Local Error Estimates for Radial Basis Function Interpolation of Scattered Data

Zong–min Wu and Robert Schaback

AMS classifications: 41A05, 41A63, 41A25, 65D05

Keywords: Multivariate interpolation, orders of convergence, Kriging, Multiquadrics, thin plate splines.

Abstract

Introducing a suitable variational formulation for the local error of scattered data interpolation by radial basis functions $\phi(r)$, the error can be bounded by a term depending on the Fourier transform of the interpolated function f and a certain "Kriging function", which allows a formulation as an integral involving the Fourier transform of ϕ. The explicit construction of locally well–behaving admissible coefficient vectors makes the Kriging function bounded by some power of the local density h of data points. This leads to error estimates for interpolation of functions f whose Fourier transform $\hat{f}$ is "dominated" by the nonnegative Fourier transform $\hat{\psi}$ of $\psi(x) = \phi(\|x\|)$ in the sense $\int |\hat{f}|^2 \hat{\psi}^{-1} dt < \infty$. Approximation orders are arbitrarily high for interpolation with Hardy multiquadrics, inverse multiquadrics and Gaussian kernels. This was also proven in recent papers by Madych and Nelson, using a reproducing kernel Hilbert space approach and requiring the same hypothesis as above on $\hat{f}$, which limits the practical applicability of the results. This work uses a different and simpler analytic technique and allows to handle the cases of interpolation with $\phi(r) = r^s$ for $s \in I\!R$, $s > 1$, $s \notin 2I\!N$, and $\phi(r) = r^s \log r$ for $s \in 2I\!N$, which are shown to have accuracy $\mathcal{O}(h^{s/2})$.

1 Introduction

Radial basis function interpolation to scattered data $(x_i, f_i) \in I\!R^{n+1}$ for pairwise distinct points ("centres") $x_1, \ldots, x_M \in I\!R^n$ uses a function $\phi : I\!R_{\geq 0} \to I\!R$ and the space $I\!P_q$ of polynomials on $I\!R^n$ with total order not exceeding q to construct the interpolant

$$s(x) := \sum_{i=1}^{M} a_i \phi(\|x - x_i\|) + \sum_{i=1}^{Q} b_i p_i(x) \tag{1.1}$$

via the linear system

$$\begin{aligned} \sum_{i=1}^{M} a_i \phi(\|x_j - x_i\|) \quad + \quad \sum_{i=1}^{Q} b_i p_i(x_j) = f_j, &\quad 1 \leq j \leq M \\ \sum_{j=1}^{M} a_j p_i(x_j) = 0, &\quad 1 \leq i \leq Q = \tbinom{q+n-1}{n}, \end{aligned} \tag{1.2}$$

[0]File : usr/nam/rschaba/tex/fertig/wu_rs_1_mod_2.tex, TEXed February 27, 1992, Status : 2nd revision

where $p_1, \ldots, p_Q$ is a basis of $I\!P_q$. For a wide choice of functions ϕ and polynomial orders q, including the case $q = Q = 0$, the nonsingularity of the $(M+Q) \times (M+Q)$ system (1.2), written as

$$\begin{pmatrix} A & P \\ P^T & 0 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix} \tag{1.3}$$

in matrix notation, has been established by Micchelli [8] and Powell [12]. Following [8], we assume $F(r) = \phi(\sqrt{r})$ to be conditionally strictly positive definite of order q, which implies that A is positive definite on the subset of vectors $u \in I\!R^M$ satisfying $P^T u = 0$. We handle the case of conditionally strictly negative definite functions $\phi(\sqrt{r})$ by going over to the function $-\phi(\sqrt{r})$. For $q > 0$ the additional condition

$$p(x_j) = 0,\ 1 \le j \le M \ \text{ for } \ p \in I\!P_q \ \text{ implies } \ p = 0 \tag{1.4}$$

on q and the positions of data points is required. The numerical problems of (1.3) can be overcome by preconditioning methods (see Dyn, Levin and Rippa [5], and Dyn [4]), making the radial basis function approach a promising tool for multivariate interpolation.

We consider interpolation of values $f_i = f(x_i)$ of a smooth function f on a domain $\Omega \subseteq I\!R^n$. There are no further conditions on Ω. The sampling points x_i are allowed to be irregularly distributed over Ω under the restriction (1.4). For local error estimation, we measure the "density" of centres x_j near some $x \in \Omega$ by

$$h_\rho(x) := \max_{y \in K_\rho(x)} \min_{1 \le j \le M} \|y - x_j\|$$

for some fixed $\rho > 0$, where $K_\rho(x) = \{y \in I\!R^n \mid \|x - y\| \le \rho\}$, using the Euclidean norm $\|.\|$ (see also [9], [10], [11]).

Our goal is to prove local error bounds of the following form: Given constants $\rho \in I\!R_{>0}$, $q \in I\!N_{\geq 0}$, a radial basis function ϕ and a certain function space F_ϕ to be described below, we want to show the existence of positive constants $k \ge m$, k, $m \in I\!N$, h_0, $C \in I\!R$ such that for any distribution of points $x_j \in I\!R^n$, $1 \le j \le M$, any function f from F_ϕ and any point $x \in \Omega$ with $h_\rho(x) < h_0$ the inequality

$$|s^{(\mu)}(x) - f^{(\mu)}(x)| \le c_f \cdot C \cdot h_\rho^{k-|\mu|}(x) \tag{1.5}$$

holds for the error and its μ–th derivatives for $0 \le |\mu| \le m$, where the constant c_f depends on f, ϕ, and F_ϕ only. Then we call (1.5) a **local error bound** of order k. Here and in the sequel we use the standard multi–index notation with $|\mu| := \sum_i \mu_i$ for $\mu \in I\!N^n$.

