Geometry of Covariance Matrices and Computation of Median

Le Yang*, Marc Arnaudon* and Frédéric Barbaresco†

*Laboratoire de Mathématiques et Applications, CNRS:UMR6086, Université de Poitiers, Téléport 2 - BP 30179, Boulevard Marie et Pierre Curie, 86962 Futuroscope Chasseneuil Cedex, France
†Thalès Air Systems, Unité Surface Radar, Département Stratégie, Technologie et Innovation Hameau de Roussigny, F-91407 Limours, France

Abstract. In this paper, we consider the manifold of covariance matrices of order n parametrized by reflection coefficients which are derived from Levinson's recursion of autoregressive model. The explicit expression of the reparametrization and its inverse are obtained. With the Riemannian metric given by the Hessian of a Kähler potential, we show that the manifold is in fact a Cartan-Hadamard manifold with lower sectional curvature bound -4. The explicit expressions of geodesics are also obtained. After that we introduce the notion of Riemannian median of points lying on a Riemannian manifold and give a simple algorithm to compute it. Finally, some simulation examples are given to illustrate the applications of the median method to radar signal processing.

Keywords: covariance matrices, geometric median, radar target detection
PACS: 02.40.Ky, 84.40.Xb.

INTRODUCTION

Covariance matrices are basic tools in signal processing and there are many ways to study them. A geometric method is to regard them as the manifold $\mathscr{P}_n$ consisting of all the Hermitian (or symmetric) positive definite matrices together with the canonical Riemannian metric [23] which can be derived by applying the method of information geometry to the multivariate normal distributions of zero mean. This metric [10] is also the unique metric which is invariant by the actions of the group $GL(n)$. Having this metric the machinery of Riemannian geometry can be started to investigate various properties of covariance matrices. For example, we can calculate explicitly the geodesics, the Riemannian distance, the curvature tensor and so on. The manifold $\mathscr{P}_n$ is a Cartan-Hadamard manifold having many geometric properties; an interesting property is given by the barycenter or the Fréchet mean of points which gives a notion of centrality of data lying in $\mathscr{P}_n$. The general properties of barycenters are studied in [1], [2], [3], [13] and [18]. The barycenter is also used in [21] as a basic tool to develop statistics on Riemannian manifolds. But one drawback of the barycenter is that it is not robust and it is sensitive to outliers. A good substitute is the geometric median which is an robust estimator for centrality of data points. So that medians are more preferred than barycenters in perturbed environment. We refer to [16] and [25] for some basic materials on this subject.

Stationary signal often arises in the theory of signal processing and in this case the covariance matrices are Toeplitz matrices due to the stationarity of the signal. Then it is easily seen that the manifold $\mathscr{P}_n$ is too large to study the geometry of these covariance matrices and it is necessary to work directly with the manifold $\mathscr{T}_n$ consisting of all the Toeplitz Hermitian positive definite matrices of dimension n which is a submanifold of $\mathscr{P}_n$. For calculation purposes, possibly the first problem is the choice of coordinates. To cope with this problem, [5] [6] [7] [8] introduced the reflection coefficients coordinate furnished by Levinson's recursion of autoregressive model under which the metric is diagonal [12], hence the geodesics and the Riemannian distance can be easily calculated. This gives the possibility of further study of the geometry of $\mathscr{T}_n$.

The aim of this paper is to give some basic geometric facts of the manifold $\mathscr{T}_n$ and their applications. Firstly, we give the explicit expression of the reflection coefficients as well as the inverse of this change of coordinates. Secondly, we show that with the Riemannian metric introduced in [6] this manifold is a Cartan-Hadamard manifold whose sectional curvatures are bounded below by -4. Then we calculate the explicit expressions of geodesics in this manifold. After that, we introduce the notion of Riemannian median of points lying on a Riemannian manifold and give a simple algorithm to compute it. Finally, some numerical simulations are given to illustrate the applications of the median method to radar signal processing.

