

## Stability of kernel-based interpolation

Stefano De Marchi · Robert Schaback

Received: 10 October 2007 / Accepted: 12 July 2008 /  
Published online: 12 August 2008  
© Springer Science + Business Media, LLC 2008

**Abstract** It is often observed that interpolation based on translates of radial basis functions or non-radial kernels is numerically unstable due to exceedingly large condition of the kernel matrix. But if stability is assessed in function space without considering special bases, this paper proves that kernel-based interpolation is stable. Provided that the data are not too wildly scattered, the  $L_2$  or  $L_\infty$  norms of interpolants can be bounded above by discrete  $\ell_2$  and  $\ell_\infty$  norms of the data. Furthermore, Lagrange basis functions are uniformly bounded and Lebesgue constants grow at most like the square root of the number of data points. However, this analysis applies only to kernels of limited smoothness. Numerical examples support our bounds, but also show that the case of infinitely smooth kernels must lead to worse bounds in future work, while the observed Lebesgue constants for kernels with limited smoothness even seem to be independent of the sample size and the fill distance.

**Keywords** Kernel-based interpolation · Numerical stability ·  
Lebesgue constants

**Mathematics Subject Classifications (2000)** 41A05 · 41A36 · 41A63 · 65D05

---

Communicated by Joe Ward.

S. De Marchi (✉)  
Department of Computer Science, University of Verona, Verona, Italy  
e-mail: stefano.demarchi@univr.it

R. Schaback  
Institut für Numerische und Angewandte Mathematik,  
University of Göttingen, Göttingen, Germany

### 1 Introduction

We consider the recovery of a real-valued function  $f : \Omega \rightarrow \mathbb{R}$  on some compact domain  $\Omega \subseteq \mathbb{R}^d$  from its function values  $f(x_j)$  on a scattered set  $X = \{x_1, \dots, x_N\} \subset \Omega \subseteq \mathbb{R}^d$ . Independent of how the reconstruction is done in detail, we denote the result as  $s_{f,X}$  and assume that it is a linear function of the data, i.e. it takes the form

$$s_{f,X} = \sum_{j=1}^N f(x_j)u_j \tag{1}$$

with certain continuous functions  $u_j : \Omega \rightarrow \mathbb{R}$ . To assert the stability of the recovery process  $f \mapsto s_{f,X}$ , we look for bounds of the form

$$\|s_{f,X}\|_{L_\infty(\Omega)} \leq C(X)\|f\|_{\ell_\infty(X)} \tag{2}$$

which imply that the map taking the data into the interpolant is continuous in the  $L_\infty(\Omega)$  and  $\ell_\infty(X)$  norm. Of course, one can also use  $L_2(\Omega)$  and  $\ell_2(X)$  norms above.

By putting (1) into (2), we see that we can bound the *stability constant*  $C(X)$  below as follows

$$C(X) \geq \max_{x \in \Omega} \sum_{j=1}^N |u_j(x)| =: \Lambda_X \tag{3}$$

where  $\Lambda_X$  is the *Lebesgue constant* which is the maximum of the *Lebesgue function*  $\lambda_X(x) := \sum_{j=1}^N |u_j(x)|$ .

It is a classical problem to derive upper bounds for the stability constant in (2) and for its lower bound, the Lebesgue constant  $\Lambda_X$ . As well-known in recovery by polynomials, in both the univariate and the bivariate case, there exist upper bounds for the Lebesgue function. Moreover, many authors faced the problem of finding near-optimal points for polynomial interpolation. All these near-optimal sets of  $N$  points have a Lebesgue function that behaves in the one dimensional case like  $\log(N)$  while as  $\log^2(N)$  in the two dimensional one (cf. [2] and references therein). An important example, worth mentioning, of points suitable for polynomial interpolation in the square whose Lebesgue constant grows as  $\mathcal{O}(\log^2(N))$  are the so-called *Padua-points* (see [1]).

However, stability bounds for multivariate *kernel-based* recovery processes are missing. We shall derive them as follows. Given a positive definite kernel  $\Phi : \Omega \times \Omega \rightarrow \mathbb{R}$ , the recovery of functions from function values  $f(x_j)$  on the set  $X = \{x_1, \dots, x_N\} \subset \Omega \subseteq \mathbb{R}^d$  of  $N$  different data sites can be done via interpolants of the form

$$s_{f,X} := \sum_{j=1}^N \alpha_j \Phi(\cdot, x_j) \tag{4}$$

taken from the finite-dimensional space  $V_X := \text{span} \{\Phi(\cdot, x) : x \in X\}$  of translates of the kernel, and satisfying the linear system

$$A_{\Phi,X}\alpha = \mathbf{f} \tag{5}$$

where  $A_{\Phi, X} := (\Phi(x_k, x_j))_{1 \leq j, k \leq N}$  is the *kernel matrix* and  $\mathbf{f}$  the vector of length  $N$  of data and  $\alpha$  the vector of the unknown coefficients. The case of *conditionally* positive definite kernels is similar, and we suppress details here.

