Stability of Radial Basis Function Interpolants

Robert Schaback

Abstract. The stability of the linear systems arising from scattered data interpolation problems with radial basis functions is analysed in full generality. Since lower bounds for the smallest eigenvalue of the coefficient matrix yield upper bounds for the absolute error of the RBF coefficients in terms of the absolute errors in the data, we then focus on a new and short proof of such bounds.

§1. Introduction

We shall study the stability of multivariate interpolation by conditionally positive definite radial functions of order $m \geq 0$.

Definition 1. *A univariate function*

$$\phi \;:\; \mathbb{R}_{\geq 0} \to \mathbb{R}$$

is called conditionally positive definite of order m *on* $\mathbb{R}^d$, *if for all possible choices of sets*

$$X = \{x_1, \ldots, x_N\} \subset \mathbb{R}^d$$

of N *distinct points, the quadratic form induced by the* $N \times N$ *matrix*

$$A = (\phi(\|x_j - x_k\|_2))_{1 \leq j,k \leq N} \tag{1}$$

is positive definite on the subspace

$$V := \left\{ \alpha \in \mathbb{R}^N \;:\; \sum_{j=1}^{N} \alpha_j p(x_j) = 0 \text{ for all } p \in \mathbb{P}_m^d \right\}$$

where $\mathbb{P}_m^d$ *stands for the space of* d*-variate polynomials of order not exceeding* m.

Approximation Theory X
Charles K. Chui, Larry L. Schumaker, and Joachim Stoeckler (eds.), pp. 1–8.

ISBN 0-8265-xxxx-x.

Note that $m = 0$ implies $V = \mathbb{R}^N$ because of $\mathbb{P}_m^d = \{0\}$, and then the matrix A in (1) is positive definite.

The most prominent examples of conditional positive definite radial basis functions of order m on $\mathbb{R}^d$ are

$$\begin{array}{lcllll}
\phi(r) & = & (-1)^{\lceil\beta/2\rceil} r^\beta & \beta > 0, \beta \notin 2\mathbb{N}_0 & m \geq \lceil\beta/2\rceil \\
\phi(r) & = & (-1)^{k+1} r^{2k}\log(r) & k \in \mathbb{N} & m \geq k+1 \\
\phi(r) & = & (c^2+r^2)^{\beta/2} & \beta < 0 & m \geq 0 \\
\phi(r) & = & (-1)^{\lceil\beta/2\rceil}(c^2+r^2)^{\beta/2} & \beta > 0, \beta \notin 2\mathbb{N}_0 & m \geq \lceil\beta/2\rceil \\
\phi(r) & = & e^{-\alpha r^2} & \alpha > 0 & m \geq 0 \\
\phi(r) & = & (1-r)_+^4(1+4r) & d \leq 3 & m \geq 0
\end{array}$$

See e.g. [10] for a comprehensive derivation of the properties of these functions.

Interpolation of real values $f_1, \ldots, f_N$ on a set $X = \{x_1, \ldots, x_N\}$ of N distinct scattered points of $\mathbb{R}^d$ by a radial basis function ϕ is done by solving the $(N+Q) \times (N+Q)$ system

$$\begin{array}{ccccc}
A\alpha & + & P\beta & = & f \\
P^T\alpha & + & 0 & = & 0
\end{array} \tag{2}$$

where $Q = \dim \mathbb{P}_m^d$ and

$$P = (p_i(x_j))_{1\leq j\leq N, 1\leq i\leq Q}$$

for a basis $p_1, \ldots, p_Q$ of $\mathbb{P}_m^d$. In fact, if the additional assumption

$$\text{rank } (P) = Q \leq N \tag{3}$$

holds, then the system (2) is uniquely solvable. The resulting interpolant has the form

$$s(x) = \sum_{j=1}^{N} \alpha_j \phi(\|x_j - x\|_2) + \sum_{i=1}^{Q} \beta_i p_i(x) \tag{4}$$

with the additional condition $\alpha \in V$.

