

Kernel density estimation on the Siegel space applied to radar processing

Emmanuel Chevallier, Frédéric Barbaresco, Jesus Angulo

▶ **To cite this version:**

Emmanuel Chevallier, Frédéric Barbaresco, Jesus Angulo. Kernel density estimation on the Siegel space applied to radar processing. 2016. hal-01344910v2

HAL Id: hal-01344910
https://hal.science/hal-01344910v2

Preprint submitted on 13 Jul 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Kernel density estimation on the Siegel space applied to radar processing

Emmanuel Chevallier[1], Frédéric Barbaresco[2], Jesús Angulo[3]

1: Computer science and applied math department, Weizmann institute of science ; Israel
2: Thales Air Systems; France
3: CMM-Centre de Morphologie Mathématique, MINES ParisTech; France
emmanuelchevallier1@gmail.com

July 13, 2016

Abstract

Main techniques of probability density estimation on Riemannian manifolds are reviewed in the case of the Siegel space. For computational reasons we chose to focus on the kernel density estimation. The main result of the paper is the expression of Pelletier's kernel density estimator. The method is applied to density estimation of reflection coefficients from radar observations.

Keywords: Kernel density estimation; Siegel space; Symmetric space; Radar signals

1 Introduction

Probability density estimation is a vast topic. There exists various standard techniques in the Euclidean context, such as histograms, kernel methods, or the characteristic function approach. These methods can sometimes be transposed to the case of Riemannian manifolds. However, the transposition often introduces additional computational efforts. This additional effort depends on the method used and the nature of the manifold. The Siegel space is a generalization of the hyperbolic space. It has a structure of symmetric Riemannian manifold, which enables to adapt different density estimation methods at a reasonable cost. Convergence rates of the density estimation using kernels and orthogonal series were progressively generalized to Riemannian manifolds, see [2][1]. The Siegel space appears in radar processing

in the study of Toeplitz block Toeplitz matrices, whose blocks represent covariance matrices of the signal, see [5, 6, 7]. Information geometry is now a standard framework in radar processing, see [4, 5, 6, 7, 8, 9, 10]. The information geometry on positive definite Teoplitz block Teoplitz matrices is directly related to the metric on the Siegel space, see [11]. Indeed these Toeplitz block Toeplitz matrices can be represented by a symmetric positive definite matrix and a set of coefficients laying in the Siegel disk. The metric considered on Toeplitz block Toeplitz matrices is a product metric between a metric on symmetric positive definite matrices and the Siegel disk metric, see [11]. One already encounters the problem of density estimation in the hyperbolic space for electrical impedance [3], networks [16] and radar signals [17]. In [12] was proposed a generalization of the Gaussian law on the hyperbolic space. Apart from [13], where authors propose a generalization of the Gaussian law, probability density estimation on the Siegel space has not yet been addressed. The contributions of the paper are the following. We review the main non parametric density estimation techniques on the Siegel disk. We provide some rather simple explicit expressions of the kernels proposed by Pelletier in [1]. These expressions makes the kernel density estimation the most adapted method. We present visual results of estimated densities in the simple case where the Siegel disk correspond to the Poincaré disk. The paper begins with an introduction to the Siegel space in Section 2. Section 3 reviews the main non-parametric density estimation techniques on the Siegel space. Section 4 presents an application to radar data estimation.

2 The Siegel space

This section presents facts about the Siegel space. The interested reader can find more details in [20, 18]

2.1 The Siegel upper half space

The Siegel upper half space is a generalization of the Poincaré upper half space. Let $Sym(n)$ be the space of real symmetric matrices of size $n \times n$ and $Sym_+(n)$ the set of real symmetric positive definite matrices of size $n \times n$. The Siegel upper half space is defined by

$$\mathcal{H}_n = \{Z = X + iY | X \in Sym(n), Y \in Sym_+(n)\}$$

$\mathcal{H}_n$ is equipped with the following metric

$$ds = 2.tr(Y^{-1}dZY^{-1}d\overline{Z})$$

The set of real symplectic matrices $Sp(n,\mathbb{R})$ is defined by

$$g \in Sp(n,\mathbb{R}) \Leftrightarrow g^t J g = J$$

where

$$J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$$

and I_n is the $n \times n$ identity matrix. $Sp(n,\mathbb{R})$ is a subgroup of $SL_{2n}(\mathbb{R})$, the set of $2n \times 2n$ invertible matrices of determinant 1. Let $g = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in Sp(n,\mathbb{R})$. The metric ds is invariant under the following action of $Sp(n,\mathbb{R})$,

$$g.Z = (AZ + B)(CZ + D)^{-1}.$$

This action is transitive, i.e.

$$\forall Z \in \mathcal{H}_n, \exists g \in Sp(n,\mathbb{R}), g.iI = Z.$$

The stabilizer K of iI is the set of elements g of $Sp(n,\mathbb{R})$ whose action leaves iI fixed. K is a subgroup of $Sp(n,\mathbb{R})$ called the isotropy group. We can verify that

$$K = \left\{ \begin{pmatrix} A & B \\ -B & A \end{pmatrix}, A + iB \in SU(n) \right\}.$$

