Lower Bounds for Norms of Inverses of Interpolation Matrices for Radial Basis Functions

Robert Schaback

Abstract: Interpolation of scattered data at distinct points $x_1, \ldots, x_n \in I\!R^d$ by linear combinations of translates $\Phi(\|x - x_j\|_2)$ of a radial basis function $\Phi : I\!R_{\geq 0} \to I\!R$ requires the solution of a linear system with the n by n distance matrix $A := (\Phi(\|x_i - x_j\|_2)$. Recent results of Ball, Narcowich and Ward, using Laplace transform methods, provide upper bounds for $\|A^{-1}\|_2$, while Ball, Sivakumar, and Ward constructed examples with regularly spaced points to get special lower bounds. This paper proves general lower bounds by application of results of classical approximation theory. The bounds increase with the smoothness of Φ. In most cases, they leave no more than a factor of n^{-2} to be gained by optimization of data placement, starting from regularly distributed data. This follows from comparison with results of Ball, Baxter, Sivakumar, and Ward for points on scaled integer lattices and supports the hypothesis that regularly spaced data are near-optimal, as far as the condition of the matrix A is concerned.

Keywords: Radial basis functions, multiquadrics, scattered data interpolation, lower bounds, condition numbers.

Classification: 65D05, 41A05, 41A63

1 Introduction

Let $\Phi : I\!R_{\geq 0} \to I\!R$ be a scalar ("radial") function, and let n distinct points ("centres") $x_1, \ldots, x_n \in I\!R^d$ be given, forming a set $X := \{x_1, \ldots, x_n\}$. As reported by Hardy [8] and Franke [7], interpolation of real values y_i, $1 \leq i \leq n$, at the centres x_i by linear combinations

$$s(x) := \sum_{j=1}^{n} \alpha_j \Phi(\|x - x_j\|_2),\ \alpha_j \in I\!R,\ x \in I\!R^d \tag{1.1}$$

of translates of $\Phi(\|.\|_2)$ can produce very good numerical results, if Φ is the "multiquadric" $\Phi_c(r) = (c^2 + r^2)^{1/2}$, for instance. The interpolation problem for a function of the form (1.1) requires the solution of the linear system

$$\sum_{j=1}^{n} \alpha_j \Phi(\|x_i - x_j\|_2) = y_i,\ 1 \leq i \leq n. \tag{1.2}$$

In some cases the n functions $\Phi(\| \cdot - x_j\|_2)$ are augmented by a basis of the space $I\!P_m^d$ of all d-variate polynomials of total degree less than m. Let

$$\mu := \mu(m, d) := \binom{m - 1 + d}{d},\ m \geq 1,\ \mu(0, d) := 0 \tag{1.3}$$

[0]File : campaper.tex, TEXed May 18, 1994, Status : Revised for J. Approx. Th.

denote the dimension of $I\!P_m^d$, and let a basis be given by $q_1, \dots, q_\mu$. From now on, we suppress the dependence of μ on m and d to keep the notation simple. For $m > 0$ we additionally assume $q_1, \dots, q_\mu$ to be linearly independent over the set $X = \{x_1, \dots, x_n\}$ of centres, which implies $n \geq \mu$. Then the $(n + \mu)$ by $(n + \mu)$ matrix

$$A := \begin{pmatrix} \Phi(\|x_i - x_j\|_2) & q_k(x_i) \\ q_k(x_j) & 0 \end{pmatrix}_{1 \leq i,j \leq n,\ 1 \leq k \leq \mu} \tag{1.4}$$

which occurs in the generalization

$$\begin{array}{lclcll} \sum_{j=1}^{n} \alpha_j \Phi(\|x_i - x_j\|_2) & + & \sum_{k=1}^{\mu} \beta_k q_k(x_i) & = & y_i, & 1 \leq i \leq n \\ \sum_{j=1}^{n} \alpha_j q_k(x_j) & + & 0 & = & 0, & 1 \leq k \leq \mu \end{array} \tag{1.5}$$

of the system (1.2) is called a (generalized) distance matrix. It is nonsingular in the cases

$$\begin{array}{lcll} \Phi(r) & = & (c^2 + r^2)^\beta, & \beta,\ c > 0,\ \beta \notin 2Z\!\!\!Z,\ m \geq 0 \\ & & & \text{(Multiquadrics for } \beta = 1/2) \\ \Phi(r) & = & (c^2 + r^2)^{-\beta}, & \beta,\ c > 0,\ m \geq 0 \\ & & & \text{(Inverse Multiquadrics for } \beta = 1/2) \\ \Phi(r) & = & r^{2\beta} \log r, & \beta = m - d/2 > 0,\ d \in 2I\!N \\ & & & \text{(Thin-plate splines)} \\ \Phi(r) & = & \log(c^2 + r^2), & c > 0,\ m \geq 0 \\ \Phi(r) & = & r^{2\beta}, & \beta = m - d/2,\ m > d/2,\ d \in 2I\!N - 1, \text{ or} \\ & & & \beta \in (0,1),\ d \in 2I\!N - 1,\ m \geq 0, \text{ or} \\ & & & \beta = 1/2,\ n \geq 2,\ m \geq 0 \\ \Phi(r) & = & \exp(-\alpha r^2), & \alpha > 0,\ m \geq 0, \text{ (Gaussians)}. \end{array}$$

This follows from the work of Micchelli [9] and Schoenberg [15] (see also Dyn [5] and Powell [14] for highly useful surveys of known results). Furthermore, condition numbers of A were often observed to be quite large. Special preconditioning strategies for solving the system (1.5) were supplied by Dyn, Levin, and Rippa [6] in a variety of special cases.