REFLECTION COEFFICIENTS PARAMETRIZATION

Let $\mathscr{T}_n$ be the set of Toeplitz Hermitian positive definite matrices of order n. It is an open submanifold of $\mathbf{R}^{2n-1}$. Each element $R_n \in \mathscr{T}_n$ can be written as

$$R_n = \begin{bmatrix} r_0 & \overline{r}_1 & \dots & \overline{r}_{n-1} \\ r_1 & r_0 & \dots & \overline{r}_{n-2} \\ \vdots & \ddots & \ddots & \vdots \\ r_{n-1} & \dots & r_1 & r_0 \end{bmatrix}.$$

For every $1 \leq k \leq n-1$, the upper left $(k+1)$-by-$(k+1)$ corner of R_n is denoted by R_k. It is associated to a k-th order autoregressive model whose Yule-Walker equation is

$$\begin{bmatrix} r_0 & \overline{r}_1 & \dots & \overline{r}_k \\ r_1 & r_0 & \dots & \overline{r}_{k-1} \\ \vdots & \ddots & \ddots & \vdots \\ r_k & \dots & r_1 & r_0 \end{bmatrix} \begin{bmatrix} 1 \\ a_1^{(k)} \\ \vdots \\ a_k^{(k)} \end{bmatrix} = \begin{bmatrix} P_k \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

where $a_1^{(k)}, \dots, a_k^{(k)}$ are the optimal prediction coefficients and $P_k = \det R_{k+1} / \det R_k$ is the mean squared error.

The last optimal prediction coefficient $a_k^{(k)}$ is called the k-th reflection coefficient and is denoted by μ_k. It is easily seen that $\mu_1, \dots, \mu_{n-1}$ are uniquely determined by the matrix R_n. Moreover, the classical Levinson's recursion [24] gives that $|\mu_k| < 1$. Hence, by letting $P_0 = r_0$, we obtain a map between two submanifolds of $\mathbf{R}^{2n-1}$:

$$\varphi : \quad \mathscr{T}_n \longrightarrow \mathbf{R}_+^* \times \mathbf{D}^{n-1}, \quad R_n \longmapsto (P_0, \mu_1, \dots, \mu_{n-1}),$$

where $\mathbf{D} = \{z \in \mathbf{C} : |z| < 1\}$ is the unit disc of the complex plane.

This map is a diffeomorphism, so that it is a reparametrization of covariance matrices in terms of reflection coefficients. It will be seen that under these coordinates the Riemannian metric of $\mathscr{T}_n$ has a very simple form and lots of calculations can be simplified. Hence it is necessary to obtain the explicit expressions of φ and its inverse. On the one hand, the closed form of the reflection coefficients can be easily obtained by Cramer's rule:

$$\mu_k = (-1)^k \frac{\det S_k}{\det R_k}, \quad \text{where} \quad S_k = R_{k+1}\begin{pmatrix} 2,\dots,k+1 \\ 1,\dots,k \end{pmatrix}$$

is the submatrix of R_{k+1} obtained by deleting the first row and the last column. On the other hand, if $(P_0, \mu_1, \dots, \mu_{n-1}) \in \mathbf{R}_+^* \times \mathbf{D}^{n-1}$, then by using the method of Schur complement [26], its inverse image R_n under φ can be calculated by the following algorithm:

$$r_0 = P_0, \quad r_1 = -P_0\mu_1,$$

$$r_k = -\mu_k P_{k-1} + \alpha_{k-1}^T J_{k-1} R_{k-1}^{-1} \alpha_{k-1}, \quad 2 \le k \le n-1,$$

where

$$\alpha_{k-1} = \begin{bmatrix} r_1 \\ \vdots \\ r_{k-1} \end{bmatrix}, \quad J_{k-1} = \begin{bmatrix} 0 & \dots & 0 & 1 \\ 0 & \dots & 1 & 0 \\ & \dots & & \\ 1 & \dots & 0 & 0 \end{bmatrix} \quad \text{and} \quad P_{k-1} = P_0 \prod_{i=1}^{k-1} (1 - |\mu_i|^2).$$

