JOURNAL OF MULTIVARIATE ANALYSIS **63**, 35–46 (1997)
ARTICLE NO. MV971690

# Geodesic Estimation in Elliptical Distributions

Maia Berkane* and Kevin Oden*

*Harvard University*

and

Peter M. Bentler†

*University of California, Los Angeles*



Elliptical distributions include the normal, the Student $t$, the contaminated normal, and many other distributions. Let $X$ be a $p$-variate vector elliptically distributed; if the density exists, it is written as

$$f_X(x, \mu, V) = |V|^{-1/2} h((x-\mu)' V^{-1}(x-\mu))$$

where $\mu(p \times 1)$ is the location vector and $V(p \times p)$ is a positive definite scatter matrix. $h$ is a function independent of $\mu$ and $V$. Hence $f$ depends on $x$ only through the quadratic form $(x-\mu)' V^{-1}(x-\mu)$. Let $U=(X-\mu)' V^{-1}(X-\mu)$, then it has been shown that the c.d.f of $u$ is

$$f_U(u) = \begin{cases} \dfrac{\pi^{p/2}}{\Gamma(p/2)} u^{p/2-1} h(u) & \text{if} \quad u > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Atkinson and Mitchell (1981) and Mitchell (1988) derived geodesic estimation and testing for some classes of distributions on the basis of Rao's metric (Rao, 1945). In particular, Mitchell (1988) investigated the manifold of univariate elliptical distributions. Geodesic estimation and testing in the multivariate case have not been very accessible, mainly because one has to

Received February 29, 1996.

* Supported by NSF Summer Grant DMS 9504834. Maia Berkane is now at Donaldson, Lufkin, and Jenrette Securities Corporation and Kevin Oden is at Goldman Sachs.
† Supported by Grant DA01070 from the National Institute on Drug Abuse.



0047-259X/97 $25.00

face very complicated systems of differential equations characterizing the geodesic curves. In general, no closed form solutions for these systems are available. In this paper, we investigate the class of multivariate elliptical distributions. We find that a change of coordinates reduces the system of geodesic equations to a very simple form admitting an appealing solution that is nothing but a straight line. The distance between two elliptical distributions with equal location and different scatter matrices is then calculated. This geodesic distance which is based on the information metric, comes as an addition to the list of geodesic distances given by Rao (1987). We then derive a geodesic discrepancy function for use in covariance structure analysis. The estimator of the structure parameter is shown to have desirable properties. A test statistic is then built upon this discrepancy function and shown to be asymptotically distributed as $\chi^2$.

## 2. METRIC IN THE CLASS OF MULTIVARIATE NORMAL DISTRIBUTIONS

James (1973) derived a geodesic distance between two multivariate normal distributions with equal mean vectors. He first used a result of Maass showing that the metric differential form $(ds)^2 = \mathrm{tr}(Y^{-1}dYY^{-1}dY)$ on the space of $p \times p$ positive definite matrices $Y$, which is a cone in $\mathbf{R}^{p(p+1)/2}$, is invariant under congruent transformation

$$Y \to LYL'$$

where $L$ is a $p \times p$ nonsingular matrix. This metric happened to coincide with the information metric in the manifold of $p$-variate normal distributions with known mean vector. He then showed that the desired geodesic distance between two multivariate normal distributions with covariances $\Sigma_1$ and $\Sigma_2$ is

$$\left( \sum_{i=1}^{p} (\ln y_i)^2 \right)^{1/2},$$

where the $y_i$ are the roots of the determinental equation

$$\det(\Sigma_1 - y\Sigma_2) = 0.$$

Suppose we have $n$ replications of the random vector $X$ and assume $X$ to be normally distributed with mean vector 0 and covariance matrix $\Sigma$. Suppose that $\Sigma$ is structured by a $q \times 1$ parameter vector $\theta$ (for example,

$X$ results from a factor analysis model or a latent variable model). Let $S$ be the sample covariance matrix; then James suggests using the statistic

$$d^2 = \frac{n}{2} \sum_{i=1}^{p} (\ln y_i)^2,$$

where $y_i$ are the latent roots of the determinental equation

$$\det(S - y\Sigma) = 0,$$

as a test of hypothesis on the structure of $\Sigma$. Swain (1975) compared several distance functions for estimating structural parameters in covariance structures analysis; among them was $d^2$ used above. He showed that the geodesic estimation leads to the same asymptotic sampling properties as the maximum likelihood estimation (m.l.e.) procedure. We will now consider the class of elliptical distributions admitting a c.d.f. and derive geometric properties of this class.

