Available online at www.sciencedirect.com
ScienceDirect
Journal of Approximation Theory 161 (2009) 645–655
JOURNAL OF Approximation Theory
www.elsevier.com/locate/jat

A Newton basis for Kernel spaces

Stefan Müller, Robert Schaback*

Institut für Numerische und Angewandte Mathematik, Universität Göttingen, Lotzestraße 16-18, D–37073 Göttingen, Germany

Received 16 January 2008; received in revised form 27 October 2008; accepted 31 October 2008
Available online 10 December 2008

Communicated by Martin Buhmann

Abstract

It is well known that representations of kernel-based approximants in terms of the standard basis of translated kernels are notoriously unstable. To come up with a more useful basis, we adopt the strategy known from Newton's interpolation formula, using generalized divided differences and a recursively computable set of basis functions vanishing at increasingly many data points. The resulting basis turns out to be orthogonal in the Hilbert space in which the kernel is reproducing, and under certain assumptions it is complete and allows convergent expansions of functions into series of interpolants. Some numerical examples show that the Newton basis is much more stable than the standard basis of kernel translates.

Keywords: Radial basis functions; Interpolation; Scattered data; Kernels; Condition

1. Stability of evaluation of interpolants

We consider multivariate interpolation on a set $X := \{x_0, \ldots, x_n\}$ of scattered data locations $x_0, \ldots, x_n$ in some bounded domain $\Omega \subset \mathbb{R}^d$. Given values $f(x_0), \ldots, f(x_n)$ of a real-valued function f there, we want to reconstruct f by a linear combination

$$s_{X,f}(x) := \sum_{j=0}^{n} \alpha_j w_j(x) \tag{1}$$

of certain basis functions $w_0, \ldots, w_n$ on Ω. The coefficients $\alpha_0, \ldots, \alpha_n$ result from solving a linear system

* Corresponding author.
E-mail address: schaback@math.uni-goettingen.de (R. Schaback).

0021-9045/$ - see front matter
doi:10.1016/j.jat.2008.10.014

$$\sum_{j=0}^{n} \alpha_j w_j(x_k) = f(x_k), \quad 0 \le k \le n$$

with the coefficient matrix $A_{X,w} = (w_j(x_k))_{0\le j,k\le n}$ which we assume to be invertible.

We now look at the norm of the interpolation projector taking the data vector

$$f_X := (f(x_0), \dots, f(x_n))^T \in \mathbb{R}^{n+1}$$

into the interpolant as an element of $C(\Omega)$ under the L_∞ norm. We get

$$\begin{aligned} \|s_{X,f}\|_\infty &\le \sum_{j=0}^{n} |\alpha_j| \|w_j\|_\infty \\ &\le \|\alpha\|_\infty \sum_{j=0}^{n} \|w_j\|_\infty \\ &= L_{X,w}\|\alpha\|_\infty \\ &\le L_{X,w}\|A_{X,w}^{-1}\|_{\infty,\infty}\|f_X\|_\infty \end{aligned} \tag{2}$$

with the generalized *Lebesgue constant*

$$L_{X,w} = \sum_{j=0}^{n} \|w_j\|_\infty.$$

Note that this way of bounding the interpolation operator is basis-dependent. But since we assume that actual calculations proceed via the coefficients α_j, the above argument describes the error behavior when evaluating the interpolant (1). In fact, the absolute error of evaluating (1) on a machine with precision ϵ will have a worst-case bound

$$\epsilon \sum_{j=0}^{n} |\alpha_j| \|w_j\|_\infty \le \epsilon L_{X,w} \|A_{X,w}^{-1}\|_{\infty,\infty} \|f_X\|_\infty.$$

This means that the instability of evaluation using the basis functions w_j and the formula (1) can be measured by the quantity

$$S_{X,w} := L_{X,w}\|A_{X,w}^{-1}\|_{\infty,\infty}. \tag{3}$$

Note that this is not the condition of the interpolation process as a whole, as considered in the early papers of W. Gautschi [5]. We plan to treat the Gautschi condition in a forthcoming paper.

