ELSEVIER

Journal of Computational and Applied Mathematics 145 (2002) 319–334

JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

www.elsevier.com/locate/cam

A distance between elliptical distributions based in an embedding into the Siegel group

Miquel Calvo *, Josep M. Oller[1]

Departament d'Estadistica, Facultat de Biologia, Diagonal 645, Universitat de Barcelona, 08028 Barcelona, Spain

Received 2 September 2000; received in revised form 20 August 2001

Abstract

This paper describes two different embeddings of the manifolds corresponding to many elliptical probability distributions with the informative geometry into the manifold of positive-definite matrices with the Siegel metric, generalizing a result published previously elsewhere. These new general embeddings are applicable to a wide class of elliptical probability distributions, in which the normal, *t*-Student and Cauchy are specific examples. A lower bound for the Rao distance is obtained, which is itself a distance, and, through these embeddings, a number of statistical tests of hypothesis are derived. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Information metric; Rao distance; Siegel geometry; Elliptical distributions

1. Introduction

The study of the family of elliptical distributions forms the basis of a generalized multivariate analysis, see for instance [9] or [15]. This family includes most importantly the multivariate normal distribution, all the Pearson-type VII distributions including the multivariate Student-*t* and Cauchy, among many other interesting distributions.

For purposes of statistical and data analysis, the distance concept has been proved to be a very useful tool, see for instance [18,5,10]. Related to a convenient introduction of a distance, several authors have described differential geometric methods and, in particular, have studied the information

* Corresponding author. Tel.: +34-93-402-15-90; fax: +34-93-411-17-33.

E-mail addresses: calvo@bio.ub.es (M. Calvo), oller@bio.ub.es (J.M. Oller).

[1] This work is supported by DGICYT grant (Spain), PB96-1004-C02-01 and 1997SGR00188.

PII: S0377-0427(01)00584-2

metric for parametric families of probability distributions. The Riemannian distance corresponding to this metric is known as the *Rao distance*, which possesses interesting properties. One of the most remarkable of these is its invariance under admissible transformations of the parameters. Several relevant contributions in this field include Amari [1], Atkinson and Mitchell [2], Burbea [4], Burbea and Rao [5], Burbea and Oller [6], Oller [16] and Oller and Corcuera [17] among many others. Unfortunately, a general closed form of the Rao distance for the elliptical family has not as yet been obtained, except for a number of specific configurations on the parameters of the implied densities, see [13,14,3].

In a previous paper, Calvo and Oller [7] introduced an embedding of the manifold of the multivariate normal densities with informative geometry into the manifold of the positive-definite matrices. Besides the geometrical interest of the embedding, perhaps the main obtained result, from applied point of view, was the closed form of the induced Riemannian distance. The distance in the Siegel group gives a close lower bound of the Rao distance. We proposed using this so-called Siegel distance as a substitute in applied data analysis for the unknown expression of the Rao distance, see [8]. In Calvo and Oller [7] general statistical tests were also derived starting from purely geometrical considerations, and were compared with classical results obtained using likelihood ratio criteria.

The need for a general distance for the family of elliptical distributions is reflected in several recent papers, see for instance [3], or, for the normal case, see [12]. In this paper we propose an extension of our earlier result for the multivariate normal case to a wide class of elliptical probability distributions, which includes those cases mentioned at the beginning of this introduction. Some statistical tests are also derived.

2. Preliminary considerations

In this section we summarize some known results of the information metric for the elliptical distributions. We begin by introducing some notation. Let $M_{n\times m}(\mathbb{R})$ be the space of all $n \times m$ real matrices, $S_n(\mathbb{R})$ stands for the subspace of symmetric matrices of $M_{n\times n}(\mathbb{R})$, $\mathrm{GL}_n(\mathbb{R})$ denotes the group of regular matrices of $M_{n\times n}(\mathbb{R})$ and $P_n(\mathbb{R})$ the sub-set of positive-definite matrices in $S_n(\mathbb{R})$. If an n-dimensional random vector X has a density, with respect to the Lebesgue measure, given by

$$f(x;\mu,\Sigma) = \frac{\Gamma(n/2)}{\pi^{n/2}}|\Sigma|^{-1/2}F((x-\mu)'\Sigma^{-1}(x-\mu)), \quad x \in \mathbb{R}^n,$$

where $\mu=(\mu_i) \in M_{n\times 1}(\mathbb{R})$ and $\Sigma=(\sigma_{ij}) \in P_n(\mathbb{R})$ are the parameters, and F is a function $F:\mathbb{R}^+ \to \mathbb{R}^+$ satisfying

$$\int_0^\infty t^{n/2-1}F(t)\,\mathrm{d}t = 1,$$

it is said to have an *elliptical distribution* with parameters μ and Σ. If the random vector X has expectation, we have $E(X)=\mu$, as in the multivariate normal model. The covariance of X is $\mathrm{cov}(X)=c_F\Sigma$, provided its existence, where

$$c_F = \frac{1}{n}\int_0^\infty t^{n/2}F(t)\,\mathrm{d}t.$$

The sub-family of n-dimensional elliptical distributions defined by the function F can be noted as $\mathscr{E}_{F,n}(\mu,\Sigma)$. Let us assume two additional conditions:

$$a = \frac{4}{n}\int_0^\infty t^{n/2}\left(\frac{\mathrm{d}\log F(t)}{\mathrm{d}t}\right)^2 F(t)\,\mathrm{d}t < \infty,$$

$$b = \frac{4}{n(n+2)}\int_0^\infty t^{n/2+1}\left(\frac{\mathrm{d}\log F(t)}{\mathrm{d}t}\right)^2 F(t)\,\mathrm{d}t < \infty. \tag{1}$$

For example, choosing F as

$$F(t) = \frac{\Gamma(m)}{\Gamma(n/2)\Gamma(m-n/2)}\frac{s^{-n/2}}{(1+t/s)^m} \quad \text{with } m > \frac{n}{2}, \quad s > 0, \tag{2}$$

we have the Pearson-type VII class of distributions, see [9]. For any member of the class in (2) of elliptical distributions, the values for a and b in (1) are

$$a = \frac{(2m-n)m}{(m+1)s}, \quad b = \frac{m}{m+1}. \tag{3}$$

In the particular case of the multivariate Student-t distribution, we have $m=(n+v)/2$ and $s=v$, where v are the degrees of freedom of the t-distribution. The values of a and b simplify now to

$$a = b = \frac{n+v}{n+v+2}. \tag{4}$$

To obtain the multivariate Cauchy distribution, take $v=1$. In this case, the values are $a=b=(n+1)/(n+3)$.