The interpolant of (4), as in classical polynomial interpolation, can also be written in terms of *cardinal functions*  $u_j \in V_X$  such that  $u_j(x_k) = \delta_{j,k}$ . Then, the interpolant (4) takes the usual *Lagrangian* form (1).

The reproduction quality of kernel-based methods is governed by the *fill distance* or *mesh norm*

$$h_{X, \Omega} = \sup_{x \in \Omega} \min_{x_j \in X} \|x - x_j\|_2 \tag{6}$$

describing the geometric relation of the set  $X$  to the domain  $\Omega$ . In particular, the reproduction error is small if  $h_{X, \Omega}$  is small.

Unfortunately the kernel matrix  $A_{\Phi, X}$  is ill-conditioned if the data locations come close, i.e. if the *separation distance*

$$q_X = \frac{1}{2} \min_{\substack{x_i, x_j \in X \\ x_i \neq x_j}} \|x_i - x_j\|. \tag{7}$$

is small. Then the coefficients of the representation (4) get very large even if the data values  $f(x_k)$  are small, and simple linear solvers will fail.

As a final introductory element, we recall that the fill distance (6) and the separation distance (7) are two fundamental ingredients for standard error and stability estimates for multivariate interpolants, and they will be also of importance here. The inequality  $q_X \leq h_{X, \Omega}$  will hold in most cases, but if points of  $X$  nearly coalesce,  $q_X$  can be much smaller than  $h_{X, \Omega}$ , causing instability of the standard solution process. Point sets  $X$  are called *quasi-uniform* with *uniformity constant*  $\gamma > 1$ , if the inequality

$$\frac{1}{\gamma} q_X \leq h_{X, \Omega} \leq \gamma q_X$$

holds. Later, we shall consider arbitrary sets with different cardinalities, but with uniformity constants bounded above by a fixed number. Note that  $h_{X, \Omega}$  and  $q_X$  play an important role in finding good points for radial basis function interpolation, as recently studied in [3, 5, 9].

## 2 Main results

To generate interpolants, we allow conditionally positive definite translation-invariant kernels

$$\Phi(x, y) = K(x - y) \text{ for all } x, y \in \mathbb{R}^d, \quad K : \mathbb{R}^d \rightarrow \mathbb{R}$$

which are reproducing in their “native” Hilbert space  $\mathcal{N}$  which we assume to be norm-equivalent to some Sobolev space  $W_2^\tau(\Omega)$  with  $\tau > d/2$ . The kernel will then have a Fourier transform satisfying

$$0 < c(1 + \|\omega\|_2^2)^{-\tau} \leq \hat{K}(\omega) \leq C(1 + \|\omega\|_2^2)^{-\tau} \tag{8}$$

at infinity. This includes, for example, Poisson radial functions (cf. [7, 8]), Sobolev/Matérn kernels and Wendland’s compactly supported kernels (cf. e.g. [12]). It is well-known that under the above assumptions the interpolation problem is uniquely solvable, and the space  $V_X$  is a subspace of Sobolev space  $W_2^\tau(\Omega)$ .

In what follows, we assume that the constants are dependent on the space dimension, the domain, and the kernel, and the assertions hold for all sets  $X$  of scattered data locations with sufficiently small fill distance  $h_{X,\Omega}$ .

Our main result is the following theorem.

**Theorem 1** *The classical Lebesgue constant for interpolation with  $\Phi$  on  $N = |X|$  data locations  $X = \{x_1, \dots, x_N\}$  in a bounded domain  $\Omega \subseteq \mathbb{R}^d$  satisfying an inner cone condition has a bound of the form*

$$\Lambda_X \leq C\sqrt{N} \left( \frac{h_{X,\Omega}}{q_X} \right)^{\tau-d/2}.$$

For quasi-uniform sets with bounded uniformity  $\gamma$ , this simplifies to

$$\Lambda_X \leq C\sqrt{N}.$$

Each single cardinal function is bounded by

$$\|u_j\|_{L_\infty(\Omega)} \leq C \left( \frac{h_{X,\Omega}}{q_X} \right)^{\tau-d/2}, \tag{9}$$

which, in the quasi-uniform case, simplifies to

$$\|u_j\|_{L_\infty(\Omega)} \leq C. \tag{10}$$

For the  $L_2$  norm,

$$\|u_j\|_{L_2(\Omega)} \leq C \left( \frac{h_{X,\Omega}}{q_X} \right)^{\tau-d/2} h_{X,\Omega}^{d/2} \tag{11}$$