Some very interesting upper bounds for the spectral norm $\|A^{-1}\|_2$ of A were given by Ball [1], and Narcowich and Ward [10] [11] [12], using Laplace transform methods. These bounds are in terms of the "separation distance"

$$h := \min_{1 \leq i \neq j \leq n} \|x_i - x_j\|_2 \tag{1.6}$$

and hold for arbitrary placements of centres x_j, $1 \leq j \leq n$. Recently, a paper by Ball, Sivakumar, and Ward [2] derived lower bounds for $\|A^{-1}\|_2$ for special regular choices of centres by a similar technique. Furthermore, Baxter [4] has investigated the case of centres on subsets of the integer lattice. Using the Toeplitz structure of A, he derived bounds for $\|A^{-1}\|_2$ for a wide class of radial basis functions including multiquadrics. His bounds are asymptotically optimal in the sense that their behaviour for $n \to \infty$ is best possible up to a constant factor.

Since the existing lower bounds of $\|A^{-1}\|_2$ only hold for data on finite regular grids, the value of $\|A^{-1}\|_2$ might be decreased by irregular placements of centres.

This paper complements the work of Ball, Baxter, Narcowich, Sivakumar, and Ward by producing **general lower** bounds for $\|A^{-1}\|_2$, independent of the distribution of centres and independent of the separation distance, via a completely different approach. Combined with the other results, they provide a bound for the possible gain by optimizing the placement of centres with respect to the condition of A. The findings of this paper are roughly of the form

$$\|A^{-1}\|_2 \geq \frac{1}{nE(2(d!n/2)^{1/d}-1, K, \Phi)},$$

where $E(\ell, K, \Phi)$ denotes the error of best Chebyshev approximation by polynomials of degree less than ℓ to the function $\Phi(\sqrt{r})$ on the interval $[0, K^2]$ defined via the diameter

$$K := \max_{1\leq i,j\leq n} \|x_i - x_j\|_2 \tag{1.7}$$

of the data set in the Euclidean norm. This relates the bound to the smoothness of Φ and the dimension d of the space, and in a very remarkable way indeed: the bound tends to be smaller, if the smoothness of the radial basis functions is decreased. The results can be generalized to other matrix norms than the spectral norm to which the other bounds in the literature are confined.

2 Basic Results

We assume that the radial function Φ, the $n \geq 2$ centres $x_j \in X := \{x_1, \ldots, x_n\}$, the space dimension d, and the polynomial order m are such that the matrix A in (1.4) is nonsingular. For an arbitrary polynomial $p \in I\!P^1_\ell$ we define the matrix

$$A_p := \begin{pmatrix} p(\|x_i - x_j\|_2^2) & q_k(x_i) \\ q_k(x_j) & 0 \end{pmatrix}_{1\leq i,j\leq n,\ 1\leq k\leq \mu} \tag{2.1}$$

as an approximation to A. Note that A and A_p differ only in their upper n by n submatrices with entries $\Phi(\|x_i - x_j\|_2)$ and $p(\|x_i - x_j\|_2^2)$, respectively. Thus a good approximation of $\Phi(\sqrt{r})$ by a polynomial $p \in I\!P^1_\ell$ on the set T of all real values $t_{ij} := \|x_i - x_j\|_2^2$ will produce a good approximation of A by A_p.

We now pick two arbitrary norms $\|.\|_r$ and $\|.\|_s$ on $I\!R^{n+\mu}$ and define the usual operator norm

$$\|B\|_{r,s} := \sup \left\{ \|Bx\|_r / \|x\|_s \mid x \in I\!R^{n+\mu} \setminus \{0\} \right\}. \tag{2.2}$$

By the theorem of Weierstrass, there will be some $p \in I\!P^1_\ell$ for ℓ large enough such that

$$\|A - A_p\|_{s,r} \cdot \|A^{-1}\|_{r,s} < 1 \tag{2.3}$$

holds. We now assert that (2.3) is a sufficient condition for A_p to be nonsingular. In fact, for arbitrary vectors x, y with $A_p x = y$ we find

$$\begin{aligned} \|x\|_r &= \|A^{-1}Ax\|_r = \|A^{-1}(y + (A - A_p)x)\|_r \\ &\leq \|A^{-1}\|_{r,s}\|y\|_s + \|A^{-1}\|_{r,s}\|A - A_p\|_{s,r}\|x\|_r, \end{aligned}$$

and reordering yields a bound for $\|A_p^{-1}\|_{r,s}$.

Now, whenever A_p is nonsingular, the $n+\mu$ polynomials

$$\begin{array}{ll} p(\|\cdot - x_j\|_2^2), & 1 \le j \le n = |X| \\ q_k(\cdot), & 1 \le k \le \mu \end{array} \tag{2.4}$$

span the n-dimensional space $C(X)$, because the n first rows or columns of A_p are linearly independent. For any univariate polynomial $p \in I\!P_\ell^1$ we can define a subspace

$$P_p := \left\{ \sum_{k=1}^{\mu} \beta_k q_k + \sum_{t=1}^{n} \gamma_t p(\|\cdot - x_t\|_2^2) \ : \ \sum_{t=1}^{n} \gamma_t q_k(x_t) = 0,\ 1 \le k \le \mu \right\} \tag{2.5}$$

of the space $I\!P_{2\ell-1}^d$, and whenever A_p is nonsingular, we have $\dim P_p \ge n$, because the polynomials in (2.4) span the n-dimensional space $C(X)$. If we define

$$\mu^*(\ell) := \mu^*(X, \ell, m, d) := \max_{p \in \boldsymbol{P}_\ell^1} \dim P_p, \tag{2.6}$$

we conclude that

$$\mu^*(\ell) \ge |X| = n \tag{2.7}$$

holds, if ℓ is large enough to satisfy (2.3) for some $p \in I\!P_\ell^1$.

We now turn this argument upside down. Clearly the function $\mu^*(\ell)$ is (weakly) monotonic in ℓ and has the obvious bounds

$$\mu \le \mu^*(\ell) \le |X| + \mu.$$

Our argument above implies existence of the maximum in the definition of

$$\ell^* := \ell^*(X, m, d) := \max \{\ell \ge 0 \mid \mu^*(\ell) < |X|\}.$$

This yields

$$\mu^*(\ell^*) < |X| = n, \tag{2.8}$$

which will be needed in the proof of

Theorem 2.1 *The inverse of A, if it exists, satisfies*

$$\|A^{-1}\|_{r,s}^{-1} \le \inf \left\{ \|A - A_p\|_{s,r} \mid p \in I\!P_{\ell^*(X,m,d)}^1 \right\} \tag{2.9}$$

where A_p is defined by (2.1).