Let us look at two typical cases. If we use a symmetric positive definite kernel $K : \Omega\times\Omega \to \mathbb{R}$ and the basis

$$w_j := K(\cdot, x_j), \quad 0 \le j \le n,$$

it is well known [8,7] that the smallest eigenvalue of $A_{X,w}$ becomes very small if n becomes large, even if the data points are placed nicely, and the effect becomes worse when the smoothness of the kernel is increased. This instability has been observed by plenty of authors, and there were many attempts to overcome it. For instance, *local* Lagrange bases have been successfully used for certain preconditioning techniques [6,1,2].

But let us look at an opposite case guided by the cited papers. Theoretically, one can go over to a full Lagrange basis $u_0, \dots, u_n$ of the space

$$U_{X,K,n} := \text{Span}\,\{K(\cdot, x_0), \dots, K(\cdot, x_n)\} \tag{4}$$

satisfying $u_j(x_k) = \delta_{jk}, 0 \le j, k \le n$. Then one has $A_{X,u} = I$ and the instability is governed solely by the classical *Lebesgue constant*

$$L_{X,u} := \sum_{j=0}^{n} \|u_j\|_\infty.$$

The paper [3] proves that this constant grows only like $\mathcal{O}(\sqrt{n})$ for reasonably distributed interpolation points and any fixed smoothness of the kernel.

These two examples show that the interpolants behave well in function space though the coefficients in the standard basis tend to be intolerably large in absolute value. This was also observed by many authors. On the other hand, the Lagrange basis is an example with much better stability behavior, but it is hard to calculate.

Consequently, this paper constructs a new type of basis halfway between the Lagrange case and the standard kernel basis. We shall do this by mimicking the Newton interpolation formula. In terms of classical polynomial interpolation, this means that we prefer the Newton form of the interpolant over solving the linear system with a Vandermonde matrix or using the Lagrange basis. As a byproduct, we get an orthogonal basis in the "native" Hilbert space in which the kernel is reproducing, and we can show that the basis is complete, if infinitely many data locations are reasonably chosen. The stability properties of the new basis are shown to lie right between those of the standard and the Lagrange bases, and some numerical examples support our theory.

2. Newton bases

As is well-known, polynomial interpolation to a real-valued function f on $\mathbb{R}$ using values on $n + 1$ data locations

$$x_0 < x_1 < \cdots < x_n$$

on the real line can be done by Newton's formula

$$p_n(x) = \sum_{j=0}^{n} \underbrace{[x_0, \ldots, x_j]f}_{:=\lambda_j(f)} \underbrace{\prod_{i=0}^{j-1}(x - x_i)}_{:=v_j(x)}$$

where $[x_0, \ldots, x_j]f$ stands for the divided difference of order j applied to f at the data locations $x_0, \ldots, x_j$. Note that this takes the form

$$p_n(x) = \sum_{j=0}^{n} \lambda_j(f) v_j(x) \tag{5}$$

splitting the formula into a sum of products of an f-independent basis function v_j and an f-dependent data functional $\lambda_j(f)$, quite like any other quasi-interpolation formula. This representation has the characteristic properties

$$\begin{aligned} v_j(x_i) &= 0, \quad 0 \le i < j \\ v_j(x_j) &\ne 0, \quad 0 \le j \\ \lambda_j(v_i) &= 0, \quad 0 \le i < j \end{aligned} \tag{6}$$

and the simple error representation

$$f(x) - p_n(x) = v_{n+1}(x)[x, x_0, \dots, x_n]f \quad \text{for all } x \in \mathbb{R}.$$

We now turn to general multivariate interpolation on a set $X := \{x_0, \dots, x_n\}$ of scattered data locations $x_0, \dots, x_n$ in some bounded domain $\Omega \subset \mathbb{R}^d$, and we assume a continuous symmetric positive definite kernel K to be given on Ω.

