# Dual Bases and Discrete Reproducing Kernel Hilbert Spaces: A Unified Framework for RBF and MLS Approximation

G. E. Fasshauer$^{(1)}$

**Abstract:** *We present a new framework of dual bases and discrete reproducing kernel Hilbert spaces that can be used to describe both least squares radial basis function approximation and (polynomial) moving least squares approximation. Standard radial basis function interpolation is included as a special case.*

*Keywords:* Radial basis functions, moving least squares, dual bases, reproducing kernel Hilbert spaces.

## 1 Introduction

**Definition 1.1** *Let* $\mathcal{H}$ *be a real Hilbert space of functions* $f : \Omega \to \mathbb{R}$*. A function* $K : \Omega \times \Omega \to \mathbb{R}$ *is called* reproducing kernel for $\mathcal{H}$ *if*

1. $K(\cdot,x) \in \mathcal{H}$ *for all* $x \in \Omega$*,*
2. $f(x) = \langle f, K(\cdot,x)\rangle_{\mathcal{H}}$ *for all* $f \in \mathcal{H}$ *and all* $x \in \Omega$*.*

## 2 Radial Basis Function Approximation

In the literature on radial basis function (RBF) approximation, usually an expansion of the form

$$Pf(x) = \sum_{j=1}^{n} c_j \Phi_j(x), \quad x \in \mathbb{R}^s,$$

is considered. Here the basis functions $\Phi_j$ are generated by a single basic function $\phi$ which is composed with the Euclidean norm, and then shifted to the centers $x_j$, $j = 1, \ldots, n$, i.e., $\Phi_j(x) = \phi(\|x - x_j\|_2)$. Common choices for $\phi$ include the Gaussian, (inverse) multiquadric, polyharmonic splines, or Wendland's compactly supported RBFs. The coefficients $c_j$ are found by solving a system of linear equations of the form $Ac = f$ generated

---

[1] Illinois Institute of Technology, Department of Applied Mathematics, Chicago, IL 60616, U.S.A. (fass@amadeus.math.iit.edu).

by enforcing a set of conditions ensuring interpolation of $f$ by $Pf$ on the set $\mathcal{X} = \{x_i\}_{i=1}^n$. Therefore, the matrix $A$ has entries

$$A_{ij} = \phi(\|x_i - x_j\|_2),$$

and the vectors $c$ and $f$ have components $c_j$, $j = 1, \ldots, n$, and $f(x_i)$, $i = 1, \ldots, n$, respectively.

The function $\Phi(x,y) = \phi(\|x-y\|_2)$ is now viewed as the reproducing kernel of a space $\mathcal{H}$ (which usually is only a pre-Hilbert space, and therefore has to be completed). It then follows from well-known results that the expansion $Pf$ provides the optimal interpolant to the function $f$ in the sense that it has minimum norm in $\mathcal{H}$ (see [1, 7, 8] for more details).

We present a different approach to defining a reproducing kernel Hilbert space associated with $Pf$. This discussion is motivated by the paper [2, 4]. However, our view is slightly different than the one presented in those papers.

First, we generalize the approximation problem to also include the case of discrete least squares approximation. We still sample the given function $f$ on the set $\mathcal{X}$, but now introduce a second set $\Xi = \{\xi_i\}_{i=1}^m$ at which we center the basis functions. Usually we will have $m \leq n$, and the case $m = n$ recovers the traditional interpolation setting. Therefore,

$$Qf(x) = \sum_{j=1}^{m} c_j \Phi_j(x), \quad x \in \mathbb{R}^s, \tag{1}$$

and the coefficients $c_j$ can be found by minimizing $\|Qf - f\|^2$, where the norm