A symmetric space is a Riemannian manifold where the reversal of the geodesics is well defined and is an isometry. Formally, $exp_p(u) \mapsto exp_p(-u)$ is an isometry for each p on the manifold, where u is a vector in the tangent space at p, and exp_p the Riemannian exponential application at p. In other words, the symmetry around each point is an isometry. $\mathcal{H}_n$ is a symmetric space, see [18]. The structure of a symmetric space can be studied through its isometry group and the Lie algebra of its isometry group. The present work will make use of the Cartan and Iwasawa decompositions of the Lie algebra of $Sp(n,\mathbb{R})$. Let $\mathfrak{sp}(n,\mathbb{R})$ be the Lie algebra of $Sp(n,\mathbb{R})$. Given A, B and C three real n by n matrices, let us write $\begin{pmatrix} A & B \\ C & -A^t \end{pmatrix} = (A,B,C)$. We have that

$$\mathfrak{sp}(n,\mathbb{R}) = \{(A,B,C) | B \text{ and } C \text{ symmetric}\}$$

The Cartan decomposition of $\mathfrak{sp}(n,\mathbb{R})$ is given by

$$\mathfrak{sp}(n,\mathbb{R}) = \mathfrak{t} \oplus \mathfrak{p}$$

where

$$\mathfrak{t} = \{(A,B,-B) | B \text{ symmetric and } A \text{ skew-symmetric}\}$$

$$\mathfrak{p} = \{(A, B, B) | A, B, \text{ symmetric}\} \tag{1}$$

The Iwasawa decomposition is given by

$$\mathfrak{sp}(n, \mathbb{R}) = \mathfrak{t} \oplus \mathfrak{a} \oplus \mathfrak{n}$$

where

$$\mathfrak{a} = \{(H, 0, 0) | H \text{ diagonal}\}$$

$$\mathfrak{n} = \{(A, B, 0) | A \text{ upper triangular with 0 on the diagonal }, B \text{ symmetric}\}$$

It can be shown that

$$\mathfrak{p} = \cup_{k \in K} Ad_k(\mathfrak{a}) \tag{2}$$

where Ad is the adjoint representation of $Sp(n, \mathbb{R})$.

2.2 The Siegel disk

The Siegel disk $\mathcal{D}_n$ is the set of complex matrices $\{Z | I - Z^* Z \geq 0\}$ where $\geq$ stands for the Loewner order. Recall that for A and B to Hermitian matrices, $A \geq B$ according to the Loewner order means that $A - B$ is positive definite. The transformation

$$Z \in \mathcal{H}_n \mapsto (Z - iI)(Z + iI)^{-1} \in \mathcal{D}_n$$

is an isometry between the Siegel upper half space and the Siegel disk. Let $C = \begin{pmatrix} I & -iI \\ I & iI \end{pmatrix}$. The application $g \in Sp(n, \mathbb{R}) \mapsto CgC^{-1}$ identifies the set of isometries of $\mathcal{H}_n$ and of $\mathcal{D}_n$. Thus, it can be shown that a matrix $g = \begin{pmatrix} A & B \\ \overline{A} & \overline{B} \end{pmatrix} \in Sp(n, \mathbb{C})$ acts isometrically on $\mathcal{D}_n$ by

$$g.Z = (AZ + B)(\overline{A}Z + \overline{B})^{-1}$$

where $\overline{A}$ stands for the conjugate of A. The point iI in $\mathcal{H}_n$ is identified with the null matrix noted 0 in $\mathcal{D}_n$. Let $Z \in \mathcal{D}_n$. There exists P a diagonal matrix with decreasing positive real entries and U a unitary matrix such that $Z = UPU^t$. Let $\tau_1 \geq ... \geq \tau_n$ be such that

$$P = \begin{pmatrix} th(\tau_1) & & \\ & \ddots & \\ & & th(\tau_n) \end{pmatrix}.$$

Let

$$A_0 = \begin{pmatrix} ch(\tau_1) & & \\ & \ddots & \\ & & ch(\tau_n) \end{pmatrix}, B_0 = \begin{pmatrix} sh(\tau_1) & & \\ & \ddots & \\ & & sh(\tau_n) \end{pmatrix}$$

and

$$g_Z = \begin{pmatrix} U & 0 \\ 0 & \overline{U} \end{pmatrix} \cdot \begin{pmatrix} A_0 & B_0 \\ A_0 & B_0 \end{pmatrix}$$

.

It can be shown that

$$g_Z \in Sp(n, \mathbb{C}) \text{ and } g_Z.0 = Z \tag{3}$$

We provide now a correspondence between the elements of $\mathcal{D}_n$ and the elements of $\mathfrak{p}$ defined in Eq. 1. Let

$$H_Z = \begin{pmatrix} \tau_1 & & & & & & \\ & \ddots & & & & & \\ & & \tau_n & & & & \\ & & & -\tau_1 & & & \\ & & & & \ddots & & \\ & & & & & -\tau_n \end{pmatrix} \in \mathfrak{a} \tag{4}$$

and

$$a_Z = \begin{pmatrix} e^{\tau_1} & & & & & & \\ & \ddots & & & & & \\ & & e^{\tau_n} & & & & \\ & & & e^{-\tau_1} & & & \\ & & & & \ddots & & \\ & & & & & e^{-\tau_n} \end{pmatrix} \in A = exp(\mathfrak{a}).$$

It can be shown that there exists $k \in K$ such that

$$Cexp(Ad_k(H_Z))C^{-1}.0 = Z$$

or equivalently

$$Cka_Z kC^{-1}.0 = Z.$$

Recall that Eq. 2 gives $Ad_k(H) \in \mathfrak{p}$ and $kak \in exp(\mathfrak{p})$. The distance between Z and 0 in $\mathcal{D}_n$ is given by

$$d(0, Z) = \left(2 \sum \tau_i^2\right)^{1/2}, \tag{5}$$

see [18] page 292.

