A Family of Anisotropic Distributions on the Hyperbolic Plane

Emmanuel Chevallier[(✉)]

Weizmann Institute of Science, Rehovot, Israel
emmanuelchevallier1@gmail.com

Abstract. Most of the parametric families of distributions on manifold are constituted of radial distributions. The main reason is that quantifying the anisotropy of a distribution on a manifold is not as straightforward as in vector spaces and usually leads to numerical computations. Based on a simple definition of the covariance on manifolds, this paper presents a way of constructing anisotropic distributions on the hyperbolic space whose covariance matrices are explicitly known. The approach remains valid on every manifold homeomorphic to vector spaces.

1 Introduction

Probability density estimation on Riemannian manifolds is the subject of several recent studies. The different approaches can be separated into two categories, the parametric and non-parametric ones. The context of Riemannian manifolds brings difficulties of two kinds. Firstly, the theoretical results about distributions and the convergence of estimators known for random variables valued in $\mathbb{R}^n$ have to be adapted to the case of random variables valued in Riemannian manifolds, see [1–3,8,9,12–15]. Secondly, the construction of probability distribution and of density estimators should require a reasonable amount of computational complexity, see [8,12,13,16–18]. A generalization of the Gaussian distribution on manifolds was proposed in [8]. Although the expression of the proposed law is hard to compute on general manifolds, expressions of radial Gaussians on symmetric spaces can be found in [12–14]. On isotropic spaces, an isotropic density is simply a radial density. The anisotropy of a density can be evaluated with the notion of covariance proposed in [8].

In this paper, we are interested in the construction of anisotropic distributions on the hyperbolic space. The problem of anisotropic normal distributions on manifold have been addressed in [19] through anisotropic diffusion. The construction is valid on arbitrary manifolds but requires important computations. The hyperbolic space is a very particular Riemannian manifold: it is at the same time isotropic and diffeomorphic to a vector space. These two specificities significantly ease the construction of probability distributions and probability density estimators. Generally, it is difficult to control the covariance of a distribution on a Riemannian manifold, e.g. the covariance of the Gaussian law proposed in [8]. We propose a simple way of constructing distributions whose covariance is fully

controlled. The method is derived from the density kernel proposed by [1]. These distributions can be used in the non parametric kernel density estimator but also to design mixture models for parametric density estimation.

The paper is organised as follows. Section 2 is a very brief introduction to the hyperbolic plane. Section 3 reviews some general facts about probabilities on Riemannian manifolds. Section 4 describes how to built anisotropic density functions on the hyperbolic space.

2 The Hyperbolic Space

The hyperbolic geometry results of a modification of the fifth Euclid's postulate on parallel lines. In two dimensions, given an line D and a point $p \notin D$, the hyperbolic geometry is an example where there are at least two lines going through p, which do not intersect D. Let us consider the open unit disk of the Euclidean plane endowed with the Riemannian metric:

$$ds_{\mathbb{D}}^2 = 4\frac{dx^2 + dy^2}{(1 - x^2 - y^2)^2} \tag{1}$$

where x and y are the Cartesian coordinates. The unit disk $\mathbb{D}$ endowed with $ds_{\mathbb{D}}$ is called the Poincaré disk and is a model of the two-dimensional hyperbolic geometry. The construction is generalized to higher dimensions. Let ISO be the isometry group of $\mathbb{D}$. It can be shown that:

- $\mathbb{D}$ is homogeneous: $\forall p, q \in \mathbb{D}, \exists \phi \in ISO, \phi(p) = q$, points are indistinguishable.
- $\mathbb{D}$ is isotropic: for any couple of geodesics γ_1 and γ_2 going through a point $p \in \mathbb{D}$, there exists $\phi \in ISO$ such that $\phi(p) = p$ and $\phi(\gamma_1) = \gamma_2$. In other words, directions are indistinguishable.
- the Riemannian exponential applications are bijective.
- $\mathbb{D}$ has a constant negative curvature.

Let x denote the coordinates of elements of $T_p\mathbb{D}$ in an orthogonal basis. x is mapped to a point on $\mathbb{D}$ by the Riemannian exponential application noted exp_p and form thus a chart of $\mathbb{D}$. This chart is called an exponential chart at the point p.

