Hermite-Birkhoff Interpolation of Scattered Data by Radial Basis Functions

Zong–min Wu

Abstract:
For Hermite-Birkhoff interpolation of scattered multidimensional data by radial basis functions ϕ, existence and characterization theorems and a variational principle are proved. Examples include $\phi(r) = r^b$, Duchon's thin-plate splines, Hardy's multiquadrics, and inverse multiquadrics.

Keywords:
Multivariate Interpolation, Hermite-Birkhoff Data, Radial Basis Functions.

AMS classification: 41A05, 41A63.

1 Introduction

For scattered data $(x_i, f_i) \in I\!R^{n+1}$, radial basis function interpolation methods, as surveyed in [1] and [4], use a "radial" function $\phi : R \to R$ to construct the interpolant g satisfying $Lg = f$ as

$$g(x) = a^T L_t \phi(\|t - x\|) \tag{1.1}$$

where $f = (f_1, ..., f_m)^T$, $Lg = (g(x_1), ..., g(x_m))^T$. L_t means L operated on the variable t. more generally,

$$g(x) = a^T L_t \phi(\|x - t\|) + \sum_{|\alpha| \le q} \mu_\alpha x^\alpha \tag{1.2}$$

where the norm $\|.\|$ is Euclidean, the coefficients μ_α and monomials x^α are written in standard multivariate notation. Furthermore, the condition

$$a^T L x^\alpha = 0 \ \text{ for all } \ |\alpha| \le q \tag{1.3}$$

usually is required to make use of the extra degrees of freedom. Micchelli [3] and Powell [4] have shown the existence of the interpolation (1.1) and (1.2) for standard radial basis functions such as

$$\phi(x) = \|x\|^b, \quad \|x\|^{2b} \ln \|x\|, \quad (c + \|x\|^2)^{-b} \ \text{ etc.}$$

Using the notation $A = (\phi(\|x_j - x_k\|))_{1 \le j,k \le m}$, $D = (x_j^\alpha)_{1 \le j \le m, |\alpha| \le q}$, the interpolation (1.1) and (1.2) is possible if the matrix

$$\hat{A} \ = \begin{pmatrix} A & D \\ D^T & 0 \end{pmatrix}$$

is nonsingular or if the quadratic form $x^T A x$ is positive definite on the subspace of all $x \in R^m$ with $D^T x = 0$ ("conditionally" positive definite, see [3]). The solution can be written as

$$g(x) = L_t^T \phi(\|x - t\|) A^{-1} f \tag{1.4}$$

in case of (1.1), and in the form

$$g(x) = S^T(x) \hat{A}^{-1} \begin{pmatrix} f \\ 0 \end{pmatrix} \tag{1.5}$$

with

$$S^T(x) = (L_t^T \phi(\|x - x_1\|), 1, \ldots, x^\alpha, \ldots) \tag{1.6}$$

in the general case (1.2).

Here, we want to construct a radial basis function interpolant for Hermite-Birkhoff data. To this end, we assume

$$L^T = (L_1, \ldots, L_m) : C_{\hat{q}}(R^n) \to R^m$$

to be a linear operator, where $L_j f$ are linear combinations of evaluations of derivatives of f of order at most $\hat{q}$ at some of the centers x_j. We then want to construct an interpolant $g(x)$ of the form (1.2) to match the data. i.e.:

$$Lg = Lf.$$

The problem coincides with ordinary radial basis function interpolation if

$$L^T f = (f(x_1), f(x_2), \ldots, f(x_m))$$

with $\hat{q} = 0$.