Proof: Take any polynomial $p \in I\!P_{\ell^*}^1$ and assume A_p to be nonsingular. Then the $n+\mu$ polynomials (2.4) span the space $C(X)$. Since they are in the space P_p occurring in (2.6), the inequality (2.8) is violated. Thus A_p must be singular for all $p \in I\!P_{\ell^*}^1$. But then the inequality (2.3) cannot hold because it would imply the nonsingularity of A_p, as was shown above. This proves the assertion. □

Theorem 2.1 relates lower bounds for $\|A^{-1}\|_{r,s}$ to a somewhat peculiar matrix-valued approximation problem. The error matrix is zero except for the upper left n by n submatrix of $A - A_p$ with entries

$$\Phi(\|x_i - x_j\|_2) - p(\|x_i - x_j\|_2^2),\ 1 \le i, j \le n.$$

Thus it involves an approximation of $\Phi(\sqrt{r})$ by polynomials of order at most ℓ^* on the set T of all real values $t_{ij} := \|x_i - x_j\|_2^2$ such that the function values are arranged in $n \times n$ matrix form and such that the approximation error is measured via the matrix norm $\|.\|_{s,r}$ from (2.2). The actual form of the approximation problem is thus determined by the matrix norm chosen. For the spectral norm $\|A - A_p\|_{2,2}$ we get a rather nasty approximation problem, but for the norm

$$\|B\|_{\infty,1} = \max_{1\le i,j\le n} |b_{ij}| \tag{2.10}$$

we can use Chebyshev approximation on T. Here, we interpreted the norms $\|.\|_r$ and $\|.\|_s$ of (2.2) as the usual L_r and L_s norms, but this is by no means mandatory. Studying linear approximation problems for peculiar norms like the spectral matrix norm may be of independent interest in approximation theory.

We want to give good asymptotic lower bounds for $\|A^{-1}\|_{r,s}$ in case $n \to \infty$, and we want to compare the bounds with those of Ball, Baxter, Narcowich, Sivakumar, and Ward for the spectral norm $\|A^{-1}\|_{2,2}$. We saw the latter to be rather inconvenient for our approach while good asymptotics are mainly available for Chebyshev approximation on real intervals. Therefore we try to get as far as we can with best Chebyshev approximation to the function $\Phi(\sqrt{r})$ on the real interval $[0, K^2]$ by polynomials of order ℓ. Note that this is equivalent to a best Chebyshev approximation of order $2\ell - 1$ to the function $\Phi(|r|)$ on $[-K, +K]$, which, by uniqueness and symmetry, must be a polynomial of maximal order ℓ in r^2. Then we denote the error by

$$\begin{aligned} E(\ell) := E(\ell, K, \Phi) &:= \min_{p\in \boldsymbol{P}_\ell^1} \max_{0\le r\le K^2} |\Phi(\sqrt{r}) - p(r)| \\ &= \min_{p\in \boldsymbol{P}_\ell^1} \|\Phi(|r|) - p(r^2)\|_{\infty,[-K,K]}, \end{aligned}$$

writing the Chebyshev norm of continuous functions on a real interval $[a, b]$ by $\|.\|_{\infty,[a,b]}$.

Theorem 2.2 *The inverse of A, if it exists, satisfies*

$$\|A^{-1}\|_{1,\infty} \ge \frac{1}{E(\ell^*)} \tag{2.11}$$

and

$$\|A^{-1}\|_{2,2} \ge \frac{1}{n \cdot E(\ell^*)}. \tag{2.12}$$

Proof: Inequality (2.11) readily follows from Theorem 2.1 and (2.10). Then we conclude (2.12) simply from (2.11) and

$$\|B\|_{1,\infty} \le n \cdot \|B\|_{2,2}$$

for arbitrary n by n matrices B. □

Unfortunately, our approach yields a factor of n^{-1} in the bound for the spectral norm; the comparisons at the end of the paper will indicate that we often seem to miss the actual behaviour of $\|A^{-1}\|_{2,2}$ for regular data asymptotically by just this factor. But there appear to be no other handy links between the spectral norm for matrices and the Chebyshev norm for the matrix entries. The factor does not arise if we use (2.11) or replace E in (2.12) by the (unknown) error of best approximation by polynomials in the spectral norm.

Theorem 2.2 relates lower bounds of $\|A^{-1}\|$ to the error of best Chebyshev approximations to $\Phi(|r|)$ by polynomials of order ℓ^* in the variable r^2, and we shall see below that ℓ^* behaves approximately like at least $n^{1/d}$ for large n. Altogether, our lower bounds will grow astronomically with n, if $\Phi(|r|)$ can be extended to an entire function in the complex plane $\mathbb{C}$ (e.g.: for Gaussians). They will be of polynomial growth, if $\Phi(|r|)$ has only finitely many continuous derivatives, which is the case for thin–plate splines and positive non-even rational powers of r. However, they will still grow exponentially for all $\Phi(|r|)$ which are analytic in $\mathbb{C}$ around $[0, K]$, e.g.: for multiquadrics and inverse multiquadrics, if c is fixed. To get around this, c must tend to zero for n tending to infinity, moving the singularity of Φ towards the real axis. The details are worked out in the rest of the paper, but due to classical Jackson–Bernstein theorems the lower bounds for $\|A^{-1}\|$ are directly related to the smoothness of Φ: they get larger, if the smoothness increases.

3 Auxiliary Results

The remaining task now is to evaluate $E(\ell^*(X, m, d), K, \Phi)$ for the functions listed in the introduction. Since the quantities $E(\ell, K, \Phi)$ can be estimated using classical results of approximation theory (this will be done in the following section), we first prove something about $\mu^*(\ell)$ and ℓ^* as defined in the beginning of the preceding section. For this, we suppose X, m, and d to be fixed, and we first look for an upper bound for μ^* from (2.6).