RIEMANNIAN METRIC OF $\mathscr{T}_n$

From now on, we regard $\mathscr{T}_n$ as a Riemannian manifold whose metric, which is introduced in [6] by the Hessian of the Kähler potential (see [4] for the definition of a Kähler potential) $\Phi(R_n) = -\ln(\det R_n) - n\ln(\pi e)$, is given by

$$ds^2 = n\frac{dP_0^2}{P_0^2} + \sum_{k=1}^{n-1} (n-k) \frac{|d\mu_k|^2}{(1-|\mu_k|^2)^2},$$

where $(P_0, \mu_1, \dots, \mu_{n-1}) = \varphi(R_n)$. With this metric the space $\mathbf{R}_+^* \times \mathbf{D}^{n-1}$ is just the product of the Riemannian manifolds $(\mathbf{R}_+^*, ds_0^2)$ and $(\mathbf{D}, ds_k^2)_{1 \le k \le n-1}$, where $ds_0^2 = n(dP_0^2/P_0^2)$ and $ds_k^2 = (n-k)|d\mu_k|^2/(1-|\mu_k|^2)^2$. The latter is just $n-k$ times the classical Poincaré metric of $\mathbf{D}$. Hence $(\mathbf{R}_+^* \times \mathbf{D}^{n-1}, ds^2)$ is a Cartan-Hadamard manifold whose sectional curvatures K verify $-4 \le K \le 0$. The Riemannian distance between two different points x and y in $\mathbf{R}_+^* \times \mathbf{D}^{n-1}$ is given by

$$d(x,y) = \left(n\sigma(P,Q)^2 + \sum_{k=1}^{n-1} (n-k)\tau(\mu_k, \nu_k)^2 \right)^{1/2},$$

where $x=(P,\mu_1,\dots,\mu_{n-1})$, $y=(Q,\nu_1,\dots,\nu_{n-1})$,

$$\sigma(P,Q)=|\ln(\frac{Q}{P})| \quad \text{and} \quad \tau(\mu_k,\nu_k)=\frac{1}{2}\ln\frac{1+|\frac{\nu_k-\mu_k}{1-\mu_k\nu_k}|}{1-|\frac{\nu_k-\mu_k}{1-\mu_k\nu_k}|}.$$

GEODESICS IN $\mathscr{T}_n$

The geodesic from x to y in $\mathscr{T}_n$ parametrized by arc length is given by

$$\gamma(s,x,y)=(\gamma_0(\frac{\sigma(P,Q)}{d(x,y)}s),\gamma_1(\frac{\tau(\mu_1,\nu_1)}{d(x,y)}s),\dots,\gamma_1(\frac{\tau(\mu_{n-1},\nu_{n-1})}{d(x,y)}s)),$$

where γ_0 is the geodesic in $(\mathbf{R}_+^*,ds_0^2)$ from P to Q parametrized by arc length and for $1\le k\le n-1$, γ_k is the geodesic in $(\mathbf{D},ds_k^2)$ from μ_k to ν_k parametrized by arc length. More precisely,

$$\gamma_0(t)=Pe^{t\,\mathrm{sign}(Q-P)},$$

and for $1\le k\le n-1$,

$$\gamma_k(t)=\frac{(\mu_k+e^{i\theta_k})e^{2t}+(\mu_k-e^{i\theta_k})}{(1+\bar{\mu}_k e^{i\theta_k})e^{2t}+(1-\bar{\mu}_k e^{i\theta_k})}, \quad \text{where} \quad \theta_k=\arg\frac{\nu_k-\mu_k}{1-\bar{\mu}_k\nu_k}.$$

Particularly,

$$\gamma'(0,x,y)=(\gamma_0'(0)\frac{\sigma(P,Q)}{d(x,y)},\gamma_1'(0)\frac{\tau(\mu_1,\nu_1)}{d(x,y)},\dots,\gamma_{n-1}'(0)\frac{\tau(\mu_{n-1},\nu_{n-1})}{d(x,y)}).$$