2 Kriging

In order to motivate our approach, we shall first review the background of Kriging interpolation [2] loosely, delaying the exact theory to the next section. Roughly speaking, Kriging interpolation considers the function f as a random function of the form

$$f = P + R,$$

where $P = \sum_{|\alpha| \leq q} A_\alpha x^\alpha$ is a polynomial with real valued random variables A_α and a random function R satisfying

$$ER = 0,\ ER(x)R(y) = \phi(\|x - y\|),$$

where E is the expectation. Now, in the classical case $L^T f = (f_1, \ldots, f_m)$, the function g is constructed as

$$g(x) = \sum_{j=1}^{m} \lambda_j(x) f_j \tag{2.1}$$

to minimize the error variance

$$E(g - f)^2 \tag{2.2}$$

under the constraint

$$E(g - f) = 0, \tag{2.3}$$

meaning that g is an unbiased estimator for f. From (2.3) and $f = P + R$ we get

$$\sum_{j+1}^{m} \lambda_j(x)(\sum_{|\alpha| \leq q} EA_\alpha x_j^\alpha) - \sum_{|\alpha| \leq q} EA_\alpha x^\alpha = 0$$

The equation must hold for any EA_α. Thus

$$\sum_{j=1}^{m} \lambda_j(x) x_j^\alpha = x^\alpha. \tag{2.4}$$

Under suitable assumptions, (2.2) is minimized under the constraint (2.4). Then with Lagrange multipliers, we get the necessary condition

$$\hat{A} \begin{pmatrix} \lambda(x) \\ \mu(x) \end{pmatrix} = S(x). \tag{2.5}$$

If $\hat{A}$ is nonsingular, the solution $g(x)$ in the form of (2.1) is just the same as the radial basis interpolation (1.5). And it is a best linear unbiased estimation (BLUE) for f, if we know the data $f(x_j)$.
In this paper we want to use the information in Lf to construct a BLUE for the Hermite-Birkhoff case, generalizing results from [6]. To do this, we formally define

$$g(x) = \lambda^T(x)Lf \tag{2.6}$$

and minimize the error variance

$$E(g-f)^2 \tag{2.7}$$

under the constraint

$$E(g-f) = 0. \tag{2.8}$$

From [6] the identity

$$EL_jR(x)L_kR(y) = L_{jx}L_{ky}\phi(\|x-y\|)$$

follows under proper assumptions on $R(x)$ and $\phi(\|x\|)$. Here, the indices j and k denote components of L, while x and y stand for the variables for which L must be applied. Then is easy to derive

$$g(x) = S^T\hat{A}^{-1}\begin{pmatrix} Lf \\ 0 \end{pmatrix} \tag{2.9}$$

if

$$S^T = (L_t^T\phi(\|t-x\|), X)$$

$$\hat{A} = \begin{pmatrix} L_sL_t^T\phi(\|t-s\|) & LX \\ (LX)^T & 0 \end{pmatrix}$$

where $X = (1, \ldots, x^\alpha, \ldots)$ is a vector containing the monomials from (2.4). If P is absent, X can be absent, too. The system is solvable, if $\hat{A}$ is nonsingular. We discuss the nonsingularity of $\hat{A}$ for general radial basis functions in the next chapter.

3 Main Theorem

We first assume the functionals L_j to be linear independent on $C_{\hat{q}}(R^n)$. If necessary, some of the L_j are dropped from L. For ordinary radial basis interpolation, linear independence of the L_j means that the centres x_j are pairwise distinct.

Furthermore, we denote the generalized Fourier transform of $\phi(\|x\|)$ on R^n by $\hat{\phi}(t)$ and assume $\hat{\phi}(t) > 0$ for all $t \in R^n \setminus \{0\}$.

Theorem 3.1 : *If*

$$\int_{R^n} |\lambda^T L_x e^{i<t,x>}|^2 \hat{\phi}(t) dt \text{ exists for } \lambda \text{ with } \lambda^T(LX) = 0 \tag{3.1}$$

and if the nondegeneracy assumption

$$\text{rank}(LX) = \dim(X) \tag{3.2}$$

holds, then the matrix $\hat{A}$ is nonsingular.