$$\|f\|^2 = \sum_{i=1}^{n} (f(x_i))^2$$

is induced by the discrete inner product

$$\langle f, g \rangle = \sum_{i=1}^{n} f(x_i) g(x_i).$$

Equivalently, we can solve the Gram system

$$Gc = f_\Phi,$$

where $G$ has entries

$$G_{jk} = \langle \Phi_j, \Phi_k \rangle = \sum_{i=1}^{n} \Phi_j(x_i)\Phi_k(x_i) = \sum_{i=1}^{n} \Phi(x_i, \xi_j)\Phi(x_i, \xi_k), \quad j,k = 1, \ldots, m,$$

and the right-hand side vector consists of the projections of $f$ onto the basis functions, i.e.,

$$f_\Phi = [\langle f, \Phi_1 \rangle, \ldots, \langle f, \Phi_m \rangle]^T = [\langle f, \Phi(\cdot, \xi_1) \rangle, \ldots, \langle f, \Phi(\cdot, \xi_m) \rangle]^T.$$

Of course, this is also equivalent to $G = A^T A$ and $f_\Phi = A^T f$, where $A$ is the rectangular analogue of the interpolation matrix mentioned above.

We are interested in the space of functions $\mathcal{H}$ generated by (1) for all possible coefficient vectors $c$. The reproducing kernel of this space will be defined as

$$K(x,y) = \sum_{j=1}^{m} \lambda_j(x) \Phi_j(y),$$

where the functions $\lambda_j$, $j = 1, \ldots, m$, form a *dual basis* for $\mathcal{H}$. The dual basis is found by solving the Gram system

$$G\lambda = \Phi, \tag{2}$$

where $G$ is as above, and the vectors $\lambda$ and $\Phi$ collect the dual basis functions and basis functions, respectively.

Note that the dual basis allows us to see the coefficients $c_j$ in a new light. In fact, we know that $c = G^{-1} f_\Phi$, and also $\lambda = G^{-1}\Phi$. Therefore,

$$f_\lambda = G^{-1} f_\Phi = c,$$

where, as above, the notation $f_\lambda$ stands for

$$f_\lambda = \left[\langle f, \lambda_1\rangle, \ldots, \langle f, \lambda_m\rangle\right]^T.$$

This shows us that the coefficients in the expansion (1) with basis functions $\Phi_j$ are obtained by projecting the data onto the dual basis.

The functions $K(\cdot, x_i)$ can be shown to be the generating functions of a quasi-interpolant, i.e.,

$$Qf(x) = \sum_{i=1}^{n} f(x_i) K(x, x_i).$$

In fact, in the case of interpolation they form a *cardinal basis*.

Using the definition of $K$ along with the dual basis we can derive

$$\begin{aligned} Qf(x) &= \sum_{i=1}^{n} f(x_i) K(x, x_i) \\ &= \sum_{i=1}^{n} f(x_i) \sum_{j=1}^{m} \lambda_j(x) \Phi_j(x_i) \\ &= \sum_{j=1}^{m} \underbrace{\left[\sum_{i=1}^{n} f(x_i) \Phi_j(x_i)\right]}_{=\langle f, \Phi_j\rangle = d_j} \lambda_j(x). \end{aligned}$$

This shows us that the coefficients in the dual expansion are given as projections of the data onto the original basis functions $\Phi_j$.

Since $K$ is a reproducing kernel, i.e., $f(x) = \langle f, K(\cdot, x)\rangle_{\mathcal{H}}$ for all $x \in \Xi$, we have the three equivalent representations

$$Qf(x) = \sum_{j=1}^{m} \langle f, \lambda_j\rangle \Phi_j(x) = \sum_{j=1}^{m} \langle f, \Phi_j\rangle \lambda_j(x) = \sum_{i=1}^{n} \langle f, K(\cdot, x_i)\rangle_{\mathcal{H}} K(x, x_i),$$

where the inner product in $\mathcal{H}$ is defined as

$$\begin{aligned} \langle f, g\rangle_{\mathcal{H}} &= \sum_{i=1}^{n} \sum_{\ell=1}^{n} f(x_i) g(x_\ell) K(x_i, x_\ell) \\ &= \sum_{j=1}^{m} \langle f, \lambda_j\rangle \langle g, \Phi_j\rangle. \end{aligned}$$

Moreover, we can establish the bi-orthogonality relationship

$$\langle \lambda_k, \Phi_j\rangle = \delta_{jk}$$

of the basis and dual basis.