Let $v=(v_0,v_1,\dots,v_{n-1})$ be a tangent vector in $T_x(\mathbf{R}_+^*\times\mathbf{D}^{n-1})$, then the geodesic starting from x with velocity v is given by

$$\zeta(t,x,v)=(\zeta_0(t),\zeta_1(t),\dots,\zeta_{n-1}(t)),$$

where ζ_0 is the geodesic in $(\mathbf{R}_+^*,ds_0^2)$ starting from P with velocity v_0 and for $1\le k\le n-1$, ζ_k is the geodesic in $(\mathbf{D},ds_k^2)$ starting from μ_k with velocity v_k. More precisely,

$$\zeta_0(t)=Pe^{\frac{v_0}{P}t},$$

and for $1\le k\le n-1$,

$$\zeta_k(t)=\frac{(\mu_k+e^{i\theta_k})e^{\frac{2|v_k|t}{1-|\mu_k|^2}}+(\mu_k-e^{i\theta_k})}{(1+\bar{\mu}_k e^{i\theta_k})e^{\frac{2|v_k|t}{1-|\mu_k|^2}}+(1-\bar{\mu}_k e^{i\theta_k})}, \quad \text{where} \quad \theta_k=\arg v_k.$$

COMPUTING RIEMANNIAN MEDIAN

In this section, we give some basic facts about the Riemannian median which are needed to estimate Doppler ambience and we refer to [25] for more details of mathematical treatment of this subject.

Let M be a complete Riemannian manifold with Riemannian distance d. We consider some different points $p_1,\dots,p_N$ in M, which are contained in an open ball $B(a,\rho)$ centered at a with a finite radius ρ. Let δ and Δ be respectively a lower and an upper bound of sectional curvatures in $\bar{B}(a,\rho)$. The injectivity radius of $\bar{B}(a,\rho)$ is denoted by $\mathrm{inj}(\bar{B}(a,\rho))$. Furthermore, we assume that the radius of the ball verifies

$$\rho < \min\Big\{ \frac{\pi}{4\sqrt{\Delta}}, \frac{\mathrm{inj}\,(\bar{B}(a,\rho))}{2} \Big\},$$

where if $\Delta \leq 0$, then $\pi/(4\sqrt{\Delta})$ is interpreted as $+\infty$.

If the points $p_1,\dots,p_N$ are not contained in a single geodesic, then the function

$$f: \qquad \bar{B}(a,\rho) \longrightarrow \mathbf{R}_+, \qquad x \longmapsto \frac{1}{N}\sum_{i=1}^{N} d(x,p_i)$$

has a unique minimum point $\mathbf{m}$, which is called the Riemannian median of $p_1,\dots,p_N$. Furthermore, $\mathbf{m}$ is characterized by the condition that for every $x \in \bar{B}(a,\rho)$,

$$x = \mathbf{m} \Longleftrightarrow \begin{cases} |H(x)| \leq 1/N, & \text{if } x \in \{p_1,\dots,p_N\}; \\ H(x) = 0, & \text{otherwise}, \end{cases}$$

where $H(x)$ is a tangent vector of M at x defined by

$$H(x) = \frac{1}{N}\sum_{\substack{1\leq i\leq N \\ p_i \neq x}} \frac{-\exp_x^{-1} p_i}{d(x,p_i)}.$$

As a particular case of the above characterization, we obtain that

$$H(x) = 0 \Longrightarrow x = \mathbf{m}.$$

In order to compute Riemannian medians in practical cases, we use the following subgradient algorithm:

$$\textbf{Initializing} \qquad x_1 \in \bar{B}(a,\rho),$$

$$\textbf{Do} \qquad x_{k+1} = \exp_{x_k}\Big(-t_k \frac{H(x_k)}{|H(x_k)|}\Big),$$

$$\textbf{While} \qquad H(x_k) \neq 0.$$

To ensure the convergence of the above algorithm to the Riemannian median $\mathbf{m}$ of points $p_1,\dots,p_N$, we can choose the stepsizes $(t_k)_k$ in a small interval $(0,\varepsilon)$ and such that

$$\lim_{k\to\infty} t_k = 0 \quad \text{and} \quad \sum_{k=1}^{\infty} t_k = +\infty.$$

Moreover, the constant ε can be computed explicitly as long as δ and Δ are known. For example, in the case of the manifold $\mathscr{T}_n$ of covariance matrices, we know that $\delta = -4$ and $\Delta = 0$, so one may choose $\varepsilon = 1/4$.