!!Negative curvature of the Siegle space!!

3 Non parametric density estimation on the Siegel space

Let Ω be a space, endowed with a σ-algebra and a probability measure p. Let X be a random variable $\Omega \to \mathcal{D}_n$. The Riemannian measure of $\mathcal{D}_n$ is called vol and the measure on $\mathcal{D}_n$ induced by X is noted μ_X. We assume that μ_X has a density, noted f, with respect to vol, and that the support of X is a compact set noted $Supp$. Let $(x_1, .., x_k) \in \mathcal{D}_n^k$ be a set of draws of X.

The Dirac measure in point a is defined as:

$$\delta_a(U) : \begin{cases} 1 \text{ if } a \in U \\ 0 \text{ if } a \in \complement U \end{cases}$$

Let $\mu_k = \frac{1}{k} \sum_{i=1}^{k} \delta_{x_i}$ denotes the empirical measure of the set of draws. This section presents four non-parametric techniques of estimation of f from the set of draws $(x_1, .., x_k)$. The estimated density at x in $\mathcal{D}_n$ is noted $\hat{f}_k(x) = \hat{f}(x, x_1, ..., x_k)$. The relevance of a density estimation technique depends on several aspects. When the space allows it, the estimation technique should equally consider each directions and locations. This leads to an isotropy and a homogeneity condition. In the kernel method for instance, a kernel function K_{x_i} is placed at each observation x_i. Firstly, in order to treat directions equally, the function K_{x_i} should be invariant under the isotropy group of x_i. Secondly, for another observation x_j, functions K_{x_i} and K_{x_j} should be similar up to the isometries that send x_i on x_j. These considerations strongly depend on the geometry of the space: if the space is not homogeneous and the isotropy group is empty, these indifference principles have no meaning. Since the Siegel space is symmetric, it is homogeneous and has a non empty isotropy group. Thus the density estimation technique should be chosen accordingly. The convergence of the different estimation techniques is widely studied. Results were first obtained in the Euclidean case, and are gradually extended to the probability densities on manifold, see [2, 1, 3]. The last relevant aspect, is computational. Each estimation technique has its own computational framework, which presents pro and cons given the different applications. For instance, the estimation by orthogonal series presents an initial pre-processing, but provides a fast evaluation of the estimated density in compact manifolds.

3.1 Histograms

The histogram is the simplest density estimation method. Given a partition of the space $\mathcal{D}_n = \cup_i A_i$, the estimated density is given by

$$\hat{f}(x \in A_i) = \frac{1}{vol(A_i)} \sum_{j=1}^{k} 1_{A_i}(x_j)$$

where 1_{A_i} stands for the indicator function of A_i. Following the considerations of the previous sections, the elements of the partition should firstly be as isotropic as possible, and secondly as similar as possible to each other. Regarding the problem of histograms, the case of the Siegel space is similar to the case of the hyperbolic space. There exist various uniform polygonal tilings on the Siegel space that could be used to compute histograms. However, there are ratio $\lambda \in \mathbb{R}$ for which there is no homothety. Thus it is not always possible to scale the size of the bins to a given set of draws of the random variable. Modifying the scale of the density estimation requires then a change of the structure of the tiling. Thus the study of histograms has not been deepened.

3.2 Orthogonal series

The estimation of the density f can be made out of the estimation of the scalar product between f and a set of orthonormal functions $\{e_j\}$. The most standard choice for $\{e_j\}$ is the eigenfunctions of the Laplacian. When the variable X takes its values in $\mathbb{R}^n$, this estimation technique becomes the characteristic function method. When the underlying space is compact, the spectrum of the Laplacian operator is countable, while when the space is non-compact, the spectrum is uncountable. In the first case, the estimation of the density f is made through the estimation of a sum, while in the second case is made through the estimation an integral. In practice, the second situation present a larger computational complexity. Unfortunately, eigenfunctions of the Laplacian operator are known on $\mathcal{D}_n$ but not on compact sub-domains. For this reason the study of this method has not been deepened.

3.3 Kernels

Let $\mathcal{K} : \mathbb{R}_+ \to \mathbb{R}_+$ be a map which verifies the following properties:
i) $\int_{\mathbb{R}^d} \mathcal{K}(||x||)dx = 1$,
ii) $\int_{\mathbb{R}^d} x\mathcal{K}(||x||)dx = 0$,
iii) $\mathcal{K}(x > 1) = 0$,
iv) $sup(\mathcal{K}(x)) = \mathcal{K}(0)$.

