Mercer Expansions for Translation-Invariant Kernels

by Stephen Crowley

October 20, 2024

Abstract

An extension of the method for deriving orthonormal expansions for kernels $K(t-s)$ of translation-invariant Gaussian processes detailed by Tronarp and Karvonen[?] is derived by showing selecting the orthonormal base of L^2 to be such that $K(t-s)=\sum_{n=0}^{\infty}\varphi_m(t-s)$ converges uniformly. That is, instead of just choosing any orthnormal basis of L^2, a basis whose partial sums uniformly converge to the kernel itself is constructed.

Table of contents

1 Introduction 1

1.1 Construction of orthonormal bases 2

1.2 On Mercer expansions 4

2 Summary of expansions 5

2.1 Gaussian kernel 5

3 Expansion of the Gaussian kernel 6

1 Introduction

[?]Let Ω be a vector space.

Definition 1. *A symmetric positive-semidefinite kernel* $r\colon\Omega\times\Omega\to\mathbb{R}$ *is* translation-invariant *if* $r(t,u)=K(t-u)$ *for some* $K\colon\Omega\to\mathbb{R}$ *and* $\forall t,u\in\Omega$.

Each positive-semidefinite kernel induces a unique reproducing kernel Hilbert space (RKHS), $\mathscr{H}_r(\Omega)$, which is equipped with an inner product $\langle\cdot,\cdot\rangle_r$ and the associated norm $\|\cdot\|_r$ [?]. Any kernel that induces a seperable infinite-dimensional RKHS $\mathscr{H}_r(\Omega)$ has an orthonormal basis $\{\psi_m\}_{m\in I}$ for some countably infinite index set I (e.g., $I=\mathbb{N}$) and that the kernel admits the pointwise convergent orthonormal expansion

$$r(t,u)=\sum_{m\in I}\psi_m^*(t)\,\psi_m(u)\quad\forall t,u\in\Omega \tag{1}$$

where z^* denotes the complex conjugate of $z \in C$ If Ω is a compact subset of $\mathbb{R}^d$ and r is continuous, the expansion (1) converges uniformly[?].

1.1 Construction of orthonormal bases

Let $|z|$ denote the modulus of $z \in \mathbb{C}$ and recall that z^* is the complex conjugate. The spaces $\mathscr{L}_2(\mathbb{R})$ and $\mathscr{L}_2(\mathbb{R}, 1/2\,\pi)$ consist of all square-integrable functions $f\colon \mathbb{R} \to \mathbb{C}$ and are equipped with the inner products

$$\langle f, g\rangle_{\mathscr{L}_2(\mathbb{R})} = \int_{-\infty}^{\infty} f^*(t)\, g(t) \mathrm{d}t \tag{2}$$

and

$$\langle f, g\rangle_{\mathscr{L}_2(\mathbb{R}, 1/2\pi)} = \frac{\int_{-\infty}^{\infty} f^*(t)\, g(t) \mathrm{d}t}{2\,\pi} \tag{3}$$

The Fourier transform and the corresponding inverse transform for any integrable or square-integrable function f are defined as

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} f(t)\, e^{-i\omega t} \mathrm{d}t \tag{4}$$

and

$$f(t) = \frac{\int_{-\infty}^{\infty} \hat{f}(\omega)\, e^{i\omega t} \mathrm{d}\omega}{2\,\pi} \tag{5}$$

The Fourier transform defines an isometry between $\mathscr{L}_2(\mathbb{R})$ and $\mathscr{L}_2(\mathbb{R}, 1/2\,\pi)$ via the Plancherel theorem

$$\int_{-\infty}^{\infty} f^*(t)\, g(t) \mathrm{d}t = \frac{\int_{-\infty}^{\infty} \hat{f}^*(\omega)\, \hat{g}(\omega) \mathrm{d}\omega}{2\,\pi} \tag{6}$$

The functions f and $\hat{f}$ are referred to as the spatiotemporal and spectral representations, respectively. The $\mathscr{H}_r(\mathbb{R})$-orthonormal expansions are derived from the following rather straight-forward theorem. Let I be a countably infinite index set, typically either $\mathbb{N}$ or $\mathbb{Z}$.

Theorem 2. [Construction of orthonormal bases] *Let the translation-invariant symmetric positive-definite kernel kernel $K \in C(\mathbb{R}) \cap \mathscr{L}_1(\mathbb{R})$ be*

$$r(t, u) = K(t - u) = \int_{-\infty}^{\infty} S(\omega) e^{i\omega h} \mathrm{d}h \tag{7}$$

where its corresponding spectral density is

$$S(\omega) = \frac{\int_{-\infty}^{\infty} K(x) e^{-i\omega x} \mathrm{d}h}{2\pi} \tag{8}$$

and $\{\varphi_m\}_{m \in I}$

$$\int_0^{\infty} \varphi_m(x) \varphi_n(x) \mathrm{d}x = \delta_{n,m} = \begin{cases} 1 & m = n \\ 0 & m \neq n \end{cases} \tag{9}$$

is an orthonormal basis of $\mathscr{L}_2(\mathbb{R})$. *Now, let*

$$h(x) = \int_{-\infty}^{\infty} \sqrt{S(\omega)} e^{i\omega h} \mathrm{d}h \tag{10}$$

be the Fourier transform of the square root of the spectral density

$$\hat{h}(\omega) = \sqrt{S(\omega)} = \frac{\int_{-\infty}^{\infty} h(x) e^{-i\omega x} \mathrm{d}x}{2\pi} \tag{11}$$