Remark 3.1 : For ordinary radial basis interpolation, equation (3.2) of nondegeneracy means that the data points $\{x_j\}$ are not on a polynomial surface of order at most q. Otherwise the matrix D in $\hat{A}$ is not of full rank and $\det \hat{A}$ vanishes. But (3.2) is only needed for uniqueness. One can always put some additional pseudo-data into L, if unicity is required. For example,

$$L_{m+\alpha} f = f(x_0 + \alpha h) \text{ with } |\alpha| \le q$$

would let part of the matrix LX be a multidimensional Vandermonde matrix, and (3.2) will be satisfied.

Proof of Theorem 3.1: Assume $\hat{A} \begin{pmatrix} \lambda \\ \mu \end{pmatrix} = 0$ or

$$L_s L_t^T \phi(\|t-s\|)\lambda + LX\mu = 0 \tag{3.3}$$

$$(LX)^T \lambda = 0. \tag{3.4}$$

Putting (3.4) into (3.3) we have

$$\lambda^T L_s L_t^T \phi(\|t-s\|)\lambda = 0. \tag{3.5}$$

First we show $\lambda = 0$ by contradiction. If $\lambda \neq 0$, then $\lambda^T L_s \neq 0$ on $C_q(R^n)$, and the function

$$h(t) := \lambda^T L_s e^{i<t,s>},$$

being analytic on R^n, can not vanish identically, because otherwise $\lambda^T L_s$ would vanish on all functions in $C_q(R^n)$ which are images of the Fourier transformation. If we use the delta sequence $c_n M^n e^{-M^2(x-(v-u))^2/4}$, we get

$$\begin{aligned}
& \lambda^T L_u L_v^T \lambda \phi(\|v-u\|) \\
= \; & \lim_{M\to\infty} \int \lambda^T L_u L_v^T \lambda (c_n M^n e^{-M^2(x-(v-u))^2/4}) \phi(\|x\|) dx \\
= \; & \lim_{M\to\infty} \int e^{-t^2/M^2} |\lambda^T L_s e^{i<t,s>}|^2 \hat{\phi}(t) dt \\
= \; & \int |\lambda^T L_s e^{i<t,s>}|^2 \hat{\phi}(t) dt > 0.
\end{aligned}$$

since $c_n M^n e^{-M^2(x-(v-u))^2/4}$ are test functions whose Fourier transforms are $e^{-t^2/M^2} e^{i<t,v-u>}$. But this contradicts (3.5). Therefore $\lambda = 0$ holds and we get $LX\mu = 0$ from (3.3). Finally, (3.2) implies $\mu = 0$.

■

Now we have the nonsingularity of $\hat{A}$ and (2.9) is well-defined.

Theorem 3.2 *: The function $g(x)$, as defined by* (2.9)*, is an interpolation.*

Proof: Using (2.9), it is easy to calaulate

$$Lg = L_x(L_t^T \phi(\|t-x\|), X) \begin{pmatrix} L_s L_t^T(\|t-s\|) & LX \\ (LX)^T & 0 \end{pmatrix}^{-1} \begin{pmatrix} Lf \\ 0 \end{pmatrix}$$

$$= (L_x L_t^T \phi(\|t-x\|), LX) \begin{pmatrix} L_s L_t^T(\|t-s\|) & LX \\ (LX)^T & 0 \end{pmatrix}^{-1} \begin{pmatrix} Lf \\ 0 \end{pmatrix} = Lf.$$

■

In this way we can interpolate scattered Hermite-Birkhoff data using radial basis functions. The interpolant $g(x)$ can be expressed by (2.9) or

$$g(x) = \sum_{j=1}^{m} a_j L_{jt}\phi(\|t - x\|) + \sum_{|\alpha| \leq q} b_\alpha x^\alpha \tag{3.6}$$

with

$$Lg(x) = Lf \text{ and } \sum_{j=1}^{m} a_j L_j x^\alpha = 0. \tag{3.7}$$

Theorem 3.3 : *The interpolation preserves polynomials of order up to q.*

Proof: It is trivial from the uniqueness. ■

Remark: The polynomial reproduction of the multiquadric interpolation is only for the infinite grid data. Now it can be hold for scattered finite data. In fact if we want to preserve a linear space, say $\hat{X}$, by the radial basis interpolation. We can trivially select the space X to contain $\hat{X}$. In the "thin plate" spline case, for example, the space X must contain $\{1, x, y\}$ (see Example 4.2) and we can choose

$$X = \{1, x, y, \sin(x), \cos(x)\},$$

and the interpolation will preserve the space X.

