A Newton Basis for Kernel Spaces

Stefan Müller [1] and Robert Schaback

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

Abstract

In this paper we present a strategy for overcoming the ill–conditioning of linear systems arising from radial basis function or kernel techniques. 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.

Key words: Radial basis functions, interpolation, scattered data, kernels, condition
AMS Classification: 41A05,41063, 41065, 65D05, 65D15

1 Stability of Evaluation of Interpolants

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

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

[1] Supported by Deutsche Forschungsgemeinschaft via Graduiertenkolleg 1023

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

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

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

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

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

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

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

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

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

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 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

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

it is well–known [8,7] that the smallest eigenvalue of $A_{X,w}$ gets very small if n gets large, even if the data points are placed nicely, and the effect gets 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 $u_0, \ldots, u_n$ of the space

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

satisfying $u_j(x_k) = \delta_{jk}, \ 0 \leq j, k \leq n$. 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 $\sqrt{n}$ 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 untolerably large in absolute value. This was also observed by many authors. Within certain limits, the quality of the interpolant as a function is not seriously affected by the instability of the basis or the bad condition of the matrix $A_{X,w}$ of the linear system.

Consequently, one should look for better bases.

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 basis, 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

$$x_0 < x_1 < \ldots < x_n$$

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 in 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

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

and the simple error representation

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

We now turn to general multivariate interpolation on a set $X := \{x_0, \ldots, x_n\}$ of scattered data locations $x_0, \ldots, x_n$ in some bounded domain $\Omega \subset \mathbb{R}^d$, and we assume a symmetric positive definite kernel K to be given on Ω. In view of (6), we 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* $X_n := (x_j)_{j=0}^n$ *for the kernel* K *by*

$$\begin{aligned} v_j(x_i) &= 0,\ 0 \le i < j \le n \\ v_j(x_j) &= 1,\ 0 \le j \le n \end{aligned} \tag{7}$$

and the requirement

$$v_j \in U_{X,K,j} := \ \mathit{Span}\ \{K(\cdot, x_0), \ldots, K(\cdot, x_j)\},\ 0 \le j \le n\,. \tag{8}$$

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 the v_j.*

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*

$$f(x_j) = \sum_{k=0}^{j} \lambda_k(f) v_k(x_j)\,, \quad 0 \le j \le n\,. \tag{9}$$

For convenience we use the notation

$$f_j(x) := \sum_{k=0}^{j} \lambda_k(f) v_k(x),\ 0 \le j \le n. \tag{10}$$

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 stucture of (9).*

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*

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

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

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

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

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

Proof:

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

□

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 do the calculation also in a direct and straightforward way using a representation

$$
\beta_{jj}v_j(x) = K(x, x_j) - \sum_{k=0}^{j-1} \beta_{jk}v_k(x)\,, \quad \beta_{jk} \in \mathbb{R},\, 0 \le k \le j \le n, \tag{11}
$$

and applying $v_j(x_i) = \delta_{ij}$, $0 \le i \le j$ from (7). The result is

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

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

$$
(K(x_i, x_j))_{ij} = \begin{pmatrix} \beta_{11} & & 0 \\ \vdots & \ddots & \\ \beta_{j1} & \cdots & \beta_{jj} \end{pmatrix} \begin{pmatrix} v_1(x_1) & \cdots & v_1(x_j) \\ & \ddots & \vdots \\ 0 & & v_j(x_j) \end{pmatrix}. \tag{12}
$$

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

3 Orthogonality

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

$$p(x) := \sum_{j=0}^{n} \alpha_j K(x, x_j) \tag{13}$$

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 (13) 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

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

This reproduction formula 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:*

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

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

Consequently, (7) and (8) imply orthogonality between the functions of the Newton basis.

Corollary 7 *Using the definition (2.1) we have*

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

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 (11) one can read off that*

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

holds, using Corollary (7) and the reproduction formula.

