and v_i are independent with $\xi_1, \dots, \xi_n$, A unknown. In this situation, the author has been able to prove an analogue of Theorem 2 with a much more complicated asymptotic distribution. Alternatively for $G = \mathrm{SO}(3)$, and $H = \{I\}$ or SO(2), the author has found more complicated test statistics with asymptotic $\chi^2(3)$ or $\chi^2(2)$ distributions, respectively. When the latter procedures are applied to the data of this example, with its very concentrated error distributions, the values of the χ^2 statistics agree with those reported above to four significant figures. The author will report on these results at a later date.

In this analysis the points (u_i, v_i) are believed to have been simultaneously coincident approximately 5.27 million years ago. In fact, geologists have dated a sequence of anomalies going back in excess of 100 million years. The general practice has been to choose a time interval (which may be shorter than the span of the data), assume a constant axis and speed of rotation over the chosen interval, and to fit them to all intersections from the chosen interval by the process described in the introduction.

If we define $\mathrm{SSE}(A) = \Sigma|v_i - \hat{A}u_i| = 2n - 2nr(A)$ we see that the distribution of

$$\frac{\mathrm{SSE}(A) - \mathrm{SSE}(\hat{A}(\mathrm{SO}(3)))}{\mathrm{SSE}(\hat{A}(\mathrm{SO}(3)))|2n}$$

is not asymptotically $\chi^2(3)$ as one might assume. It is rather asymptotically $c_2/(c_0(1 - c_0))\chi^2(3)$. Nevertheless, for extremely concentrated error distributions, such as those of the above example, $c_2/(c_0(1 - c_0))$ is, to very close approximation, equal to 1.

4. Confidence regions for the orthogonal matrix A. Although Theorem 2(b) can be used to produce confidence regions for the unknown orthogonal matrix A, the author believes that Theorem 1 is better suited for this purpose.

If G is a closed subgroup of $\mathrm{O}(p)$ and it is known a priori that $A \in G$, let $\chi^2_{1-\alpha}$ be the appropriate critical point of the χ^2 distribution with $\dim G$ degrees of freedom. Let

$$\mathscr{C} = \left\{\phi(H)|H \in L(G) \text{ and } -\mathrm{tr}(H^2\Sigma) < \frac{c_2}{nc_0^2}\chi^2_{1-\alpha}\right\}.$$

Since $\phi(-H) = \phi(H)^t$, it is easy to see that the required confidence region is $\hat{A}_n(G)\mathscr{C}$.

Alternatively, we might wish to express our confidence region in one form of $\hat{A}_n(G)$ followed by a small perturbation. In this case, since $\phi(AHA^t) = A\phi(H)A^t$, the confidence region is $\mathscr{C}'\hat{A}_n(G)$ where

$$\mathscr{C}' = \left\{\phi(H) | H \in L(G) \text{ and } -\operatorname{tr}(H^2\Sigma') < \frac{c_2\chi^2_{1-\alpha}}{nc_0^2}\right\}$$

and $\Sigma' = \hat{A}_n(G)\Sigma\hat{A}_n(G)^t$.

The following alternative definition of the exponential map ϕ might be helpful. If H is skew-symmetric, an orthogonal matrix O can be found so that

$$O^tHO = \text{block diagonal} \begin{bmatrix} 0 & -\theta_1 \\ \theta_1 & 0 \end{bmatrix} \cdots \begin{bmatrix} 0 & -\theta_k \\ \theta_k & 0 \end{bmatrix}.$$

Here $k = [p/2]$ and an additional diagonal entry of 0 needs to be added if p is odd. Then $\phi(H) = \phi(OO^tHOO^t) = O\phi(O^tHO)O^t = OAO^t$ where

$$A = \text{block diagonal} \begin{bmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{bmatrix} \cdots \begin{bmatrix} \cos\theta_k & -\sin\theta_k \\ \sin\theta_k & \cos\theta_k \end{bmatrix}$$

and an additional diagonal entry of 1 needs to be added if p is odd.

Asymptotic confidence regions of minimum volume will be achieved if the u_ι can be chosen so that $\Sigma = (1/p)I$. This can be done by using a uniform random point generator on S^p or, if $n = pr$ for some r, by replicating r times any orthogonal basis of Euclidean p space. In that case given two matrices A and B of G define a distance function on G by:

$$d(A, B) = \theta_1^2 + \cdots + \theta_k^2 \quad \text{if } \det A^tB = 1 \text{ and } A^tB \text{ has eigenvalues } e^{\iota\theta_1}, e^{-\iota\theta_1}, \dots, e^{\iota\theta_k}, e^{-\iota\theta_k}$$
$$\text{(together with } +1 \text{ if } p \text{ is odd) where } k = [p/2] \text{ and } -\pi < \theta_\iota \le \pi,$$

$$d(A, B) = \infty \quad \text{if } \det A^tB = -1.$$

It follows from Theorem 1 and the above alternate description of ϕ that $(2nc_0^2/pc_2)d(A_0, \hat{A}_n(G))$ is asymptotically $\chi^2(\dim G)$.