## 3. METRIC IN THE CLASS OF ELLIPTICAL DISTRIBUTIONS, FIXED FUNCTIONAL FORM

We will fix the class of the elliptical distribution by fixing the functional form of the c.d.f. For instance, assume that the density distribution of the random vector $X$ exists and is of the form

$$f_X(x) = |V|^{1/2}\, h((x-\mu)'\, V^{-1}(x-\mu))$$

where $h$ is fixed, $h(u) = \exp -u/2$ if the class is that of the normal distributions, $h(u) = c(1(u/v)^{-(p+v)/2})$ if the class is that of the $t$ distributions with $v$ degrees of freedom, and so on. First suppose that $\mu$ is known and we could set it to zero. Let

$$M = \{ f_X(x) = |V|^{1/2}\, h(x' V^{-1} x),\ v = \text{vecs}(V) \in \mathbf{R}^{\mathbf{p(p+1)/2}} \},$$

where vecs($V$) is the vector of nonredundent elements of $V$. $M$ is a differentiable manifold of dimension $p(p+1)/2$. We will derive the information metric (Rao's metric) on $M$. Other metrics can be found; for instance, Burbea and Rao (1982) introduce the "entropy differential metric" based on the Hessian of the $\phi$-entropy functional, where $\phi$ is some real-valued $C^2$-function defined on an interval in $[0, \infty)$. However, the information metric seems to be the natural metric on statistical manifolds (Amari 1984).

Let $w = (\partial \ln h)/(\partial u)$ and $y = (\partial^2 \ln h)/(\partial u^2)$. The logarithm of the likelihood function is then

$$\mathscr{L}(x, \mu, V) = -\tfrac{1}{2} \ln |V| + \ln h((x-\mu)' V^{-1}(x-\mu)),$$

and its first two differentials are

$$\begin{aligned} d\mathscr{L} &= -\tfrac{1}{2} \operatorname{tr}(V^{-1}\, dV) - w \operatorname{tr}(V^{-1}\, dVV^{-1}xx') \\ d^2\mathscr{L} &= \tfrac{1}{2} \operatorname{tr}(V^{-1}\, dVV^{-1}\, dV) + y(\operatorname{tr}(V^{-1}\, dVV^{-1}xx'))^2 \\ &\quad + 2w \operatorname{tr}(V^{-1}\, dVV^{-1}\, dVV^{-1}xx'). \end{aligned}$$

Define $Z$ and $T$ by

$$Z = V^{-1/2}(X-\mu), \qquad T = \frac{Z}{\|Z\|},$$

then $\|Z\|^2 = U$ and $U$ and $T$ are independent. $T$ is uniformly distributed on the unit sphere, with moments easily derived (see for example Mitchell, 1989). The expected value $E(-d^2\mathscr{L})$ is then given by

$$\begin{aligned} E(-d^2\mathscr{L}) &= -\tfrac{1}{2} \operatorname{tr}(V^{-1}\, dVV^{-1}dV) - E(y(\operatorname{tr}(V^{-1}\, dVV^{-1}xx'))^2) \\ &\quad - 2 \operatorname{tr}(V^{-1}\, dVV^{-1}\, dVV^{-1}E(wxx')). \end{aligned}$$

But the expectation in the third term on the right-hand side of the above equation is

$$E(wxx') = V^{1/2}E(tt')\, E(uw)\, V^{1/2} = V^{1/2}(1/p)\, I_p(-p/2)\, V^{1/2}$$

because $E(t_i^2) = 1/p$ and $E(uw) = -p/2$. Hence

$$-2 \operatorname{tr}(V^{-1}\, dVV^{-1}\, dVV^{-1}E(wxx')) = \operatorname{tr}(V^{-1}\, dVV^{-1}\, dV)$$

and the second term is, after some manipulations using properties of trace and vec operations,

$$\begin{aligned} E(y \operatorname{tr}(V^{-1}\, dVV^{-1}xx'))^2) &= \operatorname{tr}((V^{-1/2} \otimes V^{-1/2})(dV \otimes dV)(V^{-1/2} \otimes V^{-1/2}) \\ &\quad \times E(\operatorname{vec}(tt')(\operatorname{vec}(tt'))')\, E(u^2 y)), \end{aligned}$$

but

$$E(\operatorname{vec}(tt')(\operatorname{vec}(tt'))') = \frac{1}{p(p+2)} (2(I_p \otimes I_p) + \operatorname{vec}(I_p)(\operatorname{vec}(I_p))').$$