To derive simple results, the crude bound

$$\mu^*(\ell) \le \dim I\!P^d_{2\ell-1} = \mu(2\ell - 1, d)$$

for $2\ell > m$ suffices. It follows easily from (2.5) and uses the general definition of $\mu(\cdot, d)$ from (1.3). But there is an improvement of this bound:

Lemma 3.1 *Let p be a univariate polynomial of order $\le \ell$, where $\ell \ge m+1$. Then the subspace P_p, as defined in (2.5), is a subspace of $I\!P^d_{2\ell-1-m}$ and has dimension at most*

$$\dim I\!P^d_\ell + \dim I\!P^d_{\ell-1} - \dim I\!P^d_m = \mu(\ell, d) + \mu(\ell - 1, d) - \mu(m, d). \tag{3.1}$$

Furthermore,

$$\mu^*(\ell) \le \mu(\ell, d) + \mu(\ell - 1, d) - \mu(m, d).$$

Proof: By expansion of the translate

$$p_y(x) := p(\|x - y\|_2^2) = \sum_{k=0}^{\ell-1} \alpha_k \|x - y\|_2^{2k}$$

of p into sums of products of simpler terms we get the representation

$$\begin{aligned} p_y(x) &= \sum_{k=0}^{\ell-1} \alpha_k (x^T x - 2x^T y + y^T y)^k \\ &= \sum_{j=0}^{\ell-1} (-2x^T y)^j \sum_{i=0}^{\ell-j-1} (x^T x)^i \sum_{k=i+j}^{\ell-1} \alpha_k \binom{k}{j} \binom{k-j}{k-j-i} (y^T y)^{k-j-i}. \end{aligned}$$

Thus an arbitrary function taken from P_p in (2.5) can be written as

$$\begin{aligned}
& \sum_{k=1}^{\mu} \beta_k q_k(x) + \sum_{t=1}^{n} \gamma_t p(\|x - x_t\|^2) \\
= & \sum_{k=1}^{\mu} \beta_k q_k(x) \\
+ & \sum_{t=1}^{n} \gamma_t \sum_{j=0}^{\ell-1} (-2x^T x_t)^j \sum_{i=0}^{\ell-j-1} (x^T x)^i \sum_{k=i+j}^{\ell-1} \binom{k}{j} \binom{k-j}{k-j-i} \alpha_k (x_t^T x_t)^{k-j-i} \\
= & \sum_{k=1}^{\mu} \beta_k q_k(x) \\
+ & \sum_{j=0}^{\ell-1} \sum_{i=0}^{\ell-j-1} (x^T x)^i \sum_{k=i+j}^{\ell-1} \binom{k}{j} \binom{k-j}{k-j-i} \alpha_k \sum_{t=1}^{n} \gamma_t (-2x^T x_t)^j (x_t^T x_t)^{k-j-i}
\end{aligned}$$

and we note that the sum over t yields zero whenever $j + 2(k - i - j) \le m - 1$. Thus the range of i and j can be restricted to $m + j + 2i \le 2\ell - 2$, and the function will be contained in the subspace

$$Q = I\!P_m^d + \text{ span } \left\{ (x^T y)^j (x^T x)^i \;\middle|\; \begin{array}{c} 0 \le j \le \ell - 1,\ 0 \le i \le \ell - 1 - j \\ m + j + 2i \le 2\ell - 2,\ y \in I\!R^d \end{array} \right\}$$

of $I\!P_{2\ell-1-m}^d$. But the number of spanning functions can be further reduced. We assert that Q lies in the subspace

$$\widetilde{Q} = I\!P_\ell^d + \text{ span } \left\{ (x^T y)^j (x^T x)^{\ell-1-j} \mid m \le j \le \ell - 2,\ y \in I\!R^d \right\}. \tag{3.2}$$

The terms $(x^T x)^i (x^T y)^j$ for $2i + j \le \ell - 1$ lie directly in $I\!P_\ell^d$. Those with $i + j = \ell - 1$ that are not in $I\!P_\ell^d$ will have $j \le \ell - 2$, and the restriction $m + j + 2i \le 2\ell - 2$ implies $j \ge m$. Thus these terms clearly lie in $\widetilde{Q}$. For the remaining cases we can define $r := \ell - 1 - i - j > 0$ and split the terms as

$$(x^T x)^i (x^T y)^j = (x^T x)^{i-r} (x^T x)^r (x^T y)^j.$$

We now use that each factor $(x^T x)^r (x^T y)^j$, being a homogeneous polynomial in $I\!P_{2r+1+j}^d$, can be represented via terms of type $(x^T y)^{2r+j}$. This leads to terms of the form $(x^T x)^{i-r} (x^T y)^{2r+j}$, and these are in $\widetilde{Q}$, as an easy check of the exponents reveals.

We now have to bound the dimension of $\widetilde{Q}$ from above. If H_j^d denotes the space of homogeneous polynomials on $I\!R^d$ of order j, then the space $I\!P_\ell^d$ is representable as

$$\begin{aligned}
I\!P_\ell^d &= \text{ span } \left\{ (x^T y)^j \mid 0 \le j \le \ell - 1,\ y \in I\!R^d \right\} \\
&= \text{ span } \left\{ H_j^d \mid 1 \le j \le \ell \right\},
\end{aligned} \tag{3.3}$$

and we shall consider the excess of $\widetilde{Q}$ over $I\!P_\ell^d$. The representation (3.2) can be rearranged as

$$\widetilde{Q} = I\!P_\ell^d + \text{ span } \left\{ H_{j+1}^d (x^T x)^{\ell-1-j} \mid m \le j \le \ell - 2 \right\},$$

and we can repeat the argument used for (3.3) to conclude that the dimension of the second part is bounded from above by dim $I\!P_{\ell-1}^d -$ dim $I\!P_m^d$. This proves (3.1), and the second assertion of the lemma is obvious because $\widetilde{Q}$ contains $I\!P_m^d$ for $m \le \ell$. □