It is a basic fact of kernel-based methods [8,7] that functions of the form

$$p(x) := \sum_{j=0}^{n} \alpha_j K(x, x_j) \tag{7}$$

have a norm given by

$$\|p\|_K^2 := \sum_{j,k=0}^{n} \alpha_j \alpha_k K(x_j, x_k)$$

which arises from the inner product

$$\left(\sum_{j=1}^{n} \alpha_j K(\cdot, x_j), \sum_{k=1}^{m} \beta_k K(\cdot, y_k) \right)_K := \sum_{j=1}^{n} \sum_{k=1}^{m} \alpha_j \beta_k K(x_j, y_k).$$

Under this inner product, the span of functions (7) can be completed to form a "native" Hilbert space $\mathcal{N}$ for the given kernel, and the kernel is "reproducing" in $\mathcal{N}$ in the sense

$$g(x) = \big(g, K(x, \cdot)\big)_K \quad \text{for all } x \in \mathbb{R}^d,\ g \in \mathcal{N}. \tag{8}$$

For later use, we remark that boundedness of Ω and continuity of K imply that the native space $\mathcal{N}$ is continuously embedded in the space of continuous functions via

$$|g(x)| \le C\|g\|_K \quad \text{for all } g \in \mathcal{N}, x \in \Omega \tag{9}$$

with a positive constant C. This follows from the reproduction equation (8) via

$$|g(x)| \le \|g\|_K \sqrt{K(x, x)} \quad \text{for all } x \in \Omega, g \in \mathcal{N}$$

and (8) also implies continuity of all $g \in \mathcal{N}$ since

$$|g(x) - g(y)|^2 \le \|g\|_K^2 \left(K(x, x) - 2K(x, y) + K(y, y)\right) \quad \text{for all } x, y \in \Omega.$$

In view of (6), we now define a basis for the space (4) via "triangular" Lagrange conditions.

Definition 2.1. We define the **Newton basis** $\{v_j\}_{j=0}^n$ on the sequence $X_n := \big(x_j\big)_{j=0}^n$ for the kernel K by

$$\begin{aligned} v_j(x_i) &= 0, \quad 0 \le i < j \le n \\ v_j(x_j) &= 1, \quad 0 \le j \le n \end{aligned} \tag{10}$$

and the requirement

$$v_j \in U_{X,K,j} := \text{Span}\,\{K(\cdot, x_0), \dots, K(\cdot, x_j)\}, \quad 0 \le j \le n. \tag{11}$$

Remark 1. The functions v_j are well defined because of the positive definiteness of the kernel K [8,7]. From the definition one can also see easily the linear independence of v_j. Unlike the Lagrange basis, adding new data locations does not require a recalculation of the basis functions.

Definition 2.2. For $f \in \mathcal{N}$ we define the **coefficient functionals** $\lambda_j(f), 0 \le j \le n$ similar to (5) recursively by the equation

$$f(x_j) = \sum_{k=0}^{j} \lambda_k(f) v_k(x_j), \quad 0 \le j \le n. \tag{12}$$

For convenience we use the notation

$$f_j(x) := \sum_{k=0}^{j} \lambda_k(f) v_k(x), \quad 0 \le j \le n. \tag{13}$$

Remark 2. A permutation of the points in X will change the functionals $\lambda_j(f), 0 \le j \le n$. But for a given sequence of points these functionals are unique due to the recursive structure of (12).

Remark 3. If we use the uniqueness of the representation in the special case $f = v_i$ we get the third equation of (6) in the strengthened form

$$\lambda_j(v_i) = \delta_{ij}, \quad 0 \le i \le j.$$

Lemma 4. *The functions f_j have the interpolation property*

$$f_j(x_k) = f(x_k), \quad 0 \le k \le j.$$

Proof. This follows directly for $j = 0$ and then by induction from

$$f_j(x) = \lambda_j(f) v_j(x) + f_{j-1}(x) \quad \text{and}$$
$$v_j(x_k) = 0, \quad 0 \le k < j. \quad \square$$

Lemma 5. *The coefficient functionals $\lambda_j(f)$ can be computed by the equations*

$$\lambda_0(f) = f_0(x_0),$$
$$\lambda_j(f) = f(x_j) - f_{j-1}(x_j), \quad 1 \le j \le n.$$

Proof.

$$\begin{aligned} f(x_j) &= f_j(x_j) \\ &= \lambda_j(f) v_j(x_j) + \sum_{k=0}^{j-1} \lambda_k(f) v_k(x_j) \\ &= \lambda_j(f) + f_{j-1}(x_j). \quad \square \end{aligned}$$