Another interesting family can be obtained when F has the following form:

$$F(t) = \frac{sr^{(2m+n-2)/2s}}{\Gamma((2m+n-2)/2s)}\,t^{m-1}\exp(-rt^s) \quad \text{with } r,s>0, m>1-\tfrac{n}{2} \tag{5}$$

For the members of this family (5), the values for a and b in (1) are:

$$a = \frac{r^{1/s}}{n}\,\frac{\Gamma((2m+n-4)/2s)}{\Gamma((2m+n-2)/2s)}\{(n-2)^2+2s(2m+n-4)\},$$

$$b = \frac{n^2+2ns+4s(m-1)}{n^2+2n}.$$

In the particular case of $m=1$, $s=1$ and $r=1/2$, that is,

$$F(t) = \frac{1}{2^{n/2}\Gamma(n/2)}\,\mathrm{e}^{-t/2} \quad t \geqslant 0,$$

we have the non-singular *multivariate normal* distribution and the constants are $a=b=c_F=1$.

To close this section, let us briefly review a number of results of the *information metric* for the elliptical distributions. We note with $\Theta = M_{n\times 1}(\mathbb{R})\times P_n(\mathbb{R})$ the parametric space corresponding to the family of *elliptical distributions* parameterized by $\mu=(\mu_i)\in M_{n\times 1}(\mathbb{R})$ and $\Sigma=(\sigma_{ij})\in P_n(\mathbb{R})$. Note that Θ is the *common* parameter space of any sub-family $\mathscr{E}_{F,n}(\mu,\Sigma)$. The introduction of the information metric in Θ (see [14,16]) equips Θ with a *Riemannian manifold* structure. The expression of the line element, at any point $\theta=(\mu,\Sigma)$ is

$$ds^2 = a\,d\mu'\Sigma^{-1}\,d\mu + \frac{b}{2}\operatorname{tr}\{(\Sigma^{-1}d\Sigma)^2\} + \frac{(b-1)}{4}\operatorname{tr}^2(\Sigma^{-1}d\Sigma), \tag{6}$$

where a and b, defined in (1), can be shown to satisfy

$$a>0 \quad \text{and} \quad b>\frac{n}{n+2}.$$

Moreover, the geodesic equations are given by

$$\ddot{\mu} - \dot{\Sigma}\Sigma^{-1}\dot{\mu} = 0,$$

$$\ddot{\Sigma} + \frac{a}{b}\dot{\mu}\dot{\mu}' - \frac{a(b-1)}{2b^2+nb(b-1)}\dot{\mu}'\Sigma^{-1}\dot{\mu}\,\Sigma - \dot{\Sigma}\Sigma^{-1}\dot{\Sigma} = 0.$$

A first integration of the above system can be achieved, giving

$$\dot{\mu} = \Sigma v,$$

$$\dot{\Sigma} = \Sigma\left(W - \frac{a}{b}v\mu' + \frac{a(b-1)}{2b^2+nb(b-1)}\mu' vI\right), \tag{7}$$

where $v\in M_{n\times 1}(\mathbb{R})$ and $W\in S_n(\mathbb{R})$ are integration constants. Although a general solution of (7) has not yet been obtained, it is possible to solve these equations when $v=0$, obtaining the geodesics easily and also the Riemannian distance induced on the sub-manifold $\Theta_{\mu_0}\equiv\{(\mu,\Sigma)\in\Theta:\mu=\mu_0 \textit{ constant}\}$, see for instance [3,16].

3. The embedding of Θ

With the introduction of the differential metric in $P_{n+1}(\mathbb{R})$:

$$ds^2 = \tfrac{1}{2}\operatorname{tr}\{(\Psi^{-1}d\Psi)^2\},$$

where $\Psi\in P_{n+1}(\mathbb{R})$, $P_{n+1}(\mathbb{R})$ has a Riemannian manifold structure. This metric defines a well-known geometry, originally studied by Siegel [19] in the Hermitian matrix set. In the context of the informative geometry, it was studied by James [11], Burbea [4] and in a paper directly related to the present one by Calvo and Oller [7]. In fact, for our purposes we need to slightly modify the above-related metric to

$$ds^2 = \frac{b}{2}\operatorname{tr}\{(\Psi^{-1}\,d\Psi)^2\}, \quad b\in\mathbb{R}^+, \tag{8}$$

where b is the constant introduced in (1). The reason for its choice will be clarified later. This metric has the same geodesic equations as the original metric defined by Siegel. The geodesic equations and its solution for a geodesic starting at $\Psi(0)=\Psi_1$ are, respectively,

$$\dot{\Psi}(s) = \Psi(s)H, \quad \Psi(s) = \Psi_1\exp(Hs), \tag{9}$$

where H is a constant matrix such that $\Psi H = H'\Psi$. Let us define the norm for any $A \in M_{n\times n}(\mathbb{R})$ as

$$\|A\| = \sqrt{\operatorname{tr}(AA')}.$$

Then, the *Riemannian distance* between two points Ψ_1 and $\Psi_2 \in P_{n+1}(\mathbb{R})$ is

$$d(\Psi_1;\Psi_2) = \sqrt{\frac{b}{2}}\|\ln(\Psi_1^{-1/2}\Psi_2\Psi_1^{-1/2})\| = \left(\frac{b}{2}\sum_{i=1}^{n+1}\ln^2\lambda_i\right)^{1/2}, \tag{10}$$

where the ln in the first part of the expression stands for the matrix natural logarithm. In the second, λ_i are the eigenvalues of $\Psi_1^{-1}\Psi_2$ or $\Psi_1^{-1/2}\Psi_2\Psi_1^{-1/2}$. The metric (8) is invariant under the action of $\mathrm{GL}_n(\mathbb{R})$ given by

$$\Psi \mapsto P'\Psi P, \quad P \in \mathrm{GL}_n(\mathbb{R}).$$

The following lemma prepares the embedding of Θ in $P_{n+1}(\mathbb{R})$.