Let $p \in \mathcal{D}_n$. Generally, given a point p on a Riemannian manifold, exp_p defines an injective application only on a neighborhood of 0. The Siegel space is a non-compact symmetric space and has thus only negative sectional curvatures. Thus, exp_p is injective on the whole space. When the tangent space $T_p\mathcal{D}_n$ is endowed with the local scalar product,

$$||u|| = d(p, exp_p(u))$$

where $||.||$ is the Euclidean distance associated to the local scalar product and $d(.,.)$ is the Riemannian distance. The corresponding Lebesgue measure of $T_p\mathcal{D}_n$ is noted Leb_p. Let $exp_p^*(Leb_p)$ denote the push-forward measure of Lep_p by exp_p. The function θ_p defined by:

$$\theta_p : q \mapsto \theta_p(q) = \frac{dvol}{dexp_p^*(Leb_p)}(q), \tag{6}$$

is the density of the Riemannian measure of $\mathcal{D}_n$ with respect to the Lebesgue measure Leb_p after identification of $\mathcal{D}_n$ and $T_p\mathcal{D}_n$ induced by exp_p, see Fig.1.

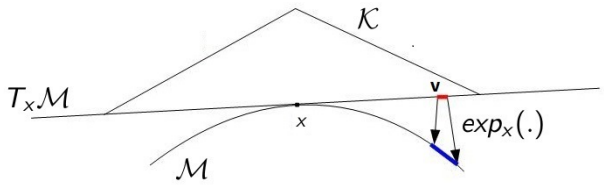


Figure 1: $\mathcal{M}$ is a Riemannian manifold, $T_x\mathcal{M}$ is its tangent space at x. The exponential application induces a volume change θ_x between $T_x\mathcal{M}$ and $\mathcal{M}$.

Given $\mathcal{K}$ and a positive radius r, the estimator of f proposed by [1] is defined by:

$$\hat{f}_k = \frac{1}{k}\sum_i \frac{1}{r^n}\frac{1}{\theta_{x_i}(x)}\mathcal{K}\left(\frac{d(x,x_i)}{r}\right). \tag{7}$$

The corrective factor $\theta_{x_i}(x)^{-1}$ is necessary since the kernel $\mathcal{K}$ originally integrates to one with respect to the Lebesgue measure, instead of the Riemannian measure. It can be noted that this estimator is the usual kernel estimator in the case of Euclidean space. When the curvature of the space is negative, which is the case of the Siegel space, the distribution placed over each sample x_i has x_i as intrinsic mean. The following theorem provides convergence rate of the estimator. It is a direct adaptation of theorem 3.1 of [1].

Theorem 3.1. *Let $(\mathcal{M}, g)$ be a Riemannian manifold of dimension n and μ its Riemannian volume measure. Let X be a random variable taking its values in a compact subset C of $(\mathcal{M}, g)$. Let $0 < r \leq r_{inj}$, where r_{inj} is the infimum of the injectivity radius on C. Assume the law of X has a twice differentiable density f with respect to the Riemannian volume measure. Let $\hat{f}_k$ be the estimator defined in Eq. 7. The exist a constant C_f such that*

$$\int_{x\in\mathcal{M}} E_{x_1,\ldots,x_k}[(f(x) - \hat{f}_k(x))^2]d\mu \leq C_f(\frac{1}{kr^n} + r^4) \tag{8}$$

If $r \sim k^{\frac{-1}{n+4}}$,

$$\int_{x\in\mathcal{M}} E_{x_1,...,x_k}[(f(x) - \hat{f}_k(x))^2]d\mu = O(k^{-\frac{4}{n+4}}) \tag{9}$$

Proof. See appendix A. □

Since the Siegel space has negative sectional curvatures, $r_{inj} = +\infty$. It can be easily verified that for an isometry α we have:

$$\hat{f}_k(x, x_1, ..., x_k) = \hat{f}_k(\alpha(x), \alpha(x_1), ..., \alpha(x_k)).$$

Each location and direction are processed as similarly as possible.

In order to obtain the explicit expression of the estimator one must first have the explicit expression of the Riemannian exponential, its inverse, and of the function θ_p, see Eq. 6-7. These expressions are difficult and sometimes impossible to obtain on general Riemannian manifolds. In the case of the Siegel space, the symmetric structure makes the computation possible. Since the space is homogeneous, the computation can be made at the origin $iI \in \mathcal{H}_n$ or $0 \in \mathcal{D}_n$ and transported to the whole space. In the present work, the random variable lays in $\mathcal{D}_n$. However in the literature the Cartan and Iwasawa decompositions are usually given for the isometry group of $\mathcal{H}_n$. Thus our computation starts in $\mathcal{H}_n$ before moving to $\mathcal{D}_n$.

