Optimal Recovery in Translation–invariant Spaces of Functions

R. Schaback
[1]Institut für Numerische und Angewandte Mathematik
Universität Göttingen
Lotzestr. 16–18
D–37083 Göttingen
E-mail: schaback@namu01.gwdg.de

Dedicated to Ted Rivlin on the occasion of his 70th birthday

Starting from optimal recovery (in the sense of Micchelli, Rivlin, and Winograd) of functions in reproducing kernel Hilbert spaces from function values at scattered data points, we show that any continuously embedded translation–invariant Hilbert subspace H of $L_2(\mathbb{R}^d)$ that allows continuous point evaluation is necessarily principal, i.e. it is the native space of a positive definite function Φ. If, in addition, H is invariant under orthogonal transformations of $\mathbb{R}^d$, the function Φ necessarily is a positive definite radial basis function. This serves to show that positive definite radial basis functions and their native spaces arise very naturally and are by no means exotic theoretical constructions.

1 Introduction

Let H be a real Hilbert space of real–valued functions on some domain Ω, and assume that point–evaluation functionals are continuous, which is a very reasonable assumption for applications. Then for any $x \in \Omega$ there is a Riesz representer $\lambda_x \in H$ for the point evaluation functional δ_x at x, i.e.

$$\delta_x f = f(x) = (\lambda_x, f)_H \tag{1.1}$$

for all $f \in H$, $x \in \Omega$. Now the function

$$K(x,y) := (\lambda_x, \lambda_y)_H = \lambda_y(x) = K(y,x) \tag{1.2}$$

on $\Omega \times \Omega$ is a reproducing kernel for H, because (1.1) can be rewritten as

$$f(x) = (K(x, \cdot), f(\cdot))_H \tag{1.3}$$

for all $x \in \Omega$, $f \in H$. There is a well–established theory of reproducing–kernel Hilbert spaces (see e.g. [2],[12] and applications in [3], [5], [6], and [8]), but we concentrate here on its implications for the theory of multivariate interpolation as an optimal recovery process in the sense of Micchelli, Rivlin, and Winograd.

To this end we additionally assume that any number N of different point evaluation functionals is linearly independent in the dual of H. This has several equivalent formulations, where $X = \{x_1, \ldots, x_N\} \subseteq \Omega$ always denotes an arbitrary set of N distinct points in Ω:

- The functions $\lambda_{x_1}, \ldots, \lambda_{x_N}$ are linearly independent in H.
- Any interpolation problem on X is uniquely solvable by functions from H.
- The matrix $A_X = (\lambda_{x_j}, \lambda_{x_k})_{1 \leq j,k \leq N} = (K(x_j, x_k))_{1 \leq j,k \leq N}$ is positive definite.

The third version allows to forget about reproducing kernel Hilbert spaces by using the following notion:

Definition .1
A real–valued symmetric function K on $\Omega \times \Omega$ is **positive definite**, *if all matrices of the form*

$$A_X = (K(x_j, x_k))_{1 \leq j,k \leq N} \tag{1.4}$$

are positive definite where N and $X = \{x_1, \ldots, x_N\} \subseteq \Omega$ vary.

In section 3 we shall address the inverse question whether any symmetric positive definite function on $\Omega \times \Omega$ is a kernel function on some Hilbert space H of functions on Ω. See Schoenberg [15] for positive definite functions arising in the problem of isometric embedding of metric spaces into Hilbert spaces.

The problem of optimal recovery of a function $f \in H$ from data $f(x_1), \ldots, f(x_N)$ on some set $X = \{x_1, \ldots, x_N\}$ can be rephrased as the minimization problem

$$\|s\|_H = \underset{\substack{s|_X = f|_X \\ s \in H}}{\text{Min!}} \ .$$

By standard arguments the solution $s_{f,X}$ of this problem can be shown to exist uniquely, and to be of the form

$$s_{f,X} = \sum_{j=1}^{N} a_j \lambda_{x_j} = \sum_{j=1}^{N} a_j K(x_j, \cdot) \tag{1.5}$$

where the coefficents a_j are obtained by solving the system

$$\sum_{j=1}^{N} K(x_j, x_k) a_j = f(x_k), \qquad 1 \le k \le N \tag{1.6}$$

with the positive definite matrix A_X of (1.4).

