Geometric Learning with Positively Decomposable Kernels

Nathaël Da Costa (University of Tübingen)
Cyrus Mostajeran (NTU Singapore)
Juan-Pablo Ortega (NTU Singapore)
Salem Said (Université Grenoble-Alpes)

Introduction

Classical kernel methods are based on positive-definite (PD) kernels that enable learning in reproducing kernel Hilbert spaces (RKHS). For non-Euclidean data spaces, the construction and computation of positive definite kernels can be challenging. In this work, we propose the use of reproducing kernel Krein space (RKKS) based methods, which require only kernels that admit a positive decomposition.

Contributions:

- We establish conditions under which a kernel is positively decomposable.
- We show access to the decomposition is not required to learn in RKKS.
- We show that invariant kernels admit a positive decomposition on homogeneous spaces under tractable regularity assumptions.
- We strengthen theoretical foundations for RKKS-based methods in general.

Learning with positive definite kernels

Definition: *A kernel on a set X is a Hermitian map $k : X \times X \to \mathbb{C}$. A kernel is said to be positive definite (PD) if for all $N \in \mathbb{N}$, $x_1, \ldots, x_N \in X$ and all $c_1, \ldots, c_N \in \mathbb{C}$,*

$$\sum_{i=1}^{N} \sum_{j=1}^{N} \overline{c_i} c_j k(x_i, x_j) \geq 0$$

i.e. the matrix $\big(k(x_i, x_j)\big)_{i,j}$ is Hermitian positive semidefinite.

A positive-definite kernel k on a set X implies the existence of an embedding $\phi : X \to \mathscr{H}$ into a reproducing kernel Hilbert space (RKHS) such that $k(x, y) = \langle \phi(x), \phi(y) \rangle$.

A learning problem associated to a positive-definite kernel given some data D:

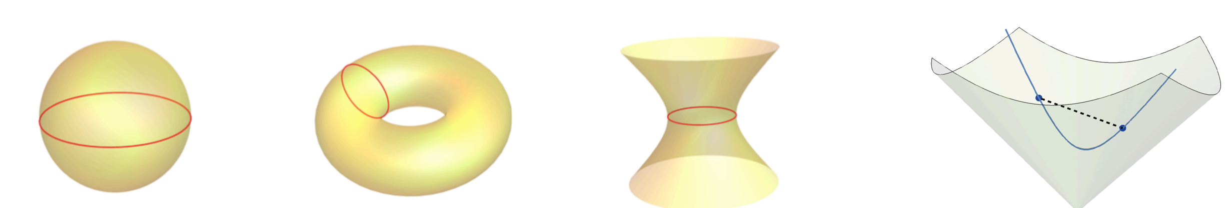
$$\min_{f \in \mathscr{H}} \quad \mathscr{L}(f, D) + \lambda \|f\|_{\mathscr{H}} \quad \text{such that} \quad f \in \Omega(D)$$

Under some mild conditions on $\mathscr{L}$ and Ω, this problem can be converted into a finite dimensional linear algebra problem by the representer theorem.

Positive definite kernels on manifolds

In a Euclidean space $X = \mathbb{R}^n$, many examples of positive-definite (PD) kernels are known. The most well-known PD kernel is the radial basis function or Gaussian kernel $k(x, y) = e^{-\lambda \|x-y\|^2}$ where $\lambda > 0$. In non-Euclidean spaces, the Gaussian kernel $k(x, y) = e^{-\lambda d(x,y)^2}$ is generally **not** PD.

Theorem: *If X is a complete Riemannian manifold, the Gaussian kernel is positive definite for all $\lambda > 0$ if and only if X is isometric to $\mathbb{R}^n$* [1].



As a response, several research groups have developed systematic tools for the generation of PD kernels on Riemannian manifolds using techniques from stochastic PDEs [2] and harmonic analysis on Riemannian symmetric spaces [3,4]. While powerful, these methods rely on advanced mathematical tools and can face computational challenges in certain settings.

Positively decomposable kernels

Definition: *An indefinite inner product on a real vector space $\mathscr{K}$ is a Hermitian bilinear map that is non-degenerate. A complex vector space $\mathscr{K}$ equipped with an indefinite inner product $\langle \cdot, \cdot \rangle$ is called a Krein space if it can be written as the algebraic direct sum*

$$\mathscr{K} = \mathscr{H}_+ \oplus \mathscr{H}_-$$

such that $\mathscr{H}_{+/-}$ equipped with $\langle \cdot, \cdot \rangle_{+/-} := \pm \langle \cdot, \cdot \rangle$ is a Hilbert space and $\langle f_+, f_- \rangle = 0$ for all $f_+ \in \mathscr{H}_+$ and $f_- \in \mathscr{H}_-$.

Definition: *A kernel k is called positively decomposable if $k = k_+ - k_-$ for some positive definite kernels k_+ and k_-.*

A positively decomposable kernel k on a set X implies the existence of an embedding $\phi : X \to \mathscr{K}$ into a reproducing kernel Krein space (RKKS) such that $k(x, y) = \langle \phi(x), \phi(y) \rangle$.