Convergence of interpolation on regular grids has been studied extensively by Buhmann (see e.g: [1] for a comprehensive treatment) and others. For the case of scattered data Duchon [3] treated the thin–plate spline case $\phi(r) = r^2 \log r$, while Jackson [6] proved a general, but non–quantitative convergence result. The dissertation of Wu [13] of 1986 contained a rather general Hilbert space theory for Kriging and related radial basis function methods in the case of a single variable, including error estimates and convergence results. Recently, Madych and Nelson [9], [10] developed a reproducing kernel Hilbert space approach to get error bounds of arbitrarily high order for multiquadrics and inverse multiquadrics, i.e.: for $\phi(r) = (c^2 + r^2)^{s/2}$ with $c \in I\!R_{>0}$, $s \in I\!R$, $s \notin 2I\!N$, $s > -n$ and $q > s/2$ if $s > 0$.

This paper is based on the approach of Wu [13] and generalizes it to the multivariate case. There are some background connections to the work of Madych and Nelson which will be explained at the appropriate places. Like Madych and Nelson we require the restrictive condition $\int |\hat{f}|^2 \hat{\psi}^{-1} dt < \infty$ for $\psi(x) := \phi(\|x\|)$ and show that k in (1.5) can be arbitrarily large for Gaussians $\phi(r) = \exp(-cr^2)$, multiquadrics $\phi(r) = \sqrt{r^2 + c^2}$, and inverse multiquadrics $\phi(r) = 1/\sqrt{r^2 + c^2}$, $c > 0$. For these cases our error bounds are equivalent to those of Madych and Nelson. But our technique also yields the convergence order $k = s/2$ for $m < s/2 < q$ in case of the radial basis functions $\phi(r) = r^s$ with $s \in I\!R_{>0}$, $s \notin 2I\!N$, and $\phi(r) = r^s \log r$ with $s \in 2I\!N$. Compared to the work of Madych and Nelson the methods of this paper are somewhat simpler and more direct, but do not provide or require additional information about the Hilbert space background. They generally imply that the order k in (1.5) is attained, if the Fourier transform $\hat{\psi}$ satisfies

$$0 < \hat{\psi}(t) \leq \mathcal{O}(\|t\|^{-n-2k})$$

for $\|t\| \to \infty$.

2 Variational formulation

Since the equations (1.2) are solvable under the hypotheses of the preceding section, there is a Lagrange–type representation

$$s(x) = \sum_{j=1}^{M} f(x_j) u_j(x), \; u_j(x_i) = \delta_{ij}, \; 1 \leq i, j \leq M, \tag{2.1}$$

of the solution. Introducing vectors

$$\begin{aligned} R(x) &:= (\phi(\|x - x_1\|), \ldots, \phi(\|x - x_M\|))^T \\ S(x) &:= (p_1(x), \ldots, p_Q(x))^T, \end{aligned}$$

we assert

Theorem 2.2 *The vector $U(x) := (u_1(x), \ldots, u_M(x))^T$ formed by the values of the Lagrange basis functions $u_1, \ldots, u_M$ of (2.1) at $x \in I\!R^n$ coincides with the solution $U_*(x)$ of the conditional minimization problem*

$$\min\{U^T A U - 2U^T R(x) + \phi(0) \mid U \in I\!R^M, P^T U = S(x)\}. \tag{2.3}$$

Proof: If (2.3) is solved by Lagrange multiplier techniques, and if the solution is written in matrix form, there exists a vector $V_*(x) = (v_1(x), \ldots, v_Q(x))^T$ of Lagrange multipliers such that

$$\begin{pmatrix} A & P \\ P^T & 0 \end{pmatrix} \begin{pmatrix} U_*(x) \\ V_*(x) \end{pmatrix} = \begin{pmatrix} R(x) \\ S(x) \end{pmatrix} \tag{2.4}$$

holds for the solution $U_*(x)$ of (2.3), which is unique since (2.4) has the same coefficient matrix as (1.3). For a single $x = x_j$, $1 \leq j \leq M$, the right-hand side of (2.4) coincides with the j-th column of the coefficient matrix. Since the system is uniquely solvable, the vector $U_*(x_j)$ must coincide with $U(x_j)$, the j-th unit vector. But as the components of $U_*(x)$ are of the form (1.1), we have $U(x) = U_*(x)$ for all $x \in I\!R^n$. $\square$

Remarks: Cardinal interpolants like the u_j defined in (2.1) normally are constructed via the system (1.3) with data δ_{ij} for u_j. But (1.3) would yield the coefficients of u_j, not the vector $U(x)$ of values $u_1(x), \ldots, u_M(x)$ at a fixed $x \in I\!R^n$. Furthermore, the value $\phi(0)$ in (2.3) will be required later, but could here be replaced by any other constant. While the approach of Madych and Nelson poses a variational problem in an infinite-dimensional Hilbert space, our approach uses the finite–dimensional conditional minimization problem (2.3).

We now do the same thing for derivatives. If ϕ is differentiable of order $|\mu|$ on $(0,\infty)$ and of order $2|\mu|$ around zero, and if we define $\psi(x) = \phi(||x||)$, the solution $U_*^{(\mu)}(x)$ of the problem

$$\min \left\{ U^T A U - 2U^T R^{(\mu)}(x) + \psi^{(2\mu)}(0) \;\middle|\; \begin{array}{c} U \in I\!R^M, \\ P^T U = S^{(\mu)}(x) \end{array} \right\} \tag{2.5}$$

uniquely exists and satisfies a nonsingular linear equation system like (2.4), which formally coincides with the system

$$\begin{pmatrix} A & P \\ P^T & 0 \end{pmatrix} \begin{pmatrix} U^{(\mu)}(x) \\ V^{(\mu)}(x) \end{pmatrix} = \begin{pmatrix} R^{(\mu)}(x) \\ S^{(\mu)}(x) \end{pmatrix} \tag{2.6}$$

obtained by differentiation of (2.4). Thus the derivatives $U^{(\mu)}$ of $U(x)$ and $V^{(\mu)}$ of $V(x)$ exist and satisfy (2.6), while $U^{(\mu)}(x)$ coincides with the solution $U_*^{(\mu)}(x)$ of (2.5) by the same argument as above. It might be worth noting that (2.6) and (2.5) always imply polynomial reproduction up to order q in the sense

$$\sum_{j=1}^{M} u_j^{(\mu)}(x) \cdot p(x_j) = p^{(\mu)}(x) \tag{2.7}$$

for all $p \in I\!P_q$. The special choice of the additive constant in (2.5) will become apparent in (3.2); it will make the minimum value nonnegative, which is required for the following