Lemma 3.1. *For a fixed* $\alpha, \eta \in \mathbb{R}$, $\eta > 0$, *any* $\Psi \in P_{n+1}(\mathbb{R})$ *can be written as*

$$\Psi = |\Sigma|^{\alpha}\begin{pmatrix} \Sigma + \beta\eta^2\mu\mu' & \beta\eta\mu \\ \beta\eta\mu' & \beta \end{pmatrix}, \quad \beta \in \mathbb{R}^+, \quad \mu \in M_{n\times 1}(\mathbb{R}), \quad \Sigma \in P_n(\mathbb{R}).$$

Conversely, any matrix of the above-stated form is symmetric and positive definite. For any $\Psi \in P_{n+1}(\mathbb{R})$ *it can be verified that*

$$ds^2 = \frac{b}{2}\Bigg[((n+1)\alpha^2 + 2\alpha)\operatorname{tr}^2(\Sigma^{-1}d\,\Sigma) + \operatorname{tr}\{(\Sigma^{-1}d\,\Sigma)^2\}$$
$$+\, 2\beta\,\eta^2\,d\mu'\,\Sigma^{-1}\,d\mu + 2\alpha\operatorname{tr}(\Sigma^{-1}d\,\Sigma)\frac{d\beta}{\beta} + \left(\frac{d\beta}{\beta}\right)^2\Bigg].$$

Proof. Any

$$\Psi = \begin{pmatrix} \Psi_{11} & \Psi_{12} \\ \Psi_{12}' & \Psi_{22} \end{pmatrix}$$

can be written as the above-stated form taking

$$\eta\mu = \Psi_{22}^{-1}\Psi_{12},$$
$$\beta = |\Psi_{11} - \Psi_{22}^{-1}\Psi_{12}\Psi_{12}'|^{-\alpha/(n\alpha+1)}\Psi_{22},$$
$$\Sigma = |\Psi_{11} - \Psi_{22}^{-1}\Psi_{12}\Psi_{12}'|^{-\alpha/(n\alpha+1)}(\Psi_{11} - \Psi_{22}^{-1}\Psi_{12}\Psi_{12}').$$

To prove the converse, decompose Ψ as

$$\Psi = |\Sigma|^{\alpha}\begin{pmatrix} \Sigma & 0 \\ 0' & 0 \end{pmatrix} + |\Sigma|^{\alpha}\beta\begin{pmatrix} \eta\mu \\ 1 \end{pmatrix}(\eta\mu' \quad 1)$$

and note that $\Sigma \in P_n(\mathbb{R})$. To prove the third part, observe that

$$d|\Sigma| = |\Sigma| \operatorname{tr}(\Sigma^{-1} d\Sigma)$$

and therefore computing

$$\Psi^{-1} = |\Sigma|^{-\alpha} \begin{pmatrix} \Sigma^{-1} & -\eta\Sigma^{-1}\mu \\ -\eta\mu'\Sigma^{-1} & \beta^{-1} + \eta^2\mu'\Sigma^{-1}\mu \end{pmatrix}, \tag{11}$$

$$\Psi^{-1} d\Psi = \alpha \operatorname{tr}(\Sigma^{-1} d\,\Sigma) I_{n+1} + \begin{pmatrix} \Sigma^{-1}(d\Sigma + \beta\eta^2 d\mu\,\mu') & \beta\eta\Sigma^{-1}d\mu \\ -\eta\mu'\Sigma^{-1}(d\Sigma + \beta\eta^2 d\mu\,\mu') + \eta\, d\mu' + \frac{\eta}{\beta}\mu' d\beta & \frac{1}{\beta}d\beta - \beta\eta^2\mu'\Sigma^{-1}d\mu \end{pmatrix} \tag{12}$$

apply the metric (8) and note that $\operatorname{tr}(Axy') = y'Ax$ for $\forall x, y \in M_{n\times 1}(\mathbb{R})$ and $\forall A \in M_{n\times n}(\mathbb{R})$. □

Let us now define the following class of maps from Θ to $P_{n+1}(\mathbb{R})$:

$$f_{\alpha,\eta} : \Theta \to P_{n+1}(\mathbb{R}),$$

$$(\mu, \Sigma) \mapsto |\Sigma|^{\alpha} \begin{pmatrix} \Sigma + \eta^2\mu\mu' & \eta\mu \\ \eta\mu' & 1 \end{pmatrix} \tag{13}$$

for $\alpha \in \mathbb{R}$, $\eta \in \mathbb{R} - \{0\}$. As is shown in the following theorem, these maps permit the embedding of the parameter space of many sub-families $\mathscr{E}_{F,n}(\mu, \Sigma)$ into a hyper-surface of $P_{n+1}(\mathbb{R})$. This sub-manifold has one dimension less than the full manifold $P_{n+1}(\mathbb{R})$.

Theorem 3.2. *Let $f_{\alpha,\eta}$ be a map as defined in* (13). *The following results hold*:

1. *$f_{\alpha,\eta}$ is a diffeomorphism of Θ onto $f_{\alpha,\eta}(\Theta)$.*
   $f_{\alpha,\eta}(\Theta)$ is a $((n+1)(n+2)/2 - 1)$-dimensional sub-manifold of $P_{n+1}(\mathbb{R})$.
2. *For the elliptical distributions such that b in* (1) *satisfies*

   $$b \geqslant \frac{n+1}{n+3}$$

   defining the constants

   $$\gamma = \sqrt{\frac{a}{b}}, \quad \alpha_j = \frac{-1 + (-1)^j \sqrt{1 + (n+1)(b-1)/2b}}{n+1}, \quad j = 1, 2,$$

   we have that
3. *The Riemannian metric defined in* (8) *induced on the sub-manifolds $f_{\alpha_j,\gamma}(\Theta)$, with $j = 1, 2$, can be expressed as*

   $$ds^2 = a\, d\mu'\, \Sigma^{-1} d\mu + \frac{b}{2} \operatorname{tr}\{(\Sigma^{-1} d\,\Sigma)^2\} + \frac{b-1}{4} \operatorname{tr}^2(\Sigma^{-1} d\,\Sigma).$$

4. *Θ is isometric to every $f_{\alpha_j,\gamma}(\Theta)$, with $j = 1, 2$.*
5. *$f_{\alpha_j,\gamma}(\Theta)$, $j = 1, 2$, are non-geodesic sub-manifolds of $P_{n+1}(\mathbb{R})$.*
6. *$f_{\alpha_j,\gamma}(\Theta_{\mu_0})$, with $j = 1, 2$, are $n(n+1)/2$-dimensional geodesic sub-manifolds of $P_{n+1}(\mathbb{R})$.*

Proof. Prove (1) noting that $f_{\alpha,\eta}$ is a one to one mapping to $f_{\alpha,\eta}(\Theta)$ and it is C^∞. (2) is trivially checked. For (3) observe that at $f_{\alpha,\eta}(\mu, \Sigma)$ expression (8) becomes

$$ds^2 = \frac{b}{2}\,\mathrm{tr}\{(f_{\alpha,\eta}^{-1}(\mu,\Sigma)\,d f_{\alpha,\eta}(\mu,\Sigma))^2\},$$

then apply Lemma 3.1, taking into account that $\beta = 1$ and $d\beta = 0$, and replace η and α by their values stated before in the theorem to identify the ds^2 element to expression (6). With this result every $f_{\alpha_j,\gamma}$ is an isometry and Θ is isometric to $f_{\alpha_j,\gamma}(\Theta)$.