§2. Stability

To investigate the numerical stability of the system (2), we replace α, β, f by perturbations of the original quantities and get

$$\begin{array}{ccccc}
(\Delta\alpha)^T A(\Delta\alpha) & + & 0 & = & (\Delta\alpha)^T\Delta f \\
P^T(\Delta\alpha) & + & 0 & = & 0
\end{array} \tag{5}$$

Since A is positive definite on the subspace $V = \ker P^T$, there are positive eigenvalues $\Lambda \geq \lambda$ such that

$$\Lambda\|\alpha\|_2^2 \geq \alpha^T A\alpha \geq \lambda\|\alpha\|_2^2 \text{ for all } \alpha \in V = \ker P^T. \tag{6}$$

We can insert this into (5) to get

$$\|\Delta\alpha\|_2 \le \frac{1}{\lambda}\|\Delta f\|_2$$

to bound the absolute error $\Delta\alpha$ of α by the absolute error Δf in the data vector f. If the solution α is nonzero, we have that $f - P\beta$ is nonzero, and a similar argument combining (2) and (6) yields

$$\frac{\|\Delta\alpha\|_2}{\|\alpha\|_2} \le \frac{\Lambda}{\lambda}\frac{\|\Delta f\|_2}{\|f - P\beta\|_2}$$

for the relative error. Thus the standard L_2 theory of numerical stability applies to the RBF part of (2). The condition is given by the ratio Λ/λ, while stability of the absolute error is dominated by λ.

The stability of the calculation of β follows the lines of the stability theory for discrete L_2 polynomial approximation, because we have

$$\beta = (P^T P)^{-1} P^T (f - A\alpha). \tag{7}$$

This means that the polynomial part can be calculated from the residuals of the RBF data via the standard operator $(P^T P)^{-1}P^T$ of discrete L_2 polynomial approximation. The absolute error of the residuals can be bounded by

$$\|\Delta f - A\Delta\alpha\|_2 \le \|\Delta f\|_2(1 + \frac{\Lambda}{\lambda}),$$

and we can see how the condition of the RBF part enters into the stability theory for the polynomial part: the upper bounds have to be multiplied by $1 + \frac{\Lambda}{\lambda}$.

To derive stability bounds for practical use, one needs upper bounds for Λ and lower bounds for λ.

§3. Upper Bounds for Eigenvalues

For bounded radial basis functions one can get crude upper bounds for Λ via Gerschgorin's theorem. In fact, if we normalize ϕ to satisfy

$$1 = \phi(0) \ge \phi(r) \text{ for all } r \in [0, \infty),$$

then

$$|1 - \Lambda| \le N - 1,$$

which is not too bad for standard applications and compared to the bad behavior of λ to become apparent later. In particular, this bound is independent of the data locations and the smoothness of ϕ, which have a strong influence on lower bounds for λ.

A somewhat more general argument works for cases in which we have a convolution representation

$$\phi(\|x-y\|_2) = \int_{\mathbb{R}^d} \Psi(x-t)\Psi(y-t)dt. \tag{8}$$

This is actually true for all smooth unconditionally positive definite and Fourier–transformable functions, because the Fourier transform of $\Psi(x)$ can be obtained via the square root of the (nonnegative) Fourier transform of $\phi(\|\cdot\|_2)$. If we assume (8), then

$$\begin{aligned}
\alpha^T A\alpha &= \int_{\mathbb{R}^d} \left(\sum_{j=1}^N \alpha_j \Psi(x_j-t)\right)^2 dt \\
&\leq \left(\sum_{j=1}^N \alpha_j^2\right) \int_{\mathbb{R}^d} \sum_{j=1}^N \Psi(x_j-t)^2 dt \\
&= N\phi(0)\sum_{j=1}^N \alpha_j^2.
\end{aligned}$$

For general positive definite functions $\Phi(x,y)$ we know that

$$\Phi(x,y) = (\Phi(x,\cdot),\Phi(y,\cdot))_{\mathcal{H}}$$

holds, where $\mathcal{H}$ is the native Hilbert space for Φ (see e.g. [9] for details). Then the above argument takes the form

$$\begin{aligned}
\alpha^T A\alpha &= \left\|\sum_{j=1}^N \alpha_j \Phi(x_j,\cdot)\right\|_{\mathcal{H}}^2 \\
&\leq \left(\sum_{j=1}^N \alpha_j^2\right)\left(\sum_{j=1}^N \Phi(x_j,x_j)\right).
\end{aligned} \tag{9}$$