4 Special radial basis functions

Example 4.1 : For $\phi(x) = e^{-bx^2}$ or $\phi(x) = (c+x^2)^{-b}$ with $b > 0$ the function ϕ is in C_∞. For any L the matrix $L_s L_t^T \phi(\|t-s\|)$ is well-defined. The Fourier transform of $\phi(x)$ is positive, regular and decays exponentially at infinity. Therefore (3.1) is always satisfied. The interpolation uniquely exists if (3.2) is satisfied.

Example 4.2 : Consider the radial basis functions

$$\phi(x) = \|x\|^{2b}, \;\; b \notin N, \text{ and}$$

$$\phi(x) = \|x\|^{2b} \ln \|x\|, \;\; b \in N.$$

Since $\phi \in C_{2b}$ the matrix $L_s L_t^T \phi(\|t-s\|)$ exists for $\hat{q} < b$. The generalized Fourier transform of ϕ is

$$\hat{\phi}(t) = \|t\|^{-n-2b}$$

or

$$\hat{\phi}(t) = c_1\|t\|^{-n-2b} + c_2 \Delta^b \delta(x),$$

respectively. Condition (3.1) is satisfied, if X contains a space of polynomials of order up to b. Then the interpolation (2.9) is uniquely solvable, if (3.2) is satisfied.

Example 4.3 : The function $\phi(x) = (c+x^2)^b$ for $b > 0$ $b \notin I\!N$ is in C_∞. For any L the matrix $L_s L_t^T \phi(\|t-s\|)$ is well-defined. The function $F(t) = (c+t)^b$ has the properties that, the $F^{(k}(t)$ is complete monotone, if we define $k = [b]+1$. From [3] with $F_\epsilon(t) = F(t+\epsilon)$, we get

$$F_\epsilon(t) - \sum_{l=0}^{k-1} \frac{F_\epsilon^{(l)}(0)}{l!} t^l = \int_0^\infty \frac{e^{-\epsilon\sigma}}{\sigma^k} \left\{ e^{-t\sigma} - \sum_{l=0}^{k-1} \frac{(-t\sigma)^l}{l!} \right\} d\mu(\sigma)$$

If $\lambda^T L x^\alpha = 0$ for $|\alpha| < k$, then

$$\lambda^T L_t L_s^T \lambda F_\epsilon(\|t-s\|^2)$$

$$= \int_0^\infty \frac{e^{-\epsilon\sigma}}{\sigma^k} \lambda^T L_t L_s^T \lambda e^{-\|t-s\|^2\sigma} d\mu(\sigma)$$

$$= \int_0^\infty \frac{e^{-\epsilon\sigma}}{\sigma^k} (\lambda^T L_t L_s^T \lambda \int e^{i<t-s,v>} e^{-\|v\|^2/4\sigma} dv) d\mu(\sigma)$$

$$= \int_0^\infty \frac{e^{-\epsilon\sigma}}{\sigma^k} (\int |\lambda^T L_t e^{i<t,v>}|^2 e^{-\|v\|^2/4\sigma} dv) d\mu(\sigma)$$

Because of the Fubini's Theorem

$$\lambda^T L_t L_s^T \lambda F_\epsilon(\|t-s\|^2) = \int |\lambda^T L_t e^{i<t,v>}|^2 \hat{\phi}_\epsilon(v) dv$$

where $\hat{\phi}_\epsilon(v) = \int_0^\infty \frac{e^{-\epsilon\sigma}}{\sigma^k} e^{\frac{\|v\|^2}{4\sigma}} d\mu(\sigma) > 0$ Furthermore using the continnuity of F and the monotone convergence

$$\lambda^T L_t L_s^T \lambda F(\|t-s\|^2) = \int |\lambda_t^T e^{i<t,v>}|^2 \hat{\phi}_0(v) dv > 0$$