Definition 2.8 *Let $F(r) = \phi(\sqrt{r})$ be conditionally strictly positive definite of order q, and assume $\phi \in C^{|\mu|}(0,\infty)$, $\phi \in C^{2|\mu|}$ around zero for $\mu \in I\!N_{\geq 0}^n$. Then, for any distribution of centres $x_1, \ldots, x_M$ satisfying* (1.4), *the nonnegative function function $\kappa_q^{(\mu)}$ defined by*

$$\left(\kappa_q^{(\mu)}(x)\right)^2 := \min \left\{ U^T A U - 2U^T R^{(\mu)}(x) + \psi^{(2\mu)}(0) \,\middle|\, U \in K_q^{(\mu)}(x) \right\}$$

with the set

$$K_q^{(\mu)}(x) := \left\{ U = (u_1, \ldots, u_M)^T \in I\!R^M \;\middle|\; \sum_{j=1}^{M} u_j p(x_j) = p^{(\mu)}(x) \quad \text{for all } p \in I\!P_q \right\} \tag{2.9}$$

of admissible vectors is called the **Kriging function** *at x.*

As we shall see later, the Kriging function is the norm of the representer of the interpolation error functional on a reproducing-kernel Hilbert space. However, the next steps will directly prove the following facts:

a) The interpolation error can be bounded by the Kriging function,

b) The Kriging function can be expressed by a nonnegative integral.

Furthermore, section 5 will show that

c) there are admissible vectors from $K_q^{(\mu)}(x)$ for (2.5) which allow the integral (and thus the Kriging function) to be bounded by some power of $h_\rho(x)$.

The first two steps will involve Fourier transforms which are not classically feasible unless we produce sufficiently "nice" admissible vectors $U \in K_q^{(\mu)}(x)$ for (2.5). From here on, we consider μ, q, and x as being fixed and do not always indicate dependence on these symbols in the notation.

3 Fourier transforms

If the hypotheses of Definition 2.8 hold and $\psi(x) = \phi(\|x\|)$ is an absolutely integrable function with a nicely behaving nonnegative Fourier transform $\hat{\psi}$ satisfying

$$\psi(y) = \phi(\|y\|) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{i\langle y,t\rangle} \hat{\psi}(t)dt, \quad y \in \mathbb{R}^n, \tag{3.1}$$

we can use identities like

$$\begin{aligned} & \sum_{j,k=1}^{M} w_j \overline{w_k} \phi(\|x_j - x_k\|) \\ = & \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \sum_{j,k=1}^{M} w_j \overline{w_k} e^{i\langle x_j - x_k, t\rangle} \hat{\psi}(t)dt \\ = & \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} |\sum_{j=1}^{M} w_j e^{i\langle x_j,t\rangle}|^2 \hat{\psi}(t)dt \end{aligned}$$

and

$$\begin{aligned} \frac{d^\mu}{dx^\mu}\phi(\|x - x_j\|) &= \frac{d^\mu}{dx^\mu}\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} e^{i\langle x-x_j,t\rangle}\hat{\psi}(t)dt \\ &= \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} (it)^\mu e^{i\langle x-x_j,t\rangle}\hat{\psi}(t)dt \end{aligned}$$

to get

$$\begin{aligned} & U^T AU - 2U^T R^{(\mu)}(x) + \psi^{(2\mu)}(0) \\ = & \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \left| \sum_{j=1}^{M} u_j e^{i\langle x_j,t\rangle} - (it)^\mu e^{i\langle x,t\rangle} \right|^2 \hat{\psi}(t)dt \end{aligned} \tag{3.2}$$

for arbitrary $U = (u_1, \ldots, u_M) \in \mathbb{R}^M$, expressing the Kriging function via an integral. This explains the choice of $\psi^{(2\mu)}(0)$ as the additive constant in (2.5).

For most of the interesting radial basis functions ϕ, however, we have to use **generalized** Fourier transforms in (3.1) and (3.2). If (3.1) is interpreted as a generalized Fourier transform (e.g.: in the sense of [7]), the same interpretation applies to (3.2). Fortunately, we can circumvent these peculiarities, because the function

$$g_U(t) := \sum_{j=1}^{M} u_j e^{i\langle x_j,t\rangle} - (it)^\mu e^{i\langle x,t\rangle} \tag{3.3}$$

(for U, x_j, x, and μ fixed) has special properties which make the integral (3.2) well-defined in the classical sense for most of the interesting cases, provided that $U = (u_1, \ldots, u_M)^T$ lies in the admissible set $K_q^{(\mu)}(x)$ of (2.9). To prove this, we start with

Lemma 3.4 *For every admissible vector $U \in K_q^{(\mu)}(x)$ the function $g_U(t)$ has the property*

$$|g_U(t)| \le \left\{ \begin{array}{ll} \mathcal{O}(\|t\|^q) & \textit{for } \|t\| \to 0 \\ \mathcal{O}(\|t\|^{|\mu|}) & \textit{for } \|t\| \to \infty \end{array} \right\}. \tag{3.5}$$

Proof. Let

$$e^x = p_q(x) + x^q r_q(x),\ p_q \in I\!\!P_q,\ |r_q(x)| \le e^{|x|}$$

be the Taylor expansion of e^x to order q, and let $U \in K_q^{(\mu)}(x)$ be admissible for (2.5). For $q > 0$ we use (2.9) to get

$$\begin{aligned} \sum_{j=1}^M u_j p_q(i\langle x_j - x, t\rangle) &= \left(\frac{d^\mu}{dy^\mu}\right)_{y=x} p_q(i\langle y - x, t\rangle) \\ &= (it)^\mu p_q^{(\mu)}(0) \\ &= \left\{ \begin{array}{ll} (it)^\mu & 0 \le |\mu| \le q-1 \\ 0 & \text{else} \end{array} \right\}. \end{aligned}$$