Now we are looking for a way to calculate the v_j. Later, we shall see that the basis has some hidden orthogonality property, but we can also do the calculation in a direct and straightforward way using a representation

$$\beta_{jj} v_j(x) = K(x, x_j) - \sum_{k=0}^{j-1} \beta_{jk} v_k(x), \qquad \beta_{jk} \in \mathbb{R}, \quad 0 \le k \le j \le n, \tag{14}$$

and applying $v_j(x_i) = \delta_{ij}, 0 \le i \le j$ from (10). The result is

$$\beta_{ji} = K(x_i, x_j) - \sum_{k=0}^{i-1} \beta_{jk} v_k(x_i), \quad 0 \le i \le j \le n,$$
$$\beta_{ji} = 0, \quad \text{for } i > j.$$

One can store the β_{jk} and the $v_j(x_k)$ together in a matrix or compute them directly via LR-decomposition.

$$\left(K(x_i, x_j)\right)_{ij} = \begin{pmatrix} \beta_{11} & & 0 \\ \vdots & \ddots & \\ \beta_{j1} & \cdots & \beta_{jj} \end{pmatrix} \begin{pmatrix} v_1(x_1) & \cdots & v_1(x_j) \\ & \ddots & \vdots \\ 0 & & v_j(x_j) \end{pmatrix}. \tag{15}$$

But we do not claim that the above calculation is the best possible.

3. Orthogonality

The reproduction formula (8) proves

Theorem 6. *For $p \in \mathcal{N}$, $p(x) := \sum_{j=0}^{n} \alpha_j K(x, x_j)$, the following orthogonality relation holds:*

$$(p, g)_K = 0 \quad \text{for all } g \in \mathcal{N} \quad \text{with } g(x_j) = 0, 0 \le j \le n.$$

Proof.

$$(p, g)_K = \sum_{j=0}^{n} \alpha_j \left(K(\cdot, x_j), g\right)_K = \sum_{j=0}^{n} \alpha_j \underbrace{g(x_j)}_{=0} = 0. \quad \square$$

Consequently, (10) and (11) imply orthogonality between the functions of the Newton basis.

Corollary 7. *Using Definition 2.1 we have*

$$(v_j, v_k)_K = 0, \quad 0 \le k < j \le n.$$

Proof. The proof follows directly from Theorem 6 together with

$$v_k \in \text{Span}\,\{K(\cdot, x_0), \ldots, K(\cdot, x_k)\}$$

and $v_j(x_i) = 0$, for $0 \le i < j$. $\square$

Remark 8. The functions $v_j, 0 \le j \le n$, are not orthonormal. However from (14) one can read off that

$$\begin{aligned} \|v_j\|_K^2 &= \left(K(\cdot, x_j) - \sum_{k=0}^{j-1} \beta_{jk} v_k, v_j\right) / \beta_{jj} \\ &= \left(K(\cdot, x_j), v_j\right) / \beta_{jj} \\ &= v_j(x_j) / \beta_{jj} \\ &= 1/\beta_{jj} \end{aligned}$$

holds, using Corollary 7 and the reproduction formula.

From Corollary 7 and Definition (13) we see that $\lambda_j(f)\|v_j\|_K$ is the jth expansion coefficient of f_j in the orthogonal basis $\{v_k\}_{k=0}^{n}$. Therefore we can conclude:

Theorem 9. *The coefficients $\lambda_j(f), 0 \le j \le n$ have the representations*

$$\lambda_j(f) = \left(f_j, \frac{v_j}{\|v_j\|_K^2} \right)_K = \left(f, \frac{v_j}{\|v_j\|_K^2} \right)_K = \left(f_n, \frac{v_j}{\|v_j\|_K^2} \right)_K, \quad 0 \le j \le n$$

for all functions $f \in \mathcal{N}$.