The case of unbounded radial basis functions is somewhat more complicated. It comprises the radial basis functions with positive minimal order of conditional positive definiteness, for instance the multiquadrics $\phi(r) = \sqrt{r^2+c^2}$ or thin–plate splines $\phi(r) = r^2 \log r$. In contrast to the above upper bound, which did not use the additional condition $\alpha \in V = \ker P^T$, we now have to rely on the latter. If we insert (7) into the first equation in (2), we get

$$(A - P(P^TP)^{-1}P^TA)\alpha = f - P(P^TP)^{-1}P^Tf.$$

Due to $P(P^TP)^{-1}P^T\alpha = 0$, this system can be written as

$$RAR^T\alpha = Rf,$$

where $R := I - P(P^TP)^{-1}P^T = R^T$ is the operator that maps discrete data into their residuals after least–squares approximation by polynomials from $\mathbb{P}_m^d$. The matrix $B := RAR^T$ will now have a significantly better behavior than A, as far as upper bounds are concerned, because it can be interpreted as an interpolation matrix of an unconditionally positive definite non–radial function (see §6 of [9]). In particular, we can apply (9) to this new function, and we get a numerically computable upper bound on the largest eigenvalue. Further details are suppressed here. We summarize:

Theorem 2. *For any positive definite radial basis function ϕ, the largest eigenvalue Λ of the matrix A as defined in (1) is bounded above by $N\phi(0)$. For non–radial positive definite basis functions the bound takes the form*

$$\sum_{j=1}^{N} \Phi(x_j, x_j),$$

and the case of positive order of conditional positive definiteness can be reduced to the positive definite non–radial case by matrix transformations.

§4. Lower Bounds for Eigenvalues

The work of Ball [1][2], Narcowich, Sivakumar, and Ward [5][6] already contains lower bounds for the smallest eigenvalue λ of A, and these bounds are near–optimal due to [7]. However, the proofs are complicated, and we want to provide a much shorter though less general argument, which can be transferred to expansion kernels [4]. It relies on the existence of positive definite functions with compact support, which were not available before 1995 due to Wu [13] and Wendland [12].

The idea is to perturb the matrix A on the diagonal by subtracting from a conditionally positive definite radial function ϕ of order m some positive definite radial function ψ with small support, such that $\phi - \psi$ still is conditionally positive (semi)definite of order m. If we write A_ϕ when A in (1) is based on ϕ, we then get

$$\alpha^T A_\phi \alpha = \alpha^T A_{\phi-\psi}\alpha + \alpha^T A_\psi \alpha \geq \alpha^T A_\psi \alpha = \psi(0)\|\alpha\|_2^2 \text{ for all } \alpha \in V$$

if the support of ϕ is smaller than the minimal distance

$$q := \min_{1\leq i<j\leq N} \|x_i - x_j\|_2 \tag{10}$$

between two different data points, and then $A_\psi = \psi(0)I$.

Theorem 3. *Let ϕ be a conditionally positive definite function of order m. If ψ is a positive definite radial basis function with support in $[0, q]$ with q from (10) such that $\phi - \psi$ is conditionally positive definite of order at least m, then $\psi(0)$ is a lower bound for the smallest eigenvalue of A_ϕ as defined in (1).*

Thus we get $\psi(0)$ as a lower bound for λ, whenever we can find a ψ with support in $[0, q]$ such that ψ and $\phi-\psi$ are positive definite. Of course, we would like to take the maximal $\psi(0)$ under these conditions, but the corresponding optimization problem still is an open challenge.

To be more specific, we confine ourselves to conditionally positive definite radial basis functions with a radial generalized Fourier transform $\widehat{\phi}$ satisfying

$$\widehat{\phi}(r) \geq c_\infty r^{-d-\beta} \text{ for all } r > 1. \tag{11}$$

Note that this places an upper bound on the smoothness of ϕ, and thus it rules out infinitely differentiable cases like the Gaussian and the multiquadrics. Furthermore, it implies by arguments from [11] that the standard $L_\infty(\Omega)$ error bounds for interpolation of functions in the native space $\mathcal{H}$ cannot be better than of order $\mathcal{O}(h^{\beta/2})$, where h is the data density

$$h := \sup_{y \in \Omega} \min_{1 \leq j \leq N} \|x_j - y\|_2.$$

By the Uncertainty Principle in [8], the optimal lower bounds of eigenvalues λ have the form $\mathcal{O}(q^\beta)$ for small q, and this is what we want to recover by our new technique.

Theorem 4. *Let ϕ be a conditionally positive definite radial basis function whose Fourier transform has at most the decay (11). Then the smallest eigenvalues of the matrices A in (1) have a lower bound of the form*

$$\lambda \geq c q^\beta$$

for all data sets with $q \leq 1$.