The Killing form on the Lie algebra $\mathfrak{sp}(n, \mathbb{R})$ of the isometry group of $\mathcal{H}_n$ induces a scalar product on $\mathfrak{p}$. This scalar product can be transported on $exp(\mathfrak{p})$ by left multiplication. This operation gives $exp(\mathfrak{p})$ a Riemannian structure. It can be shown that on this Riemannian manifold, the Riemannian exponential at the identity coincides with the group exponential. Furthermore,

$$\begin{array}{rccc} \phi & : & exp(\mathfrak{p}) & \rightarrow & \mathcal{H}_n \\ & & g & \mapsto & g.iI \end{array} \tag{10}$$

is a bijective isometry, up to a scaling factor. Since the volume change θ_p is invariant under rescaling of the metric, this scaling factor has no impact. Thus $\mathcal{H}_n$ can be identified with $exp(\mathfrak{p})$ and $exp_{iI\in\mathcal{H}_n}$ with $exp_{|\mathfrak{p}}$. The expression of the Riemannian exponential is difficult to obtain in general, however it boils down to the group exponential in the case of symmetric spaces. This is the main element of the computation of θ_p. The Riemannian volume measure on $exp(\mathfrak{p})$ is noted vol'. Let

$$\begin{array}{rccc} \psi & : & K \times \mathfrak{a} & \rightarrow & \mathfrak{p} \\ & & (k, H) & \mapsto & Ad_k(H) \end{array}$$

Let $\mathfrak{a}^+$ be the diagonal matrices with strictly decreasing positive eigenvalues. Let Λ^+ be the set of positive roots of $\mathfrak{sp}(n,\mathbb{R})$ as described in [18] page 282,

$$\Lambda^+ = \{e_i + e_j, i \leq j\} \cup \{e_i - e_j, i < j\}$$

where $e_i(H)$ is the i-th diagonal term of the diagonal matrix H. Let $C_c(E)$ be the set of continuous compactly supported function on the space E. In [19] is given page 73 that for all $t \in C_c(\mathfrak{p})$, there exists $c_1 > 0$ such that

$$\int_{\mathfrak{p}} t(Y)dY = c_1 \int_K \int_{\mathfrak{a}^+} t(\psi(k,H)) \prod_{\lambda \in \Lambda^+} \lambda(H) dk dH, \tag{11}$$

where dY is a Lebesgue measure on the coefficients of Y. Let $\tilde{\mathfrak{p}} = \psi(K \times \mathfrak{a}^+)$. $\lambda \in \Lambda^+$ never vanishes on $\mathfrak{a}^+$ and $\mathfrak{p} \setminus \tilde{\mathfrak{p}}$ has a null measure. Thus

$$\int_{\tilde{\mathfrak{p}}} t(Y) \prod_{\lambda \in \Lambda^+} \frac{1}{\lambda(H_Y)} dY = c_1 \int_K \int_{\mathfrak{a}^+} t(Ad_k(H)) dk dH, \tag{12}$$

where H_Y is the point in $\mathfrak{a}^+$ such that there exists k in K such that $\psi(k, H_Y) = Y$. Calculation in [19] page 73 also gives that for all $t \in C_c(\mathfrak{p})$, there exists $c_2 > 0$ such that

$$\int_{Sp(n,\mathbb{R})} t(g)dg = c_2 \int_K \int_{\mathfrak{a}^+} \int_K t(k_2.exp(Ad_{k_1}(H)))J(H)dk_1 dH dk_2, \tag{13}$$

where dg is the Haar measure on $Sp(n,\mathbb{R})$ and

$$\begin{aligned} J(H) &= \prod_{\lambda \in \Lambda^+} e^{\lambda(H)} - e^{-\lambda(H)} \\ &= 2^{|\Lambda^+|} \prod_{\lambda \in \Lambda^+} \sinh(\lambda(H)). \end{aligned}$$

Thus for all $t \in C_c(Sp(n,\mathbb{R})/K)$,

$$\int_{Sp(n,\mathbb{R})/K} t(x)dx = c_2 \int_K \int_{\mathfrak{a}^+} t(exp(Ad_k(H)))J(H)dk dH, \tag{14}$$

where dx is the invariant measure on $Sp(n,\mathbb{R})/K$. After identifying $Sp(n,\mathbb{R})/K$ and $exp(\mathfrak{p})$ the Riemannian measure on $exp(\mathfrak{p})$ coincides with the invariant measure on $Sp(n,\mathbb{R})/K$. Thus for all $t \in C_c(exp(\mathfrak{p}))$,

$$\int_{exp(\mathfrak{p})} t(x)dvol' = c_2 \int_K \int_{\mathfrak{a}^+} t(exp(Ad_k(H)))J(H)dk dH. \tag{15}$$

Using the notation H_Y of Eq. 12,

$$\int_{\tilde{\mathfrak{p}}} t(exp(Y))J(H_Y)\prod_{\lambda\in\Lambda^+}\frac{1}{\lambda(H_Y)}dY = c_1\int_K\int_{\mathfrak{a}^+} t(exp(Ad_k(H)))J(H)dkdH. \tag{16}$$

Combining Eq. 15 and Eq. 16 we obtain that exists c_3 such that

$$\int_{\tilde{\mathfrak{p}}} t(exp(Y))\prod_{\lambda\in\Lambda^+}\frac{\sinh(\lambda(H_Y))}{\lambda(H_Y)}dY = c_3\int_{exp(\mathfrak{p})} t(x)dvol'. \tag{17}$$

The term $\frac{\sinh(\lambda(H))}{\lambda(H)}$ can be extended by continuity on $\mathfrak{a}$ thus

$$\int_{\mathfrak{p}} t(exp(Y))\prod_{\lambda\in\Lambda^+}\frac{\sinh(\lambda(H_Y))}{\lambda(H_Y)}dY = c_3\int_{exp(\mathfrak{p})} t(x)dvol'. \tag{18}$$