# 3 Moving Least Squares Approximation

In the moving least squares (MLS) literature there exist already two interpretations: the discrete (moving) least squares formulation, and the Backus-Gilbert formulation which corresponds to constrained quadratic optimization. For an overview of these two approaches see, e.g., [3].

The main difference between the RBF and the MLS settings lies in the definition of the discrete inner product. In the MLS method we require positive weight functions $W_i$ which we will interpret in a way similar to RBFs, i.e., $W_i(x) = W(x,x_i)$, with the points $x_i$ coming from the same set $\mathcal{X}$ as earlier. Now, the discrete inner product is defined as

$$\langle f,g\rangle_W = \sum_{i=1}^{n} f(x_i)g(x_i)W(x,x_i). \tag{3}$$

We point out that this is a significant difference since this inner product depends on the fixed (evaluation) point $x$. This will cause the Gram matrix as well as the coefficients $c_j$ to depend on $x$ as well.

In the moving least squares method one usually takes a (multivariate) polynomial basis for the approximation, i.e.,

$$Qf(x) = \sum_{j=1}^{m} c_j(x)p_j(x), \qquad x \in \mathbb{R}^s.$$

The formalism for determining the coefficients $c_j$ is still the same as before, i.e., we need to solve the $x$-dependent Gram system

$$Gc = f_p,$$

where $G$ has entries

$$G_{jk} = \langle p_j, p_k\rangle_W = \sum_{i=1}^{n} p_j(x_i)p_k(x_i)W(x,x_i), \quad j,k = 1,\ldots,m,$$

and the right-hand side vector consists of the projections of $f$ onto the basis functions, i.e.,

$$f_p = \left[\langle f,p_1\rangle,\ldots,\langle f,p_m\rangle\right]^T.$$

In the Backus-Gilbert approach one assumes a quasi-interpolant of the form

$$Qf(x) = \sum_{i=1}^{n} f(x_i)\Psi_i(x).$$

The generating functions $\Psi_i(x) = \Psi(x,x_i)$ are found by minimizing

$$\frac{1}{2}\sum_{i=1}^{n} \Psi_i^2(x)\frac{1}{W(x,x_i)}$$

subject to the polynomial reproduction constraints

$$\sum_{i=1}^{n} p(x_i)\Psi_i(x) = p(x), \quad \text{for all } p \in \mathcal{P}_d^s,$$

where $\mathcal{P}_d^s$ is the space of $s$-variate polynomials of total degree at most $d$ which has dimension $m = \binom{s+d}{d}$.

This kind of problem is solved with the aid of Lagrange multipliers $\lambda_j(x)$. The generating functions $\Psi_i$ are then given as

$$\Psi_i(x) = W(x,x_i) \sum_{j=1}^{m} \lambda_j(x) p_j(x_i),$$

where the Lagrange multipliers are obtained by solving the Gram system

$$G\lambda = p \tag{4}$$

with $p = [p_1(x), \ldots, p_m(x)]^T$. Since the entries of both $p$ and $G$ depend on the point $x$, so will the Lagrange multipliers.

By comparing (4) with (2) we can immediately recognize that the Lagrange multipliers play the role of the dual basis. Moreover, the generating functions $\Psi_i$ define the reproducing kernel, i.e.,

$$K(x,y) = \Psi(x,y) = W(x,y) \sum_{j=1}^{m} \lambda_j(x) p_j(y).$$

Within our new framework we now have the more general formulas (just as in the case of RBFs)

$$Qf(x) = \sum_{j=1}^{m} \langle f, \lambda_j \rangle_W p_j(x) = \sum_{j=1}^{m} \langle f, p_j \rangle_W \lambda_j(x) = \sum_{i=1}^{n} \langle f, K(\cdot, x_i) \rangle_{\mathcal{H}} K(x, x_i)$$

which hold for arbitrary (linearly independent) polynomials.