Then the convolutions of the orthonormal basis $\mathscr{L}_2(\mathbb{R})$ *with* h

$$\psi_m(t) = \int_{-\infty}^{\infty} h\,(t - \tau)\, \varphi_m(\tau) \mathrm{d}\tau \tag{12}$$

form an orthonormal basis of $\mathscr{H}_r(\mathbb{R})$ *and the kernel* r *has the pointwise convergent expansion*

$$r(t, u) = \sum_{m \in I} \psi_m^*(t)\, \psi_m(u) \quad \forall t, u \in \mathbb{R} \tag{13}$$

Proof. The fact that $r(\cdot,\cdot)$ is symmetric positive-definite implies that the spectral density $S(\omega)$ is real-valued and positive [?, Theorem 6.11] therefore the convolution theorem yields

$$\widehat{\mathcal{H} f}(\omega) = \hat{h}(\omega)\, \hat{f}(\omega) \tag{14}$$

$\forall h$ such that

$$|\hat{h}(\omega)| = \sqrt{S(\omega)} > 0 \forall \omega \in \mathbb{R} \tag{15}$$

where the convolution operator $\mathcal{H}: \mathscr{L}_2(\mathbb{R}) \to \mathscr{L}_2(\mathbb{R})$ is defined via

$$(\mathcal{H} f)(t) = \int_{-\infty}^{\infty} h\,(t - \tau)\, f(\tau) \mathrm{d}\tau \quad \forall t \in \mathbb{R} \tag{16}$$

By the standard characterisation (see [?] or [?, Theorem 10.12]) of the RKHS of a translation-invariant kernel,

$$\langle f, g \rangle_r = \frac{\int_{-\infty}^{\infty} \frac{\hat{f}^*(\omega)\,\hat{g}(\omega)}{\hat{\Phi}(\omega)} \mathrm{d}\omega}{2\pi} \quad \forall f, g \in \mathscr{H}_r(\mathbb{R}) \tag{17}$$

For any $f, g \in \mathscr{L}_2(\mathbb{R})$ the convolution theorem and Plancherel theorem thus give

$$\begin{aligned} \langle \mathcal{H} f, \mathcal{H} g \rangle_r &= \frac{\int_{-\infty}^{\infty} \frac{|\hat{h}(\omega)|^2 \hat{f}^*(\omega)\,\hat{g}(\omega)}{\hat{\Phi}(\omega)} \mathrm{d}\omega}{2\pi} \\ &= \frac{\int_{-\infty}^{\infty} \hat{f}^*(\omega)\,\hat{g}(\omega) \mathrm{d}\omega}{2\pi} \\ &= \langle f, g \rangle_{\mathscr{L}_2(\mathbb{R})} \end{aligned} \tag{18}$$

which shows that $\mathcal{H}$ is an isometry between $\mathscr{L}_2(\mathbb{R})$ and $\mathscr{H}_r(\mathbb{R})$. It follows from (17) that the inverse Fourier transform

$$(\mathcal{H}^{-1} f)(t) = \frac{\int_{-\infty}^{\infty} \frac{\hat{f}(\omega)}{\hat{h}(\omega)} e^{i\omega t} \mathrm{d}\omega}{2\pi} \quad \forall t \in \mathbb{R} \tag{19}$$

defines the inverse of $\mathcal{H}$. Therefore $\mathcal{H}$ and its inverse are constitute an isometric isomorphism and thus maps orthonormal basis of $\mathscr{L}_2(\mathbb{R})$ and $\mathscr{H}_r(\mathbb{R})$ to each other. [?, Section 2.6] Therefore, the kernel has a pointwise convergent expansion of the form (13) for every orthonormal basis of $\mathscr{H}_r(\mathbb{R})$ [?] □

To obtain the spatiotemporal basis functions ψ_m using Theorem 2 either the convolution $\int_{-\infty}^{\infty} h(t-\tau)\varphi_m(\tau)\mathrm{d}\tau$ or the inverse Fourier transform of $\hat{h}(\omega)\,\hat{\varphi}_m(\omega)$ has to be computed. It is therefore necessary to select a basis of $\mathscr{L}_2(\mathbb{R})$ for which either of these operations can be done in closed form.