Let dY be the Lebesgue measure corresponding to the metric. Then the exponential application does not introduce a volume change at $0 \in \mathfrak{p}$. Since $H_0 = 0$ and $\frac{\sinh(\lambda(H))}{\lambda(H)} \underset{H\to 0}{\longrightarrow} 1$, we have $c_3 = 1$. Let log denote the inverse of the exponential application. We have

$$\frac{dlog^*(vol')}{dY} = \prod_{\lambda\in\Lambda^+}\frac{\sinh(\lambda(H_Y))}{\lambda(H_Y)}.$$

Since ϕ from Eq. 10 is an isometry up to a scaling factor, if $Y \in \mathfrak{p}$ and $C\phi(exp(Y))C^{-1} = exp_0(u \in T_0\mathcal{D}_n)$, then

$$\frac{dlog^*(vol)}{dLeb_0}(u) = \frac{dlog^*(vol')}{dY}(Y),$$

where Leb_0 refers to the Lebesgue measure on the tangent space $T_0\mathcal{D}_n$ as in Eq. 6. Given $Z \in \mathcal{D}_n$, H_Z from Eq. 4 verifies $C\phi(exp(Ad_k(H_Z)))C^{-1} = Z$ for some k in K. Thus

$$\theta_0(Z) = \frac{dlog^*(vol')}{dY}(Ad_k(H_Z)) = \prod_{\lambda\in\Lambda^+}\frac{\sinh(\lambda(H_Z))}{\lambda(H_Z)}.$$

We have then

$$\theta_0(Z) = \prod_{i<j}\frac{\sinh(\tau_i-\tau_j)}{\tau_i-\tau_j}\prod_{i\leq j}\frac{\sinh(\tau_i+\tau_j)}{\tau_i+\tau_j}$$

where the (τ_i) are described in section 2.2. Given $Z_1, Z_2 \in \mathcal{D}_n$,

$$\theta_{Z_1}(Z_2) = \theta_0(g_{Z_1}^{-1}.Z_2)$$

where $g_{Z_1}^{-1}$ is defined in Eq. 3. It is thus possible to use the density estimator defined in Eq. 7. Indeed,

$$\frac{1}{\theta_{Z_1}(Z_2)}\mathcal{K}(d(Z_1,Z_2)) = \prod_{i<j}\frac{\tau_i-\tau_j}{\sinh(\tau_i-\tau_j)}\prod_{i\leq j}\frac{\tau_i+\tau_j}{\sinh(\tau_i+\tau_j)}\mathcal{K}((2\sum\tau_i^2)^{1/2}) \tag{19}$$

where the (τ_i) are the diagonal elements of $H_{g_{Z_1}^{-1}.Z_2}$.

4 Application to radar processing

4.1 Radar data

In STAP radar processing, the signal is formed by a succession of matrix X representing the realization of a temporal and spatial process. Let $\mathcal{B}_{n,m}^+$ be the set of positive definite block Teoplitz matrices composed of $n \times n$ blocks of $m \times m$ matrices (PD BT). For a stationary signal the autocorrelation matrix R is PD BT, see [11]. Authors of [11] proposed a generalization of Verblunsky coefficients and defined a parametrization of PD BT matrices,

$$\begin{array}{ccc} \mathcal{B}_{n,m}^+ & \rightarrow & Sym^+ \times \mathbb{D}_n^{m-1} \\ \mathcal{R} & \mapsto & (P_0, \Omega_1, .., \Omega_{m-1}) \end{array} \tag{20}$$

in which the metric induced by the Kähler potential is the product metric of an affine invariant metric on Sym^+ and $m-1$ times the metric of the Siegel disk, up to a scaling factor. Among other references, positive definite block Teoplitz matrices have been studied in the context of STAP-radar processing in [5, 6, 7].

4.2 Experimental results

In this section we show density estimation results of the marginal parameters Ω_k. For the sake of visualization, only the Siegel disk $\mathbb{D}_1$ is considered. Recall that $\mathbb{D}_1$ coincides with the Poincaré disk. The results are partly extracted from the conference paper [17]. Data used in the experimental tests are radar observations from THALES X-band Radar, recorded during 2014 field trials campaign at Toulouse Blagnac Airport for European FP7 UFO study (Ultra-Fast wind sensOrs for wake-vortex hazards mitigation). Data are representative of Turbulent atmosphere monitored by radar. Fig. 2-3 illustrate the density estimation of six coefficients on the Poincaré unit disk. Fig. 2 correspond to a clear environment and Fig. 3 to a rainy environment. The densities are individually re-scaled for visualization purposes. For each