When $p = 3$ and $G = \mathrm{SO}(3)$, the general element of $L(\mathrm{SO}(3))$ is of the form

$$H = \begin{bmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{bmatrix} \tag{4}$$

and it can be shown that $\phi(H)$ is right-hand rule rotation of $\sqrt{t_1^2 + t_2^2 + t_3^2}$ radians around the axis $(t_1^2 + t_2^2 + t_3^2)^{-1/2}[t_1 \quad t_2 \quad t_3]^t$.

If we identify $L(\mathrm{SO}(3))$ with R^3 by identifying an H in the form (4) above with $[t_1, t_2, t_3]^t$, we get the following equivalent description ψ: $R^3 \to \mathrm{SO}(3)$ of the exponential map: If $x \in R^3$, let $\theta = |x|$ and $\xi = x/|x|$; then $\psi(x)$ is right-hand rule rotation of θ radians around the axis ξ. In terms of ψ, the regions $\mathscr{C}$ and $\mathscr{C}'$

above become

$$\mathscr{C} = \left\{ \psi(x) | x^t (I - \Sigma) x < \frac{c_2}{n c_0^2} \chi^2_{1-\alpha} \right\},$$

$$\mathscr{C}' = \left\{ \psi(x) | x^t (I - \hat{A} \Sigma \hat{A}^t) x < \frac{c_2}{n c_0^2} \chi^2_{1-\alpha} \right\},$$

where $\hat{A} = \hat{A}_n(\mathrm{SO}(3))$. As before, our confidence regions become $\hat{A}\mathscr{C}$ or $\mathscr{C}'\hat{A}$.

EXAMPLE. We continue with the example of the previous section. We have $\hat{c}_2 = 0.5812 \times 10^{-6}$, $\hat{c}_0 = 1.0000$, and estimating Σ by

$$\hat{\Sigma} = \frac{1}{n} \sum_i u_i u_i^t = \begin{bmatrix} 0.3568 & 0.4532 & 0.1325 \\ 0.4532 & 0.5924 & 0.1733 \\ 0.1325 & 0.1733 & 0.0508 \end{bmatrix}$$

we have

$$\hat{A} \Sigma \hat{A}^t = \begin{bmatrix} 0.3451 & 0.4470 & 0.1401 \\ 0.4470 & 0.5961 & 0.1872 \\ 0.1401 & 0.1872 & 0.0589 \end{bmatrix}.$$

Using $\chi^2 = 7.81$, the 95% critical point of a $\chi^2(3)$ distribution, we have that $\mathscr{C}'$ consists of all $\psi([x_1 \quad x_2 \quad x_3]^t)$ satisfying

$$0.345 x_1^2 + 0.596 x_2^2 + 0.0589 x_3^2 + 0.894 x_1 x_2 + 0.280 x_1 x_3$$
$$+ 0.374 x_2 x_3 < 0.413 \times 10^{-6}$$

and the 95% confidence region for A is any rotation of the form $\hat{A}$ (rotation of 2.38° around 25.31° N, 24.29° E) followed by any rotation in $\mathscr{C}'$. For example we could follow $\hat{A}$ by a rotation around 0° N latitude, 90° E longitude ($[0, 1, 0]^t$) of at most $((0.413 \times 10^{-6})/0.596)^{1/2} = 0.832 \times 10^{-3} = 0.048°$.

The eigenvectors of $I - \hat{A} \Sigma \hat{A}^t$ are

$$[0.5857, 0.7733, 0.2428]^t, \qquad [0.8099, -0.5465, -0.2131]^t,$$

and

$$[0.0322, -0.3214, 0.9464]^t$$

with corresponding eigenvalues 0.00668, 0.99337, and 0.99995. Thus the largest rotation in $\mathscr{C}'$ is $((0.413 \times 10^{-6})/0.668 \times 10^{-2})^{1/2} = 0.451°$ around an axis 14.05° N, 52.86° E ($= [0.5857, 0.7733, 0.2428]^t$).