Given a reference point p the point of polar coordinates (r, α) of the hyperbolic space is defined as the point at distance r of p on the geodesic with initial direction $\alpha \in \mathbb{S}^1$. Since the hyperbolic space is isotropic, the expression of the metric in polar coordinates only depends on r,

$$ds^2 = dr^2 + \sinh(r)^2 d\alpha^2, \tag{2}$$

see [10,11].

3   Distributions on $\mathbb{D}$

3.1   Densities

The metric of a Riemannian manifold provides a measure of volumes *vol*. In a chart, if G is the matrix of the metric, the density of *vol* with respect to the Lebesgue measure of the chart is

$$\frac{dvol}{dLeb} = |det(\sqrt{G})|$$

where $\sqrt{G}$ is the matrix square root of G. Let μ be a measure on $\mathcal{M}$. If μ has a density f with respect to the Lebesgue measure of a chart, then the density with respect to the Riemannian volume measure is given by

$$\frac{d\mu}{dvol} = \frac{d\mu}{dLeb}\frac{dLeb}{dvol} = \frac{1}{|det(\sqrt{G})|}f. \tag{3}$$

3.2   Intrinsic Means

Given a distribution μ, the variance at p be defined by

$$\sigma^2(p) = \int_{\mathbb{D}} d(p,.)^2 d\mu.$$

When the variance is finite everywhere, its minima are called mean points. The hyperbolic space is a Cartan-Hadamar manifold, that is to say it is complete, simply connected and of negative curvature. On Cartan-Hadamar manifolds, when the variance is finite everywhere, the mean exists and is unique, see [8] corollary 2. It is achieved at p such that

$$\int_{T_p\mathbb{D}} x d\tilde{\mu} = 0,$$

where $\tilde{\mu}$ is the image of the measure μ by the inverse of the exponential application at p.

3.3   Covariance on Manifold

The covariance of a random vector is the matrix formed by the covariance of its coordinates. In a vector space the coordinates of a vector are given in terms of projection on the corresponding axis. On a Riemannian manifold the notions of projection on a geodesic usually do not lead to explicit expressions. Even if it does not conserve all the properties of the covariance of vectors, when possible, the simplest generalisation to manifolds is to take the Euclidean covariance after lifting the distribution on a tangent space by the inverse of the exponential map, see [8]. Since on the hyperbolic space the exponential application a bijection, it

is always possible to lift distributions on tangent spaces. Given a distribution μ and a orthogonal basis of $T_p\mathbb{D}$, the covariance at $p \in \mathbb{D}$ is thus defined as

$$\Sigma_p(\mu) = \int_{T_p\mathbb{D}} xx^t d\tilde{\mu}$$

This definition of covariance was used to define a notion of principal geodesic analysis on manifolds in [20]. It can be noted that the covariance at the point p is a point in $T\mathbb{D} \otimes T\mathbb{D}$.

4   Constructing Anisotropic Distributions

The author of [8] proposes a generalization of Gaussian distributions on manifolds as the distribution that maximizes the entropy given its barycenter and covariance. This generalization leads to a density of the form,

$$N_{(p,\Gamma)}(exp_p(x)) = k.\exp\left(-\frac{x^t \Gamma x}{2}\right)$$

Given p and the covariance matrix Σ_p, the main difficulties are to obtain expressions of the normalizing factor k and of the concentration matrix Γ. Since hyperbolic space is homogenous, k and Γ only depend on the matrix Σ_p. The expression of k and Γ when Σ_p is a (positive) multiple of the identity matrix can be found in [12]. However, it is difficult to obtain these relations when Σ_p is not diagonal.

It might be interesting to define parametric families of distributions whose means and covariances can easily be controlled, even if they do not verify the same statistical properties as the Gaussian distributions. Let $K : \mathbb{R}_+ \to \mathbb{R}_+$ be a function such that,

i. $\int_{\mathbb{R}^2} K(\|y\|)\,dy = 1$
ii. $\int_{\mathbb{R}^2} \|y\|^2 K(\|y\|)\,dy = 2$

Given Γ a symmetric positive definite matrix, we have then

$$\int_{\mathbb{R}^2} \frac{1}{\sqrt{det(\Gamma)}} K(\sqrt{x^t \Gamma^{-1} x})dx = 1.$$