while for quasi-uniform data locations they behave like

$$\|u_j\|_{L_2(\Omega)} \leq Ch_{X,\Omega}^{d/2}. \tag{12}$$

*Proof* Let us start by bounding the  $u_j$ . Letting  $\Psi \in C^\infty$ , having support in the unit ball and such that  $\Psi(0) = 1$ ,  $\|\Psi\|_{L_\infty(\Omega)} = 1$  (i.e. a ‘‘bump’’ function). We notice that

$$|u_j(x)| \leq \left| \Psi \left( \frac{x - x_j}{q_X} \right) \right| + \left| u_j(x) - \Psi \left( \frac{x - x_j}{q_X} \right) \right|.$$

Since the interpolant  $I_X \Psi \left( \frac{\cdot - x_j}{q_X} \right)$  to  $\Psi \left( \frac{x - x_j}{q_X} \right)$  on  $X$  is  $u_j$ , by using standard error estimates (cf. [13, Corollary 11.33]), we get

$$\|u_j\|_{L_\infty(\Omega)} \leq 1 + \left\| I_X \Psi \left( \frac{\cdot - x_j}{q_X} \right) - \Psi \left( \frac{\cdot - x_j}{q_X} \right) \right\|_{L_\infty(\Omega)} \leq 1 + Ch_{X,\Omega}^{\tau-d/2} \left\| \Psi \left( \frac{\cdot}{q_X} \right) \right\|_{\mathcal{N}}. \tag{13}$$

For the  $L_2$  norm, we obtain the inequality

$$\|u_j\|_{L_2(\Omega)} \leq q_X^{d/2} \|\Psi\|_{L_2(\Omega)} + Ch_{X,\Omega}^\tau \left\| \Psi \left( \frac{\cdot}{q_X} \right) \right\|_{\mathcal{N}}. \tag{14}$$

Hence, we simply need to estimate the native space norm of  $\Psi \left( \frac{\cdot}{q_X} \right)$ .

$$\begin{aligned} \left\| \Psi \left( \frac{\cdot}{q_X} \right) \right\|_{\mathcal{N}}^2 &\leq C \int |q_X^d \hat{\Psi}(q_X \omega)|^2 (1 + |\omega|^2)^\tau d\omega \\ &\leq Cq_X^d \int |\hat{\Psi}(t)|^2 \left( 1 + \left| \frac{t}{q_X} \right|^2 \right)^\tau dt \\ &\leq Cq_X^{d-\tau/2} \int |\hat{\Psi}(t)|^2 (1 + |t|^2)^\tau dt \leq C_1 q_X^{d-\tau/2} \|\Psi\|_{L_2}^2. \end{aligned}$$

Thus, the estimates (9)–(12) easily follow.

Finally we give the claimed bound for the Lebesgue constant. Let

$$p_{f,X}(x) = \sum_{j=1}^N f(x_j) \Psi \left( \frac{x - x_j}{q_X} \right)$$

be the interpolant of the function  $f$  to  $X$  written in terms of translates of the function  $\Psi$ . Then

$$\|I_X p_{f,X}\|_{L_\infty(\Omega)} \leq \|p_{f,X}\|_{L_\infty(\Omega)} + \|I_X p_{f,X} - p_{f,X}\|_{L_\infty(\Omega)}.$$

The first term is bounded by  $\|f\|_{\ell_\infty(X)}$ , since  $p_{f,X}$  is a sum of functions with nonoverlapping supports. For the second term, since  $p_{f,X} \in \mathcal{N}$  we get

$$\|I_X p_{f,X} - p_{f,X}\|_{L_\infty(\Omega)} \leq Ch_{X,\Omega}^{\tau-d/2} \|p_{f,X}\|_{\mathcal{N}}.$$

Then, it remains to estimate  $\|p_{f,X}\|_{\mathcal{N}}$ . For  $\tau \in \mathbb{N}$ , we have

$$\begin{aligned} \|p_{f,X}\|_{\mathcal{N}} &\leq C \left( \sum_{|\alpha| \leq \tau} \|D^\alpha p_{f,X}\|_{L_2}^2 \right)^{1/2} \\ &\leq C \left( \sum_{|\alpha| \leq \tau} \sum_{i=1}^N |f(x_i)|^2 \left\| D^\alpha \Psi \left( \frac{x - x_j}{q_X} \right) \right\|_{L_2}^2 \right)^{1/2} \\ &\leq C \left( \sum_{|\alpha| \leq \tau} \sum_{i=1}^N |f(x_i)|^2 q_X^{d-2\tau} \|D^\alpha \Psi\|_{L_2}^2 \right)^{1/2} \\ &\leq C q_X^{d-2\tau} \|\Psi\|_{W_2^\tau} \left( \sum_{i=1}^N |f(x_i)|^2 \right)^{1/2} \leq C q_X^{d-2\tau} \|\Psi\|_{W_2^\tau} \sqrt{N} \|f\|_{\ell_\infty(X)}. \end{aligned}$$