Then (3.3) yields

$$\begin{aligned} g_U(t)e^{-i\langle x,t\rangle} &= \sum_{j=1}^M u_j e^{i\langle x_j - x, t\rangle} - (it)^\mu \\ &= \sum_{j=1}^M u_j p_q(i\langle x_j - x, t\rangle) - (it)^\mu \\ &\quad + \sum_{j=1}^M u_j (i\langle x_j - x, t\rangle)^q r_q(i\langle x_j - x, t\rangle) \\ &= \sum_{j=1}^M u_j (i\langle x_j - x, t\rangle)^q r_q(i\langle x_j - x, t\rangle) \end{aligned}$$

for $0 \le |\mu| \le q-1$. This proves (3.5), if (3.3) is used directly for $\|t\| \to \infty$. The cases $q = 0$ and $|\mu| \ge q$ now are easy.

With Lemma 3.4 and some additional assumptions on $\hat{\psi}$ the integral in (3.2) can now be shown to exist classically:

Theorem 3.6 *Let the generalized Fourier transform of $\psi(x) = \phi(\|x\|)$ exist and coincide with a continuous function $\hat{\psi}$ on $R^n \setminus \{0\}$ satisfying*

$$0 < \hat{\psi}(t) \le c \left\{ \begin{array}{ll} ||t||^{-n-s_0} & \textit{for } \|t\| \to 0 \\ ||t||^{-n-s_\infty} & \textit{for } \|t\| \to \infty \end{array} \right\} \tag{3.7}$$

with constants $c \in I\!\!R_{>0}$, s_0, $s_\infty \in I\!\!R$, where we additionally assume

$$2|\mu| < s_\infty \ \textit{ and } \ s_0 < 2q. \tag{3.8}$$

Then for all $U \in K_q^{(\mu)}(x)$ we have (3.2) with a well-defined integral.

Proof: From Lemma 3.4, (3.8), and (3.7) we get the classical existence of the integral in (3.2). With (3.3) and the theorem on monotone convergence, (3.2) equals

$$\begin{aligned} &\lim_{m\to\infty} \frac{1}{(2\pi)^n} \int_{R^n} |g_U(t)|^2 e^{-||t||^2/m^2} \hat{\psi}(t)dt \\ &= \lim_{m\to\infty} \frac{1}{(2\pi)^n} \int_{\boldsymbol{R}^n} \hat{G}_m(z)\psi(z)dz \end{aligned} \tag{3.9}$$

with the test functions

$$G_m(t) = |g_U(t)|^2 e^{-||t||^2/m^2}$$

using the definition of the generalized Fourier transform for tempered distributions. Furthermore, the Fourier transform $\hat{G}_m(z)$ of G_m can be explicitly calculated up to a constant σ_{mn} as

$$\begin{aligned} & m^n \sum_{j,k} u_j u_k e^{-||z-(x_j-x_k)||^2 m^2/4} \\ - \quad & 2m^n(-1)^n \sum_k u_k D^\mu (e^{-||z-(x-x_k)||^2 m^2/4}) \\ + \quad & m^n D^{2\mu}(e^{-||z||^2 m^2/4}), \end{aligned}$$

where D denotes differentiation with respect to z. Insertion into (3.9) yields (3.2), using the properties of the delta sequence $\sigma_{mn} m^n e^{-||z||^2 m^2/4}$. □

The reproducing kernel Hilbert space approach of Madych and Nelson [9] [10] uses a different method of regularization of integrals of the form (3.2): they consider a space of test functions modulo L_2–orthogonality to polynomials and have to go all through a specific theory of distributions to make their variational problem well–defined. Note that in our approach there are no problems with the variational formulation; the specific integral representation needed elaboration. We now can write the Kriging function as an integral:

Theorem 3.10 *Under the assumptions of Theorem 3.6 the Kriging function has the representation*

$$\left(\kappa_q^{(\mu)}(x)\right)^2 = \min_{U\in K_q^{(\mu)}(x)} \frac{1}{(2\pi)^n} \int_{\boldsymbol{R}^n} \left| \sum_{j=1}^M u_j e^{i\langle x_j,t\rangle} - (it)^\mu e^{i\langle x,t\rangle} \right|^2 \hat{\psi}(t)dt \tag{3.11}$$

with $K_{\mu,q}(x)$ *defined as in* (2.9)*, and the integral exists in the classical sense.* □

4 Error bounds

If the data (x_j, f_j) stem from a smooth and absolutely integrable real-valued function f on $I\!R^n$ with a nicely behaving Fourier transform $\widehat{f}$ satisfying

$$f(x) = \frac{1}{(2\pi)^n} \int_{\boldsymbol{R}^n} e^{i\langle x,t\rangle} \hat{f}(t)dt, \quad x \in I\!R^n,$$

we can use the solution $U^{(\mu)}(x) = (u_1^{(\mu)}(x), \ldots, u_M^{(\mu)}(x))^T$ of (2.6), which coincides with the derivatives of the Lagrange interpolation functions from (2.1), to find the error representation

$$|s^{(\mu)}(x) - f^{(\mu)}(x)|^2 = \left| \frac{1}{(2\pi)^n} \int_{\boldsymbol{R}^n} \left(\sum_{j=1}^M u_j^{(\mu)}(x) e^{i\langle x_j,t\rangle} - (it)^\mu e^{i\langle x,t\rangle} \right) \hat{f}(t)dt \right|^2 \tag{4.1}$$

of the interpolant s to f in the form (2.1), as far as it is feasible to take derivatives. Note that the bracketed function in (4.1) can be viewed as the Fourier transform of the representer of the error functional in some Hilbert space. It also is a special instance of a function of the form (3.3), but with optimal coefficients with respect to the minimization problem (2.5). The right-hand side of (4.1) is very similar to the Kriging function, and to relate the two we assume

$$c_f^2 := \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} |\hat{f}(t)|^2(\hat{\psi}(t))^{-1}dt < \infty \tag{4.2}$$

in addition to the hypotheses of Definition 2.8 and Theorem 3.6. Then, using the Cauchy–Schwarz inequality,

$$\begin{aligned} &|s^{(\mu)}(x) - f^{(\mu)}(x)|^2 \leq \\ &\leq \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\left|\sum_{j=1}^{M} u_j^{(\mu)}(x)e^{i\langle x_j,t\rangle} - (it)^\mu e^{i\langle x,t\rangle}\right|^2 \hat{\psi}(t)dt \cdot \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}|\hat{f}(t)|^2(\hat{\psi}(t))^{-1}dt \qquad (4.3)\\ &= (\kappa_q^{(\mu)}(x))^2 \cdot c_f^2, \end{aligned}$$

and the error of interpolation is pointwise bounded by the Kriging function.