Proof. Since the v_j are orthogonal we get from $f_j(x) = \sum_{k=0}^{j} \lambda_k(f) v_k(x)$ the equation

$$\lambda_j(f)\|v_j\|_K = \left(f_j, \frac{v_j}{\|v_j\|_K} \right)_K.$$

The second and third equations of the theorem follow from

$$0 = (f_j - f)(x_k) = (f_j - f_n)(x_k), \quad 0 \le k \le j \le n$$

and Theorem 6. □

Furthermore, we get from Parseval's identity together with $\left(f_n, \frac{v_j}{\|v_j\|_K}\right)_K = 0$ for $j > n$ the equation

$$\|f_n\|_K^2 = \sum_{j=0}^{n} \left(f_n, \frac{v_j}{\|v_j\|_K} \right)_K^2 = \sum_{j=0}^{n} \left(f, \frac{v_j}{\|v_j\|_K} \right)_K^2 = \sum_{j=0}^{n} \lambda_j^2(f)\|v_j\|_K^2.$$

Since the interpolants to functions f from the native space always are norm-minimal [8,7], we get

$$\|f_n\|_K^2 = \sum_{j=0}^{n} \lambda_j^2(f)\|v_j\|_K^2 \le \|f\|_K^2$$

proving that one can take the limit $n \to \infty$ without problems, if there are infinitely many points.

4. Stability

But before we consider completeness questions and $n \to \infty$ in detail, we want to show a bound like (2) for the Newton basis.

Theorem 10. *For representation (5) there is the bound*

$$\sum_{j=0}^{n} |\lambda_j(f)||v_j(x)| \le C\sqrt{n+1}\|f\|_K \quad \text{for all } f \in \mathcal{N} \tag{16}$$

using the constant from (9).

Proof.

$$\begin{aligned}\sum_{j=0}^{n} |\lambda_j(f)||v_j(x)| &\le C \sum_{j=0}^{n} |\lambda_j(f)|\|v_j\|_K \\ &\le C\sqrt{n+1}\sqrt{\sum_{j=0}^{n} \lambda_j^2(f)\|v_j\|_K^2} \\ &\le C\sqrt{n+1}\|f\|_K. \quad \square\end{aligned}$$

The above result shows that both the coefficients and the functions in the representation of the interpolant by the Newton basis cannot grow exceedingly fast for $n \to \infty$. However, this does not mean that the actual values $\lambda_j(f)$ and $v_j(x)$ are calculated stably. As in the standard Newton representation of polynomial interpolants, the calculation of divided differences from purely pointwise data is necessarily unstable.

For sufficiently dense and well-distributed data in bounded domains, and for kernels with finite smoothness, we have uniform boundedness of $\|v_j\|_\infty$ because each such function is part of a Lagrange basis [3]. Under such conditions, the Lagrange basis also satisfies a bound like (16) due to [3]. Applying Theorem 9, the divided difference functionals $\lambda_j(f)$ then have bounds

$$|\lambda_j(f)| = \frac{(f_n, v_j)_K}{\|v_j\|_K^2} \le \frac{\|f_n\|_K}{\|v_j\|_K} \le \frac{\|f\|_K}{\|v_j\|_K},$$

but these bounds are weaker than the summability implied by (16).

5. Convergence and completeness

As we saw before, it is no problem to let n tend to infinity, but one cannot expect to have a good reproduction quality of interpolants without making further assumptions on the placement of the data locations.

Theorem 11. *Let $x_0, x_1, \ldots$ be an infinite sequence of data locations in Ω which is asymptotically dense, i.e. which has each point of Ω as an accumulation point. If for all $n \ge 0$ the function f_n is the kernel-based interpolant on the points $x_0, \ldots, x_n$ to some fixed given function $f \in \mathcal{N}$, there is norm convergence*

$$\|f - f_n\|_K \to 0, \quad n \to \infty. \tag{17}$$

Furthermore, the orthogonal system consisting of the Newton basis functions v_j is complete in the native Hilbert space $\mathcal{N}$ of the kernel, and each function $f \in \mathcal{N}$ can be represented as

$$f = \sum_{j=0}^{\infty} \frac{(f, v_j)_K}{\|v_j\|_K^2} v_j$$

in the sense of convergence in $\mathcal{N}$ and uniform convergence in Ω.

This result will surely have applications elsewhere, because it is a first case of an orthogonal expansion of functions from reproducing kernel Hilbert spaces into a convergent series of interpolants.