This concludes the proof. □

But the Lebesgue constants are only upper bounds for the stability constant in function space. In fact, we can do better:

**Corollary 2** *Interpolation on sufficiently many quasi-uniformly distributed data is stable in the sense of*

$$\|s_{f,X}\|_{L_\infty(\Omega)} \leq C (\|f\|_{\ell_\infty(X)} + \|f\|_{\ell_2(X)}) \tag{15}$$

and

$$\|s_{f,X}\|_{L_2(\Omega)} \leq C h_{X,\Omega}^{d/2} \|f\|_{\ell_2(X)} \tag{16}$$

with a constant  $C$  independent of  $X$ .

*Proof* The results easily follow from Theorem 1. □

*Remarks*

1. Note that, in the right-hand side of the inequality (16), the  $\ell_2$  norm is the norm weighted by the cardinality of  $X$ , i.e. a properly scaled discrete version of the  $L_2$  norm.
2. The assumption (8) is crucial and, as we shall show below, we are not able to extend the results to kernels with infinite smoothness, such as the Gaussian. The next section will provide examples showing that similar results are not possible for kernels with infinite smoothness.
3. All the previous results can be proved also by using a *sampling inequality* (cf. [14, Theorem 2.6]), as shown in the note [4].

## 2.1 Examples

We ran a series of examples for uniform grids on  $[-1, 1]^2$  and increasing numbers  $N$  of data locations. Interested readers may refer to the link

<http://profs.sci.univr.it/~demarchi/RBFStability/LebesgueRBF.zip>

where a complete description of these examples is given as well as with illustrative pictures and all the M-files used to produce them.

**Acknowledgements** This work has been supported by the CNR-DFG exchange programme, 2006 and by the GNCS-Indam funds 2007. The authors thank the referees for some useful suggestions, including different, but only slightly shorter proofs for Theorem 1.

## References

1. Bos, L., Caliari, M., De Marchi, S., Vianello, M., Xu, Y.: Bivariate Lagrange interpolation at the Padua points: the generating curve approach. *J. Approx. Theory* **143**, 15–25 (2006)
2. Caliari, M., De Marchi, S., Vianello, M.: Bivariate polynomial interpolation on the square at new nodal sets. *Appl. Math. Comput.* **165**(2), 261–274 (2005)
3. De Marchi, S.: On optimal center locations for radial basis interpolation: computational aspects. *Rend. Sem. Mat. Torino* **61**(3), 343–358 (2003)
4. De Marchi, S., Schaback, R.: Stability constants for kernel-based interpolation processes, R.R. 59/08. Department of Computer Science, University of Verona (2008)
5. De Marchi, S., Schaback, R., Wendland, H.: Near-optimal data-independent point locations for radial basis function interpolation. *Adv. Comput. Math.* **23**(3), 317–330 (2005)
6. Driscoll, T.A., Fornberg, B.: Interpolation in the limit of increasingly flat radial basis functions. *Comput. Math. Appl.* **43**, 413–422 (2002)
7. Flyer, N.: Exact polynomial reproduction for oscillatory radial basis functions on infinite lattices. *Comput. Math. Appl.* **51**(8), 1199–1208 (2004)
8. Fornberg, B., Larsson, E., Wright, G.: A new class of oscillatory radial basis functions. *Comput. Math. Appl.* **51**(8), 1209–1222 (2004)
9. Iske, A.: Optimal distribution of centers for radial basis function methods. Tech. Rep. M0004, Technische Universität München (2000)
10. Madych, W.R.: An estimate for multivariate interpolation II. *J. Approx. Theory* **142**(2), 116–128 (2006)
11. Schaback, R.: Multivariate interpolation by polynomials and radial basis functions. *Constr. Approx.* **21**, 293–317 (2005)
12. Schaback, R., Wendland, H.: Characterization and construction of radial basis function. In: Dyn, N., Leviatan, D., Levin, D., Pinkus, A. (eds.) *Multivariate Approximation and Applications*, pp. 1–24. Cambridge University Press (2001)
13. Wendland, H.: Scattered data approximation. In: *Cambridge Monographs on Applied and Computational Mathematics*, vol. 17 (2005)
14. Wendland, H., Rieger, C.: Approximate interpolation with applications to selecting smoothing parameters. *Numer. Math.* **101**, 643–662 (2005)