For (5), combine the expression of the geodesics in Θ expressed in (7) with the isometry now established in expression (13). The resulting expression will correspond to the embedding of the geodesic curve in Θ into $f_{\alpha_j,\gamma}(\Theta)$. This results in

$$\begin{aligned}\Psi^{-1}\dot{\Psi} &= \alpha(\mathrm{tr}\, W - 2a\kappa\mu' v) I_{n+1} \\ &\quad + \begin{pmatrix} W + a(b-1)b^{-1}\kappa\mu' v I_n & \gamma v \\ -\gamma\mu' W + a(b-1)b^{-1}\kappa\gamma\mu' v\mu' + \gamma v'\Sigma & -\gamma^2\mu' v \end{pmatrix},\end{aligned}$$

where $\kappa = (2b + n(b-1))^{-1}$. Observe that $\Psi^{-1}\dot{\Psi}$ is not constant, and by expression (9) of the geodesics in $P_{n+1}(\mathbb{R})$, it can be concluded that the geodesic equations are not equivalent.

Finally, for (6), follow the same steps as in (5), but now with $\mu = \mu_0$, $v = 0$ and thus $\dot{\mu} = 0$. The expression simplifies to

$$\Psi^{-1}\dot{\Psi} = (\alpha\,\mathrm{tr}\, W)\, I_{n+1} + \begin{pmatrix} W & 0 \\ -\gamma\mu_0' W & 0 \end{pmatrix}.$$

Now $\Psi^{-1}\dot{\Psi}$ is also constant and the geodesic equations are the same. □

Let us consider the particular case of the Pearson-type VII class of distributions defined in (2). From (3), the distributions which verify the required condition $b \geqslant (n+1)/(n+3)$ are the densities such that

$$m \geqslant (n+1)/2.$$

Some examples of such densities are the multivariate-t ($m = (n+v)/2$; $s = v$) and the Cauchy distributions. In the latter case, where $m = (n+1)/2$, $s = 1$, and $b = a = (n+1)/(n+3)$ there is only one possible embedding: $\alpha_1 = \alpha_2 = -(n+1)^{-1}$.

For the family defined in (5), the required condition for embedding the density is restricted to the members such that

$$m \geqslant \frac{n}{2s(n+3)} - \frac{n}{2} + 1.$$

As a specific case, the multivariate normal ($m=s=1$; $r=1/2$) can be embedded in $P_{n+1}(\mathbb{R})$ with $\gamma=1$ and

$$\alpha_1 = \frac{-2}{n+1} \quad \text{or} \quad \alpha_2 = 0. \tag{14}$$

The embedding corresponding to the second root, with $\alpha=0$ and $\gamma=1$, was proposed in Calvo and Oller [7].

Corollary 3.3. *Let θ_1 and θ_2 be two points of Θ. If ρ is the Rao distance between them and d_j is the Riemannian distance between $f_{\alpha_j,\gamma}(\theta_1)$ and $f_{\alpha_j,\gamma}(\theta_2)$ in $P_{n+1}(\mathbb{R})$, then $\rho \geqslant d_j$ for $j=1,2$. If $\theta_1,\theta_2 \in \Theta_{\mu_0}$ then $d_j=\rho$.*

Proof. By Theorem 3.2(4), Θ is isometric to $f_{\alpha_j,\gamma}(\Theta)$. Then, ρ is also the Riemannian distance, induced by the Siegel metric (8), between $f_{\alpha_j,\gamma}(\theta_1)$ and $f_{\alpha_j,\gamma}(\theta_2)$ on $f_{\alpha_j,\gamma}(\Theta)$. By (5), the geodesic distance restricted there is greater than or equal to the geodesic distance on the complete manifold $P_{n+1}(\mathbb{R})$. For the second part, note that by Theorem 3.2(6) Θ_{μ_0} and $f_{\alpha_j,\gamma}(\Theta\mu_0)$ are geodesic sub-manifolds on Θ and $P_{n+1}(\mathbb{R})$, respectively; therefore, $d_j=\rho$. □

Observe that the above-stated corollary, in particular, allows us to affirm that the distance obtained in [3] or [16] is the Rao distance for the full sub-family $\mathscr{E}_{F,n}(\mu,\Sigma)$, *which coincides* with the Rao distance restricted to the sub-model corresponding to Θ_{μ_0}.

Corollary 3.4. *Let $\theta_1,\theta_2 \in \Theta$. If $d_m = \max(d_1(\theta_1;\theta_2); d_2(\theta_1;\theta_2))$, then d_m is a distance and $\rho \geqslant d_m$.*

Notice that with the aid of (14) and Corollary 3.4 the lower bound for the Rao distance introduced in [7] is improved.

Let us now consider the following transformation in the parameter space Θ:

$$\theta=(\mu,\Sigma) \mapsto \bar{\theta}=(\bar{\mu},\bar{\Sigma})=(Q\mu+c, Q\Sigma Q'), \quad c\in\mathbb{R}^n, \quad Q\in \mathrm{GL}_n(\mathbb{R}),$$

which corresponds to an affine transformation on the random variables of the form $\bar{X}=Q'X+c$. The expression for $f_{\alpha_j,\gamma}(\bar{\theta})=\bar{\Psi}$ is

$$\bar{\Psi} = |\bar{\Sigma}|^{\alpha_j}\begin{pmatrix} \bar{\Sigma}+\gamma^2\bar{\mu}\bar{\mu}' & \gamma\bar{\mu} \\ \gamma\bar{\mu}' & 1 \end{pmatrix},$$

which also can be written as

$$\bar{\Psi} = |Q|^{2\alpha_j}\,|\Sigma|^{\alpha_j}\begin{pmatrix} Q' & \gamma c \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \Sigma+\gamma^2\mu\mu' & \gamma\mu \\ \gamma\mu' & 1 \end{pmatrix}\begin{pmatrix} Q & 0 \\ \gamma c' & 1 \end{pmatrix}.$$

With the latter expression it can be easily proved that for any $\theta_1=(\mu_1,\Sigma_1)$ and $\theta_2=(\mu_2,\Sigma_2)$, the distances d_j are invariant for affine transformations on the random variables:

$$d_j(\theta_1,\theta_2)=d_j(\bar{\theta}_1,\bar{\theta}_2), \quad j=1,2.$$

Obviously, the same property is also applicable to the distance d_m.