For completeness, we add some of the well–known additional properties of this optimal recovery scheme (see [3], [5], [6], and [8], for instance). If x is another point of Ω, one can ask for the best reconstruction of $f(x)$ from known information $f(x_1), \ldots, f(x_N)$ by linear quasi–interpolation formulas of the form

$$q_f(x) := \sum_{j=1}^{N} f(x_j) \cdot u_j(x).$$

where, since x is fixed, we can view $u_j(x)$ just as real numbers. The error is representable as

$$f(x) = \left(\lambda_x - \sum_{j=1}^{N} u_j(x) \lambda_{x_j}, f \right)_H$$

and has the bound

$$|f(x) - q_f(x)| \le \|f\|_H \cdot \left\| \lambda_x - \sum_{j=1}^{N} u_j(x) \lambda_{x_j} \right\|_H$$

which can be optimized by a best approximation of λ_x by $\lambda_{x_1}, \ldots, \lambda_{x_N}$. The unique solution is given by $u_j^*(x)$ satisfying

$$\sum_{j=1}^{N} K(x_j, x_k) u_j^*(x) = K(x, x_k), \qquad 1 \le k \le N$$

and it is evident that $u_1^*, \ldots, u_N^*$, now viewed as functions in H, provide the Lagrange reformulation of the interpolant we already know:

$$s_{f,X} = \sum_{j=1}^{N} f(x_j) u_j^* .$$

Thus the norm–minimal recovery process also minimizes the pointwise error.

In addition, $s_{f,X}$ is the best approximation to f from the span S_X of $K(x_1, \cdot), \ldots, K(x_N, \cdot)$, since the representation (1.5) implies

$$(s_{h,X}, g)_H = \sum_{j=1}^{N} a_j g(x_j)$$

for all $g, h \in H$, and we find

$$(s_{h,X}, f - s_{f,X})_H = 0$$

for all $s_{h,X} \in S_X$.

These three simultaneous optimality properties make interpolation via positive definite functions in the sense of (1.5) a very useful tool for recovery of functions, provided that the system (1.6) can be solved at reasonable computational cost. This occurs if the matrix A_X of (1.4) is sparse and well–conditioned (see [13], [14] for overviews), or if special techniques like multipole expansions are applied in order to increase computational efficiency (see [1],[11]).

The next section will demonstrate how certain invariance properties of the Hilbert space H affect the kernel function K, while section 3 will show that any positive definite function K arises as a kernel of some Hilbert space. Finally, in section 4 we specialize to the case of translation–invariant Hilbert subspaces S of $L_2(\mathbb{R}^d)$ and show that any such space S is a reproducing kernel Hilbert space which can alternatively be written as a space of functions whose L_2 Fourier transforms are in a weighted L_2 space. The connection to the general setting is that $K(x, y)$ takes the form $K(x, y) = \Phi(x - y)$ for $\Phi : \mathbb{R}^d \to \mathbb{R}$ with $\Phi(-\cdot) = \Phi(\cdot)$ such that $\hat{\Phi}$ is nonnegative and $w = 1/\hat{\Phi}$ is the weight function in the alternative representation

$$S = \{f \in L_2(\mathbb{R}^d) : \int |\hat{f}(\omega)|^2 w(\omega) d\omega < \infty\}$$

that arises in the well–established theory of radial basis functions.

2 Invariance properties

We now assert that any invariance property of H carries over to K. More precisely, let us assume that the space H is invariant under a group $\mathcal{T}$ of transformations on the domain Ω, i.e.

$$(f(T\cdot), g(T\cdot))_H = (f, g)_H$$

for all $T \in \mathcal{T}$, $f, g \in H$, where all $T \in \mathcal{T}$ are mappings $\Omega \to \Omega$ such that $f(T\cdot) \in H$ for all $f \in H$, $T \in \mathcal{T}$. Then we have

$$\begin{aligned}
\lambda_x(Ty) &= (\lambda_x, \lambda Ty)_H \\
&= (\lambda_y, \lambda_x(T\cdot))_H \\
&= (\lambda_y(T^{-1}\cdot), \lambda_x)_H
\end{aligned}$$

for all $x, y \in \Omega$, $T \in \mathcal{T}$, such that

$$\lambda_{Ty}(Tz) = \lambda_y(z) = K(y, z) = K(Ty, Tz) \tag{2.1}$$

for all $y, z \in \Omega$, $T \in \mathcal{T}$, proving invariance of K with respect to $\mathcal{T}$. We now consider four important examples.