Equation (3.1) is satisfied if X contains a space of polynomials of order up to b. From [2] the interpolation using ϕ is equivalent to the interpolation using ϕ_1. The interpolation is uniquely solvable if (3.2) is satisfied. If $b < 1$ especially, with the analogous discuss as Lemma 3.2 of [3] . The polynomial term can be omited as ordinary radial basis interpolation.

5 Variational principle

Assume there is a polynomial $P(t)$ on R^n with values in R^l satisfying

$$< P(t), P(t) > \hat{\phi}(t) = 1, \tag{5.1}$$

where $<,>$ is the inner product in R^l. This defines a generalized differential operator G satisfying

$$\widehat{Gf(x)} = P(t)\hat{f}(t) \text{ or } G = P(-i\nabla) \tag{5.2}$$

via Fourier transforms. We have

Theorem 5.1 *: If g is the Hermite-Birkhoff interpolant from the previous sections, then*

$$\int < Gg, Gg > dx = \int |\lambda^T L_s e^{i<s,t>}|^2 \hat{\phi}(t) dt < \infty. \tag{5.3}$$

If g_1 is another interpolant such that Fourier transforms can be taken of $h := g - g_1$, then

$$\int < Gg, Gh > dx = 0 \text{ or}$$

$$\int < Gg, Gg > dx \leq \int < Gg_1, Gg_1 > dx. \tag{5.4}$$

This shows the optimality of the interpolation with respect to the integral (5.3).

Proof: From (2.9) we get

$$g(x) = \lambda^T L_s \phi(s - x) + \mu^T X \text{ with } \lambda^T L X = 0,$$

where X is the basis of the space of polynomials of total degree not exceeding $b/2$. If $\hat{\phi}(t) = \mathcal{O}(\|t\|^{-b-n})_{t\to 0}$, for instance, P can not contain polynomials of degree lower than $(n + b)/2$. Thus $GX = 0$. Furthermore, we have

$$\lambda^T L_s e^{i<t,s>} = \mathcal{O}(t^{1+b/2})_{t\to 0}$$

and follows from the Parseval formula

$$\begin{aligned}
\infty &> \int |\lambda^T L_s e^{-i<t,s>}|^2 \hat{\phi}(t) dt \\
&= \int < P(t)\lambda^T L_s \hat{\phi}(t) e^{-i<t,s>}, P(t)\lambda^T L_s \hat{\phi}(t) e^{-i<t,s>} > dt \\
&= \int < G\lambda^T L_s \phi(\|s-x\|), G\lambda^T L_s \phi(\|s-x\|) > dx \\
&= \int < Gg, Gg > dx.
\end{aligned}$$

To prove the second part of the theorem we find

$$\begin{aligned}
&\int < Gg, Gh > dx \\
&= \int < G_x \lambda^T L_s \phi(\|s-x\|), Gh > dx \\
&= \int < P(t)\lambda^T L_s \hat{\phi}(t) e^{-i<s,t>}, P(t)\hat{h}(t) > dt \\
&= \int \lambda^T L_s e^{-i<s,t>} \hat{h}(t) dt \\
&= \lambda^T L h = 0.
\end{aligned}$$

■

Example 5.1 : If $\phi(\|x\|) = \|x\|^b$ with n and b odd, then $\hat{\phi}(t) = \|t\|^{-n-b}$ and $G = \nabla^{(n+b)/2}$.
The case $n = 1, b = 3$ yields cubic spline interpolation, minimizing the integral

$$\int (g'')^2 dx.$$

It is the natural spline for data $(f_0, \ldots, f_m)$,and is the complete spline for data $(f_0', f_0, \ldots, f_m, f_m')$.