Let $b_h = -E(u^2w^2)/p(p+2)$; then

$$E(-d^2\mathscr{L}) = 2b_h \mathrm{tr}(V^{-1}\,dVV^{-1}\,dV) + (b_h - \tfrac{1}{4})\,\mathrm{tr}((V^{-1}\otimes V^{-1})(dV\otimes dV)).$$

Let $H$ be an orthogonal $p\times p$ matrix such that $H'VH = \Lambda$ where $\Lambda$ is a diagonal matrix of eigenvalues of $V$; then James showed (1973) that

$$\mathrm{tr}(V^{-1}\,dVV^{-1}\,dV) = \sum_{i=1}^{p} \frac{(d\lambda_i)^2}{\lambda_i^2} + \sum_{i<j}^{p} \frac{(\lambda_i - \lambda_j)^2}{\lambda_i\lambda_j}(d\theta_{ij})^2$$

where $d\theta_{ij}$ are the differential forms of the skew symmetric matrix $H'dH$. On the other hand,

$$\begin{aligned}\mathrm{tr}(V^{-1}\,dV) &= \mathrm{tr}(H'\Lambda^{-1}H[(dH')\,\lambda H + H'(d\Lambda)\,H + H'\Lambda H]\\ &= \mathrm{tr}(H\,dH' + \Lambda^{-1}(d\Lambda) + H'\,dH)\\ &= \mathrm{tr}(\Lambda^{-1}(d\Lambda))\\ &= \sum_{i=1}^{p} \frac{d\lambda_i}{\lambda_i}\end{aligned}$$

since $\mathrm{tr}(H'dH + HdH') = 0$ (due to the equality $H'H = I$). Theorem 3.1 follows

THEOREM 3.1. *The information metric differential form in $M$ is*

$$ds^2 = 2b_h\,\mathrm{tr}(V^{-1}\,dVV^{-1}\,dV) + (b_h - \tfrac{1}{4})\,\mathrm{tr}(V^{-1}\,dV)\,\mathrm{tr}(V^{-1}\,dV).$$

*By diagonalization of $V$, this metric becomes*

$$\begin{aligned}ds^2 = \left(3b_h - \frac{1}{4}\right)\sum_{i=1}^{p}\frac{d\lambda_i^2}{\lambda_i^2} + 2\left(b_h - \frac{1}{4}\right)\sum_{i<j}^{p}\frac{d\lambda_i\,d\lambda_j}{\lambda_i\lambda_j}\\ + 2b_h\sum_{i<j}^{p}\frac{(\lambda_i-\lambda_j)^2}{\lambda_i\lambda j}(d\theta_{ij})^2.\end{aligned}$$

We can use this metric differential form for computing the distance between two elliptical distributions in the same class differing only in the scatter matrix. This distance is identified with the distance between the two scatter matrices. In other words, it is

$$d(V_1, V_2) = \int_{V_1}^{V_2} ds.$$

THEOREM 3.2. *Let $H$ be a nonsingular matrix which reduces $V_1$ to $I_p$ and $V_2$ to a diagonal matrix with diagonal elements the eigenvalues of $V_1 V_2^{-1}$. The distance between two elliptical distributions with mean zero and scatter matrices $V_1$ and $V_2$ is*

$$\int_{V_1}^{V_2} ds = \frac{1}{|H|^{2p}} \int_{I_p}^{\Lambda} \sqrt{\left(3b_h - \frac{1}{4}\right) \sum_{i=1}^{p} \frac{(d\lambda_i)^2}{\lambda_i^2} + 2\left(b_h - \frac{1}{4}\right) \sum_{i<j}^{p} \frac{d\lambda_i \, d\lambda j}{\lambda_i \lambda_j}}.$$