Let $\bar{p}$ be a point in $\mathbb{D}$. Set an orthonormal basis of the tangent space $T_{\bar{p}}\mathbb{D}$ and consider the distribution $\nu_{\bar{p},\Gamma}$ on $T_{\bar{p}}\mathbb{D}$ whose density with respect to the Lebesgue measure of $T_{\bar{p}}\mathbb{D}$ is given by $\frac{1}{\sqrt{det(\Sigma)}} K(\sqrt{x^t \Gamma^{-1} x})$, where x and Γ are expressed in the reference basis. Let $\mu_{\bar{p},\Gamma} = exp_{\bar{p}*}(\nu_{\bar{p},\Gamma})$ be the pushforward measure of $\nu_{\bar{p},\Gamma}$ by the Riemannian exponential at $\bar{p}$.

Theorem 1. $\bar{p}$ *is the unique mean of* $\mu_{\bar{p},\Gamma}$.

Proof. It can be checked that $\mu_{\overline{p},\Gamma}$ has a finite variance everywhere. Moreover,

$$\int_{T_{\overline{p}}\mathbb{D}} \sqrt{\Gamma^{-1}}x \frac{1}{\sqrt{\det \Gamma}} K(\sqrt{x^t \Gamma^{-1} x})\, dx = 0.$$

The integrability of the function can be deduced from i and ii and the nullity from its symmetry. Therefore according to Sect. 3.2 $\overline{p}$ is the unique mean of $\mu_{\overline{p},\Gamma}$.

Theorem 2. *The covariance $\Sigma_{\overline{p}}$ of $\mu_{\overline{p},\Gamma}$ at $\overline{p}$ and the concentration matrix Γ are equal.*

Proof. In the reference basis, making use of ii with the change of variables $y = \sqrt{\Gamma^{-1}}x$

$$\begin{aligned}
\Sigma_{\overline{p}} &= \int_{\mathbb{R}^2} xx^t \frac{1}{\sqrt{\det(\Gamma)}} K(\sqrt{x^t \Gamma^{-1} x})\, dx \\
&= \Gamma^{1/2} \int_{\mathbb{R}^2} yy^t K(\sqrt{y^t y})\, dy\, \Gamma^{1/2} \\
&= \Gamma^{1/2} \left(\int_{\mathbb{R}} \int_0^{2\pi} r^2 \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}^t K(r) r\, dr\, d\theta \right) \Gamma^{1/2} \\
&= \Gamma^{1/2} \left(\frac{1}{2} \int_{\mathbb{R}} r^2 I K(r) 2\pi r\, dr \right) \Gamma^{1/2} \\
&= \Gamma^{1/2} I \left(\frac{1}{2} \int_{\mathbb{R}^2} \|y\|^2 K(\|y\|)\, dy \right) \Gamma^{1/2} \\
&= \Gamma.
\end{aligned}$$

The tangent space $T_{\overline{p}}\mathbb{D}$ endowed with the reference basis provides a parametrization of the hyperbolic space. By definition, the density of $\mu_{\overline{p},\Gamma}$ in this parametrization is given by $\frac{1}{\sqrt{det(\Sigma)}} K(\sqrt{x^t \Sigma^{-1} x})$. In order to obtain the density with respect to the Riemannian measure this term should be multiplied by the density of the Lebesgue measure of the parametrization with respect to the Riemannian measure, see Eq. 3. In an adapted orthonormal basis of $T_{\overline{p}}\mathbb{D}$, Eq. 2 leads to the following expression of the matrix of the metric,

$$G = \begin{pmatrix} 1 & 0 \\ 0 & \frac{\sinh(r)^2}{r^2} \end{pmatrix}.$$

Thus,

$$det(\sqrt{G}) = \frac{\sinh(r)}{r}.$$

Equation 3 leads to the density ratio,

$$\frac{dx}{dvol}(x) = \frac{\|x\|}{\sinh(\|x\|)},$$

where dx is the Lebesgue measure induced by the reference basis. Recall that in this parametrization, the Euclidean norm of x is the distance between $exp_{\overline{p}}(x)$ and $\overline{p}$. The density of $\mu_{\overline{p},\Gamma}$ with respect to the Riemannian measure is given by

$$f(exp_{\overline{p}}(x)) = \frac{||x||}{\sinh(||x||)\sqrt{det(\Sigma)}} K\left(\sqrt{x^t \Sigma^{-1} x}\right).$$