1.2 On Mercer expansions

Let Ω be a subset of $\mathbb{R}^d$ and $w: \Omega \to [0, \infty)$ a weight function. The Hilbert space $\mathscr{L}_2(\Omega, w)$ is equipped with the inner product

$$\langle f, g \rangle_{\mathscr{L}_2(\Omega, w)} = \int_{\Omega} f^*(t)\, g(t)\, w(t) \mathrm{d}t \tag{20}$$

and consists of all functions $f\colon \mathbb{R} \to \mathbb{C}$ for which the corresponding norm is finite. Suppose that the kernel r is continuous and define the integral operator

$$\mathcal{T}_{r,w} f = \int_{-\infty}^{\infty} r(\cdot, u)\, f(u)\, w(u) \mathrm{d}u \tag{21}$$

Under certain assumptions, Mercer's theorem [?] states that (i) $\mathcal{T}_{r,w}$ has continuous eigenfunctions $\{\vartheta_m\}_{m=0}^{\infty}$ and corresponding positive non-increasing eigenvalues $\{\mu_m\}_{m=0}^{\infty}$ which tend to zero, (ii) $\{\vartheta_m\}_{m=0}^{\infty}$ are an orthonormal basis of $\mathscr{L}_2(\Omega, w)$, and (iii) $\{\sqrt{[b]\mu_m}\vartheta_m\}_{m=0}^{\infty}$ is an orthonormal basis of $\mathscr{H}_r(\Omega)$. Consequently, the kernel has the pointwise convergent *Mercer expansion*

$$r(t, u) = \sum_{m=0}^{\infty} \mu_m \vartheta_m^*(t) \vartheta_m(u) \quad \forall t, u \in \Omega \tag{22}$$

While Mercer's theorem and the eigenvalues of $\mathcal{T}_{r,w}$ constitute a powerful tool for understanding topics such as optimal approximation in $\mathscr{L}_2(\Omega, w)$-norm (e.g., [?, Corollary 4.12] and [?, Section 2.4]) and improved approximation orders in subsets of $\mathscr{H}_r(\Omega)$ [?, Section 11.5], both in theoretical research and practical applications there is often no reason to prefer a Mercer expansion (22) over a generic RKHS-orthonormal expansion (1). For example, the Karhunen–Loève theorem is merely a special case of a more general result that a Gaussian process with covariance kernel r can be expanded in terms of any orthonormal basis of $\mathscr{H}_r(\Omega)$ [?, Chapter III].

Constructing a Mercer expansion by first identifying a convenient weight and then finding the eigendecomposition of the integral operator (21) can be rather involved. **What makes Theorem 2 convenient is therefore that it does *not* require that the expansion be Mercer for some weight.** However, identifying a weight w for which the basis function ψ_m constructed via Theorem 2 are $\mathscr{L}_2(\mathbb{R}, w)$-orthogonal shows that the expansion is Mercer because the $\mathscr{L}_2(\mathbb{R}, w)$-normalised versions of ψ_m are the eigenfunctions of $\mathcal{T}_{r,w}$.

2 Summary of expansions

This section summarises the expansions that we derive using Theorem 2. Each expansion converges pointwise for all $t, u \in \mathbb{R}$. All expansions are for kernels with unit scaling. Expansions of arbitrary scalings, λ, may be obtained by considering the kernel $r(\lambda t, \lambda u)$, for which the corresponding basis functions are $\psi_m(\lambda t)$.

2.1 Gaussian kernel

Expansions for the Gaussian kernel are derived in Section 3. The Gaussian kernel is

$$r(t, u) = e^{-\frac{(t-u)^2}{2}} \tag{23}$$

The functions

$$\psi_m(t) = \sqrt{\frac{2\sqrt{2}}{6^m\, m!3}}\, e^{-\frac{t^2}{3}}\, \mathrm{H}_m\!\left(\frac{2\,t}{\sqrt{3}}\right) \quad \forall m \in \mathbb{N}_0 \tag{24}$$

form an orthonormal basis of the RKHS and the kernel has the expansion

$$r(t,u) = \sum_{m=0}^{\infty} \psi_m(t)\, \psi_m(u) \tag{25}$$

for all $t, u \in \mathbb{R}$. This expansion is a special case of the well-known Mercer expansion of the Gaussian kernel [?, Section 12.2.1]. The basis functions (24) are orthogonal in $\mathscr{L}_2(\mathbb{R}, w_\alpha)$ for the weight function

$$w_\alpha(t) = \frac{\alpha}{\sqrt{\pi}}\, e^{-\alpha^2 t^2} \tag{26}$$

with

$$\alpha = \sqrt{\frac{[b]2}{3}} \tag{27}$$

3 Expansion of the Gaussian kernel

The Gaussian kernel and its Fourier transform are

$$r(t,u) = e^{-\frac{(t-u)^2}{2}} \quad \text{and} \quad S(\omega) = \sqrt{2\,\pi}\; e^{-\frac{\omega^2}{2}} \tag{28}$$

A square-root is

$$\hat{h}(\omega) = \sqrt{S(\omega)} = (2\,\pi)^{1/4}\, e^{-\frac{\omega^2}{4}} \tag{29}$$

so that taking the inverse Fourier transform gives the function h in Theorem 2 as

$$h(t) = 2^{1/4}\, \pi^{-1/4}\, e^{-t^2} \tag{30}$$