This argument is feasible, whenever the Fourier transform $\widehat{f}$ of f allows all intermediate integrals to exist classically. To cope with generalized Fourier transforms, we proceed as in the previous section:

Definition 4.4 *A function $f : \mathbb{R}^n \to \mathbb{R}$ is* **dominated** *by a radial basis function ϕ satisfying* (3.7) *and* (3.8) *on $R^n \setminus \{0\}$, iff f has a generalized Fourier transform $\hat{f}$ coinciding on $R^n \setminus \{0\}$ with a continuous function satisfying* (4.2) *for $\psi(x) = \phi(\|x\|)$.*

Remarks. The set F_ϕ of functions dominated by ϕ may be completed to form a Hilbert space with inner product

$$(f_1, f_2)_\phi = \int_{\mathbb{R}^n} \hat{f}_1(t)\overline{\hat{f}_2(t)}(\hat{\psi}(t))^{-1}dt$$

which was thoroughly studied by Madych and Nelson in [9], [10]. This is why the rather restrictive condition (4.2) occurs there, too, while (4.2) does not appear in [3] and [6]. We do not want to make direct use of Hilbert space properties here, but proceed directly to an estimate of $\kappa_q^{(\mu)}(x)$ in the next section. The application of the Cauchy-Schwarz inequality to (4.1) allows the error bound to be factored into a term c_f depending on f, but not on the data distribution, while the other factor consists of the Kriging function, which is independent of f but incorporates the sampling points. This is how the special form of (1.5) is obtained, and this is why we need (4.2). Of course, equation (4.3) shows that the Kriging function $\kappa_q^{(\mu)}(x)$ is nothing else than the norm of the representer of the error functional on Madych and Nelson's Hilbert space. But knowledge of this fact does not improve the situation; a tight upper bound on $\kappa_q^{(\mu)}(x)$ is still to be constructed. This can be done by inserting special admissible vectors $U \in K_q^{(\mu)}(x)$ into (3.11). The next section will perform such a construction. We summarize the results of this section:

Theorem 4.5 *If f is in the space F_ϕ of functions dominated by a radial basis function ϕ satisfying* (3.8) *and* (3.7), *then the interpolation error can be bounded by*

$$|s^{(\mu)}(x) - f^{(\mu)}(x)| \leq \kappa_q^{(\mu)}(x) \cdot c_f, \tag{4.6}$$

where c_f is given by (4.2).

Proof: If $\tilde{g}$ is the bracketed function in (4.1), then $\tilde{g}\sqrt{\hat{\psi}}$ and $\hat{f}/\sqrt{\hat{\psi}}$ are in $L^2(I\!R^n)$. Thus the right-hand side of (4.1) is well-defined and (4.3) implies (4.6). $\square$

5 Construction of admissible vectors for local error bounds

We start with a fundamental perturbation lemma which bridges the gap between regular and scattered data sets:

Lemma 5.1 *For arbitrary constants $\rho \in I\!R_{>0}$, $k \in I\!N$ there exist positive real constants h_0, c_1, c_2 such that for any distribution of scattered centres $x_i \in \Omega$, $1 \le i \le M$ and arbitrary points $x \in \Omega$ around which the local density of data points satisfies*

$$h_\rho(x) := \max_{y \in K_\rho(x)} \min_{1 \le i \le M} \|y - x_i\| \le h_0 \tag{5.2}$$

and all $\mu \in I\!N^n_{\ge 0}$ with $0 \le |\mu| \le k-1$ there is a vector $\tilde{U} := (\tilde{u}_1^{(\mu)}(x), \ldots, \tilde{u}_M^{(\mu)}(x))^T$ in the admissible set $K_k^{(\mu)}(x)$ of (2.9) such that the inequalities

$$\|x_j - x\| \le c_1 h_\rho(x) \ \text{ for all } j,\ 1 \le j \le M \ \text{ with } \tilde{u}_j^{(\mu)}(x) \ne 0 \tag{5.3}$$

$$\sum_{j=1}^{M} |\tilde{u}_j^{(\mu)}(x)| \le c_2 h_\rho^{-|\mu|}(x) \tag{5.4}$$

hold.

Proof. The matrix $R_k = (\alpha^\mu)_{0 \le |\mu|, |\alpha| \le k-1}$ is a multidimensional form of the classical Vandermonde matrix; it is nonsingular because interpolation with polynomials of total order not exceeding k is possible on the set $S_k^n := \{\alpha \in I\!N^n_{\ge 0} \mid 0 \le |\alpha| \le k-1\}$. Here and in the rest of this section the n-dimensional multi–indices α, β, μ, ν, and σ will always vary in S_k^n. When occurring as indices in matrices, we assume a fixed (e.g.: lexicographic) ordering, and the first index at the brackets in a matrix notation like $B = (b_{\mu\beta})_{\mu,\beta}$ always indicates the row index. We use ρ, k, and n to define constants N, c_1, c_2, h_0 by

$$\begin{array}{llll} \text{(a)} & N & := & \max\left(1, 2(k-1)(2k-1)^k \binom{n+k-1}{n} n^{k/2} \|R_k^{-1}\|_\infty\right), \\ \text{(b)} & c_1 & := & 1 + N(k-1)\sqrt{n}, \\ \text{(c)} & c_2 & := & 2(k-1)! \binom{n+k-1}{n}^2 \|R_k^{-1}\|_\infty, \\ \text{(d)} & h_0 & := & \dfrac{\rho}{N(k-1)}, \end{array} \tag{5.5}$$