*Proof.* Since there is no change in $H$ (we are using the same $H$ to transform both matrices $H'V_1H = I_p$ and $H'V_2H = \Lambda$) we could compute the distance between $I_p$ and $\Lambda$ and then divide this distance by the determinant of the Jacobian of the transformation:

$$\int_{I_p}^{\Lambda} ds = \int_{V_1}^{V_2} |H' \otimes H| \, ds.$$

But $|H' \otimes H| = |H|^{2p}$ (Magnus and Neudecker, 1988). The shortest distance would be along the geodesic curve $\lambda_i(t) = f_i(t)\, i = 1, p$ for some $f_i$, therefore we will try to find this curve in the next section. Note that the metric is of hyperbolic type.

Swain (1975) suggests using the "geodesic distance"

$$d(V_1, V_2) = \left( \sum_{i=1}^{p} (\ln \lambda_i)^2 \right)^{1/2}$$

to estimate $V_2$, given $V_1$ ($V_1$ could be derived from the data) when the distribution of the population is not known. However, it is not clear that the quantity is even a distance, let alone a geodesic if the distribution of the population is unknown. Geodesics are defined for smooth manifolds and depend on the metric. If the metric is not the natural metric for the manifold, it is not clear what the geodesic is really minimizing. The shortest distance with respect to one metric is not necessarily the shortest distance with respect to the natural metric. To make things clear, assume that we have the unit sphere on $\mathbf{R}^3$. The geodesic distance between the north and south poles is half the length of the great circle on the sphere (which is $\pi$) when we use the metric induced on the sphere, which is the proper metric on the sphere. However, if we use the metric in $\mathbf{R}^3$, the geodesic distance is 2, and the geodesics in $\mathbf{R}^3$ are straight lines, can one travel along straight lines on the sphere?

The same idea applies here where we are considering the space of elliptical distributions with fixed functional form. If we used the information metric defined for the manifold of normal distributions, we might be changing the natural geometric structure of the nonnormal elliptical space and the distance derived from this metric will not really reflect the natural length between two points in the space.

## 4. DERIVATION OF THE GEODESIC CURVE

Noting by $g_{ij}$ the elements of the information matrix, the positive definite quadratic differential form based on the elements of this information matrix

$$ds^2 = \sum_{i,\, j=1}^{p} g_{ij}\, d\lambda_i\, d\lambda_j,$$

but we have found $ds^2$ to be

$$\left(3b_h - \frac{1}{4}\right) \sum_{i=1}^{p} \frac{(d\lambda_i)^2}{\lambda_i^2} + 2\left(b_h - \frac{1}{4}\right) \sum_{i<j}^{p} \frac{d\lambda_i\, d\lambda j}{\lambda_i \lambda_j}$$

when we use the same orthogonal transformation to send $V_1$ to identity and $V_2$ to the diagonal matrix of eigenvalues of $V_1^{-1} V_2$. Consequently,

$$g_{ii}(\lambda) = \frac{(3b_h - (1/4))}{\lambda_i^2}, \qquad i = 1, p,$$

$$g_{ij}(\lambda) = \frac{2(b_h - (1/4))}{\lambda_i \lambda_j}, \qquad i < j,$$

and the geodesic is the solution of the system of differential equations

$$\ddot{\lambda}_k(t) + \sum_{i,\, j=1}^{p} \Gamma_{ij}^{k}\, \dot{\lambda}_i(t)\, \dot{\lambda}_j(t) = 0 \qquad k = 1 \cdot p,$$

where $\dot{\lambda}_i$ and $\ddot{\lambda}_i$ are the first and second derivatives of $\lambda_i$ with respect to $t$ and $\Gamma_{ij}^{k}$ are the Christoffel symbols defined by

$$\Gamma_{ij}^{k} = \frac{1}{2} \sum_{l} g^{kl} \left( \frac{\partial g_{jl}(\lambda)}{\partial \lambda_i} + \frac{\partial g_{il}(\lambda)}{\partial \lambda_j} - \frac{\partial g_{ij}(\lambda)}{\partial \lambda_l} \right)$$

with $g^{kl}$ being the $(k, l)$ element of the inverse of the information matrix $g$. But these expressions simplify, in our situation, to