In case $n = 3, b = 1$ the interpolation minimizes the integral

$$\int (g_{xx}^2 + g_{yy}^2 + g_{zz}^2 + 2g_{xy}^2 + 2g_{yz}^2 + 2g_{zx}^2) dxdydz.$$

Example 5.2 : For $\phi(x) = e^{-\alpha\|x\|^2}$ we have

$$\hat{\phi}^{-1}(t) = e^{\|t\|^2/4\alpha} = \sum_{k=0}^{\infty} \|t\|^{2k}/k!(4\alpha)^k.$$

Since the results of Theorem 5.1 can easily be generalized to hold for $l = \infty$, radial basis interpolation with Gaussian kernels minimizes the integral

$$\int \sum_{k=0}^{\infty} \frac{< \nabla^k g, \nabla^k g >}{k!(4\alpha)^k} dx.$$

Example 5.3 : For $\phi(x) = e^{-\alpha\|x\|_1}$ and $\hat{\phi}(t) = \prod_{j=1}^{n}(1/(1+\alpha^2 t_j^2))$, radial basis interpolation with ϕ minimizes the integral

$$\int f^2 + \alpha^2(f')^2 dx \ \text{ for } \ n = 1 \ \text{ and}$$

$$\int f^2 + \alpha^2((f_x)^2 + (f_y)^2) + \alpha^4(f_{xy})^2 dxdy \ \text{ for } \ n = 2.$$

Remark: Equation (5.1) is equivalent to $G^*G\phi(x) = \delta(x)$ in the sense of distributions. The theory, as developed so far, can take any solution ϕ of this equation, and does not necessarily require a radial function(see [5] for the case of interpolation on a rectangle using techniques of Fourier series). If G is a differential operator of the form (5.2), we can construct an interpolation to minimize the integral

$$\int < Gg, Gg > dx, \tag{5.5}$$

by choosing $\phi(x)$ as the generalized Fourier transform of $< P(t), P(t) >^{-1}$. Normally ϕ is not radial, if $< P(t), P(t) >$ is not radial. But we can always use (2.9) to interpolate the data, and the interpolation will have the optimality property (5.4).

Example 5.4 : To construct an interpolation minimizing the integral

$$\int (g_{xx})^2 + (g_{xxy})^2 dxdy, \tag{5.6}$$

we can take $G^T g = (g_{xx}, g_{xxy})$ and $< P, P >= t_1^4(1 + t_2^2)$. Then

$$\phi(x, y) = |x|^3 e^{-|y|}.$$

If we use this function to construct an interpolant via (2.9), the integral (5.6) will be minimized.

Acknowledgment: I would like to thank Prof. R. Schaback to suggest the topic and for help to formulate the final version of this paper.

References

[1] Dyn, N., Interpolation of scattered data by radial functions, in: *Topics in Multivariate Approximation*, C.K. Chui, L.L. Schumaker, and F.I. Utreras (eds.) Academic Press, 1987

[2] Matheron, G., The intrinsic random functions and their applications, Adv. appl. Prob.**5** (1973) 439-468.

[3] Micchelli, C.A., Interpolation of scattered data: distance matrices and conditionally positive definite functions, Constructive Approximation **2**, 11-22

[4] Powell, M.J.D., Radial basis functions for multivariable interpolation: a review, in *Numerical Analysis 1987*, D.F. Griffiths and G.A. Watson (eds.), Longman Scientific & Technical (Harlow), 223–241

[5] Schaback, R., Interpolation and Approximation von Funktionen mehrerer Veränderlicher. Vorlesungsmanuskript 1982 Göttingen, FRG.

[6] Wu, Zong–min, Die Kriging–Methode zur Lösung mehrdimensionaler Interpolationsprobleme, Dissertation, 1986 Göttingen, FRG.

[7] Wu, Zong-min,and Schaback,R., Local Error Estimates for radial basis function interpolation of scattered data, preprint.

Author's address:

Zong–min Wu, Dept. of Mathematics, Fudan University, Shanghai, 200433 People's Republic of China