where $\|.\|_\infty$ is the row-sum matrix norm. Now let the distribution of centres be dense enough to guarantee (5.2) for $x \in I\!R^n$. Then near all of the points $x + Nh\alpha$ with $\alpha \in S_k^n$, which are in the ball $K_\rho(x) := \{y \in I\!R^n \mid \|y - x\| \le \rho\}$ by (d) of (5.5), there must be points $z_\alpha \in X$ with

$$\|z_\alpha - x - Nh\alpha\| \le h := h_\rho(x). \tag{5.6}$$

Now we get (5.3) for precisely these points z_α, if we use (b) of (5.5) for

$$\begin{aligned} \|z_\alpha - x\| &\leq h + \|Nh\alpha\|_2 \\ &\leq h(1+N(k-1)\sqrt{n}) = c_1 h. \end{aligned} \tag{5.7}$$

We now want to bound the difference of the two Vandermonde matrices $V := ((\alpha \cdot N \cdot h)^\mu)_{\mu,\alpha}$ and $Z := ((z_\alpha - x)^\mu)_{\mu,\alpha}$ with respect to regular and scattered data, respectively, using (5.6). By a straightforward expansion of the monomial t^μ we find

$$|t^\mu - (t+y)^\mu| \leq |\mu| \; \|y\|_\infty \left(\max_{0\leq\tau\leq 1} (\|t+\tau y\|_\infty) \right)^{|\mu|-1}$$

and get the bound

$$|(z_\alpha - x)^\mu - (\alpha N h)^\mu| \leq |\mu| h((c_1 + (k-1)N)h)^{|\mu|-1} \tag{5.8}$$

for the matrix elements of $Z - V$. The scaled regular Vandermonde matrix $V = ((\alpha N h)^\mu)_{\mu,\alpha}$ is expressible as

$$V = ((\alpha N h)^\mu)_{\mu,\alpha} = (\delta_{\mu\alpha} \cdot (Nh)^{|\mu|})_{\mu,\alpha} \cdot R_k \tag{5.9}$$

via the unscaled (integer) Vandermonde matrix R_k. We now use (5.8) and (5.9) for

$$\begin{aligned} \|V^{-1}(Z-V)\|_\infty &= \|R_k^{-1} \cdot (\delta_{\mu\alpha}(Nh)^{-|\mu|})_{\mu,\alpha}(Z-V)\|_\infty \\ &\leq \|R_k^{-1}\|_\infty \max_\mu (Nh)^{-|\mu|} |\mu| h^{|\mu|} (c_1 + (k-1)N)^{|\mu|-1} \binom{n+k-1}{n} \\ &\leq \|R_k^{-1}\|_\infty \binom{n+k-1}{n} \frac{1}{N} \max_\mu |\mu| \; \left(\frac{c_1}{N} + k - 1\right)^{|\mu|-1}. \end{aligned} \tag{5.10}$$

With (a) and (b) of (5.5) we find

$$\begin{aligned} \frac{c_1}{N} + k - 1 &= \tfrac{1}{N} + (k-1)(1+\sqrt{n}) \\ &\leq 1 + (k-1)(1+\sqrt{n}) \\ &\leq k + (k-1)\sqrt{n} \\ &\leq (2k-1)\sqrt{n} \end{aligned}$$

and continue (5.10) to get

$$\|V^{-1}(Z-V)\|_\infty \leq \|R_k^{-1}\|_\infty \binom{n+k-1}{n} \frac{1}{N} \cdot (k-1)(2k-1)^k n^{k/2} \leq \frac{1}{2}. \tag{5.11}$$

By Neumann's series, $B := Z^{-1}V$ exists and its norm is bounded by 2. Some elementary matrix calculations give

$$(\delta_{\alpha\mu}(Nh)^{|\mu|})_{\mu,\alpha} = (ZBR_k^{-1})_{\mu,\alpha}$$

and we can define

$$\tilde{u}_i^{(\mu)}(x) := \left\{ \begin{array}{cc} \mu! \left(BR_k^{-1}\right)_{\alpha,\mu} (Nh)^{-|\mu|} & \text{if } \; x_i = z_\alpha \\ 0 & \text{otherwise} \end{array} \right\}.$$

Note that the values $\tilde{u}_i^{(\mu)}(x)$ vanish except for those points x_i for which we have proven (5.7). This finishes (5.3). We now assert

$$\sum_{j=1}^{M} \tilde{u}_i^{(\mu)}(x)p(x_j) = p^{(\mu)}(x)$$

for all $p \in I\!P_k$, which is equivalent to $\tilde{U} := (\tilde{u}_1^{(\mu)}(x), \dots, \tilde{u}_M^{(\mu)}(x))^T \in K_k^{(\mu)}(x)$ of (2.9) and to

$$\sum_{j=1}^{M} \tilde{u}_i^{(\mu)}(x)p(x_j - x) = p^{(\mu)}(0) \tag{5.12}$$

for all $p \in I\!P_k$. Representing an arbitrary $p \in I\!P_k$ in the monomial basis as

$$p(y) = \sum_{|\nu|<k} p_\nu y^\nu$$

we use the elements $Z_{\nu,\alpha} = (z_\alpha - x)^\nu$ of the Vandermonde matrix Z to get

$$p(z_\alpha - x) = \sum_\nu p_\nu (z_\alpha - x)^\nu = \sum_\nu Z_{\nu,\alpha} p_\nu$$

and prove (5.12) via

$$\begin{aligned}
p^{(\mu)}(0) &= \mu! p_\mu = \mu!(Nh)^{-|\mu|} \sum_{|\sigma|<k} p_\sigma \delta_{\mu,\sigma} (Nh)^{|\sigma|} \\
&= \mu!(Nh)^{-|\mu|} \sum_\sigma p_\sigma \left(ZBR_k^{-1}\right)_{\sigma,\mu} \\
&= \mu!(Nh)^{-|\mu|} \sum_\alpha \left(BR_k^{-1}\right)_{\alpha,\mu} \sum_\sigma Z_{\sigma,\alpha} p_\sigma \\
&= \mu!(Nh)^{-|\mu|} \sum_\alpha \left(BR_k^{-1}\right)_{\alpha,\mu} p(z_\alpha - x) \\
&= \sum_{j=1}^{M} \tilde{u}_j^{(\mu)}(x) p(x_j - x).
\end{aligned}$$