As for RBFs we obtain orthogonality of the basis polynomials to the dual basis functions, i.e.,

$$\langle \lambda_k, p_j \rangle_W = \delta_{jk}.$$

The inner product in $\mathcal{H}$ is given by

$$\langle f, g \rangle_{\mathcal{H}} = \sum_{i=1}^{n} \sum_{\ell=1}^{n} f(x_i) g(x_\ell) K(x_i, x_\ell).$$

## 3.1 Plots of Basis-Dual Basis Pairs

To obtain plots of a typical Gaussian basis function and dual basis function for the RBF framework we let $\mathcal{X} = \Xi$ be the set of 13 equally spaced points in $[-5, 5]$.

Notice that the kernel is a cardinal function, i.e., equal to 1 at $\xi_7 = 0$, and equal to zero at all other $\xi_j$. In the case of true approximation, i.e., when $m < n$, the plots will look similar. However, the kernel is no longer a cardinal function – but only an approximate cardinal function (a generating function for a quasi-interpolant).

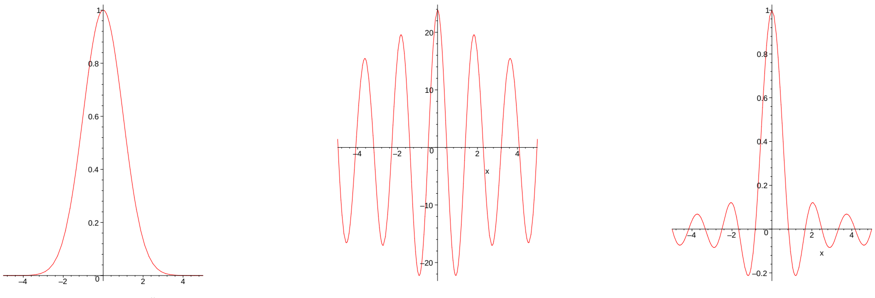

Figure 1: Plot of Gaussian basis function (left), dual basis (center), and reproducing kernel (right) centered at $\xi_7 = 0$.

# References

[1] E. W. Cheney and W. A. Light, *A Course in Approximation Theory*, Brooks/Cole, Pacific Grove, CA, 2000.

[2] T. Evgeniou, M. Pontil, and T. Poggio, Regularization networks and support vector machines, Adv. Comput. Math. 13 (2000), no. 1, 1–50.

[3] G. E. Fasshauer, Matrix-free multilevel moving least-squares methods, in *Approximation Theory, X* (St. Louis, MO, 2001), C. K. Chui, L. L. Schumaker, and J. Stöckler (eds.), Innov. Appl. Math., Vanderbilt Univ. Press, Nashville, TN, 2002, 271–281.

[4] F. Girosi, An equivalence between sparse approximation and support vector machines, Neural Computation 10 (1998), 1455–1480.

[5] S. Li and W. K. Liu, Meshfree and particle methods and their applications., Appl. Mech. Rev. 55 (2002), no. 1, 1–34.

[6] V. Maz'ya, and G. Schmidt, On quasi-interpolation with non-uniformly distributed centers on domains and manifolds, J. Approx. Theory 110 (2001), 125–145.

[7] R. Schaback, Native Hilbert spaces for radial basis functions. I, in *New Developments in Approximation Theory* (Dortmund, 1998), Internat. Ser. Numer. Math., 132, Birkhäuser, Basel, 1999, 255–282.

[8] R. Schaback, A unified theory of radial basis functions. Native Hilbert spaces for radial basis functions. II, Numerical analysis in the 20th century, Vol. I, Approximation theory, J. Comput. Appl. Math. 121 (2000), no. 1-2, 165–177.