Every rotation in $\mathscr{C}'\hat{A}$ satisfies the inequalities

$$23.60° \mathrm{N} \le \text{axis latitude} \le 27.40° \mathrm{N},$$
$$17.52° \mathrm{E} \le \text{axis longitude} \le 29.05° \mathrm{E},$$
$$2.00° \le \text{rotation angle} \le 2.78°.$$

Hence, these three inequalities are asymptotic at least 95% simultaneous confidence intervals.

The confidence region $\mathscr{C}'\hat{A}$ was reexpressed in the form: axis $\in \mathscr{A}$, $f(\text{axis}) <$ rotation angle $< g(\text{axis})$ where $\mathscr{A}$ is a subset of S^3, and f and g are real valued

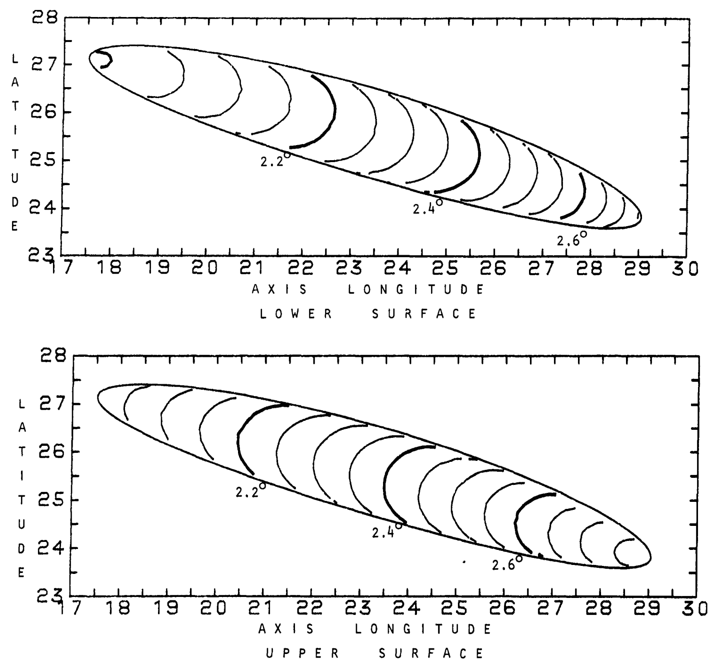

FIG. 1.

functions on $\mathscr{A}$. Points on the graph of f and g (the lower and upper surfaces, respectively) were calculated and contour maps of the upper and lower surfaces drawn using the SURFACE II package developed by the Kansas Geological Survey. These maps appear in Figure 1. Thus, for example, all rotations around the axis 24°E, 26°N with a rotation angle between 2.3 and 2.4° lie in the confidence region.

Acknowledgments. The author wishes to thank Professor Colin Ferguson of the Department of Geology, Birbeck College, University of London, and Mr. Steven Jones of the Kansas Geological Survey for their extensive help with the geophysical aspects of this paper. He would also like to thank the referee for his enthusiasm and several helpful remarks.

The author wishes to thank the State of Kansas General Research Fund for funding the preliminary work on this project. He especially wishes to thank the Statistics faculty at Simon Fraser University for their generosity in sharing their National Science and Engineering Research Council (Canada) grants during the term of the majority of the work.

REFERENCES

BILLINGSLEY, P. (1979). *Probability and Measure.* Wiley, New York.

CHASE, C. G. (1972). The N plate problem of plate tectonics. *Geophys. J. Roy. Astron. Soc.* **29** 117–122.

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

ENGEBRETSON, D., COX, A. and GORDON, R. (1984). Relative notions between oceanic plates of the Pacific basin. *J. Geophys. Res.* **89B** 10291–10310.

GOULD, A. L. (1969). A regression technique for angular variates. *Biometrics* **25** 683–700.

LABRECQUE, J. L., KENT, D. V. and CANDE, S. C. (1977). Revised magnetic polarity time scale for Late Cretaceous and Cenozoic time. *Geology* **5** 330–335.

LE PICHON, X. (1968). Sea-floor spreading and continental drift. *J. Geophys. Res.* **73** 3661–3697.

LE PICHON, X., FRANCHETEAN, J. and BONNIN, J. (1973). *Plate Tectonics.* Elsevier, New York.

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

MCKENZIE, D. P., DARRE, D. and MOLNAR, P. (1970). Plate tectonics of the Red Sea and East Africa. *Nature* **226** 243–248.

RIVEST, L. P. (1984). The bivariate Fisher–von Mises distribution. Unpublished manuscript.

SPIVAK, M. (1979). *A Comprehensive Introduction to Differential Geometry* **1**, 2nd ed. Publish or Perish, Boston.

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

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF KANSAS
LAWRENCE, KANSAS 66045