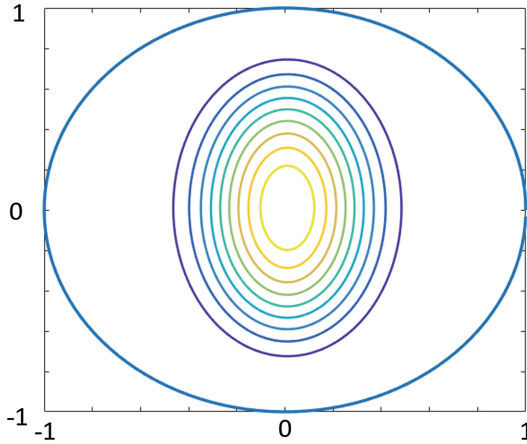Figure 1 shows the level lines when K is Gaussian.

Fig. 1. In this example $K(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2}$ and Σ has 1 and $\frac{1}{4}$ as eigenvalues. The level lines of the corresponding density f are flattened circle but are not ellipses.

5   Estimating the Mean and the Covariance

Let the function K and the distribution $\mu_{\overline{p},\Gamma}$ be as defined in Sect. 4. Given a set of draws drawn from this distribution it is important to have estimators of the two parameters: the mean and the covariance. In order to estimate the unknown parameters $(\overline{p}, \Sigma_{\overline{p}})$ given a set of independent samples $(p_1, .., p_n)$, it is usual to try to maximize the likelihood function. The log-likelihood of a set of samples is defined as

$$\mathcal{L}(p_1, .., p_n; (\hat{p}, \hat{\Sigma})) = \sum_i \log\left(\frac{||x_i||}{\sinh(||x_i||)\sqrt{det(\hat{\Sigma})}} K\left(\sqrt{x_i^t \hat{\Sigma}^{-1} x_i}\right)\right)$$

$$= \sum_i \log\left(\frac{||x_i||}{\sinh(||x_i||)\sqrt{det(\hat{\Sigma})}}\right) + log\left(K\left(\sqrt{x_i^t \hat{\Sigma}^{-1} x_i}\right)\right).$$

The major difficulty is that it is not possible to optimize the mean and the covariance separately. Thus there might not be explicit expressions of the maximum likelihood. However, the mean and the covariance have natural estimators. It is already known that the empirical barycenter is a strongly consistent estimator of the barycenter, see [21] Theorem 2.3.

Given an estimate of the barycenter, it is possible to compute the empirical covariance in the corresponding tangent plane,

$$\hat{\Sigma}_{\hat{p}} = \frac{1}{N} \sum x_i x_i^t \tag{4}$$

Using a similar construction as the Sasakian metric, see [22], the vector bundle $T\mathbb{D} \otimes T\mathbb{D}$ can be endowed with a Riemannian metric. Although we do not prove it in this paper, we are convinced that almost surely

$$d((\hat{p}, \hat{\Sigma}_{\hat{p}}), (\overline{p}, \Sigma)) \xrightarrow[n \to +\infty]{} 0,$$

where d is the Riemannian distance on $T\mathbb{D} \otimes T\mathbb{D}$.

6   Conclusion

In this paper we proposed a set of parametric families of anisotropic distributions on the hyperbolic plane. The main interest of these distributions is that the covariance matrix and concentration matrix are equal. The empirical mean and covariance provide thus simple estimators of the parameters of the distribution. Working with anisotropic distributions is expected to reduce the number of distributions used in mixture models and thus to reduce the computational complexity of the parameter estimation of the mixture models. On the one hand, our future work will focus on deriving convergence rates of the estimation of the covariance. On the other hand, we will study the use of these distributions in problems of radar signal classification.