From Corollary (7) and Definition (10) we see that $\lambda_j(f)\|v_j\|_K$ is the j-th 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*

$$
\begin{aligned}
\lambda_j(f) &= \Big(f_j, \frac{v_j}{\|v_j\|_K^2}\Big)_K \\
&= \Big(f, \frac{v_j}{\|v_j\|_K^2}\Big)_K \\
&= \Big(f_n, \frac{v_j}{\|v_j\|_K^2}\Big)_K, \ 0 \le j \le n.
\end{aligned}
$$

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

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

The second and third equation 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 $(f_n, \frac{v_j}{\|v_j\|_K})_K = 0$ for $j > n$ the equation

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

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

$$\begin{aligned}\|f_n\|_K^2 &= \sum_{j=0}^{n} \lambda_j^2(f)\|v_j\|_K^2 \\ &\leq \|f\|_K^2\end{aligned}$$

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 the representation (5) there is the bound*

$$\sum_{j=0}^{n} |\lambda_j(f)||v_j(x)| \leq C\sqrt{n+1}\|f\|_K\,, \tag{14}$$

if we assume that the native space $\mathcal{N}$ is continuously embedded into the space of continuous functions via

$$|g(x)| \leq C\|g\|_K \text{ for all } g \in \mathcal{N},\ x \in \Omega. \tag{15}$$

Proof:

$$\begin{aligned}&\sum_{j=0}^{n} |\lambda_j(f)||v_j(x)| \\ \leq\, & C\sum_{j=0}^{n} |\lambda_j(f)|\|v_j\|_K\end{aligned}$$

$$\leq C\sqrt{n+1}\sqrt{\sum_{j=0}^{n}\lambda_j^2(f)\|v_j\|_K^2}$$
$$\leq C\sqrt{n+1}\|f\|_K\,. \qquad \square$$

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. Like 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, we have uniform boundedness of $\|v_j\|_\infty$ because each such function is part of a Lagrange basis [3]. Due to Theorem 9, the divided difference functionals $\lambda_j(f)$ have bounds

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

but these bounds are weaker than the summability implied by (14).

5 Stability for the Lagrange Basis

In this section we want to show that the bound (14) holds also for the Lagrange basis. The proof again uses the fact that the elements of the basis are orthogonal with respect to $(\cdot,\cdot)_K$.

Theorem 11 *For the Lagrange basis $\{u_j^n\}_{j=0}^n$,*

$$u_j^n(x_i) = \delta_{ij}, \quad u_j^n \in \mathrm{span}\{K(\cdot, x_k) : x_k \in X_n\},\ x_i \in X_n\,, \tag{16}$$

there is the bound

$$\sum_{j=0}^{n} |f(x_j)||u_j^n(x)| \leq C\sqrt{n+1}\|f\|_K\,, \tag{17}$$

if we assume that the native space $\mathcal{N}$ is continuously embedded into the space of continuous functions via

$$|g(x)| \leq C\|g\|_K \text{ for all } g \in \mathcal{N},\ x \in \Omega. \tag{18}$$

Proof: With the definition $s_{f,X_n}(x) := \sum_{j=0}^n f(x_j)u_j^n(x)$ we get

$$
\begin{aligned}
\|f\|_K^2 \geq \|s_{f,X_n}\|_K^2 \\
&= \sum_{j=0}^{n} (s_{f,X_n}, \frac{u_j^n}{\|u_j^n\|_K})_K^2 \\
&= \sum_{j=0}^{n} (f(x_j)u_j, \frac{u_j^n}{\|u_j^n\|_K})_K^2 \\
&= \sum_{j=0}^{n} \frac{f(x_j)^2}{\|u_j^n\|_K^2} (u_j^n, u_j^n)_K^2 \\
&= \sum_{j=0}^{n} f(x_j)^2 \|u_j^n\|_K^2.
\end{aligned}
$$

Thus we can conclude

$$
\begin{aligned}
\sum_{j=0}^{n} |f(x_j)||u_j^n(x)| &\leq C \sum_{j=0}^{n} |f(x_j)| \|u_j^n\|_K \\
&\leq C\sqrt{n+1} \sqrt{\sum_{j=0}^{n} f(x_j)^2 \|u_j^n\|_K^2} \\
&\leq C\sqrt{n+1} \|f\|_K . \qquad \square
\end{aligned}
$$

6 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.