4. Some applications to hypothesis testing

As in [7], we describe the possibility of constructing statistical tests of hypothesis based on the parameters of elliptical distributions, by only taking into consideration geometrical questions. For a hypothesis testing problem such that

$$H_0: g(\theta) = 0, \quad H_1: g(\theta) \neq 0, \tag{15}$$

where g is a smooth function and $\theta \in \Theta$, the null hypothesis H_0 defines a sub-manifold $\Theta_{H_0} \subset \Theta$. If $\theta_N^* = (\mu^*, \Sigma^*)$ denotes the likelihood estimation of θ, based on a sample $x \in \mathbb{R}^{n\times N}$ of size $N > 1$, the distance from θ_N^* and Θ_{H_0} is defined as

$$d_j(\theta_N^*, \Theta_{H_0}) \equiv \inf\{d(f_{\alpha_j,\gamma}(\theta), f_{\alpha_j,\gamma}(\theta_N^*)) \colon \theta \in \Theta_{H_0}\},$$

where d is the distance on $P_{n+1}(\mathbb{R})$ defined in (10) and α_j is any of the two constants in Theorem 3.2. The critical region corresponding to the previous hypothesis testing problem is

$$W_\varepsilon \equiv \{x \in \mathbb{R}^{n\times N} \colon d_j(\theta_N^*, \Theta_{H_0}) > A_\varepsilon\},$$

where ε is the significance level of the test and the constant A_ε is chosen in such a way that

$$\mathrm{Prob}(x \in W_\varepsilon \mid H_0) \leqslant \varepsilon.$$

Note that, a priori, each of the two possible embeddings of Θ (corresponding to α_1 or α_2) give rise to *possible* different statistical tests.

4.1. Testing equality to a fixed covariance matrix

Let us consider the following hypothesis testing problem:

$$H_0: \Sigma = \Sigma_0, \quad H_1: \Sigma \neq \Sigma_0.$$

As in [7], the problem of minimizing d is more easily solved by considering that $P_{n+1}(\mathbb{R})$ is a *complete manifold*. Then, there exist $\theta_0 = (\mu_0, \Sigma_0)$ such that $d(f_{\alpha_j,\gamma}(\theta_0), f_{\alpha_j,\gamma}(\theta_N^*))$ is minimum. By the *Gauss Lemma* the geodesic curve between the two points is orthogonal to $\Theta_{H_0} \equiv \Theta_{\Sigma_0} \equiv \{(\mu, \Sigma) \in \Theta : \Sigma = \Sigma_0\}$ at $f_{\alpha_j,\gamma}(\theta_0)$. Taking the boundary conditions $\Psi(0) = f_{\alpha_j,\gamma}(\theta_0)$ and $\Psi(\rho) = f_{\alpha_j,\gamma}(\theta_N^*)$, where $\rho = d(f_{\alpha_j,\gamma}(\theta_0), f_{\alpha_j,\gamma}(\theta_N^*))$ the expression of the orthogonal geodesic in Lemma 5.3 simplifies to

$$\Psi(s) = |\Sigma_0(\Sigma_0^{-1}\Sigma^*)^{s/\rho}|^{\alpha_j} \begin{pmatrix} \Sigma_0(\Sigma_0^{-1}\Sigma^*)^{s/\rho} + \gamma^2\mu\mu' & \gamma\mu \\ \gamma\mu' & 1 \end{pmatrix}.$$

Notice that along this geodesic $\gamma\mu$ is constant, and $\theta_0 = (\mu^*, \Sigma_0)$. The element $(n+1)\times(n+1)$ is 1, and therefore the geodesic in $P_{n+1}(\mathbb{R})$ and the geodesic in $f_{\alpha_j,\gamma}(\Theta_{\Sigma_0})$ are both the same. Because the Rao distance and the Siegel distance coincide in this case, a test based on the Rao distance may produce the same test as that of the Siegel. The distance between $f_{\alpha_j,\gamma}(\theta_0)$ and $f_{\alpha_j,\gamma}(\theta_N^*)$ can be expressed as

$$d_j(\theta_N^*, \Theta_{\Sigma_0}) = \left\{ \frac{b-1}{4} \ln^2|\Sigma_0^{-1}\Sigma^*| + \frac{b}{2}\|\ln(\Sigma_0^{-1}\Sigma^*)\|^2 \right\}^{1/2}. \tag{16}$$

Observe that this expression is identical for both α_j. Because of the invariance of the Siegel distance on $P_{n+1}(\mathbb{R})$ under the action of $\mathrm{GL}_n(\mathbb{R})$, the critical region obtained, which may be expressed as

$$W_\varepsilon \equiv \left\{ x \in \mathbb{R}^{n\times N}: \frac{b-1}{4}\ln^2|\Sigma_0^{-1}\Sigma^*| + \frac{b}{2}\|\ln(\Sigma_0^{-1}\Sigma^*)\|^2 > A_\varepsilon^2 \right\},$$

is also invariant from the same transformations. The study of the distribution of the statistic is beyond the scope of this paper. We only refer the particular case of the multivariate normal family, see [7] for more details. Defining the statistic:

$$\mathscr{U} = \frac{1}{\sqrt{2}} \|\ln(\Sigma_0^{-1}\Sigma^*)\|,$$

the asymptotic distribution of $N\mathscr{U}^2$ is a χ^2 with $n(n+1)/2$ degrees of freedom. It was also noted in [7] that the test based on the Siegel distance differs from that of the likelihood ratio.

4.2. Sphericity test

Let us now consider the hypothesis testing problem:

$$H_0: \Sigma = \lambda I, \lambda > 0, \quad H_1: \Sigma \neq \lambda I.$$

The parametric space restricted under the null hypothesis will be expressed as $\Theta_{H_0} \equiv \{(\mu,\Sigma) \in \Theta : \exists \lambda > 0 \text{ with } \Sigma = \lambda I\}$. For a fixed λ the minimization problem is a particular case of (15). The nearest point to θ_N^* is $\theta_0 = (\mu^*, \lambda I)$. Expression (16) simplifies to

$$d_j(\theta_N^*, \Theta_{\lambda I}) = \left\{ \frac{(b-1)n + 2b}{4} \ln\lambda \, (n \ln\lambda - 2\ln|\Sigma^*|) + \frac{b-1}{4} \ln^2|\Sigma^*| + \frac{b}{2}\|\ln \Sigma^*\|^2 \right\}^{1/2}.$$

The minimum of the latter expression is obtained when $\lambda = |\Sigma^*|^{1/n}$, and the expression becomes

$$d_j(\theta_N^*, \Theta_{H_0}) = \sqrt{\frac{b}{2}} \, \|\ln |\Sigma^*|^{-1/n}\Sigma^*\|$$

and is also different, for the multivariate normal case, from the likelihood ratio test, the statistic test of which is

$$\lambda = \frac{|\Sigma^*|^{N/2}}{(n^{-1}\operatorname{tr}\Sigma^*)^{Nn/2}},$$

see [9] for more details.