References

1. Pelletier, B.: Kernel density estimation on Riemannian manifolds. Stat. Probab. Lett. **73**, 297–304 (2005)
2. Hendriks, H.: Nonparametric estimation of a probability density on a Riemannian manifold using Fourier expansions. Ann. Stat. **18**, 832–849 (1990)
3. Huckemann, S., Kim, P., Koo, J., Munk, A.: Mobius deconvolution on the hyperbolic plan with application to impedance density estimation. Ann. Stat. **38**, 2465–2498 (2010)
4. Barbaresco, F.: Robust statistical radar processing in Fréchet metric space: OS-HDR-CFAR and OS-STAP processing in siegel homogeneous bounded domains. In: Proceedings of the 2011 12th International Radar Symposium (IRS), Leipzig, Germany, 7–9 September 2011

5. Barbaresco, F.: Information geometry of covariance matrix: Cartan-Siegel homogeneous bounded domains, Mostow/Berger fibration and Fréchet median. In: Nielsen, F., Bhatia, R. (eds.) Matrix Information Geometry, pp. 199–256. Springer, Heidelberg (2012). doi:10.1007/978-3-642-30232-9_9

6. Barbaresco, F.: Information geometry manifold of Toeplitz Hermitian positive definite covariance matrices: Mostow/Berger fibration and Berezin quantization of Cartan-Siegel domains. Int. J. Emerg. Trends Signal Process. (IJETSP) **1**, 1–87 (2013)

7. Cannon, J.W., Floyd, W.J., Kenyon, R., Parry, W.R.: Hyperbolic Geometry, vol. 31. MSRI Publications, Cambridge (1997)

8. Pennec, X.: Intrinsic statistics on Riemannian manifolds: basic tools for geometric measurements. J. Math. Imaging Vis. **25**, 127 (2006)

9. Kim, P.T., Richards, D.St.P: Deconvolution density estimation on the space of positive definite symmetric matrices. In: Nonparametric Statistics and Mixture Models, pp. 147–168 (2011)

10. Grigoryan, A.: Heat Kernel and Analysis on Manifolds, vol. 47. American Mathematical Soc., Providence (2012)

11. Anker, J.-P., Ostellari, P.: The heat kernel on noncompact symmetric spaces. In: Lie Groups and Symmetric Spaces. AMS Transl. Ser. 2, vol. 210, pp. 27–46 (2003)

12. Said, S., Bombrun, L., Berthoumieu, Y.: New Riemannian priors on the univariate normal model. Entropy **16**(7), 4015–4031 (2014)

13. Said, S., Bombrun, L., Berthoumieu, Y., Jonathan Manton, J.: Riemannian Gaussian distributions on the space of symmetric positive definite matrices. arXiv:1507.01760 [math.ST] (2015)

14. Said, S., Hajri, H., Bombrun, L., Vemuri, B.C.: Gaussian distributions on Riemannian symmetric spaces: statistical learning with structured covariance matrices. arXiv:1607.06929 [math.ST] (2016)

15. Le Bihan, N., Flamant, J., Manton, J.H.: Density estimation on the rotation group using diffusive wavelets. arXiv:1512.06023 (2015)

16. Chevallier, E., Barbaresco, F., Angulo, J.: Probability density estimation on the hyperbolic space applied to radar processing. In: Nielsen, F., Barbaresco, F. (eds.) GSI 2015. LNCS, vol. 9389, pp. 753–761. Springer, Cham (2015). doi:10.1007/978-3-319-25040-3_80

17. Chevallier, E., Forget, T., Barbaresco, F., Angulo, J.: Kernel density estimation on the siegel space with an application to radar processing. Entropy **18**(11), 396 (2016)

18. Chevallier, E., Kalunga, E., Angulo, J.: Kernel density estimation on spaces of Gaussian distributions and symmetric positive definite matrices. SIAM J. Imaging Sci. **10**(1), 191–215 (2017)

19. Sommer, S.: Anisotropic distributions on manifolds: template estimation and most probable paths. In: Ourselin, S., Alexander, D.C., Westin, C.-F., Cardoso, M.J. (eds.) IPMI 2015. LNCS, vol. 9123, pp. 193–204. Springer, Cham (2015). doi:10.1007/978-3-319-19992-4_15

20. Fletcher, P.T., Lu, C., Pizer, S.M., Joshi, S.: Principal geodesic analysis for the study of nonlinear statistics of shape. IEEE Trans. Med. Imaging **23**(8), 995–1005 (2004)

21. Bhattacharya, R., Patrangenaru, V.: Large sample theory of intrinsic and extrinsic sample means on manifolds. I. Ann. Stat. **31**, 1–29 (2003)

22. Kappos, E.: Natural metric on tangent bundles. Master's thesis (2001)