We deal in this section with an f–independent setting. Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain with an interior cone condition. Then we assume an infinite sequence $x_0, x_1, \dots$ of quasi-uniform data locations such that the consecutive *fill distances*

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

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

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

are bounded below by

$$q_j \geq c \cdot h_j, \; j \geq 0$$

by some positive constant c. We call such sequences *quasi-uniformly space-filling* and remark that there are various ways to get such sequences, for example by a special greedy method [4].

If N points fill the domain in a quasi–uniform way, the volume of the domain must roughly be covered by N balls of radius h_N, such that

$$h_N \approx c \cdot N^{-1/d}$$

must be expected. If this is done by refinement of regular grids by a factor of $1/2$, one still gets

$$h_j \approx c \cdot 2^d \cdot j^{-1/d}. \tag{19}$$

If the kernel is such that its native Hilbert space is a subspace of $W_2^\tau(\mathbb{R}^d)$, one can expect a convergence like

$$\|f - f_n\|_\infty \leq Ch_n^{\tau-d/2}\|f\|_K \overset{(19)}{\leq} Cn^{1/2-\tau/d}\|f\|_K$$

when using interpolants f_n based on $n+1$ quasi-uniformly distributed points. Furthermore, one has norm convergence $\|f - f_n\|_K \to 0$ for $n \to \infty$. if the sequence of data points is quasi-uniformly space-filling.

Thus we have a series representation

$$f = \sum_{j=0}^{\infty} \lambda_j(f)v_j \tag{20}$$

which is at least convergent in $\|\cdot\|_K$, and our error bound shows that the partial sums are convergent in the L_∞ norm at the given rate.

These considerations prove the following

Theorem 12 *For quasi-uniformly space-filling sequences and for kernels generating "native" subspaces of $W_2^\tau(\mathbb{R}^d)$ for $\tau > d/2$, the orthogonal system consisting of the Newton basis functions v_j is complete in the native Hilbert space of the kernel, and we can represent each function there as*

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

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.

7 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. Figure 1 displays the stability constant $S_{X,w}$ of (3) for the three bases as a function of the number of data points used.

To compare conditions of interpolation matrices, see Figure 2. The Lagrange basis always has condition 1, and thus it is not displayed.

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 Figure 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`

