On Geodesic Exponential Kernels

Aasa Feragen[1], François Lauze[1], and Søren Hauberg[2]

[1]Department of Computer Science, University of Copenhagen, Denmark; [2]DTU Compute, Copenhagen, Denmark
aasa@diku.dk

This extended abstract summarizes work presented at CVPR 2015 [1].

Standard statistics and machine learning tools require input data residing in a Euclidean space. However, many types of data are more faithfully represented in general nonlinear metric spaces or Riemannian manifolds, e.g. shapes, symmetric positive definite matrices, human poses or graphs. The underlying metric space captures domain specific knowledge, e.g. non-linear constraints, which is available *a priori*. The intrinsic geodesic metric encodes this knowledge, often leading to improved statistical models.

A seemingly straightforward approach to statistics in metric spaces is to use kernel methods [3], designing exponential kernels:

$$k(x,y) = \exp\left(-\lambda(d(x,y))^q\right), \qquad \lambda, q > 0, \tag{1}$$

which only rely on geodesic distances $d(x,y)$ between observations. For $q = 2$ this gives a geodesic generalization of the *Gaussian kernel*, and $q = 1$ gives the geodesic *Laplacian kernel*. While this idea has an appealing similarity to familiar Euclidean kernel methods, we show that it is highly limited if the metric space is curved, see Table 1.

Theorem 1 *For a geodesic metric space* (X, d)*, assume that* $k(x,y) = \exp(-\lambda d^2(x,y))$ *is positive definite (PD) for all* $\lambda > 0$*. Then* (X, d) *is flat in the sense of Alexandrov.*

This is a negative result, as most metric spaces of interest are not flat. As a consequence, we show that **geodesic Gaussian kernels on Riemannian manifolds are PD for all $\lambda > 0$ only if the Riemannian manifold is Euclidean**.

Theorem 2 *Let* M *be a complete, smooth Riemannian manifold with its associated geodesic distance metric* d*. Assume, moreover, that* $k(x,y) = \exp(-\lambda d^2(x,y))$ *is PD for all* $\lambda > 0$*. Then the Riemannian manifold* M *is isometric to a Euclidean space.*

Do these negative results depend on the choice $q = 2$ in (1)?

	Extends to general	
Kernel	**Metric spaces**	**Riemannian manifolds**
Gaussian ($q = 2$)	No (only if flat)	No (only if Euclidean)
Laplacian ($q = 1$)	Yes, iff metric is CND	Yes, iff metric is CND
Geodesic exp. ($q > 2$)	Not known	No

Table 1. Overview of results: For a geodesic metric, the geodesic exponential kernel (1) is only positive definite for all $\lambda > 0$ for

Space	Distance metric	Geodesic metric?	Euclidean? metric?	CND metric?	PD Gaussian kernel?	PD Laplacian kernel?
$\mathbb{R}^n$	Euclidean metric	✓	✓	✓	✓	✓
$\mathbb{R}^n$, $n > 2$	l_q-norm $\|\cdot\|_q$, $q > 2$	✓	÷	÷	÷	÷
Sphere $\mathbb{S}^n$	classical intrinsic	✓	÷	✓	÷	✓
Real projective space $\mathbb{P}^n(\mathbb{R})$	classical intrinsic	✓	÷	÷	÷	÷
Grassmannian	classical intrinsic	✓	÷	÷	÷	÷
Sym_d^+	Frobenius	✓	✓	✓	✓	✓
Sym_d^+	Log-Euclidean	✓	✓	✓	✓	✓
Sym_d^+	Affine invariant	✓	÷	÷	÷	÷
Sym_d^+	Fisher information metric	✓	÷	÷	÷	÷
Hyperbolic space $\mathbb{H}^n$	classical intrinsic	✓	÷	✓	÷	✓
1-dimensional normal distributions	Fisher information metric	✓	÷	✓	÷	✓
Metric trees	tree metric	✓	÷	✓	÷	✓
Geometric graphs (e.g. kNN)	shortest path distance	✓	÷	÷	÷	÷
Strings	string edit distance	✓	÷	÷	÷	÷
Trees, graphs	tree/graph edit distance	✓	÷	÷	÷	÷

Table 2. For a set of popular data spaces and metrics, we record whether the metric is a geodesic metric, whether it is a Euclidean metric, whether it is a CND metric, and whether its corresponding Gaussian and Laplacian kernels are PD.

Theorem 3 *Let M be a Riemannian manifold with its associated geodesic distance metric d, and let $q > 2$. Then there is some $\lambda > 0$ so that the kernel (1) is not PD.*

The existence of a $\lambda > 0$ such that the kernel is not PD may seem innocent; however, as a consequence, the kernel bandwidth parameter cannot be learned. In contrast, the choice $q = 1$ in (1), giving a geodesic Laplacian kernel, leads to a more positive result:

Theorem 4 *i) The geodesic distance d in a geodesic metric space (X, d) is conditionally negative definite (CND) if and only if the corresponding geodesic Laplacian kernel is PD for all $\lambda > 0$.*

ii) In this case, the square root $d_{\sqrt{}}(x, y) = \sqrt{d(x, y)}$ is also a distance metric, and $(X, d_{\sqrt{}})$ can be isometrically embedded as a metric space into a Hilbert space H.

iii) The square root metric $d_{\sqrt{}}$ is not a geodesic metric, and $d_{\sqrt{}}$ corresponds to the chordal metric in H, not the intrinsic metric on the image of X in H.

The proofs rely on Schönberg's classical theorem [4], metric geometry and recent results on conditionally negative definite kernels [2]. Theoretical and empirical results on PD'ness of geodesic exponential kernels are summarized in Table 2.

References

1. A. Feragen, F. Lauze, and S. Hauberg. Geodesic exponential kernels: When curvature and linearity conflict. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, 2015.
2. J. Istas. Manifold indexed fractional fields. *ESAIM: Probability and Statistics*, 16:222–276, 1 2012.
3. B. Schölkopf and A.J. Smola. *Learning with kernels : support vector machines, regularization, optimization, and beyond*. Adaptive computation and machine learning. MIT Press, 2002.
4. I. J. Schönberg. Metric spaces and completely monotone functions. *Annals of Mathematics*, 39(4):811–841, 1938.