The proof of (17) follows from standard Hilbert space arguments using

$$\|f - f_n\|_K^2 = \|f - f_{n+m}\|_K^2 + \|f_{m+n} - f_n\|_K^2 \quad \text{for all } f \in \mathcal{N}, n, m \in \mathbb{N}$$

due to Theorem 6, and proving that the f_n form a Cauchy sequence with some limit $g \in \mathcal{N}$ with $f - g$ being orthogonal to all $K(x_j, \cdot)$. Then the reproduction equation (8) implies that $f = g$ on all data points, and on Ω by continuity. We thus get

$$f = \lim_{n\to\infty} f_n = \sum_{j=0}^{\infty} \lambda_j(f) v_j = \sum_{j=0}^{\infty} \frac{(f, v_j)_K}{\|v_j\|_K^2} v_j \tag{18}$$

which is convergent in $\|\cdot\|_K$, and then also in $\|.\|_\infty$ by (9). □

The convergence rate in (18) in the uniform norm is the one obtained for the sequence of interpolants, and thus all standard results on uniform error bounds (see e.g. [8], chapter 11) apply to the convergence rate of the series in (18).

To illustrate this, we assume that $\Omega \subset \mathbb{R}^d$ satisfies an interior cone condition and that the infinite sequence $x_0, x_1, \ldots$ is filling the domain *quasi-uniformly*. This means that the consecutive *fill distances*

$$h_j := \sup_{y\in\Omega} \min_{0\le k\le j} \|y - x_k\|_2$$

tend to zero for $j \to \infty$, and at the same time the *separation distances*

$$q_j := \min_{0\le i<k\le j} \|x_i - x_k\|_2$$

are bounded below by

$$q_j \ge c \cdot h_j, \quad j \ge 0$$

by some positive constant c. There are various ways to get such sequences, for example by a special greedy method [4].

Now for each j the volume of the domain can roughly be covered by j balls of radius h_j, such that

$$h_j \approx c \cdot j^{-1/d}$$

holds. Then there is uniform convergence guided by

$$\left\| f - \sum_{j=0}^{n} \frac{(f, v_j)_K}{\|v_j\|_K^2} v_j \right\|_\infty = \|f - f_n\|_\infty \le C h_n^{\tau - d/2} \|f\|_K \le C n^{1/2 - \tau/d} \|f\|_K$$

provided that the kernel is such that its native Hilbert space $\mathcal{N}$ is a subspace of $W_2^\tau(\mathbb{R}^d)$. This is true for all sufficiently smooth kernels, in particular those which are translation-invariant and have a d-variate Fourier transform $\hat{K}$ decaying like

$$\hat{K}(\omega) \le C(1 + \|\omega\|_2^2)^{-\tau}$$

for $\|\omega\|_2$ tending to infinity. See chapters 10 and 11 of [8] for details.

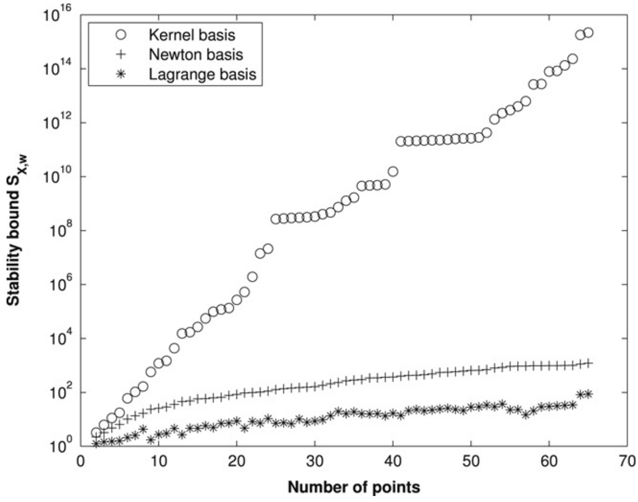

Fig. 1. Stability bound $S_{X,w}$ of (3).

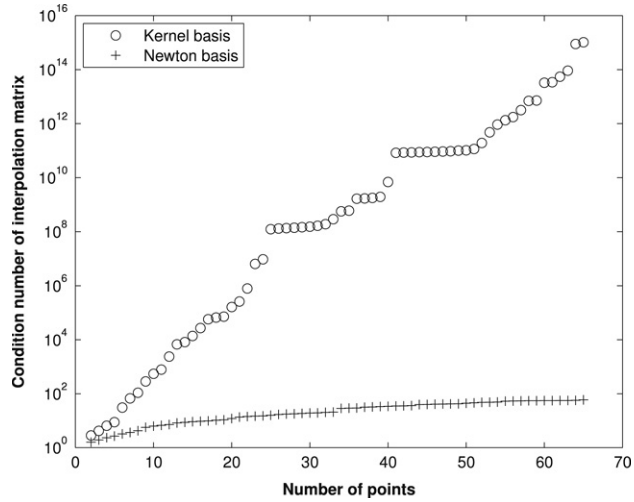

Fig. 2. Condition of interpolation matrix.