Unlike RKHS, the RKKS need not be unique

A learning problem of the form

$$\text{stabilize}_{f \in \mathscr{K}} \quad \mathscr{L}(f, D) + g(\langle f, f \rangle_{\mathscr{K}}) \quad \text{such that} \quad f \in \Omega(D)$$

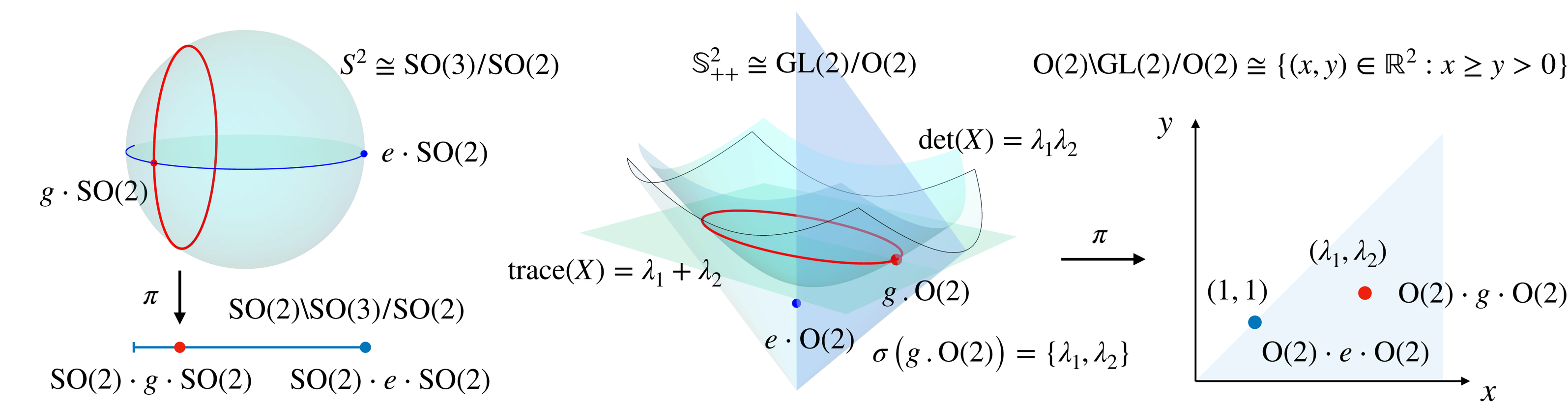
associated with a positively decomposable kernel k given data D, a strictly monotonic differentiable function $g : \mathbb{R} \to \mathbb{R}$, and suitable equality and inequality constraints $f \in \Omega(D)$ can be converted into a finite dimensional linear algebra problem through a representer theorem. Thus, positively decomposable kernels can often be used much like a PD kernel.

Invariant kernels

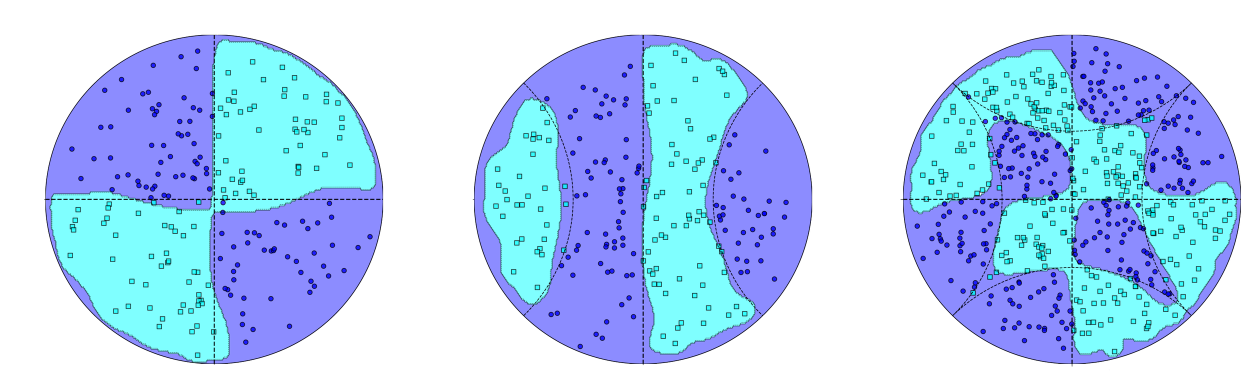
On a homogeneous manifold $X = G/H$, for a smooth kernel k to be positively decomposable it suffices that:

- k is invariant: $k(g \cdot x, g \cdot y) = k(x, y) \; \forall \, x, y \in G/H, g \in G$,
- k and its derivatives decay fast enough at infinity: $\Delta^i f \in L^2(G/H)$ for integers i, where $f = k(\cdot, o)$ for a left-invariant Riemannian metric on G/H.

Example: The Gaussian kernel is positively decomposable in reductive non-compact symmetric spaces such as hyperbolic n-space and the SPD cone!



One-to-one correspondence between positively decomposable invariant kernels on G/H and Hermitian functions on the double coset space $H \backslash G / H$



Krein SVM in Poincaré disk model of hyperbolic plane

References

[1] S Jayasumana, et al. (2015). "Kernel methods on Riemannian manifolds with Gaussian RBF kernels." IEEE Transactions on Pattern Analysis and Machine Intelligence 37(12), 2464–2477.

[2] V Borovitskiy, et al. (2020). "Matérn Gaussian processes on Riemannian manifolds." Advances in Neural Information Processing Systems, 33, 12426-12437.

[3] N. Da Costa, et al. (2025) "Invariant kernels on Riemannian symmetric spaces: a harmonic-analytic approach." SIAM Journal on Mathematics of Data Science. 2025 Jun 30;7(2):752-76.

[4] S. Said, et al. (2024) "Invariant kernels on the space of complex covariance matrices." arXiv preprint arXiv:2404.02169.