Finally, we establish (5.4) using (c) of (5.5) and

$$\begin{aligned}
\sum_{j=1}^{M} |\tilde{u}_i^{(\mu)}(x)| &\le \mu!(Nh)^{-|\mu|} \sum_\alpha \left| \left(BR_k^{-1}\right)_{\alpha,\mu} \right| \\
&\le \mu!(Nh)^{-|\mu|} \|BR_k^{-1}\|_1 \\
&\le (k-1)! h^{-|\mu|} \binom{n+k-1}{n}^2 2 \|R_k^{-1}\|_\infty,
\end{aligned}$$

where we had to go from the column-sum norm $\|.\|_1$ to the row-sum norm $\|.\|_\infty$. □

Now assume $\rho \in I\!R_{>0}$ and an integer $k \ge q$ to be chosen for Lemma 5.1 such that (5.3) and (5.4) hold, whenever (5.2) is satisfied. Then the proof technique of Lemma 3.4 yields

$$\begin{aligned}
\left| g_{\tilde{U}}(t) e^{-i\langle x,t\rangle} \right| &= \left| \sum_{j=1}^{M} \tilde{u}_j^{(\mu)} e^{i\langle x_j - x, t\rangle} - (it)^\mu \right| \\
&\le c_2 h_\rho^{-|\mu|}(x) c_1^k h_\rho^k(x) \|t\|^k e^{c_1 h_\rho(x)\|t\|}
\end{aligned}$$

for $0 \le |\mu| < k$, which specializes to

$$\left|g_{\tilde U}(t)e^{-i\langle x,t\rangle}\right| \le c_2 c_1^k e^{c_1}\|t\|^k h_\rho^{k-|\mu|}(x)$$

for $h_\rho(x)\cdot\|t\|\le 1$. Direct application of (5.4) to $g_{\tilde U}$ implies

$$\left|g_{\tilde U}(t)e^{-i\langle x,t\rangle}\right| \le c_2 h_\rho^{-|\mu|}(x) + \|t\|^{|\mu|} \le (c_2+1)\|t\|^{|\mu|}.$$

for $h_\rho(x)\cdot\|t\|\ge 1$. Both bounds can be inserted into the Kriging function to get

$$\begin{aligned}\left(\kappa_q^{(\mu)}(x)\right)^2 &\le c_2^2 c_1^{2k} e^{2c_1} h_\rho^{2k-2|\mu|}(x) K_1(2k,\hat\psi,h_\rho(x))\\ &+ (1+c_2)^2 K_2(2|\mu|,\hat\psi,h_\rho(x))\end{aligned}$$

for $0\le|\mu|<\min(k,s_\infty/2)$ with the abbreviations

$$\begin{aligned}K_1(j,z,h) &:= \frac{1}{(2\pi)^n}\int_{\|t\|\le 1/h}\|t\|^j z(t)dt,\\ K_2(j,z,h) &:= \frac{1}{(2\pi)^n}\int_{\|t\|\ge 1/h}\|t\|^j z(t)dt.\end{aligned} \tag{5.13}$$

To evaluate $K_1(2k,\hat\psi,h)$ for $\hat\psi$ satisfying Theorem 3.6, we make use of (3.8) and $s_0 < 2q \le 2k$ to get

$$\begin{aligned}K_1(2k,\hat\psi,h) &= K_1(2k,\hat\psi,h_0) + \frac{1}{(2\pi)^n}\int_{1/h_0\le\|t\|\le 1/h}\|t\|^{2k}\hat\psi(t)dt\\ &\le \mathcal{O}(1)+\mathcal{O}(h^{s_\infty-2k})\end{aligned}$$

for $h\to 0$. Likewise, we find

$$K_2(2|\mu|,\hat\psi,h) = \mathcal{O}(h^{s_\infty-2|\mu|})$$

using $2|\mu| < s_\infty$ from (3.8). We summarize:

Theorem 5.14 *Let ϕ satisfy the assumptions of Theorem 3.6, and let $\rho\in I\!R_{>0}$ be given. Then there exist positive real constants h_0 and C such that for any distribution of centres $x_i\in I\!R^n$, $1\le i\le M$, and any point $x\in I\!R^n$ with* (5.2) *the Kriging function can be bounded by*

$$\kappa_q^{(\mu)}(x) \le C h_\rho^{s_\infty/2-|\mu|}(x) \tag{5.15}$$

for $0\le|\mu|<s_\infty/2$.

Proof: Lemma 5.1 allows to pick any $k\ge s_\infty/2$ and to apply the above argument, where large values of k will lead to small values of h_0 and large values of C. Then

$$\begin{aligned}\left(\kappa_q^{(\mu)}(x)\right)^2 &\le \mathcal{O}(h_\rho^{2k-2|\mu|}(x))\cdot(\mathcal{O}(1)+\mathcal{O}(h_\rho^{s_\infty-2k}(x)))+\mathcal{O}(h_\rho^{s_\infty-2|\mu|}(x))\\ &\le \mathcal{O}(h_\rho^{2k-2|\mu|}(x))+\mathcal{O}(h_\rho^{s_\infty-2|\mu|}(x))\end{aligned}$$

yields (5.15). □

6 Special radial basis functions

All examples of this section are based on Theorems 4.5 and 5.14 when specialized to certain radial basis functions ϕ. We obtain specific error estimates of the form (1.5) valid for all functions dominated by ϕ in the sense of Definition 4.4. The attainable orders of the interpolation error will only depend on the constants s_∞ for the radial basis functions in question.