$$\Gamma^{i}_{ii} = -((3b_h - \tfrac{1}{4})\,\lambda_i^{-3}\Big)\, g^{ii} + 2(b_h - \tfrac{1}{4})\,\lambda_i^{-2} \sum_{l \neq i}^{p} g^{il}\lambda_l^{-1}$$

$$\Gamma^{j}_{ii} = -((3b_h - \tfrac{1}{4})\,\lambda_i^{-3}\Big)\, g^{ij} + 2(b_h - \tfrac{1}{4})\,\lambda_i^{-2} \sum_{l \neq i}^{p} g^{il}\lambda_l^{-1}$$

$$\Gamma^{i}_{ij} = 0, \qquad i \neq j$$

$$\Gamma^{k}_{ij} = 0, \qquad i \neq j,\, i \neq k.$$

And we obtain for the equations of the geodesic

$$\ddot{\lambda}_k(t) + \sum_{i=1}^{p} \Gamma^{k}_{ii} \dot{\lambda}_i^2(t) = 0, \qquad k = 1, ..., p.$$

Since this system of O.D.E. is of second order and nonlinear, solving it is a very challenging task. We will not attempt to do that; instead, we will make the nice and convenient following change of coordinates

$$r_i = \ln \lambda_i.$$

The entries of the metric matrix become

$$g_{ii} = 3b_h - 1/4, \qquad g_{ij} = 2(b_h - 1/4),$$

and the metric becomes

$$ds^2 = (3b_h - 1/4) \sum_{i=1}^{p} (dr_i)^2 + 2(b_h - 1/4) \sum_{i<j}^{p} dr_i\, dr_j.$$

It is easy to check that this form is positive definite. Now we can solve the geodesic equations in terms of the $r_i$, $i = 1, ..., p$, coordinates. Since the metric does not depend on the $r_i$'s at all, all of the Christoffel symbols vanish and the parametric equations of the geodesic curve are

$$\ddot{r}_k(t) = 0, \qquad k = 1, ..., p.$$

The solution of this system is straightforward,

$$r_k(t) = a_k t + b_k, \qquad k = 1, ..., p.$$

and this is the parametric equation of the straight line in $\mathbf{R}^p$. The initial conditions are $r_k(0) = 0$, and $r_k(1) = r_k$ (since we go from the identity matrix to the diagonal matrix $\Lambda$, the transformations under $r_i = \ln \Lambda_i$ changes the identity matrix to 0 and the matrix $\Lambda$ to some diagonal matrix R; hence $b_k = 0$,

$k=1, ..., p$, and $a_k=\ln \lambda_k$, $k=1, ..., p$). The geodesic distance between $V_1$ and $V_2$ then becomes

$$\int_{V_1}^{V_2} ds = \frac{1}{|H|^{2p}} \int_{I_p}^{A} \left( \left(3b_h - \frac{1}{4}\right) \sum_{i=1}^{p} \frac{(d\lambda_i)^2}{\lambda_i^2} + 2\left(b_h - \frac{1}{4}\right) \sum_{i<j}^{p} \frac{d\lambda_i \, dj}{\lambda_i \lambda_j} \right)^{1/2}$$

$$= \frac{1}{|H|^{2p}} \int_{0}^{R} \left( \left(3b_h - \frac{1}{4}\right) \sum_{i=1}^{p} (dr_i)^2 + 2\left(b_h - \frac{1}{4}\right) \sum_{i<j}^{p} dr_i \, dr_j \right)^{1/2}$$

$$= \frac{1}{|H|^{2p}} \int_{0}^{1} \left( \left(3b_h - \frac{1}{4}\right) \sum_{i=1}^{p} (\ln \lambda_i)^2 + 2\left(b_h - \frac{1}{4}\right) \sum_{i<j}^{p} (\ln \lambda_i)(\ln \lambda_j) \right)^{1/2} dt$$

$$= \frac{1}{|H|^{2p}} \left( \left(3b_h - \frac{1}{4}\right) \sum_{i=1}^{p} (\ln \lambda_i)^2 + 2\left(b_h - \frac{1}{4}\right) \sum_{i<j}^{p} (\ln \lambda_i)(\ln \lambda_j) \right)^{1/2}.$$