If we define

$$\ell^{**} := \ell^{**}(n,m,d) := \max\{\ell \geq 0 \mid \mu(\ell,d) + \mu(\ell-1,d) - \mu(m,d) < n\}, \tag{3.4}$$

we have

$$m \leq \ell^{**} \leq \ell^{*}. \tag{3.5}$$

Note that $\ell^{**}(n,m,d)$ is monotonic with respect to n. For $\ell \to \infty$ and d fixed, clearly

$$\mu(\ell,d) = \frac{\ell^d}{d!} + \mathcal{O}(\ell^{d-1}),$$

which implies the asymptotic behaviour

$$\ell^{**}(n,m,d) = \left(\frac{d!}{2}n\right)^{1/d} + \mathcal{O}(1) \tag{3.6}$$

for $n = |X| \to \infty$ and fixed values of m and d. This will be sufficient for later use, because (3.5) allows us to work with $\ell^{**}(n,m,d)$ and (3.6) instead of $\ell^{*}(X,m,d)$ for large n. Since we are not interested in cases with very small values of n, we pay no further attention to the unimportant restrictions $n \geq \max(2, \mu(m,d))$and $\ell \geq m+1$.

In the one-dimensional case we easily get

$$\begin{aligned} \mu^*(X,\ell,m,1) &= 2\ell - 1 \text{ for } 0 \leq m \leq 2\ell - 1 \\ \ell^*(X,m,1) &= \lfloor \tfrac{|X|+m}{2} \rfloor, \end{aligned}$$

while for $d = 2$ we find

$$\begin{aligned} \mu^*(X,\ell,m,2) &\leq \ell^2 \text{ for } 0 \leq m \leq \ell \\ \ell^{**}(|X|,m,2) &= \lfloor \sqrt{|X| - 1 + \tfrac{m(m+1)}{2}} \rfloor. \end{aligned}$$

4 Approximation Orders

To get bounds for $E(\ell, K, \Phi)$ we apply some classical results of approximation theory. We start with

Theorem 4.1 *Let $f \in C[a,b]$ be a real-valued function with a holomorphic extension to an ellipse in the complex plane with foci a and b, and let f have a singularity at the boundary of the ellipse. If $\rho = R(b-a)/2$ is the sum of the two half-axes of the ellipse, then*

$$\overline{\lim}_{\ell\to\infty} \sqrt[\ell-1]{\inf_{p \in \boldsymbol{P}_\ell^1} \|f - p\|_{\infty,[a,b]}} = \frac{1}{R}.$$

If, in addition, the absolute value of the real part $\Re(f)$ of f is bounded by 1 on the ellipse, and if we set $a = -1,\ b = +1$ for simplicity, we have

$$\inf_{p \in \boldsymbol{P}_\ell^1} \|f - p\|_{\infty,[-1,+1]} \leq \frac{8}{\pi \rho^\ell}. \tag{4.1}$$

Proof: The first part is due to Bernstein and can be found in [13], p. 194. The refined statement is proven in [16], p. 203. □

To cope with functions like $\Phi(r) = r^{2\beta}$ we apply

Theorem 4.2 *For any function $f \in C^k[a,b]$ with a k-th derivative in the class* $\mathrm{Lip}_M \alpha$ *we have*

$$\inf_{p \in \boldsymbol{P}_\ell^1} \|f - p\|_{\infty,[a,b]} \le \frac{c_k (b-a)^{k+\alpha}}{\ell^{k+\alpha}} M \tag{4.2}$$

where

$$c_k = 12 \frac{6^k k^k}{k!} \left(\frac{k+1}{2} \right)^\alpha .$$

If the k-th derivative is continuous, but not contained in any Lipschitz class, then still

$$\inf_{p \in \boldsymbol{P}_\ell^1} \|f - p\|_{\infty,[a,b]} \le \frac{c}{\ell^k} \tag{4.3}$$

with a constant c that depends on f, k, and $[a,b]$, but not on ℓ.

Proof: This is due to Jackson, see e.g.: [13], p.128.

5 Application to Multiquadrics

We first treat the case of functions $\Phi_c(r) = f(c^2 + r^2)$ on $[-1,1]$ for $c > 0$, where f is analytic in $\mathbb{C}$ except for a singularity at the origin. The regularity ellipse of Φ_c then has half-axes of length $\sqrt{1+c^2}$ and c, yielding $\rho = c + \sqrt{1+c^2}$. If $|\Re \Phi_c|$ is bounded by C_c on the ellipse, then (4.1) implies

$$E(\ell, 1, f(c^2 + r^2)) \le \frac{8}{\pi \rho^{2\ell - 1}} C_c,$$

if we approximate $\Phi_c(r) = f(c^2 + r^2)$ by $p \in I\!P^1_{2\ell-1}$ and write the result as a polynomial in r^2 of order at most ℓ. For functions $\Phi_c(r) = (c^2 + r^2)^\beta$ on $[-1,1]$ with $\beta > 0$ we find

$$E(\ell, 1, (c^2 + r^2)^\beta) \le \frac{8(1 + 2c^2)^\beta}{\pi \rho^{2\ell - 1}}$$

and the general case of $\Phi_c(r) = (c^2 + r^2)^\beta$ on $[-K, +K]$ is easily recovered as

$$E(\ell, K, (c^2 + r^2)^\beta) \le K^{2\beta} \frac{8(1 + 2\gamma^2)^\beta}{\pi \rho^{2\ell - 1}}$$

with the scaled quantities

$$\gamma = c/K, \ \rho = \gamma + \sqrt{1 + \gamma^2},$$

indicating that the relative size γ of c and K is crucial.