Example 1
If Ω is itself a real vector space, e.g. $\Omega = \mathbb{R}^d$, and if $\mathcal{T}$ is the group of all translations $T_x(z) = z - x$, then

$$K(y, z) = K(y - z, 0)$$

follows from setting $T = T_z$ in (2.1). Then the space S_X of functions of the form (1.5) necessarily is a space of translates of the function $\Phi(\cdot) := K(\cdot, 0)$. Thus interpolation by translates of a fixed function is a very natural (and in a multiple sense optimal) approach to recovery in the sense of Micchelli, Rivlin, and Winograd of functions in translation–invariant spaces.

Example 2
If, more generally, Ω is a topological group with unity 1, the function K takes the form

$$K(y, z) = K(y \cdot z^{-1}, 1),$$

and this approach was treated in detail by Gutzmer [7].

Example 3
If we take $\Omega = \mathbb{R}^d$ and allow Euclidean rigid–body transformations of the form

$$T_{Q,x}(y) := Qy - x$$

where Q is an arbitrary orthogonal d by d matrix, then by suitable choice of Q we get

$$K(y, z) = K(y - z, 0) = K(\|y - z\|_2 \cdot e_1, 0) =: \phi(\|y - z\|_2) \tag{2.2}$$

for all $y, z \in \mathbb{R}^d$, where e_1 is the first unit vector and $\phi : \mathbb{R}_{\geq 0} \to \mathbb{R}$ is a scalar function. Thus the space S_X now is generated by translates of a "radial basis function" ϕ.

Example 4
If we work on the sphere $\Omega = S^{d-1} \subset \mathbb{R}^d$ and again allow all orthogonal transformations on $\mathbb{R}^d$, we end up with a kernel of the special form

$$K(y, z) = \phi(y^T z), \qquad y, z \in S^{d-1} \subset \mathbb{R}^d$$

where $\phi : [-1, +1] \to I\!R$ is a real–valued function. This special case was treated in Schoenberg [16].

Note that the general definition of positive definiteness of K is applicable in all of these cases. The solution of the system (1.6) will uniquely exist if K is positive definite, and there is no need for any Hilbert space interpretation of K, if one just wants to work with functions of the form (1.5). Thus the construction of "nice" positive definite functions can be carried out independently, and this was actually done in a series of recent papers on the radial case [13], [17], [18].

3 Native spaces

But the construction of a symmetric and positive definite function K will always generate a suitable reproducing–kernel Hilbert space H such that (1.3) holds. This was already done in special situations (see the surveys [13], [14] and their references), but the technique works in general.

If $K : \Omega \times \Omega \to I\!R$ is symmetric and positive definite on some domain Ω, we first define the space

$$S = \left\{ \sum_{j=1}^{N} a_j K(x_j, \cdot) : a_j \in I\!R,\ N \in I\!N,\ x_j \in \Omega \right\}.$$

Any function from S has a unique representation

$$f_{a,X} = \sum_{j=1}^{N} a_j K(x_j, \cdot) \tag{3.1}$$

where $a = (a_1, \ldots, a_N) \in I\!R^N$ and $X = \{x_1, \ldots, x_N\}$ is a set of N **distinct** points of Ω. If we define $\lambda_x = K(x, \cdot)$ and

$$(\lambda_x, \lambda_y)_S := K(x, y)$$

for all $x, y \in \Omega$, this definition extends to a scalar product on all of S, because K is positive definite. Then (1.3) holds by definition for all functions of the form (3.1), such that S becomes a reproducing–kernel semi–Hilbert space. Its closure H can be abstractly defined, and if $i : S \to H$ is the canonical injection, then for all abstract elements $f \in H$ and all points $x \in \Omega$ one can **define** a value

$$f(x) := (i\lambda_x, f)_S$$

that turns f into a function on Ω and H into a reproducing–kernel Hilbert space of functions on Ω having K as its kernel. In this sense one can call H the "native" space for the positive definite symmetric function K. Note that this technique applies to all of the special cases described in Section 2. In particular, the positive definite functions on the topological group $SO(3)$ constructed by Gutzmer [7] generate Hilbert spaces of functions on $SO(3)$ that deserve further study.

