MATH 590: Meshfree Methods

Chapter 2 — Part 3: Native Space for Positive Definite Kernels

Greg Fasshauer

Department of Applied Mathematics
Illinois Institute of Technology

Fall 2014

Outline

In this section we will show that every positive definite kernel can indeed be associated with a reproducing kernel Hilbert space — its native space.

First, we note that the definition of an RKHS tells us that $\mathcal{H}_K(\Omega)$ contains all functions of the form

$$f = \sum_{j=1}^{N} c_j K(\cdot, \boldsymbol{x}_j)$$

provided $\boldsymbol{x}_j \in \Omega$.

Using the properties of RKHSs established earlier along with the form of f just mentioned we have that

$$\begin{aligned}\|f\|^2_{\mathcal{H}_K(\Omega)} &= \langle f, f\rangle_{\mathcal{H}_K(\Omega)} = \left\langle \sum_{i=1}^N c_i K(\cdot, \boldsymbol{x}_i), \sum_{j=1}^N c_j K(\cdot, \boldsymbol{x}_j) \right\rangle_{\mathcal{H}_K(\Omega)} \\ &= \sum_{i=1}^N \sum_{j=1}^N c_i c_j \langle K(\cdot, \boldsymbol{x}_i), K(\cdot, \boldsymbol{x}_j)\rangle_{\mathcal{H}_K(\Omega)} \\ &= \sum_{i=1}^N \sum_{j=1}^N c_i c_j K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \boldsymbol{c}^T \mathsf{K} \boldsymbol{c}.\end{aligned}$$

So — for these special types of f — we can easily calculate the Hilbert space norm of f.
In particular, if $f = s$ is a kernel-based interpolant, i.e., $\boldsymbol{c} = \mathsf{K}^{-1}\boldsymbol{y}$, then we also have

$$\|s\|^2_{\mathcal{H}_K(\Omega)} = \boldsymbol{y}^T \mathsf{K}^{-T} \mathsf{K} \mathsf{K}^{-1} \boldsymbol{y} = \boldsymbol{y}^T \mathsf{K}^{-1} \boldsymbol{y}.$$

Therefore, we define the (possibly infinite-dimensional) space of all linear combinations

$$H_K(\Omega) = \operatorname{span}\{K(\cdot, \boldsymbol{z}) : \ \boldsymbol{z} \in \Omega\} \tag{1}$$

with an associated bilinear form $\langle \cdot, \cdot \rangle_K$ given by

$$\left\langle \sum_{i=1}^{N} c_i K(\cdot, \boldsymbol{x}_i), \sum_{j=1}^{M} d_j K(\cdot, \boldsymbol{z}_j) \right\rangle_K = \sum_{i=1}^{N} \sum_{j=1}^{M} c_i d_j K(\boldsymbol{x}_i, \boldsymbol{z}_j) = \boldsymbol{c}^T \mathsf{K} \boldsymbol{d}.$$

Remark

Note that this definition implies that a general element in $H_K(\Omega)$ has the form (where $N = \infty$ is allowed)

$$f = \sum_{j=1}^{N} c_j K(\cdot, \boldsymbol{x}_j).$$

However, not only the coefficients c_j, but also the specific value of N and choice of points $\boldsymbol{x}_j$ will vary with f.

Theorem

If $K : \Omega \times \Omega \to \mathbb{R}$ is a symmetric strictly positive definite kernel, then the bilinear form $\langle \cdot, \cdot \rangle_K$ defines an inner product on $H_K(\Omega)$.

Furthermore, $H_K(\Omega)$ is a pre-Hilbert space with reproducing kernel K.

Remark

A pre-Hilbert space is an inner product space whose completion is a Hilbert space.

Proof.

$\langle\cdot,\cdot\rangle_K$ is obviously bilinear and symmetric.
We just need to show that $\langle f, f\rangle_K > 0$ for nonzero $f \in H_K(\Omega)$.
Any such f can be written in the form

$$f = \sum_{j=1}^{N} c_j K(\cdot, \boldsymbol{x}_j), \qquad \boldsymbol{x}_j \in \Omega.$$

Then

$$\langle f, f\rangle_K = \sum_{i=1}^{N}\sum_{j=1}^{N} c_i c_j K(\boldsymbol{x}_i, \boldsymbol{x}_j) > 0$$

since K is strictly positive definite.
The reproducing property follows from

$$\langle f, K(\cdot, \boldsymbol{x})\rangle_K = \sum_{j=1}^{N} c_j K(\boldsymbol{x}, \boldsymbol{x}_j) = f(\boldsymbol{x}).$$

□

Since we just showed that $H_K(\Omega)$ is a pre-Hilbert space, i.e., need not be complete, we now first form the completion $\widetilde{H}_K(\Omega)$ of $H_K(\Omega)$ with respect to the K-norm $\|\cdot\|_K$ ensuring that

$$\|f\|_K = \|f\|_{\widetilde{H}_K(\Omega)} \quad \text{for all } f \in H_K(\Omega).$$

In general, this completion will consist of equivalence classes of Cauchy sequences in $H_K(\Omega)$, so that we can obtain the native space $\mathcal{N}_K(\Omega)$ of K as a space of continuous functions with the help of the point evaluation functional (which extends continuously from $H_K(\Omega)$ to $\widetilde{H}_K(\Omega)$), i.e., the (values of the) continuous functions in $\mathcal{N}_K(\Omega)$ are given via the right-hand side of

$$\delta_{\boldsymbol{x}}(f) = \langle f, K(\cdot, \boldsymbol{x})\rangle_K, \qquad f \in \widetilde{H}_K(\Omega).$$

Remark

The technical details concerned with this construction are discussed in [Wen05].