For $\beta < 0$ the real part of $\Phi_c(r) = (c^2 + r^2)^\beta$ is unbounded on the regularity ellipse. Thus we have to use the weaker form of Theorem 4.1 to get

$$E(\ell, K, (c^2 + r^2)^\beta) \le \frac{C}{\rho^{2\ell - 1}}$$

on $[-K,+K]$ with a constant C that will depend on c, K, γ, and β, but not on ℓ. Similar estimates hold for the radial basis function $\Phi_c(r) = \log(c^2 + r^2)$.

The value ρ^j can be bounded using

$$\rho = \gamma + \sqrt{1+\gamma^2} \geq 1 + \gamma \geq (1-\gamma/2)^{-1}$$

for $0 \leq \gamma \leq 1$ to get

$$\rho^j \geq \exp(j\gamma/2).$$

Inserting everything into (2.12) we get the lower bound

$$\|A^{-1}\|_{2,2} \geq \frac{\pi \exp((2\ell^{**}-1)\gamma/2)}{8nK^{2\beta}} \left\{ \begin{array}{lll} (1+2\gamma^2)^{-\beta} & \text{for} & \beta > 0 \\ C(c,K,\gamma,\beta) & \text{for} & \beta < 0 \end{array} \right\} \tag{5.1}$$

with exponential growth for fixed $\gamma \in [0,1]$, where ℓ^{**} is given by (3.4). Note that (3.6) implies

$$\|A^{-1}\|_{2,2} \geq \mathcal{O}\left(\frac{1}{n}\exp\left(\frac{\gamma}{2}\left(2\left(\frac{d!}{2}n\right)^{1/d} - 1\right)\right)\right) \tag{5.2}$$

for $n \to \infty$ and fixed γ, d. The theory for infinite grids (see Buhmann [3]) varies c proportional to the minimal distances of centres. For a fixed finite domain in $\mathbb{R}^d$ with centres forming a dense subset, this strategy is resembled by letting γ vary like $n^{-1/d}$. Then the behaviour for $n \to \infty$ is

$$\|A^{-1}\|_{2,2} \geq \mathcal{O}\left(\frac{1}{n}\right)$$

for $\beta > 0$, because the exponential function in (5.2) gets a constant argument due to cancellation.

6 Application to other radial basis functions

For $\Phi(r) = r^{2\beta}$ with $2\beta = 2p+1$, $p \in \mathbb{N}_{\geq 0}$ we again approximate $\Phi(|r|)$ on $[-K,K]$ by polynomials in $\mathbb{P}^1_{2\ell-1}$ and take advantage of the symmetry of the best approximation. Since the derivatives up to order $2p = 2\beta - 1$ are continuous and the $k := 2p$-th derivative is in a Lipschitz class $\text{Lip}_C\alpha$ with $\alpha = 1$ and $C := (2\beta)!$ we can invoke Theorem 4.2 to get

$$E(\ell, K, \Phi) \leq \frac{c_{2p}(2K)^{2\beta}}{(2\ell-1)^{2\beta}}(2\beta)!$$

and the lower bound

$$\|A^{-1}\|_{2,2} \geq \frac{(2\ell^{**}-1)^{2\beta}}{nc_{2p}(2K)^{2\beta}(2\beta)!}$$

with the asymptotic behaviour

$$\|A^{-1}\|_{2,2} \geq \mathcal{O}\left(n^{\frac{2\beta}{d}-1}\right)$$

for $n \to \infty$.

A similar analysis holds for thin-plate splines of the form $\Phi(r) = r^{2\beta}\log r$ with $\beta \in \mathbb{N}_{\geq 1}$. Derivatives up to order $2\beta - 1$ are continuous, and for $\ell^{**} \geq \beta$ there is no problem with the polynomials that arise when taking derivatives. However, the highest continuous derivative is not contained in any Lipschitz class. Application of the restricted statement of Theorem 4.2 then implies

$$\|A^{-1}\|_{2,2} \geq C(\beta, K)\frac{(2\ell^{**}-1)^{2\beta-1}}{n} = \mathcal{O}\left(n^{\frac{2\beta-1}{d}-1}\right)$$

for $n \to \infty$ where $C(\beta, K)$ is a constant.

For Gaussians $\Phi(r) = \exp(-\alpha^2 r^2)$ we can use the Taylor expansion to get

$$E(\ell, K, \Phi) \leq \frac{(\alpha K)^\ell}{\ell!},$$

providing a disastrous lower bound on $\|A^{-1}\|_{2,2}$ for fixed values of αK if n is large.

Remark. A variation of our approach would be to approximate $\Phi(\sqrt{r})$ on $[h^2, K^2]$ for the separation distance (1.6) as used by Ball, Narcowich, and Ward. This would make any of the classical radial basis functions analytic in a neighbourhood of $[h^2, K^2]$, leading to exponential decrease of the approximation error $E(\ell, [h^2, K^2], \Phi)$ with respect to polynomials of order ℓ for $\ell \to \infty$. The basic estimate has the same form as (5.2), but a detailed analysis reveals that the exponential term of this bound gets a constant argument because of $K \geq n^{1/d}h$ and $\ell \approx n^{1/d}$. This corresponds to the cancellation in (5.2) when γ or c are scaled to decrease with the minimal separation distance.

For illustration, consider the function $(r^2)^\beta = t^\beta$ for $t = r^2 \in [h^2, K^2]$. After rescaling, it coincides with the multiquadric

$$t^\beta = \left(h^2 + s\frac{K^2-h^2}{K^2}\right)^\beta = \left(\frac{K^2-h^2}{K^2}\right)^\beta \left(\frac{h^2K^2}{K^2-h^2} + s\right)^\beta$$

with exponent β and $c = hK/\sqrt{K^2-h^2}$ for $s \in [0, K^2]$. Thus the behaviour of $\Phi(r) = r^{2\beta}$, as far as our lower bounds are concerned, is roughly the same as for multiquadrics with exponent β and a scaling of $c \approx h \approx \gamma \approx n^{-1/d}$, where the exponential in the bound (5.2) gets a constant argument due to cancellation. For multiquadrics themselves, the effect of introducing h just acts like a corresponding increase of c and does not yield any improvement.