4.3. Testing equality to a fixed mean vector

We can reduce the problem of testing equality with a fixed mean vector μ_0 to the following test:

$$H_0: \mu = 0, \quad H_1: \mu \neq 0.$$

Then, because of the invariance properties of the Siegel distance, the problem can be formulated in terms of minimizing the distance of $f_{\alpha_j,\gamma}(\theta_N^*) = f_{\alpha_j,\gamma}(y,I)$ to $\Theta_{H_0} \equiv \Theta_0 \equiv \{(\mu,\Sigma) \in \Theta\colon\ \mu = 0\}\}$, where $y = (\Sigma^*)^{-1/2}\mu^*$. The shortest geodesic is orthogonal to Θ_0 and, by Lemma 5.4, takes the form

$$\Psi(s) = \beta_0(s) \begin{pmatrix} \Sigma + \beta_1(s)vv' & \beta_2(s)v \\ \beta_2(s)v' & 1 + \beta_3(s) \end{pmatrix},$$

where $\beta_i(s)$ for $i = 0,\ldots,3$ are convenient scalar functions and where v is an arbitrary vector, and the closest point in Θ_0 to $f_{\alpha_j,\gamma}(y,I)$ is

$$f_{\alpha_j,\gamma}(0,\Sigma) = |\Sigma|^{\alpha_j} \begin{pmatrix} \Sigma & 0 \\ 0' & 1 \end{pmatrix}.$$

Renaming $z = \gamma y$, and taking into account the boundary conditions, i.e., $\Psi(0) = f_{\alpha_j,\gamma}(0,\Sigma)$ and $\Psi(\rho) = f_{\alpha_j,\gamma}(y,I)$, v and Σ must be

$$v = \tau z, \quad \Sigma = \phi(I + \xi zz') \quad \text{with } \tau \in \mathbb{R}, \quad \phi \in \mathbb{R}^+, \quad \text{and} \quad \xi > -\frac{1}{\|z\|^2}.$$

Therefore,

$$d_j(f_{\alpha_j,\gamma}(\theta_N^*), \Theta_0) = \inf\left\{ \delta_j(\theta_N^*, \phi, \xi) : \phi \in \mathbb{R}^+, \text{ and } \xi > -\frac{1}{\|z\|^2} \right\},$$

where $\delta_j(\theta_N^*, \phi, \xi) \equiv d(f_{\alpha_j,\gamma}(\theta_N^*), f_{\alpha_j,\gamma}(0, \phi(I + \xi zz')))$ and is equal to

$$\delta_j(\theta_N^*, \phi, \xi) = \sqrt{\frac{b}{2}} \left\| \ln |\phi(I + \xi zz')|^{\alpha_j} \begin{pmatrix} \phi(I + \xi zz') + zz' & -z \\ -z' & 1 \end{pmatrix} \right\|.$$

Computing the eigenvalues of both matrices, and taking into account the definition of α_j in Theorem 3.2, the above expression of $\delta_j(\theta_N^*, \phi, \xi)$ can be simplified to

$$\begin{aligned} \delta_j(\theta_N^*, \phi, \xi)^2 = {} & \frac{b(n-1)}{2} \ln^2\phi + \frac{b-1}{4} \ln^2(\phi^n(1 + \xi\|z\|^2)) \\ & + \frac{b}{2} \sum_{k=1}^{2} \ln^2 \left(1 + \frac{\phi - 1 + \|z\|^2(1 + \phi\xi)}{2} \right. \\ & \left. + (-1)^{k+1} \sqrt{\left(\frac{\phi - 1 + \|z\|^2(1 + \phi\xi)}{2} \right)^2 + \|z\|^2} \right) \end{aligned}$$

and therefore this expression is independent of any of the two α_j chosen to imbed the elliptical family. Moreover, if we let $\omega = 1 + \xi\|z\|^2$, since ω is an arbitrary positive real number, defining

$$\tilde{\delta}_j(\theta_N^*, \phi, \omega)^2 = \frac{b(n-1)}{2} \ln^2 \phi + \frac{b-1}{4} \ln^2(\phi^n \omega) + \frac{b}{4} \ln^2 \omega + b \operatorname{arg\,cosh}^2 \left(\frac{1 + \|z\|^2 + \omega}{2\sqrt{\omega}} \right),$$

we have that

$$d_j(f_{\alpha_j,\gamma}(\theta_N^*),\Theta_0)=\inf\{\tilde{\delta}_j(\theta_N^*,\phi,\omega):\ \phi,\omega\in\mathbb{R}^+\},$$

which clearly is a strictly increasing function on $\|z\|^2$, and so on in $(\mu^*)'(\Sigma^*)^{-1}\mu^*$. Therefore, the test based on the Siegel distance is equivalent to considering a critical region as

$$W_\varepsilon\equiv\{x\in\mathbb{R}^{n\times N}:\ t^2=(\mu^*)'(\Sigma^*)^{-1}\mu^*>A_\varepsilon\},$$

which, in the multivariate normal case, is the T^2-Hotelling test, equivalent to likelihood ratio criteria, see [9].

Appendix A.

Here, we will establish some other differential geometric properties of $f_{\alpha_j,\gamma}(\Theta)$, needed to prove earlier results. Hereafter, α_j will be either α_1 or α_2.

Lemma A.1. *Let N_{Σ_0} be a unitary normal field of $f_{\alpha_j,\gamma}(\Theta_{\Sigma_0})$ at the point $\Psi_{\Sigma_0}=f_{\alpha_j,\gamma}(\mu,\Sigma_0)$, where $\Theta_{\Sigma_0}=\{(\mu,\Sigma)\in\Theta:\Sigma=\Sigma_0\text{ constant}\}$. Then, N_{Σ_0} has the following form*:

$$N_{\Sigma_0}=w_j\begin{pmatrix}U & \gamma\mu\\ \gamma\mu' & 1\end{pmatrix},$$

where $U\in S_n(\mathbb{R})$ is arbitrary, and w_j is

$$w_j=\frac{\pm\sqrt{2}|\Sigma_0|^{\alpha_j}}{\sqrt{b\,(\mathrm{tr}\{(\Sigma_0^{-1}(U-\gamma^2\mu\mu'))^2\}+1)}}.$$

Proof. If N denotes the unitary, up to sign, normal field of the full manifold at $\Psi=f_{\alpha_j,\gamma}(\mu,\Sigma)$, it has to verify:

$$\langle d\Psi,N\rangle=\frac{b}{2}\,\mathrm{tr}\{\Psi^{-1}d\Psi\Psi^{-1}N\}=0, \tag{17}$$

$$\langle N,N\rangle=\frac{b}{2}\,\mathrm{tr}\{(\Psi^{-1}N)^2\}=1. \tag{18}$$

Writing

$$N=\begin{pmatrix}W & t\\ t' & w_j\end{pmatrix}$$

and applying (11) and (13)–(17), we obtain

$$\begin{aligned}\langle d\Psi,N\rangle=\frac{b}{2}|\Sigma|^{-\alpha_j}\{&\mathrm{tr}(\Sigma^{-1}d\Sigma\Sigma^{-1}W)+\gamma(\gamma w_j\mu-2t)'\Sigma^{-1}d\Sigma\Sigma^{-1}\mu\\ &+2\gamma(t-\gamma w_j\mu)'\Sigma^{-1}d\mu\\ &+\alpha_j\,\mathrm{tr}(\Sigma^{-1}d\Sigma)(\mathrm{tr}(\Sigma^{-1}W)+\gamma(\gamma w_j\mu-2t)'\Sigma^{-1}\mu+w_j)\}.\end{aligned} \tag{19}$$