In summary, we now know that the native space $\mathcal{N}_K(\Omega)$ is given by (continuous functions in) the completion of

$$H_K(\Omega) = \operatorname{span}\{K(\cdot, \boldsymbol{z}) : \ \boldsymbol{z} \in \Omega\}$$

— a not very intuitive definition of a function space.

In the special case when we are dealing with strictly positive definite (translation invariant) functions $\Phi(\boldsymbol{x} - \boldsymbol{z}) = K(\boldsymbol{x}, \boldsymbol{z})$ and when $\Omega = \mathbb{R}^d$ we get a characterization of native spaces in terms of Fourier transforms.

We present the following theorem without proof (for details see [Wen05]).

Theorem

Suppose $\Phi \in C(\mathbb{R}^d) \cap L_1(\mathbb{R}^d)$ *is a real-valued strictly positive definite function. Define*

$$\mathcal{G} = \{f \in L_2(\mathbb{R}^d) \cap C(\mathbb{R}^d) : \frac{\hat{f}}{\sqrt{\hat{\Phi}}} \in L_2(\mathbb{R}^d)\}$$

and equip this space with the bilinear form

$$\langle f, g\rangle_{\mathcal{G}} = \frac{1}{\sqrt{(2\pi)^d}} \langle \frac{\hat{f}}{\sqrt{\hat{\Phi}}}, \frac{\hat{g}}{\sqrt{\hat{\Phi}}} \rangle_{L_2(\mathbb{R}^d)} = \frac{1}{\sqrt{(2\pi)^d}} \int_{\mathbb{R}^d} \frac{\hat{f}(\omega)\overline{\hat{g}(\omega)}}{\hat{\Phi}(\omega)} d\omega.$$

Then $\mathcal{G}$ *is a real Hilbert space with inner product* $\langle \cdot, \cdot \rangle_{\mathcal{G}}$ *and reproducing kernel* $\Phi(\cdot - \cdot)$. *Hence,* $\mathcal{G}$ *is the native space of* Φ *on* $\mathbb{R}^d$, *i.e.,* $\mathcal{G} = \mathcal{N}_\Phi(\mathbb{R}^d)$ *and both inner products coincide.*
In particular, every $f \in \mathcal{N}_\Phi(\mathbb{R}^d)$ *can be recovered from its Fourier transform* $\hat{f} \in L_1(\mathbb{R}^d) \cap L_2(\mathbb{R}^d)$.

Mercer's theorem allows us to construct the native space/RKHS $\mathcal{H}_K(\Omega)$ for any continuous positive definite kernel K by representing the functions in $\mathcal{H}_K$ as infinite linear combinations of the eigenfunctions φ_n of the Hilbert–Schmidt integral operator $\mathcal{K}$, i.e.,

$$\mathcal{H}_K = \left\{ f : \ f = \sum_{n=1}^{\infty} c_n \varphi_n \right\}.$$

Thus the eigenfunctions $\{\varphi_n\}_{n=1}^{\infty}$ of $\mathcal{K}$ provide an alternative basis for $\mathcal{H}_K(\Omega)$ instead of the standard $\{K(\cdot, \boldsymbol{z}) : \ \boldsymbol{z} \in \Omega\}$.
For any fixed $\boldsymbol{x}$, the corresponding "basis transformation" is given by the Mercer series

$$K(\cdot, \boldsymbol{z}) = \sum_{n=1}^{\infty} \lambda_n \varphi_n \varphi_n(\boldsymbol{z}).$$

This shows that indeed $K(\cdot, \boldsymbol{z}) \in \mathcal{H}_K(\Omega)$.

The inner product for $\mathcal{H}_K(\Omega)$ can now be written as

$$\langle f, g\rangle_{\mathcal{H}_K(\Omega)} = \left\langle \sum_{m=1}^{\infty} c_m\varphi_m, \sum_{n=1}^{\infty} d_n\varphi_n \right\rangle_{\mathcal{H}_K(\Omega)} = \sum_{n=1}^{\infty} \frac{c_n d_n}{\lambda_n},$$

where we used the fact that the eigenfunctions are not only L_2-orthonormal, but also orthogonal in $\mathcal{H}_K(\Omega)$, i.e.,

$$\langle \varphi_m, \varphi_n\rangle_{\mathcal{H}_K(\Omega)} = \frac{\delta_{mn}}{\sqrt{\lambda_m}\sqrt{\lambda_n}}.$$