7 Comparison with other bounds

In general, the upper bounds by Ball [1] and Narcowich, and Ward [10][11][12] provide estimates of $\|A^{-1}\|_{2,2}$ from above in terms of the separation distance h of (1.6). The corresponding lower bounds of Ball, Sivakumar, and Ward [2] hold for a specific regularly distributed set of centres and thus act as strict lower bounds for the worst possible data set with prescribed separation distance h.

In contrast to this, our lower bounds provide best–case estimates of $\|A^{-1}\|_{2,2}$ from below, because they hold for every distribution of the data points, including the best possible choice, if the latter should exist. These bounds must necessarily be smaller than the lower bounds of Ball, Sivakumar, and Ward, the difference being leeway for optimizing the placement of centres.

We start with a comparison of results for multiquadrics on increasing sets of centres with separation distance 1 in $I\!R^d$. In this case, the optimal bounds of Baxter [4] for $\|A^{-1}\|_{2,2}$ must lie between our lower bounds and the upper bounds of Ball, Narcowich, and Ward. Whenever Baxter's bounds coincide with the latter, the conclusion is that regular data asymptotically realize the worst possible distribution with separation distance 1. The difference between Baxter's bounds and ours may possibly be used for optimization of placements of centres, because we do not make assumptions on the separation distance or regularity of distribution.

The optimal bound of Baxter for $\|A^{-1}\|_{2,2}$ in the case of centres on integer grids takes the form

$$\|A^{-1}\|_{2,2} \leq \frac{\pi}{4c} \frac{1}{e^{-c\pi} + \frac{1}{3}e^{-3c\pi} + \frac{1}{5}e^{-5c\pi} \dots} \approx \frac{\pi}{4c} e^{c\pi}$$

for $d = 1$ and for multiquadrics $\Phi_c(r) = (c^2 + r^2)^{1/2}$. The same value arises as the precise limit of $\|A^{-1}\|_{2,2}$ for $n \to \infty$, when the n integer points $i \in I\!R$ with $0 \leq i \leq n-1$ are taken.

Ball, Sivakumar, and Ward [2] get

$$\|A^{-1}\|_{2,2} \quad \geq \quad C\ e^{c\sqrt{d}}$$

for $n \to \infty$ and the same data distribution, where the constant C does not depend on c and d. The worst–case upper bound of Narcowich and Ward [10] is constant, too, for $n \to \infty$ and data with separation distance 1, while the dependence on c is

$$\|A^{-1}\|_{2,2} \quad \leq \quad C e^{4cd}.$$

To compare with our results, we consider **arbitrary** distributions of centres in $[0, n]$ and get

$$\|A^{-1}\|_{2,2} \quad \geq \quad \frac{\pi}{8n^2} \exp\left(\frac{n-1}{n}\frac{c}{2}\right)\left(1 + 2\frac{c^2}{n^2}\right)^{-1/2} \tag{7.1}$$

from (5.1) with exponential behaviour for $c \to \infty$ with n fixed. Thus no other choice of centres can get rid of this exponential increase of $\|A^{-1}\|_{2,2}$ with c.

The variation of (7.1) with n for fixed c is off from Baxter's optimal bound for equidistant data by only a factor of n^{-2}. Note that a factor n^{-1} may be due to our special combination of matrix norms, and that there must be at least some leeway to optimize placements of centres, which is clearly bounded by gaining a factor of n^{-2}.

We now compare our bounds with those of Ball [1] and Narcowich and Ward [10] [11] [12] and Ball, Sivakumar, and Ward [2] for irregular centres. Since these results are in terms of the separation distance which does not enter explicitly into our results, we have to make sure that the scaling is fair. Thus we can either keep the separation distance fixed and let the centres spread out into all of $I\!R^d$ when their number n tends to infinity, or we can consider large data sets of centres contained in the unit cube $[0, 1]^d$ of $I\!R^d$, letting the separation distance tend to zero when the number of centres tends to infinity. We choose the latter possibility because it is somewhat more related to possible applications. The diameter K of the sets of centres we consider will thus always be bounded by $\sqrt{d}$, and the Euclidean separation distance h for n centres in $[0, 1]^d$ will be at most

$$h \leq \sqrt{d}\left(n^{1/d} - 1\right)^{-1} \approx \sqrt{d}\, n^{-1/d}, \tag{7.2}$$

as can be easily shown by summing the volumes of sufficiently small disjoint cubes around each centre. In the following we shall simply use $h = \sqrt{d}\, n^{-1/d}$ because we are mainly interested in the case $n \to \infty$.

For multiquadrics $\Phi_c(r) = (c^2 + r^2)^{1/2}$ the results of Narcowich and Ward [10], as refined in [2] by Ball, Sivakumar, and Ward, yield

$$\|A^{-1}\|_{2,2} \le C(d)\frac{1}{h}\exp\left(2d\frac{c}{h}\right)$$

with a constant $C(d)$ not depending on c, h, and n. Due to (7.2) this estimate is always worse than

$$\|A^{-1}\|_{2,2} \le C(d)\; n^{1/d}\exp\left(2cn^{1/d}d\right)$$

if arbitrary placements of n centres in $[0,1]^d$ are allowed. Regular distribution of centres on a scaled integer lattice will yield $h = n^{-1/d}$ and

$$C(d)\; n^{1/d}\exp\left(\frac{1}{2}cn^{1/d}\sqrt{d}\right) \le \|A^{-1}\|_{2,2}$$

according to Ball, Sivakumar, and Ward [2], which has the same asymptotic behaviour for $n \to \infty$ as the best possible upper bound. Thus, as far as upper bounds in terms of the separation distance are concerned, the approach of Narcowich and Ward gives a best possible result for n regular data with $n \to \infty$.

Our approach allows arbitrary centres in $[0,1]^d$ and proves

$$C_1(d)\frac{1}{n\sqrt{d+2c^2}}\exp\left(c\frac{2(d!n/2)^{1/d} + C_2(d)}{2\sqrt{d}}\right) \le \|A^{-1}\|_{2,2}$$

with suitable constants depending on d only. In the limit $n \to \infty$, the exponential increase cannot be overcome by optimized placement of centres; there is only a factor of at most $\mathcal{O}(n^{-1-1/d})$ to be gained.