Considering now the particular case of N_{Σ_0} at $\Psi_{\Sigma_0} = f_{\alpha_j,\gamma}(\mu, \Sigma_0)$, substituting $d\Sigma$ by 0 in (19), it simplifies to

$$0 = \langle d\Psi, N_{\Sigma_0} \rangle = \gamma b\,|\Sigma|^{-\alpha_j}(t - \gamma w_j \mu)' \Sigma_0^{-1} d\mu.$$

Therefore, $W \in S_n(\mathbb{R})$ is arbitrary and can be written as $W = w_j U$ with $U \in S_n(\mathbb{R})$ also an arbitrary symmetric matrix, while t must be $t = w_j \gamma \mu$. The unitary norm condition (18) implies

$$\frac{b}{2}|\Sigma_0|^{-2\alpha_j} w_j^2 (\mathrm{tr}\{(\Sigma_0^{-1}(U - \gamma^2 \mu\mu'))^2\} + 1) = 1$$

equality which determines, up to sign, the value of w_j. □

Lemma A.2. *Let N_0 be a unitary normal vector field of $f_{\alpha_j,\gamma}(\Theta_0)$ at the point $\Psi_0 = f_{\alpha_j,\gamma}(0, \Sigma)$ where $\Theta_0 \equiv \{(\mu, \Sigma) \in \Theta : \mu = 0\}$. Then, N_0 has the following form*:

$$N_0 = c_j \begin{pmatrix} -\alpha_j \Sigma & v \\ v' & 1 + n\alpha_j \end{pmatrix},$$

where $v \in M_{n\times 1}(\mathbb{R})$ is arbitrary, and c_j, a normalizing constant, is given by

$$c_j = \frac{\pm 2|\Sigma|^{\alpha_j}}{\sqrt{(n+2)b - n + 4\,b\,v'\Sigma^{-1}v}}.$$

Proof. Compute N_0 with $\mu = d\mu = 0$ and $w_j = c_j(1 + n\alpha_j)$ in (19). Then, simplify the expression to

$$0 = \langle d\Psi, N_0 \rangle = \frac{b}{2}|\Sigma|^{-\alpha_j}\{\mathrm{tr}(\Sigma^{-1} d\Sigma \Sigma^{-1} W) + \alpha_j \,\mathrm{tr}(\Sigma^{-1} d\Sigma)\,(\mathrm{tr}(\Sigma^{-1} W) + c_j(1 + n\alpha_j))\}.$$

We have no restrictions on $t \in M_{n\times 1}$ which can be written as $t = c_j v$ with $v \in M_{n\times 1}$ and since $d\Sigma$ is an arbitrary symmetric matrix, W must be

$$W = -c_j\,\alpha_j\,\Sigma.$$

By the unitary norm condition (18), we obtain

$$\frac{b|\Sigma|^{-2\alpha_j}}{2} c_j^2 (n\alpha_j^2 + (1 + n\alpha_j)^2 + 2v'\Sigma^{-1}v) = 1$$

and c_j is obtained observing that

$$n\alpha_j^2 + (1 + n\alpha_j)^2 = \frac{(n+2)b - n}{2b}. \qquad \square$$

Lemma A.3. *Let* $\Psi_{\Sigma_0} = f_{\alpha_j,\gamma}(\mu, \Sigma_0)$. *Any orthogonal geodesic of* $f_{\alpha_j,\gamma}(\Theta_{\Sigma_0})$ *at* Ψ_{Σ_0} *has the following form*:

$$\Psi(s) = |\Sigma_0|^{\alpha_j} e^{w_j s |\Sigma_0|^{-\alpha_j}} \begin{pmatrix} \Sigma_0 \exp(w_j s |\Sigma_0|^{-\alpha_j} M) + \gamma^2 \mu\mu' & \gamma\mu \\ \gamma\mu' & 1 \end{pmatrix},$$

$$M = \Sigma_0^{-1}(U - \gamma^2 \mu\mu') - I, \quad w_j = \frac{\pm\sqrt{2}|\Sigma_0|^{\alpha_j}}{\sqrt{b\,(\mathrm{tr}\{(\Sigma_0^{-1}(U - \gamma^2\mu\mu'))^2\} + 1)}},$$

where $U \in S_n(\mathbb{R})$ *is an arbitrary symmetric matrix.*

Proof. Combining Lemma 5.1 and expression (9) we obtain

$$H = \Psi^{-1}\dot{\Psi} = w_j|\Sigma_0|^{-\alpha_j} \begin{pmatrix} \Sigma_0^{-1} & -\gamma\Sigma_0^{-1}\mu \\ -\gamma\mu'\Sigma_0^{-1} & 1 + \gamma^2\mu'\Sigma_0^{-1}\mu \end{pmatrix} \begin{pmatrix} U & \gamma\mu \\ \gamma\mu' & 1 \end{pmatrix}$$
$$= w_j|\Sigma_0|^{-\alpha_j} \begin{pmatrix} \Sigma_0^{-1}(U - \gamma^2\mu\mu') & 0 \\ \gamma\mu'(I - \Sigma_0^{-1}(U - \gamma^2\mu\mu')) & 1 \end{pmatrix}.$$

By induction it follows that

$$H^p = w_j^p |\Sigma_0|^{-p\,\alpha_j} \begin{pmatrix} (\Sigma_0^{-1}(U - \gamma^2\mu\mu'))^p & 0 \\ \gamma\mu'(I - (\Sigma_0^{-1}(U - \gamma^2\mu\mu'))^p) & 1 \end{pmatrix}$$

for any $p \in \mathbb{N}$, and the geodesic is therefore obtained by computing

$$\Psi(s) = \Psi_{\Sigma_0} \exp(Hs). \qquad \square$$

Lemma A.4. *Let* $\Psi_0 = f_{\alpha_j,\gamma}(0, \Sigma)$. *Any orthogonal geodesic of* $f_{\alpha_j,\gamma}(\Theta_0)$ *at* Ψ_0 *has the following form*:

$$\Psi(s) = |\Sigma|^{\alpha_j} e^{-sc_j\alpha_j|\Sigma|^{-\alpha_j}} \left\{ \begin{pmatrix} \Sigma & 0 \\ 0' & 1 \end{pmatrix} + \sum_{k=1}^{2} \frac{e^{sc_j\lambda_{kj}|\Sigma|^{-\alpha_j}} - 1}{v'\Sigma^{-1}v + \lambda_{kj}^2} \begin{pmatrix} vv' & \lambda_{kj}v \\ \lambda_{kj}v' & \lambda_{kj}^2 \end{pmatrix} \right\}$$

with $j = 1$ *or* 2 *and where*

$$c_j = \frac{\pm 2|\Sigma|^{\alpha_j}}{\sqrt{(n+2)b - n + 4bv'\Sigma^{-1}v}}, \quad \zeta = \sqrt{1 + (n+1)(b-1)/2b},$$

$v \in M_{n\times 1}(\mathbb{R})$ *is arbitrary and*

$$\lambda_{kj} = \frac{(-1)^j\zeta + (-1)^k\sqrt{\zeta^2 + 4v'\Sigma^{-1}v}}{2}$$

with $k = 1, 2$.

