GEODESICS CONNECTED WITH THE FISHER METRIC ON THE MULTIVARIATE NORMAL MANIFOLD

P.S. Eriksen

Institute of Electronic Systems
Department of Mathematics and Computer Science
Aalborg University Centre, Strandvejen 19
DK 9000 Aalborg, Denmark

INTRODUCTION

The Riemannian manifold of multivariate normal distributions equipped with the Fisher information metric has been studied by several people, initiating with Rao (1945). The main emphasis has been put on determining the geodesic curves and the distance between two distributions. For the univariate model, -which is the Poincaré half-plane, - this was solved by Yoshizawa (1971) and Atkinson & Mitchell (1981), and for the p-variate normal with mean zero by James (1973) and Atkinson & Mitchell (1981). In this contribution we give a representation of the geodesic in the general case.

GEODESICS

Let $M = \{(\Sigma,\mu) | \Sigma\in PD(p), \mu\in\mathbb{R}^p\}$ represent the class of p-variate normal distributions, where μ is the mean vector and Σ is the covariance matrix, which is positive definite (PD). The Fisher information defines a Riemannian metric on M. In general, the geodesic curves are solutions of the differential equation

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

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

Since M is isometric to $GA^+(p)/SO(p)$ - $GA^+(p)$ denoting positive affine transformations - we can by a translation argument restrict ourselves to describe the geodesic through $(I_p, 0)$, - and in the direction (B,x), say. By considering the parameterization $\Delta = \Sigma^{-1}$ and $\delta = \Sigma^{-1}\mu$, it is fairly easy to show that the geodesic equation (1), then can be reformulated as

$$\begin{aligned} \dot{\Delta} &= -B\Delta + x\delta^* \\ \dot{\delta} &= -B\delta + (1+\delta^*\Delta^{-1}\delta)x \\ \Delta(0) &= I, \delta(0) = 0 \end{aligned} \qquad (2)$$

Now let

$$A = \begin{Bmatrix} -B & x & 0 \\ x^* & 0 & -x^* \\ 0 & -x & B \end{Bmatrix}$$

and

$$\Lambda = \exp(At) = \sum_{n=0}^{\infty} (At)^n/n! \qquad (3)$$

$$= \begin{Bmatrix} \Delta & \delta & \Phi \\ \delta^* & \varepsilon & \gamma^* \\ \Phi^* & \gamma & \Gamma \end{Bmatrix} \qquad t \in \mathbb{R}$$

Then we have

Theorem The geodesic curve through $(I,0)$ with tangent $(-B,x)$ is given by $(\Delta(t), \delta(t))$, where $(\Delta(t), \delta(t))$ is determined by (3).

Proof It is clear that $\dot{\Lambda} = A\Lambda$, i.e.

$$\begin{Bmatrix} \dot{\Delta} & \dot{\delta} & \dot{\Phi} \\ \dot{\delta}^* & \dot{\varepsilon} & \dot{\gamma}^* \\ \dot{\Phi}^* & \dot{\gamma} & \dot{\Gamma} \end{Bmatrix} = \begin{Bmatrix} -B & x & 0 \\ x^* & 0 & -x^* \\ 0 & -x & B \end{Bmatrix} \begin{Bmatrix} \Delta & \delta & \Phi \\ \delta^* & \varepsilon & \gamma^* \\ \Phi^* & \gamma & \Gamma \end{Bmatrix} \qquad (4)$$

It follows that

$$\dot{\Delta} = -B\Delta + x\delta^*$$

$$\dot{\delta} = -B\delta + \varepsilon x$$

If $\varepsilon = 1+\delta^*\Delta^{-1}\delta$ then (Δ,δ) is seen to be a solution to (2). Since $\varepsilon(0)=1$ it clearly suffices to show that $\dot{\varepsilon}=\dot{\sigma}$, where $\sigma=\delta^*\Delta^{-1}\delta$.

By (4) we have that

$$\dot{\varepsilon} = x^*\delta - x^*\gamma$$

$$\dot{\sigma} = x^*\delta - x^*([\sigma-\varepsilon]\Delta^{-1}\delta+\Phi^*\Delta^{-1}\delta)$$

where it has been utilized that $\dot{\delta}^*=x^*\Delta-x^*\Phi^*$. It follows that $\dot{\varepsilon}=\dot{\sigma}$ if

$$\gamma = [\sigma-\varepsilon]\Delta^{-1}\delta + \Phi^*\Delta^{-1}\delta$$

This is established by the relation $\Lambda\,\Lambda^{-1} = I$, where we observe that $\Lambda^{-1}=\exp(-At)$, i.e.

$$\Lambda^{-1}=\Lambda(-t) = \begin{Bmatrix} \Gamma & \gamma & \Phi^* \\ \gamma^* & \varepsilon & \delta^* \\ \Phi & \delta & \Delta \end{Bmatrix} \qquad (5)$$

The relation $\Lambda\,\Lambda^{-1}=I$ implies that

(i) $\Delta\gamma + \varepsilon\delta + \Phi\delta = 0$

(ii) $\Delta\Phi^* + \delta\delta^* + \Phi\Delta = 0$

Clearly, (ii) is equivalent to $\Phi^*\Delta^{-1} + \Delta^{-1}\delta\delta^*\Delta^{-1} + \Delta^{-1}\Phi = 0$, so that

$$\text{(iii)} \quad \Phi^*\Delta^{-1}\delta + \sigma\Delta^{-1}\delta + \Delta^{-1}\Phi\delta = 0$$

Furthermore, (i) is equivalent to

$$\text{(iv)} \quad \gamma = -\varepsilon\Delta^{-1}\delta - \Delta^{-1}\Phi\delta$$

and it is now obvious from (iii) and (iv) that

$$\gamma = [\sigma-\varepsilon]\,\Delta^{-1}\delta + \Phi^*\Delta^{-1}\delta.$$

This completes the proof.

DISCUSSION

It is interesting to note that (5) means that $(\Delta(-t),\delta(-t)) = (\Gamma(t),\gamma(t))$, i.e. (Γ,γ) is the "point opposite" to (Δ,δ). Besides we have shown that $\varepsilon = 1+\delta^*\Delta^{-1}\delta = 1+\gamma^*\Gamma^{-1}\gamma$. Considering the representation (3) of the solution, there has been no success in seeking for an interpretation of the $p\times p$ matrix Φ.

Further problems:

i) Uniqueness of geodesic connecting two points?

ii) Determination of distance.

iii) Any relation to $SO(p,p+1)$, which is the least group containing the Λ's defined by (3)?

iv) Characterization and statistical interpretation - if any - of totally geodesic submanifolds.

REFERENCES

Atkinson, C. and Mitchell A.F.S. (1981): Rao's distance measure. *Sankhyā*, *43*, 345-365.

James, A.T. (1973): The variance information manifold and the functions on it. *Proc. 3rd Int. Symp. on Mult. Anal.*, Krishnaiah (ed.). Academic Press.

Rao, C.R. (1945): Information and the accuracy attainable in the estimation of statistical parameters. *Bulletin of the Calcutta Math. Soc.*, *37*, 81-91.

Skovgaard, L.T. (1984): A Riemannian geometry of the multivariate normal model. *Scand. J. Statist.*, *11*, 211-223.

Yoshizawa, T. (1971): A geometry of parameter space and its statistical interpretation. Memo TYH-2, Harvard University.