APPLICATIONS TO RADAR TARGET DETECTION

We give some simulated examples of the median method applied to radar target detection. Since the autoregressive spectra are closely related to the speed of targets, we shall first investigate the spectral performance of the median method. In order to illustrate the basic idea, we only consider the detection of one fixed direction. The range along this direction is subdivided into 200 lattices in which we add two targets, the echo of each lattice is modeled by an autoregressive process. The following Figure 1 gives the initial spectra of the simulation, where x axis represents the lattices and y axis represents frequencies. Every lattice is identified with a 1×8 vector of reflection coefficients which is calculated by using the regularized Burg algorithm to the original simulating data. The spectra are represented by different colors whose corresponding values are indicated in the colormetric on the right.

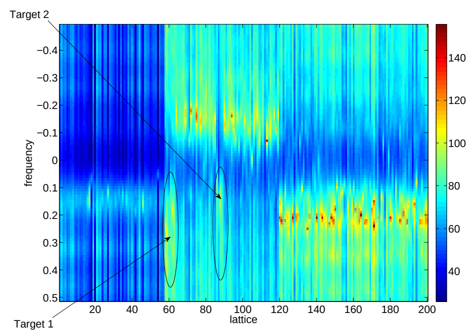

FIGURE 1. Initial spectra with two added targets

For every lattice, by using the preceding algorithm, we calculate the median of the window centered on it and consisting of 15 lattices and then we get the spectra of medians shown in Figure 2. Furthermore, by comparing it with Figure 3 which are spectra of Fréchet means, we see that in the middle of the mean spectra, just in the place where the second target appears, there is an obvious distortion. This explains that median is much more robust than mean when outliers come.

The principle of target detection is that a target appears in a lattice if the distance between this lattice and the median of the window around it is much bigger than that of the ambient lattices. The following Figure 4 shows that the two added targets are well detected by the median method, where x axis represents lattice and y axis represents the distance in $\mathscr{T}_8$ between each lattice and the median of the window around it.

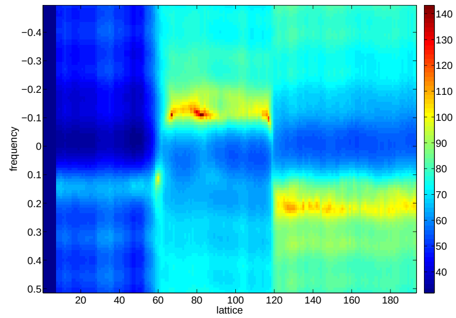

FIGURE 2. Median spectra

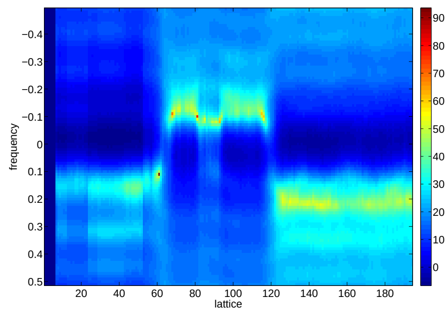

FIGURE 3. Mean spectra

ACKNOWLEDGMENTS

The authors would like to thank Thalès Air Systems for their financial supports. The first author would like to thank Guillaume Bouyt and Nicolas Charon for their generosity of providing basic simulation programs and many helpful discussions.