The following theorem follows,

THEOREM 4.1. *Suppose the first and second logarithmic derivatives of h exist. The geodesic distance between two elliptical densities with equal means and different scatter matrices $V_1$ and $V_2$ is, for the fixed functional form h,*

$$d(V_1, V_2) = \frac{1}{|H|^{2p}} \sqrt{\left(3b_h - \frac{1}{4}\right) \sum_{i=1}^{p} (\ln \lambda_i)^2 + 2\left(b_h - \frac{1}{4}\right) \sum_{i<j}^{p} (\ln \lambda_i)(\ln \lambda_j)}.$$

Since the $\lambda$'s are never 0, $d(V_1, V_2)$ is well defined. When $V_1 = V_2$, $d(V_1, V_2) = 0$. Note that when the elliptical class is that of the multivariate normal distributions then H is the identity matrix (the metric is then invariant under congruent transformations) and $b_h = \frac{1}{4}$, therefore we obtain the metric obtained by James (1973).

An immediate consequence of the above theorem is

THEOREM 4.2. *Let $\hat{V}_n$ be the m.l.e. computed from a sample of size n drawn from the c.d.f. $f(x, V, h) = |V|^{-1/2} h(x' V^{-1} x)$. Assume the population scatter matrix V to be structured by a $q \times 1$, $q < p(p+1)/2$) parameter vector $\theta$, $V = V(\theta)$. Suppose that the true value $\theta_0$ is an interior point of the parameter space $\Theta$. In addition, suppose that $\Theta$ is a compact subset of $\mathbf{R}^q$ and $V(\theta)$ is differentiable; then this structure induces an embedding*

$$N = \{ f(x, V, h) = |V(\theta)|^{-1/2} h(x' V^{-1}(\theta) x), \theta \in \mathbf{R}^q \}.$$

*Let $\lambda_1, \lambda_2, ..., \lambda_p$ be the latent roots of the determinantal equation*

$$\det(\hat{V} - \lambda V(\theta)),$$

*then the minimum geodesic distance estimator of $\theta$ is consistent, asymptotically normal and first order efficient Furthermore, the statistic*

$$d_n^2 = n \frac{1}{|H|^{2p}} \left(3b_h - \frac{1}{4}\right) \sum_{i=1}^{p} (\ln \lambda_i)^2 + 2\left(b_h - \frac{1}{4}\right) \sum_{i<j}^{p} (\ln \lambda_i)(\ln \lambda_j)$$

*is asymptotically distributed as a $\chi^2$ variable with $p(p+1)/2 - q$ degrees of freedom.*

*Proof.* The proof follows from Kass (1980) or Skovgaard (1981). The degree of freedom is $\mathrm{Dim}(M) - \mathrm{Dim}(N)$. Note that the above results can be expressed for the covariance matrix $\Sigma$ by use of the relation $\Sigma = -2\psi(0)\,V = (E(U)/p)\,V$, where $\psi(t'Vt)$ is the characteristic function of $X$ (with zero location).

*Remark.* Let $v = \mathrm{Vecs}(V)$, where Vecs refers to the vector of nonredundant elements of V stacked columnwise. Let $\Gamma$ be the asymptotic covariance matrix of the maximum likelihood estimator of $V$. $\Gamma$ is the inverse of the information matrix. From Mitchell (1989) or Kano *et al.* (1993) the entries of the information matrix can be derived as

$$2b_h \mathrm{tr}(V^{-1}I_{ij}V^{-1}I_{kl}) + \tfrac{1}{4}(4b_h - 1)\,\mathrm{tr}(V^{-1}I_{ik})\,\mathrm{tr}(V^{-1}I_{jl}),$$

where

$$I_{rs} = \begin{cases} I_{r,r} & \text{if } r = s \\ I_{r,s} + I_{s,r} & \text{if } r \neq s \end{cases}.$$