Proof. Let us define

$$P = \begin{pmatrix} \Sigma^{1/2} & 0 \\ 0' & 1 \end{pmatrix} \in M_{(n+1)\times(n+1)}(\mathbb{R}).$$

Combining Lemma 5.2 and expression (9) we obtain

$$H = \Psi^{-1}\dot{\Psi} = c_j|\Sigma|^{-\alpha_j}P^{-1}\begin{pmatrix} -\alpha_j I & \Sigma^{-1/2}v \\ v'\Sigma^{-1/2} & 1+n\alpha_j \end{pmatrix} P.$$

Therefore, since

$$\exp(Hs) = \exp\left(sc_j|\Sigma|^{-\alpha_j}P^{-1}\begin{pmatrix} -\alpha_j I & \Sigma^{-1/2}v \\ v'\Sigma^{-1/2} & 1+n\alpha_j \end{pmatrix} P\right)$$

$$= P^{-1}\exp\left(sc_j|\Sigma|^{-\alpha_j}\begin{pmatrix} -\alpha_j I & \Sigma^{-1/2}v \\ v'\Sigma^{-1/2} & 1+n\alpha_j \end{pmatrix}\right) P$$

with $P \in \mathrm{GL}_n(\mathbb{R})$ and taking into account that the eigenvalues of matrix

$$\begin{pmatrix} -\alpha_j I & \Sigma^{-1/2}v \\ v'\Sigma^{-1/2} & 1+(n+1)\alpha_j \end{pmatrix}$$

are $\lambda_{1j}-\alpha_j$, $\lambda_{2j}-\alpha_j$ and $-\alpha_j$ while their corresponding non-normalized eigenvectors are $(v'\Sigma^{-1/2}, \lambda_{kj})'$, $k = 1, 2$ and $n-1$ mutually ortogonal vectors of the form $(p', 0)'$ with p such that $p'\Sigma^{-1/2}v = 0$, respectively, we have

$$\exp\left(sc_j|\Sigma|^{-\alpha_j}\begin{pmatrix} -\alpha_j I & \Sigma^{-1/2}v \\ v'\Sigma^{-1/2} & 1+n\alpha_j \end{pmatrix}\right)$$

$$= \mathrm{e}^{-sc_j\alpha_j|\Sigma|^{-\alpha_j}}\left\{ I_{n+1} + \sum_{k=1}^{2} \frac{\mathrm{e}^{sc_j\lambda_{kj}|\Sigma|^{-\alpha_j}}-1}{v'\Sigma^{-1}v+\lambda_{kj}^2}\begin{pmatrix} \Sigma^{-1/2}vv'\Sigma^{-1/2} & \lambda_{kj}\Sigma^{-1/2}v \\ \lambda_{kj}v'\Sigma^{-1/2} & \lambda_{kj}^2 \end{pmatrix}\right\}$$

and finally, since $\Psi(s) = \Psi_0 \exp(H)$, we obtain the previously stated formulas. □

References

[1] S. Amari, Differential–Geometrical Methods in Statistics, Lecture notes in Statistics, Vol. 28, Springer, New York, 1985.

[2] C. Atkinson, A.F.S. Mitchell, Rao's distance measure, Sankhya 43 A (1981) 345–365.

[3] M. Berkane, K. Oden, P.M. Bentler, Geodesic estimation in elliptical distributions, J. Multivariate Anal. 63 (1997) 35–46.
[4] J. Burbea, Informative geometry of probability spaces, Exposition. Math. 4 (1986) 347–378.
[5] J. Burbea, C.R. Rao, Entropy differential metric, distance and divergence measures in probability spaces: a unified approach, J. Multivariate Anal. 12 (1982) 575–596.
[6] J. Burbea, J.M. Oller, The information metric for univariate linear elliptic models, Statist. Decisions 6 (1988) 209–221.
[7] M. Calvo, J.M. Oller, A distance between multivariate normal distributions based in an embedding into the Siegel group, J. Multivariate Anal. 35 (2) (1990) 223–242.
[8] M. Calvo, A. Villarroya, J.M. Oller, Non-linear biplots for multivariate normal grouped populations, in: A. Rizzi, M. Vichi, H.H. Bick (Eds.), Advances in Data Science and Classification, Springer, Berlin, 1998, pp. 343–348.
[9] K.T. Fang, Y.T. Zhang, Generalized Multivariate Analysis, Springer, New York, 1990.
[10] J.C. Gower, D.J. Hand, Biplots, Chapman & Hall, London, 1996.
[11] A.T. James, The variance information manifold and the functions on it, in: P.K. Krishnaiah (Ed.), Multivariate Analysis III, Academic Press, New York, 1973, pp. 157–169.
[12] W.J. Krzanowski, Rao's distance between normal populations that have common principal components, Biometrics 52 (1996) 1467–1471.
[13] A.F.S. Mitchell, Statistical manifolds of univariate elliptical distributions, Internat. Statist. Rev. 56 (1988) 464–467.
[14] A.F.S. Mitchell, Information matrix, skewness tensor and α-connections for the general multivariate elliptic distribution, Ann. Inst. Statist. Math. 41 (2) (1989) 289–304.
[15] R.J. Muirhead, Aspects of Multivariate Statistical theory, Wiley, New York, 1982.
[16] J.M. Oller, Some geometrical aspects of data analysis and statistics, in: Y. Dodge (Ed.), Statistical Data Analysis and Inference, Elsevier Science Publishers B.V., North-Holland, Amsterdam, 1989, pp. 41–58.
[17] J.M. Oller, J.M. Corcuera, Intrinsic analysis of the statistical estimation, Ann. Statist. 23 (5) (1995) 1562–1581.
[18] C.R. Rao, Information and accuracy attainable in the estimation of statistical parameters, Bull. Calcutta Math. Soc. 37 (1945) 81–91.
[19] C.L. Siegel, Symplectic Geometry, Academic Press, New York, 1964.