We can also verify that K is indeed the reproducing kernel of $\mathcal{H}_K$ since the Mercer series of K and the orthogonality of the eigenfunctions imply

$$
\begin{aligned}
\langle f, K(\cdot, \boldsymbol{x}) \rangle_{\mathcal{H}_K(\Omega)} &= \left\langle \sum_{m=1}^{\infty} c_m \varphi_m, \sum_{n=1}^{\infty} \lambda_n \varphi_n \varphi_n(\boldsymbol{x}) \right\rangle_{\mathcal{H}_K(\Omega)} \\
&= \sum_{n=1}^{\infty} \frac{c_n \lambda_n \varphi_n(\boldsymbol{x})}{\lambda_n} \\
&= \sum_{n=1}^{\infty} c_n \varphi_n(\boldsymbol{x}) \\
&= f(\boldsymbol{x}).
\end{aligned}
$$

Finally (cf. [Wen05]), we can also describe the RKHS $\mathcal{H}_K$ as

$$\mathcal{H}_K(\Omega) = \left\{ f \in L_2(\Omega) : \sum_{n=1}^{\infty} \frac{1}{\lambda_n} \left|\langle f, \varphi_n \rangle_{L_2(\Omega)}\right|^2 < \infty \right\}$$

with inner product

$$\langle f, g \rangle_{\mathcal{H}_K(\Omega)} = \sum_{n=1}^{\infty} \frac{1}{\lambda_n} \langle f, \varphi_n \rangle_{L_2(\Omega)} \langle g, \varphi_n \rangle_{L_2(\Omega)}, \qquad f, g \in \mathcal{H}_K(\Omega).$$

Remark

Since $\mathcal{H}_K(\Omega)$ is a subspace of $L_2(\Omega)$ this latter interpretation corresponds to the identification of the coefficients in the eigenfunction expansion of an $f \in \mathcal{H}_K(\Omega)$ with the generalized Fourier coefficients of f, i.e., $c_n = \langle f, \varphi_n \rangle_{L_2(\Omega)}$.

The theorem characterizing the native spaces of translation invariant functions on all of $\mathbb{R}^d$ shows that these spaces can be viewed as a generalization of standard Sobolev spaces.
For $m > d/2$ the Sobolev space W_2^m can be defined as (see, e.g., [AF03])

$$W_2^m(\mathbb{R}^d) = \{f \in L_2(\mathbb{R}^d) \cap C(\mathbb{R}^d) : \ \hat{f}(\cdot)(1 + \|\cdot\|_2^2)^{m/2} \in L_2(\mathbb{R}^d)\}. \quad (2)$$

Remark

One also frequently sees the definition

$$W_2^m(\Omega) = \{f \in L_2(\Omega) \cap C(\Omega) : \ D^\alpha f \in L_2(\Omega) \text{ for all } |\alpha| \le m, \ \alpha \in \mathbb{N}^d\}, \quad (3)$$

which applies whenever $\Omega \subset \mathbb{R}^d$ is a bounded domain.

Example

Using the notation $r = \|\boldsymbol{x}\|$ and modified Bessel functions of the second kind $K_{d/2-\beta}$, the Matérn kernels

$$\kappa_\beta(r) = \frac{K_{d/2-\beta}(r)}{r^{d/2-\beta}}, \qquad \beta > \frac{d}{2},$$

have Fourier transform

$$\hat{\kappa}_\beta(\|\boldsymbol{\omega}\|) = \left(1 + \|\boldsymbol{\omega}\|^2\right)^{-\beta}.$$

So it can immediately be seen that their native space is

$$\mathcal{N}_K(\mathbb{R}^d) = W_2^\beta(\mathbb{R}^d) \quad \text{with } \beta > d/2,$$

which is why some people refer to the Matérn kernels as Sobolev splines.

The native spaces for Gaussians is rather small.

Example

According to the Fourier transform characterization of the native space, for Gaussians the Fourier transform of $f \in \mathcal{N}_\Phi(\Omega)$ must decay faster than the Fourier transform of the Gaussian (which is itself a Gaussian).

The native space of Gaussians was recently characterized in [FY11] in terms of an (infinite) vector of differential operators. In fact, the native space of Gaussians is contained in the Sobolev space $W_2^m(\mathbb{R}^d)$ for any m.

It is known that, even though the native space of Gaussians is small, it contains the important class of so-called band-limited functions, i.e., functions whose Fourier transform is compactly supported.

Band-limited functions play an important role in sampling theory.

References I

[AF03] Robert Alexander Adams and John J. F. Fournier, *Sobolev Spaces*, Academic Press, 2003.

[FY11] G. E. Fasshauer and Qi Ye, *Reproducing kernels of generalized Sobolev spaces via a Green function approach with distributional operators*, Numerische Mathematik **119** (2011), 585–611.

[Wen05] H. Wendland, *Scattered Data Approximation*, Cambridge Monographs on Applied and Computational Mathematics, vol. 17, Cambridge University Press, 2005.