The dependence on c is exponential in all cases, and the exponential behaviour is eliminated, if $c \approx n^{-1/d}$ is chosen. Then, up to constants C_1, C_2 not depending on n, we have

$$C_1 n^{-1} \le \|A^{-1}\|_{2,2} \le C_2\, n^{1/d}$$

as the possible variation of $\|A^{-1}\|_{2,2}$ with the centres, optimizing from a regular distribution. Note that the factor n^{-1} occurring above was introduced by solving the matrix approximation problem of section 2 in the "wrong" norm.

For inverse multiquadrics $\Phi_c(r) = (c^2 + r^2)^{-1/2}$ the lower bound of Ball, Sivakumar, and Ward [2] for regular centres is

$$C(d)\; c\; \exp\left(\frac{1}{2}cn^{1/d}\sqrt{d}\right) \le \|A^{-1}\|_{2,2},$$

while we get

$$C(d)\frac{1}{nc}\exp\left(\frac{c}{2\sqrt{d}}\left(2\left(\frac{d!}{2}n\right)^{1/d} - 1\right)\right) \le \|A^{-1}\|_{2,2}$$

for arbitrary centres, which is smaller by only a factor $\mathcal{O}(n^{-1})$ for $n \to \infty$.

We now consider thin–plate splines with $\Phi(r) = r^2 \log r$ with $d = 2,\ m = 2$. Here, the best possible form of the upper bound by Narcowich and Ward [12] is

$$\|A^{-1}\|_{2,2} \leq C(d)n^3,$$

while we get

$$C(d)n^{-1/2} \leq \|A^{-1}\|_{2,2}.$$

Numerical experiments indicate that regular data have $\|A^{-1}\|_{2,2} = \mathcal{O}(n^{1/2})$ for $n \to \infty$. Thus we conjecture that our bound is off by a factor of at most $\mathcal{O}(n^{-1})$ from the actual behaviour of $\|A^{-1}\|_{2,2}$ for regular data.

For $\Phi(r) = r^{2\beta}$ with $\beta \in (0,1)$, the best possible upper bound by Narcowich and Ward [11] is

$$\|A^{-1}\|_{2,2} \leq C_1(d,\beta)n^{2\beta/d},$$

and regularly distributed centres yield

$$C_2(d,\beta)n^{2\beta/d} \leq \|A^{-1}\|_{2,2},$$

as was shown by Ball, Sivakumar, and Ward [2]. Our general bound is

$$C_3(d,\beta)n^{2\beta/d-1} \leq \|A^{-1}\|_{2,2}$$

for arbitrary centres and $\beta = 1/2$, containing a $\mathcal{O}(n^{-1})$ factor again.

8 Conclusions

The results of this paper support the hypothesis that regular placement of centres is a good strategy as far as minimization of the condition number of the matrix A is concerned. The theoretically possible gain by optimization of placement of centres is not more than a factor of $\mathcal{O}(n^{-2})$ or $\mathcal{O}(n^{-1})$ for $n \to \infty$ in all cases, and the proof technique indicates that the factor $\mathcal{O}(n^{-1})$ arises for technical reasons only. Possibly the bounds on the Euclidean norm of A^{-1} in the literature can be generalized to hold also for the norm $\|A^{-1}\|_{1,\infty}$, and then the factor $\mathcal{O}(n^{-1})$ can be eliminated.

Acknowledgement. The author thanks the two referees for a number of useful suggestions that helped to improve the presentation.

References

[1] Ball, K. M., Invertibility of Euclidean distance matrices and radial basis interpolation, CAT Report 201, Texas A& M University, College Station 1989

[2] Ball, K. M., Sivakumar, N., and J. D. Ward, On the sensitivity of radial basis interpolation with respect to minimal separation distance, CAT Report 240, Texas A& M University, College Station 1989

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

[4] Baxter, B.J.C., Norm estimates for inverses of Toeplitz distance matrices, preprint, Dept. of Applied Math. and Theor. Physics, Univ. of Cambridge, 1991

[5] 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

[6] Dyn. N, Levin, D., and S. Rippa, Numerical procedures for surface fitting of scattered data by radial functions, SIAM J. Sci. Stat. Comput. **7** 1986, 639–659

[7] R. Franke (1982). Scattered data interpolation: tests of some methods. Math. Comp., Vol. 38, pp. 181–200

[8] R.L. Hardy (1971). Multiquadric equations of topography and other irregular surfaces. J. Geophysical Res., Vol. 76, pp. 1905–1915

[9] Micchelli, C.A., Interpolation of scattered data: distance matrices and conditionally positive definite functions, Constructive Approximation **2**, 11-22, 1986

[10] Narcowich, F.J., and J.D. Ward, Norms of Inverses and Condition Numbers for Matrices Associated with Scattered Data, J. Approx. Th. 64, 69-94, 1991

[11] Narcowich, F.J., and J.D. Ward, Norm Estimates for the Inverses of a General Class of Scattered Data Interpolation Matrices, Preprint, Texas A & M University, 1991

[12] Narcowich, F.J., and J.D. Ward, Norm Estimates for Inverses of Scattered Data Matrices Associated with Completely Monotonic Radial Functions, Preprint, Texas A & M University, 1991

[13] Natanson, I.P., Constructive Function Theory, Vol. I, Frederick Ungar Publishing Co. New York, 1964

[14] 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

[15] Schoenberg, I. J., Metric spaces and completely monotone functions, Ann. of Math. **39** (1938) 811-841

[16] Schönhage, A., Approximationstheorie, Berlin-New York : Gruyter & Co., 1971

Author's address:

Robert Schaback, Institut für Numerische
und Angewandte Mathematik
Universität Göttingen
Lotzestr. 16–18
D – 3400 Göttingen