The functions f in H will have nice localization properties, if the value $f(x)$ at some $x \in \Omega$ is independent of values of f at points far from x. From (1.1) we can see that this requires the representer λ_x to be locally supported around x, and then (1.2) implies that $K(x,y)$ should be small if x and y are far away from each other. In the Hilbert space dual H' of H the point evaluation functionals δ_x and δ_y will then be orthogonal. So far we have left open what "near" means in Ω. The standard example is $\Omega = \mathbb{R}^d$ with Euclidean distance and Euclidean invariance of H. Then K is radial in the sense of (2.2) and ϕ should be compactly supported around the origin. Examples of such functions, consisting of a single polynomial piece each, were given in [13] and by Wu [18] and Wendland [17].

4 Translation Invariant Native Subspaces of $L_2(\mathbb{R}^d)$

To be able to apply Fourier transform techniques we now specialize to translation–invariant Hilbert subspaces H of $L_2(\mathbb{R}^d)$. We assume the embedding $A : H \to L_2(\mathbb{R}^d)$ to be continuous and point evaluation functionals to be linearly independent. Since $L_2(\mathbb{R}^d)$ itself is translation–invariant, the embedding A and its adjoint A^* are translation–invariant, i.e. they commute with the translation operator

$$(T_x f)(y) := f(y-x), \qquad x, y \in \mathbb{R}^d,\ f \in L_2(\mathbb{R}^d)$$

written as an operator on H or $L_2(\mathbb{R}^d)$. This is trivial for A and follows for A^* from

$$\begin{aligned} (A^*T_x f, T_x g)_H &= (T_x f, AT_x g)_{L_2} \\ = (T_x f, T_x A g)_{L_2} &= (f, Ag)_{L_2} \\ = (A^* f, g)_H &= (T_x A^* f, T_x g)_H \end{aligned}$$

for all $g \in H,\ f \in L_2,\ x \in \mathbb{R}^d$.

Now for all $f \in L_2$ and $x \in I\!\!R^d$ we get the representation

$$\begin{aligned}(A^*f)(x) &= (\lambda_x, A^*f)_H = (A\lambda_x, f)_{L_2}\\ &= \int_{I\!\!R^d} f(y)\lambda_x(y)dy\\ &= \int_{I\!\!R^d} f(y)\lambda_0(y-x)dy\\ &= (f * \lambda_0(-\cdot))(x)\end{aligned}$$

of A^* as a convolution with $\lambda_0(-\cdot)$. By the way, symmetry of K implies $\lambda_0(-\cdot) = \lambda_0(\cdot)$ and $\hat{\lambda}_0(\cdot) \in I\!\!R$. Taking Fourier transforms we get

$$\widehat{A^*f} = \hat{f} \cdot \hat{\lambda}_0 \tag{4.1}$$

and the scalar product on H can be rewritten on the subspace $A^*(L_2) \subseteq H$ as

$$\begin{aligned}(A^*f, A^*g)_H &= (f, AA^*g)_{L_2} = (\hat{f}, \widehat{AA^*g})_{L_2} = (\hat{f}, \widehat{A^*g})_{L_2}\\ &= \int_{I\!\!R^d} \hat{f} \cdot \overline{\hat{g}} \cdot \hat{\lambda}_0\end{aligned}$$

for all $f, g \in L_2(I\!\!R^d)$. Since any $f \in A^*(L_2)$ can be written as $f = A^*h_f$ with some $h_f \in L_2(I\!\!R^d)$,

$$\begin{aligned}(f,g)_H &= (A^*h_f, A^*h_g)_H\\ &= \int_{I\!\!R^d} \hat{h}_f \cdot \overline{\hat{h}_g} \cdot \widehat{\lambda_0}\end{aligned}$$

is a representation of the abstract inner product on H when restricted to the subspace $A^*(L_2)$ by an integral over L_2 functions. Since one can admit any L_2 function for h_f and h_g, we get positivity of $\widehat{\lambda_0}$, except for (possibly) a set of measure zero in $I\!\!R^d$. Thus we can use (4.1) to write

$$\hat{f} = \widehat{h_f}\widehat{\lambda_0} \tag{4.2}$$

for all $f \in A^*(L_2) \subseteq H \subseteq L_2$, and

$$(f,g)_H = \int_{I\!\!R^d} \frac{\hat{f} \cdot \overline{\hat{g}}}{\widehat{\lambda_0}}$$

for all $f, g \in A^*(L_2)$. This induces an intermediate Hilbert space

$$\overline{A^*(L_2)} = H_1 = \{f \in L_2(I\!\!R^d) : |\hat{f}| \setminus \sqrt{\hat{\lambda}_0} \in L_2(I\!\!R^d)\}$$

between $A^*(L_2)$ and H, but since $(A^*(L_2))^\perp = \{0\}$ in H we can identify H_1 and H. This finally proves