$I_{r,s}$ denotes the $p \times p$ matrix whose $(r, s)$th entry is 1 and 0 elsewhere. Let

$$Q_{\theta\theta'} = \frac{\partial^2(d_n^2)}{\partial\theta\,\partial\theta'}$$

$$Q_{\theta v'} = \frac{\partial^2(d_n^2)}{\partial\theta\,\partial v'},$$

where the derivatives are evaluated at the true values of $v$ and $\theta$. If $Q_{\theta\theta'}$ is positive definite, the covariance matrix $\Omega$ of the minimum geodesic distance estimator is easily derived by use of the implicit function theorem when solving for

$$F(\hat{v}, \theta) = \frac{\partial d^2(\hat{v}, \theta)}{\partial\theta} = 0,$$

see for example Shapiro (1983). It is

$$\Omega = Q_{\theta\theta'}^{-1} Q_{\theta v'} \Gamma Q'_{\theta v'} Q_{\theta\theta'}^{-1}.$$

Since $v$ and the eigenvalues $\lambda_i$ will depend on the structure parameters, the chain rule needs to be used in $Q_{\theta\theta'}$ and $Q_{\theta v'}$.

*Remark.* It is known that the geodesic distance between two multivariate normal distributions with distinct means $\mu_1$ and $\mu_2$ and the same variance $\Sigma$ is the Mahalanobis distance $(\mu_2 - \mu_1)' \Sigma^{-1}(\mu_2 - \mu_1)$. One can easily show that the geodesic distance (based on the information metric) between two elliptical distributions with locations $\mu_1$ and $\mu_2$ and the same scatter matrix $V$ is

$$d(\mu_1, \mu_2) = \frac{\sqrt{4c_h}}{|V|^{1/2}} ((\mu_2 - \mu_1)' V^{-1}(\mu_2 - \mu_1))^{1/2},$$

where $c_h = (1/p) E(Uw^2)$. This distance can be used to derive estimators of model parameters related to mean structures.

## 5. DISCUSSION

In this paper, we have exploited the geometric properties of the manifold of multivariate elliptical distributions equipped with Rao's metric (information metric) to propose a geometric method for estimating the parameters of a model on covariance matrix and testing for the fit of this model. The test statistic based on the geodesic distance between two elliptical distributions with identical means and different covariance matrices turned out to be a generalization of the test proposed by James (1973). A geodesic distance between distributions with distinct locations and scatter matrices is under study.

## REFERENCES


Amari, S. I. (1984). *Differential–Geometrical Methods in Statistics.* Springer-Verlag, New York/Berlin.

Atkinson, C., and Mitchell, A F. S. (1981). Rao's distance measure. *Sankhyã Ser. A* **43**, (3) 345–365.

Burbea, J., and Rao, C. R. (1982). Entropy differential metric, distance and divergence measures in probability spaces: A unified approach. *J. Multivariate Anal.* **12** 575–596.

James, A. T. (1973). The variance information manifold and the functions on it. In *Multivariate analysis III*, pp. 157–169. Academic Press, New York.

Kano, Y., Berkane, M., and Bentler, P. (1993). Statistical inference based on pseudo maximum liklihood estimators in elliptical distributions. *J. Amer. Statist. Accos.* **421** 135–143.

Kass, R. E. (1980). *The Riemannian Structure of Model Spaces*: *A Geometrical Approach to Inference*. Ph.D. thesis. [University of Chicago]

Maass, H. (1955). Die Bestimmung der dirichletreihen mit Grossencharakteren zu den Modulformen *n*-ten Grades. *J. Indian Math. Soc.* **19** 1–23.

Mitchell, A. F. S. (1988). Statistical manifolds of univariate elliptical distributions. *Inter. Statist. Rev.* **56** 464–467.

Mitchell, A. F. S. (1989). The information matrix, skewness tensor and $\alpha$-connections for the general multivariate elliptical distribution. *Ann. Inst. Statist. Math.* **41** 289–304.

Rao, C. R. (1945). Information and the accuracy atteinable in the estimation of statistical parameters. *Bull. Calcutta Math. Soc.* **37** 81–91.

Rao, C. R. (1987). Differential metrics in probability spaces. *Differential Geom. Statist. Inference* 217–237.

Shapiro, A. (1983). Asymptotic distribution theory in the analysis of covariance structures. *South African Statist. J.* **17** 33–81.

Skovgaard, L. T. (1981). *A Riemannian Geometry of the Multivariate Normal Model.* Research Report 81/3. Statistical Research unit, Danish medical Research Council, Danish Social Science Research Council.

Swain, A. J. (1975). A class of factor analysis estimation procedures with common asymptotic sampling properties. *Psychometrika* **40** 315–335.