6. Examples

In this section we provide numerical examples to support our theoretical results. The data points were quasi-uniformly space-filling in $[-3, 3]^2$ by the greedy method of [4]. We used the Gaussian kernel

$$K(x, y) = \exp(-\|x - y\|^2/25)$$

throughout.

The graphs show that there are big differences between the three bases (kernel, Lagrange, and Newton) as far as evaluation stability is concerned. Fig. 1 displays the stability constant $S_{X,w}$ of (3) for the three bases as a function of the number of data points used.

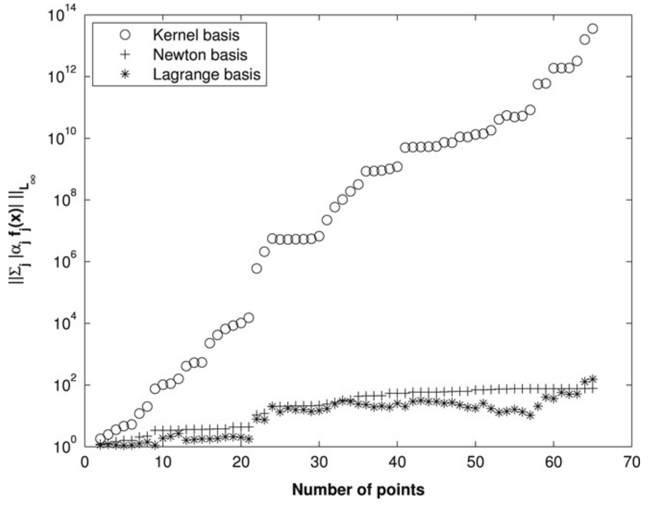

Fig. 3. Bound (2) for stability of evaluation.

To compare the conditions of interpolation matrices, see Fig. 2. The Lagrange basis always has condition 1, and thus it is not displayed. A theoretical investigation of the conditions of the matrices arising from the Newton basis is still missing.

If the MATLAB `peaks` function is interpolated, one can calculate the bound of (2) based on the available coefficients. It cannot exceed the stability constant $S_{X,w}$ up to the factor $\|f_X\|_\infty$, and Fig. 3 shows that the stability bound $S_{X,w}$ is not unrealistic.

A MATLAB© program package is available via http://www.num.math.uni-goettingen.de/schaback/research/group.html.

Acknowledgment

The first author was supported by Deutsche Forschungsgemeinschaft via Graduiertenkolleg 1023.

References

[1] B.J.C. Baxter, Preconditioned conjugate gradients, radial basis functions, and Toeplitz matrices, Comput. Math. Appl. 43 (2002) 305–318.
[2] D. Brown, L. Ling, E.J. Kansa, J. Levesley, On approximate cardinal preconditioning methods for solving PDEs with radial basis functions, Eng. Anal. Bound. Elem. 19 (2005) 343–353.
[3] St. De Marchi, R. Schaback, Stability of kernel-based interpolation, Adv. Comput. Math. (2008), in press (doi:10.1007/s10444-008-9093-4).
[4] St. De Marchi, R. Schaback, H. Wendland, Near-optimal data-independent point locations for radial basis function interpolation, Adv. Comput. Math. 23 (2005) 317–330.
[5] Walter Gautschi, The condition of polynomials in power form, Math. Comp. 33 (145) (1979) 343–352.
[6] C.T. Mouat, Fast algorithms and preconditioning techniques for fitting radial basis functions, Ph.D. Thesis, Mathematics Department, University of Canterbury, Christchurch, New Zealand, 2001.
[7] R. Schaback, H. Wendland, Kernel techniques: From machine learning to meshless methods, Acta Numer. 15 (2006) 543–639.
[8] H. Wendland, Scattered Data Approximation, Cambridge University Press, 2005.