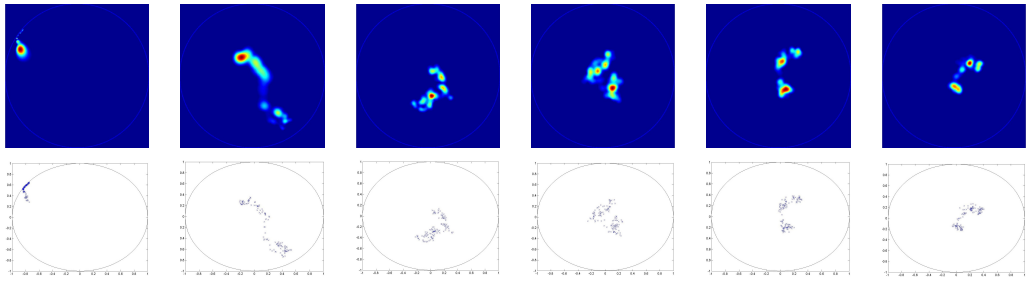


Figure 3: Estimation of the density of 6 coefficients Ω_k under rainy conditions. The expression of the used kernel is $K(x) = \frac{3}{\pi}(1-x^2)^2\mathbf{1}_{x<1}$. Densities are rescaled for visual purpose.

environment the dataset is composed of 120 draws. The densities of the coefficients Ω_k are representative of different backgrounds. These information on the background are expected to ease the detection of interesting targets.

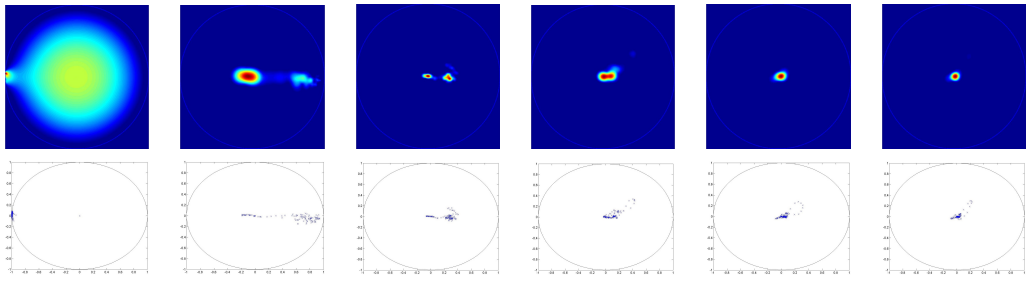


Figure 2: Estimation of the density of 6 coefficients Ω_k under clear conditions. The expression of the used kernel is $K(x) = \frac{3}{\pi}(1-x^2)^2\mathbf{1}_{x<1}$. Densities are rescaled for visual purpose.

5 Conclusion

Three non parametric density estimation techniques have been considered. The main advantage of histograms in the Euclidean context is there simplicity of use. This make histograms an interesting tool despite the fact that they do not present optimal convergence rate. On the Siegel space, histograms loses their simplicity advantage. They were thus not deeply studied. The orthogonal series density estimation also present technical disadvantages on the Siegel space. Indeed, the series become integrals which make the numerical computation of the estimator more difficult than in the

Euclidean case. On the other hand, the use of the kernel density estimator does not present major differences with the Euclidean case. The convergence rate obtained in [1] can be extended to compactly supported random variables on non compact Riemannian manifolds. Furthermore the corrective term whose computation is required to use Euclidean kernels on Riemannian manifolds turns out to have a reasonably simple expression. Our future efforts will concentrate on the use of kernel density estimation on the Siegel space in radar signal processing. We strongly believe that the estimation of the densities of the Ω_k will provide a interesting description of the different backgrounds.

The authors would like to thank Salem Said, Michal Zidor and Dmitry Gourevitch for the help they provided in the understanding of symmetric spaces and the Siegel space.

A Demonstration of theorem 3.1

Lemma A.1. *Let (M, g) be a Riemannian manifold, let C be a compact subset of M and let U be a relatively compact open subset of M containing C. Then there is a compact Riemannian manifold (M', g') such that U is an open subset of M', the inclusion $i : U \hookrightarrow M'$ is a diffeomorphism onto its image and $g' = g$ on $\overline{U}$.*

Proof. We can assume that M is not compact. Let $f : M \to \mathbb{R}$ be smooth function M which tends to $+\infty$ at infinity. Since $\overline{U}$ is compact, $f^{-1}(]-\infty, a[)$ contains $\overline{U}$ for a large enough. By Sard Theorem there exists a value $a \in \mathbb{R}$ such that $f^{-1}(a)$ contains no critical point of f and such that $f^{-1}(]-\infty, a[)$ contains $\overline{U}$. It follows that $N = f^{-1}(]-\infty, a])$ is a submanifold with boundary of M. Since f tends to $+\infty$ at infinity N is compact as well as its boundary $\partial N = f^{-1}(\{a\})$. Call M' the double of N. It is a compact manifold which contains N such that the inclusion $i : N \hookrightarrow M'$ is a diffeomorphism onto its image (see [21], Theorem 5.9 and Definition 5.10). Choose any metric g_0 on M'. Consider two open subsets W_1 and W_2 in M' and two smooth functions $f_1, f_2 : M' \to [0, 1]$ such that

$$\overline{U} \subset W_1 \subset \overline{W}_1 \subset W_2 \subset \overline{W}_2 \subset \operatorname{int} N$$

the interior of N,

$$f_1(x) = 1$$

on $\overline{W}_1$, vanishes outside of W_2, and

$$f_2(x) = 1$$

outside W_1 and vanishes in $\overline{U}$. Define g' on M' by

$$g' = f_1 g + f_2 g_0$$

on N and

$$g' = f_2 g_0$$

outside of N. Since $f_1 + f_2 > 0$, g' is positive definite everywhere on M'. Since f_1 vanishes outside of W_2, g' is smooth on M'. Finally, since $f_1 = 1$ and $f_2 = 0$ on $\overline{U}$, $g' = g$ on $\overline{U}$. □