Theorem .2
Any translation–invariant continuously embedded Hilbert subspace H of $L_2(\mathbb{R}^d)$ which admits continuous and linear independent point evaluation functionals is the native space of a positive definite L_2 function $\Phi : \mathbb{R}^d \to \mathbb{R}^d$ with $\Phi(-\cdot) = \Phi(\cdot)$ and can be written as the space

$$H = \left\{ f \in L_2(\mathbb{R}^d) : \int \frac{|\hat{f}(\omega)|^2}{\hat{\Phi}(\omega)}\, d\omega < \infty \right\},$$

while $\hat{\Phi}$ is positive almost everywhere. If H is invariant under Euclidean transformations on $\mathbb{R}^d$, then Φ is a radial function.

In view of the definition of principal shift–invariant spaces as in [4] we can interpret Theorem .2 as a general sufficient criterion for translation–invariant subspaces of $L_2(\mathbb{R}^d)$ to be principal and to be generated by a radial basis function.

References

[1] R.K. Beatson and G.N. Newsam, Fast evaluation of radial basis functions: I. Advances in the theory and applications of radial basis functions. Comput. Math. Appl. 24 (1992), no. 12, 7–19

[2] S. Bergman, The kernel function and conformal mapping, Math. Surveys, 5 (1950), Amer. Math. Soc., New York.

[3] C. de Boor and R.E. Lynch, On splines and their minimum properties, J. Math. Mech. 15 (1966), 953–969.

[4] C. de Boor and A. Ron, Fourier analysis of approximation orders from principal shift–invariant spaces, Constructive Approximation, 8 (1992) 427–462.

[5] N. Dyn, Minimal interpolation and approximation in Hilbert spaces, SIAM J. Numer. Anal. 8 (1971) 585–597.

[6] M. Golomb and H.F. Weinberger, Optimal approximation and error bounds, in *On Numerical Approximation*, R.F. Langer ed., Univ. of Wisconsin Press, Madison, 1959, pp. 117–190.

[7] T. Gutzmer, Interpolation by positive definite functions on locally compact groups with application to $SO(3)$, to appear in Results in Mathematics.

[8] J. Meinguet, Optimal approximation and error bounds in semi–normed spaces, Numer. Math., 10 (1987), 370–388.

[9] C.A. Micchelli and T.J. Rivlin, A survey of optimal recovery, in *Optimal Estimation in Approximation Theory*, C.A. Micchelli and T.J. Rivlin eds., Plenum Press, 1977, pp. 1–54.

[10] C.A. Micchelli, T.J. Rivlin and S. Winograd, Optimal recovery of smooth functions, Numer. Math., 260 (1976) 191–200.

[11] M.J.D. Powell, Truncated Laurent expansions for the fast evaluation of thin plate splines. Algorithms for Approximation, III (Oxford, 1992). Numer. Algorithms 5 (1993), no. 1–4, 99–120.

[12] S. Saito, Theory of reproducing kernels and its applications, Pitman Research Notes 189, Longman: Harlow 1988

[13] R. Schaback, Creating surfaces from scattered data using radial basis functions, in *Mathe-*

matical Methods for Curves and Surfaces, M. Daehlen, T. Lyche and L.L. Schumaker eds., Vanderbilt University Press, Nashville, TN, 1995, 477–496.

[14] R. Schaback, Multivariate interpolation and approximation by translates of a basis function, to appear in: Approximation Theory VIII, Charles K. Chui and L.L. Schumaker eds., World Scientific Publishing Co., Inc.

[15] I.J. Schoenberg, Metric spaces and positive definite functions, Trans. Amer. Math. Soc., 44 (1938) 522–536.

[16] I.J. Schoenberg, Positive definite functions on spheres, Duke Math. J., 9 (1942), 96–108.

[17] H. Wendland, Piecewise polynomial, positive definite, and compactly supported radial functions of minimal degree, Preprint, University of Göttingen, 1995.

[18] Z. Wu, Multivariate compactly supported positive definite radial functions, Preprint, University of Göttingen, 1994. To Appear in AICM.