In case of Gaussians

$$\phi(r) = e^{-\alpha r^2}, \quad \alpha \in I\!R_{>0}, \quad r \in I\!R_{\geq 0}$$

the Fourier transform

$$\hat{\psi}(y) = 2^{-n}(\alpha\pi)^{-n/2} e^{-\|y\|^2/(4\alpha)}$$

decays exponentially at infinity and we have arbitrarily large values of s_∞, giving

Theorem 6.1 *For all functions* $f : I\!R^n \to I\!R$ *having generalized Fourier transforms* $\hat{f}(t)$ *satisfying*

$$\int_{R^n} \|\hat{f}(t)\|^2 e^{\|t\|^2} dt < \infty$$

the error of interpolation by Gaussian radial basis functions has arbitrarily high local order in the sense of (1.5). □

In case of multiquadrics

$$\phi(r) = (c^2 + r^2)^{s/2} \quad s \in I\!R_{>-n}, \ s \notin 2Z\!\!Z,$$

we have generalized Fourier transforms containing a Bessel function of the second kind and decaying exponentially at infinity like in the Gaussian case, but now with a singularity of type $\|y\|^{-n-s}$ at zero. Thus again s_∞ is arbitrarily large, yielding

Theorem 6.2 *For all functions* $f : I\!R^n \to I\!R$ *having generalized Fourier transforms* $\hat{f}(t)$ *satisfying*

$$\int_{R^n} |\hat{f}(t)|^2 \|t\|^{n+s} e^{\|t\|} dt < \infty$$

the error of interpolation by multiquadric radial basis functions

$$\phi(r) = (c^2 + r^2)^{s/2} \quad s \in I\!R_{>-n}, \ s \notin 2Z\!\!Z$$

has unbounded local order in the sense of (1.5).

The above results for special radial basis functions were already obtained by Madych and Nelson [10] as a result of their reproducing kernel Hilbert space theory. Within the latter, the integral occurring in (3.2) plays an important part, as was pointed out by N. Dyn in her lucid survey [4] of Madych and Nelson's work. There the integral is split at $\|t\| = 1$ and not at $\|t\| = 1/h$, which makes it impossible to handle the cases that follow.

The radial basis functions

$$\begin{aligned} \phi(r) &= r^s, & s \in I\!R_{>1}, \ s \notin 2I\!N, \\ \phi(r) &= r^s \log r, & s \in 2I\!N, \end{aligned}$$

have generalized Fourier transforms $\hat{\psi}(t)$ with behaviour like $\|t\|^{-n-s}$ on all of $I\!R^n$. This requires $s_\infty = s, \ q > s/2$ and yields

Theorem 6.3 *For all functions $f : I\!R^n \to I\!R$ having generalized Fourier transforms satisfying*

$$\int_{\mathbb{R}^n} |\hat{f}(t)|^2 \|t\|^{n+s} dt < \infty$$

the error of interpolation by radial basis functions

$$\begin{aligned} \phi(r) &= r^s, & s \in I\!R_{>1},\ s \notin 2I\!N, \\ \phi(r) &= r^s \log r, & s \in 2I\!N, \end{aligned}$$

has local order $s/2$ in the sense of (1.5).

Remarks. The convergence rates obtained so far are quite different from those of Buhmann [1] for gridded data. Infinite orders, especially for the Gaussian kernel, are unexpected as are the rather low orders for the functions r^s and $r^s \log r$. The gridded case is fundamentally different from the scattered data case in several aspects:

- Convergence orders, as given by our technique, are mainly dominated by the behaviour of the Fourier transform of ϕ near infinity. However, similar convergence orders are obtained, if multiquadric interpolation with a fixed constant c is used on a regular grid (see Buhmann and Dyn [2]).
- Our methods require the interpolated functions to have Fourier transforms which are "dominated" by $\widehat{\phi}$ in the sense of (4.2). This requirement is not surprising, because a reasonable local convergence order cannot occur if the "high frequency part" of f exceeds that of ϕ. Interpolation on grids hides this phenomenon.
- The "dominance" requirement (4.2) imposes a rather hard restriction on f, which has no direct counterpart in the infinite grid case. Further research should try to weaken (4.2) at the expense of smaller (and possibly non–local) convergence orders.

References

[1] Buhmann, M.D., Multivariable Interpolation using Radial Basis Functions, Ph.D. dissertation, University of Cambridge, 1989.

[2] Buhmann, M.D., and N. Dyn, Error estimates for multiquadric interpolation, in *Curves and Surfaces*, P.-J. Laurent, A. LeMéhauté, L.L. Schumaker (eds.), Academic Press 1991

[3] Duchon,J., Sur L'erreur d'interpolation des fonctions de plusiers variables par les D^m-splines, RAIRO Analyse Numerique 12 (1978) 325-334.

[4] 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

[5] Dyn. N, Levin, D., and S. Rippa, Numerical procedures for surface fitting of scattered data by radial functions, SIAM J. Sci. Stat. Comput. **7**, 639–659

[6] Jackson, I.R.H., An order of convergence for some radial basis functions, IMA J. of Numer. Anal. 1987

[7] Jones, D.S., The Theory of Generalised Functions, Cambridge University Press, Cambridge 1982

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

[9] Madych, W.R., and Nelson, S.A., Multivariate Interpolation and conditionally positive definite functions, Approx. Theory and its Applications 4 (1988), 77–89

[10] Madych, W.R., and Nelson, S.A., Multivariate Interpolation and conditionally positive definite functions II, Math. of Computation 54 (1990), 211–230

[11] Madych, W.R., and Nelson, S.A., Error Bounds for Multiquadric Interpolation, in: *Approximation Theory VI: Vol 2.* C. K. Chui, L. L. Schumaker and J. D. Ward.(eds.) 413-416.

[12] 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

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

Author's addresses:

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

Robert Schaback, Institut für numerische und angewandte Mathematik, Universität Göttingen, Lotzestraße 16–18, D–3400 Göttingen, FRG