Proof: For convenience of notation, we add

$$0 < c_0 \leq \widehat{\phi}(r) \text{ for all } r \leq 1$$

to (11). From Wendland's supply of arbitrarily smooth compactly supported positive definite radial basis functions we can find some σ with support on $[0, 1]$ satisfying $\sigma(0) = 1$ and having a positive radial Fourier transform $\widehat{\sigma}$ with

$$\begin{array}{rclc} \widehat{\sigma}(r) & \leq & C_\infty r^{-d-\beta} & r > 0 \\ \widehat{\sigma}(r) & \leq & C_0 & r \leq 1. \end{array}$$

We now take $\psi(\cdot) = \epsilon\sigma(\cdot/q)$ to squeeze the support of ψ into $[0, q]$, and we maximize ϵ under the constraint

$$\widehat{\phi}(r) \geq \widehat{\psi}(r) = \epsilon q^d \widehat{\sigma}(rq) \text{ for all } r \geq 0 \tag{12}$$

which still makes $\phi - \psi$ conditionally positive semidefinite of at least the same order as ϕ, because this order is related to the order of the singularity of $\widehat{\phi}$ at zero.

We first treat the case $r > 1$, in which it suffices to guarantee (12) by

$$\widehat{\psi}(r) = \epsilon q^d \widehat{\sigma}(rq) \leq \epsilon q^d C_\infty (rq)^{-d-\beta} = \epsilon q^{-\beta} C_\infty r^{-d-\beta} \leq c_\infty r^{-d-\beta} \leq \widehat{\phi}(r)$$

by picking

$$\epsilon \leq \frac{c_\infty}{C_\infty} q^\beta.$$

The case $r \leq 1$ has $rq \leq 1$ and we can satisfy

$$\widehat{\psi}(r) = \epsilon q^d \widehat{\sigma}(rq) \leq \epsilon q^d C_0 \leq \epsilon C_0 \leq c_0 \leq \widehat{\phi}(r)$$

by taking

$$\epsilon \leq \frac{c_0}{C_0}.$$

□

Note that we could incorporate infinitely differentiable cases like the Gaussian and the multiquadrics, if we had a sufficient supply of infinitely differentiable radial basis functions with small compact supports. The case of expansion kernels suffered also from lack of positive definite functions with arbitrarily small support, but this was overcome in [4]. An application to non–radial basis functions with varying scales and shapes is in [3].

Acknowledgments. Special thanks go to Holger Wendland for proofreading.

References

1. Ball, K, Eigenvalues of Euclidean distance matrices, J. Approx. Theory **68** (1992) 74–82
2. Ball, K., N. Sivakumar, and J.D. Ward, On the sensitivity of radial basis interpolation to minimal data separation distance. Constr. Approx. **8** (1992) 401–426
3. Bozzini, M., L. Lenarduzzi, M. Rossini, and R. Schaback, Interpolation by basis functions of different scales and shapes, preprint 2001
4. Narcowich, F. J., R. Schaback, and J. D. Ward, Stability Estimates for Interpolation by Positive Definite Kernels, preprint, 2001.
5. Narcowich, F. J., N. Sivakumar, and J. D. Ward, On condition numbers associated with radial–function interpolation, J. Math. Anal. Appl. **186** (1994) 457–485
6. Narcowich, F. J., and J. D. Ward, Norm estimates for the inverses of a general class of scattered–data radial–function interpolation matrices, J. Approx. Theory **69** (1992) 84–109
7. Schaback, R., Lower bounds for norms of inverses of interpolation matrices for radial basis functions, J. Approx. Theory **79** (1994) 287–306

8. Schaback, R., Error estimates and condition numbers for radial basis function interpolation, Advances in Comp. Math. **3** (1995) 251–264

9. Schaback, R., Native Hilbert spaces for radial basis functions, in: *New Developments in Approximation Theory*, Buhmann, M.D., Mache, D. H., Felten, M., and Müller, M. W.(eds.), Birkhäuser Verlag, International Series of Numerical Mathematics **132** (1999) 255–282

10. Schaback, R. and H. Wendland, Characterization and construction of radial basis functions, in *Multivariate Approximation and Applications*, Dyn, N., D. Leviatan, D. Levin, and A. Pinkus (eds.), Cambridge University Press 2000

11. Schaback, R. and H. Wendland, Inverse and saturation theorems for radial basis function interpolation, to appear in Math. Comp.

12. Wendland, H., Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree, Advances in Comp. Math. **4** (1995) 389–396

13. Wu, Z., Multivariate compactly supported positive definite radial functions, Advances in Comp. Math. **4** (1995) 283–292

Prof. Dr. R. Schaback
Institut für Numerische und Angewandte Mathematik
Zentrum für Informatik
Lotzestraße 16–18
D-37083 Göttingen
`schaback@math.uni-goettingen.de`