We can now prove theorem 3.1. Let X be a random variable as in theorem 3.1. Following the notations of the theorem and the lemma, let $U = \{x \in \mathcal{M}, d(x, C) < r_{inj}\}$. U is open, relatively compact and contains C. Let $(\mathcal{M}', g')$ be as in the lemma. Let $\hat{f}$ and $\hat{f}'$ be the kernel density estimators defined on M and M' respectively. Theorem 3.1 of [1] provides the desired results for $\hat{f}'$. for $r \leq r_{inj}$, the support and the values on the support of $\hat{f}'$ and $\hat{f}$ coincide. Thus the desired result also hold for $\hat{f}$.

References

[1] Pelletier, B. Kernel density estimation on Riemannian manifolds. *Statistics and Probability letters, Elsevier*, **2005**, 73(3):297-304.

[2] Hendriks, H. Nonparametric estimation of a probability density on a Riemannian manifold using Fourier expansions. *The Annals of Statistics*, *1990*, 832-849.

[3] Huckemann, S.; Kim, P.; Koo, J.; Munk, A. Mobius deconvolution on the hyperbolic plan with application to impedance density estimation. *The Annals of Statistics, Institute of Mathematical Statistics*, *2010 38*(4), 2465-2498.

[4] Barbaresco, F. Robust Median-Based STAP in Inhomogeneous Secondary Data : Frechet Information Geometry of Covariance Matrices. In *2nd French-Singaporian SONDRA Workshop on EM Modeling, New Concepts and Signal Processing For Radar Detection and Remote Sensing*, Cargese France, 25-28 May 2010

[5] Barbaresco, F. Robust statistical radar processing in Frchet metric space: OS-HDR-CFAR and OS-STAP processing in siegel homogeneous bounded domains, proceedings of IRS11, in Proc. Int. Radar Conf., Leipzig, Germany, Sep. 2011.

[6] Barbaresco F. Information Geometry of Covariance Matrix: Cartan-Siegel Homogeneous Bounded Domains, Mostow/Berger Fibration and Frchet Median. In *Matrix Information Geometry*, R. Bhatia and F. Nielsen, Eds. New York, NY, USA: Springer, pp. 199256, 2012

[7] Barbaresco, F. Information geometry manifold of Toeplitz Hermitian positive definite covariance matrices: Mostow/Berger fibration and Berezin quantization of Cartan-Siegel domains, Int. J. Emerging Trends in Signal Process. (IJETSP), vol. 1, no. 3, Mar. 2013

[8] Degurse, J.F.; Savy, L.; Molinie, J.P.; Marcos, S. A Riemannian Approach for Training Data Selection in Space-Time Adaptive Processing Applications, Radar Symposium (IRS), Vol. 1, pp.319-324, 2013

[9] Degurse, J.F.; Savy, L.; Marcos, S. Information Geometry for radar detection in heterogeneous environments, Maxent'14, 21-26 Sept. 2014

[10] Barbaresco, F. Koszul Information Geometry and Souriau Geometric Temperature/Capacity of Lie Group Thermodynamics, Entropy, 16, special issue on Information Geometry, 2014

[11] Jeuris, B.; Vandebril, R. The Kahler mean of Block-Toeplitz matrices with Toeplitz structured block, *SIAM Journal on Matrix Analysis and Applications*, **2015**.

[12] Said, S.; Bombrun, L.; Berthoumieu, Y. New Riemannian Priors on the Univariate Normal Model. *Entropy,* 16,4015-4031, 2014.

[13] S. Said, . ??. *??* .

[14] Helgason, S. Differential geometry, Lie groups, and symmetric spaces. *Academic press*, **1979**.

[15] P. Kim, D. Richards. Deconvolution density estimation on the space of positive definite symmetric matrices. In *Nonparametric Statistics and Mixture Models: A Festschrift in Honor of Thomas P. Hettmansperger,* pp. 147–168, World Scientific Publishing, 2008.

[16] Asta, D.; Shalizi, C. Geometric network comparison. arXiv preprint arXiv:1411.1350, 2014.

[17] Chevallier, E.; Barbaresco, F.; Angulo, J. Probability density estimation on the hyperbolic space applied to radar processing. In *Geometric*

Science of Information. Springer International Publishing, **2015**, 753-761.

[18] Terras, A. Harmonic analysis on symmetric spaces and applications II. *Springer Science and Business Media*, **2012**.

[19] Gangolli, R.; Varadarajan, V. S. Harmonic analysis of spherical functions on real reductive groups. *Springer Science and Business Media*.

[20] Siegel, C. L. Symplectic geometry. *American Journal of Mathematics*, **1943** *65*(1), 1-86.

[21] Munkres, J.R. Elementary Differential Topology. *Annals of Mathematics Studies-54*, **1967**.