REFERENCES

1. M. Arnaudon, Espérances conditionnelles et C-martingales dans les variétés. *Séminaire de Probabilités*-XXVIII. 300-311. Lecture Notes in Maths. 1583. Springer. Berlin. 1994
2. M. Arnaudon, Barycentres convexes et approximations des martingales dans les variétés. *Séminaire de Probabilités*-XXIX. 70-85. Lecture Notes in Maths. 1613. Springer. Berlin. 1995
3. M. Arnaudon and X. M. Li, Barycenters of measures transported by stochastic flows. *Annals of probability*. vol 33. no. 4. 2005. pp 1509-1543.
4. W. Ballmann, Lectures on Kähler manifolds. ESI Lectures in Mathematics and Physics. European Math. Soc., Zürich. 2006
5. F. Barbaresco, Innovative Tools for Radar Signal Processing Based on Cartan's Geometry of SPD Matrices and Information Geometry. IEEE International Radar Conference, 2008.
6. F. Barbaresco, Interactions between Symmetric Cone and Information Geometries, ETVC'08, Springer Lecture Notes in Computer Science, vol.5416, pp. 124-163, 2009
7. F. Barbaresco and G. Bouyt, Espace Riemannien symĺętrique et géométrie des espaces de matrices de covariance : équations de diffusion et calculs de médianes, GRETSI'09 conference, Dijon, September

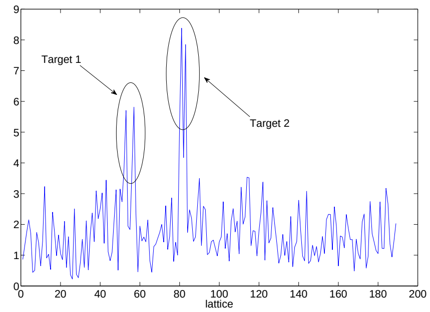

FIGURE 4. Detection by median

2009

8. F. Barbaresco, New Foundation of Radar Doppler Signal Processing based on Advanced Differential Geometry of Symmetric Spaces: Doppler Matrix CFAR and Radar Application, Radar'09 Conference, Bordeaux, October 2009
9. F. Barbaresco, Annalyse Doppler: régularisation d'un problème inverse mal posé. Support de cours
10. R. Bhatia, Positive definite matrices. Princeton university press. 2007
11. G. Bouyt, Algorithmes innovants pour la détection de cible radar. Rapport de travail d'option de Thalès, 2009
12. N. Charon, Une nouvelle approche pour la détection de cibles dans les images radar. Rapport du stage de Thalès, 2008
13. M. Emery and G. Mokobodzki, Sur le barycentre d'une probabilité dans une variété. *Séminaire de Probabilités (Strasbourg)* tome 25 (1991) p.220-233.
14. S. Fiori, Learning the Fréchet mea over the manifold of symmetric positive-definite matrices. Cogn Comput(2009)1:279-291
15. S. Fiori and T. Tanaka, An algorithm to compute averages on matrix Lie groups. IEEE Transactions on signal processing, vol. 57, No. 12, Decembre 2009.
16. P. Fletcher et al., The geometric median on riemannian manifolds with applications to robust atlas estimation. NeuroImage 45 (2009) S143-S152.
17. I. S. Iohvidov, Hankel and Toeplitz matrices and forms - Algebraic theory *Birkhäuser* 1982
18. W. S. Kendall, Probability, convexity, and harmonic maps with small image I: uniqueness and fine existence. Proc. London Math. Soc. (3) 61 (1990) 371-406
19. S. G. Krantz, Complex analysis: The geometric view point. The mathematical association of America 1990
20. B. N. Mukherjee and S. S. Maiti, On some properties of positive definite Toeplitz Matrices and their possible applications. *Linear algebra and its applications* 102: 211-240, 1988
21. X. Pennec, Intrinsic statistics on riemannian manifolds: Basic tools for geometric measurements. *Journal of Mathematical Imaging and Vision* 2006
22. A. Sahib, Espérance d'une variable aléatoire à valeur dans un espace métrique. Thèse de l'Université de Rouen (1998).
23. L. T. Skovgaard, A riemannian geometry of the multivariate normal model. *Scand J Statist* 11: 211-223. 1984
24. P. P. Vaidyanathan, The Theory of Linear Prediction. Morgan and Claypool. 2008
25. L. Yang, Riemannian median and its estimation. arXiv: 0911. 3474
26. F. Z. Zhang, The Schur complements and its applications. Springer. 2005