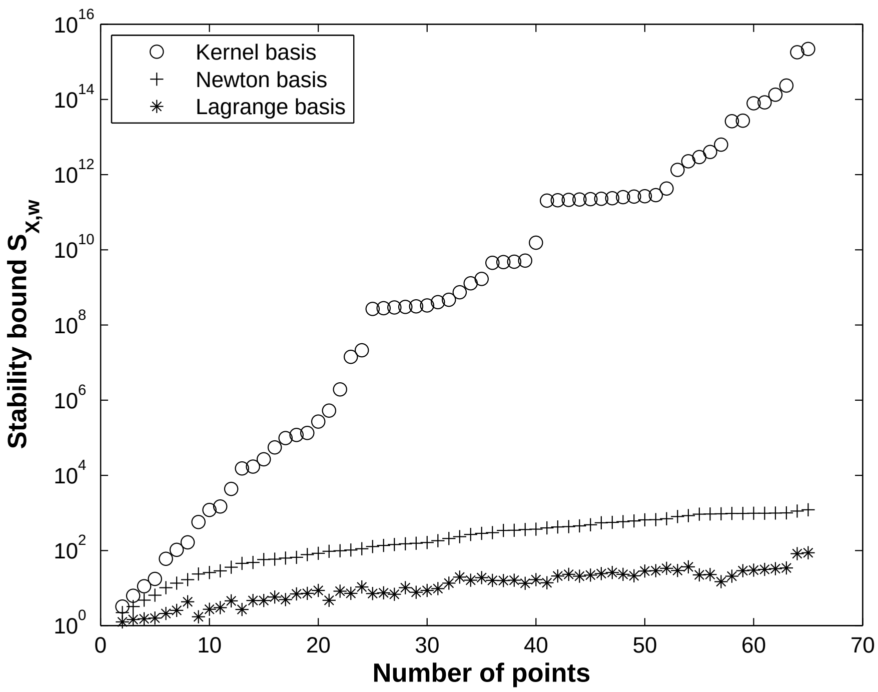

Fig. 1. Stability bound $S_{X,w}$ of (3)

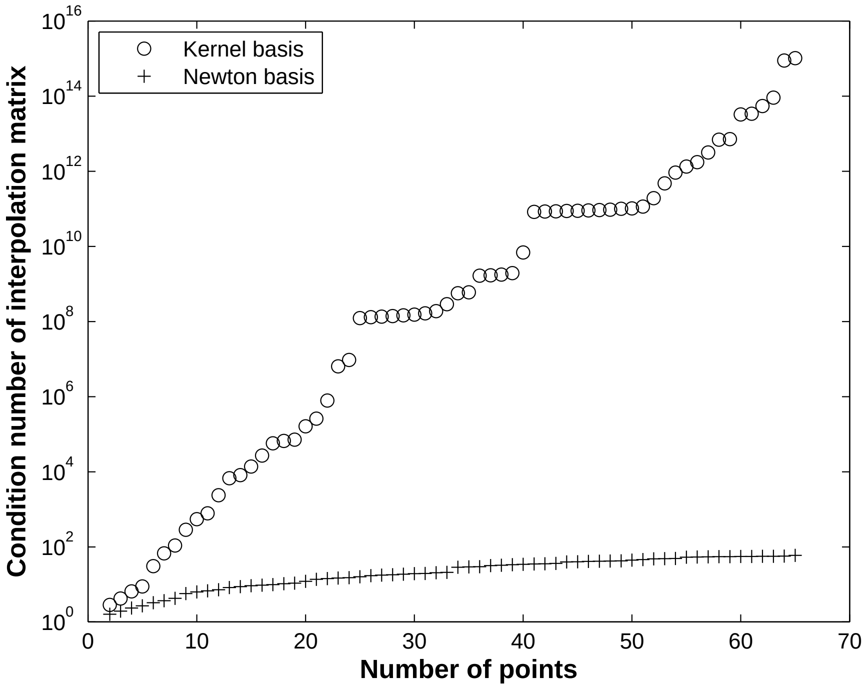

Fig. 2. Condition of interpolation matrix

References

[1] B.J.C. Baxter. Preconditioned conjugate gradients, radial basis functions, and Toeplitz matrices. *Comput. Math. Appl.*, 43:305–318, 2002.

[2] D. Brown, L. Ling, E.J. Kansa, and J. Levesley. On approximate cardinal preconditioning methods for solving PDEs with radial basis functions. *Engineering Analysis with Boundary Elements*, 19:343–353, 2005.

[3] St. De Marchi and R. Schaback. Lebesgue constants for kernel approximants. preprint, 2007.

[4] St. De Marchi, R. Schaback, and H. Wendland. Near–optimal data-independent point locations for radial basis function interpolation. *Adv. Comput. Math.*, pages 317–330, 2005.

[5] Walter Gautschi. The condition of polynomials in power form. *Math. Comp.*, 33(145):343–352, 1979.

[6] C.T. Mouat. *Fast Algorithms and Preconditioning Techniques for Fitting Radial Basis Functions.* PhD thesis, Mathematics Department, University of Canterbury, Christchurch, New Zealand, 2001.

[7] R. Schaback and H. Wendland. Kernel techniques: From machine learning to meshless methods. *Acta Numerica*, 15:543–639, 2006.

[8] H. Wendland. *Scattered Data Approximation.* Cambridge University Press, 2005.

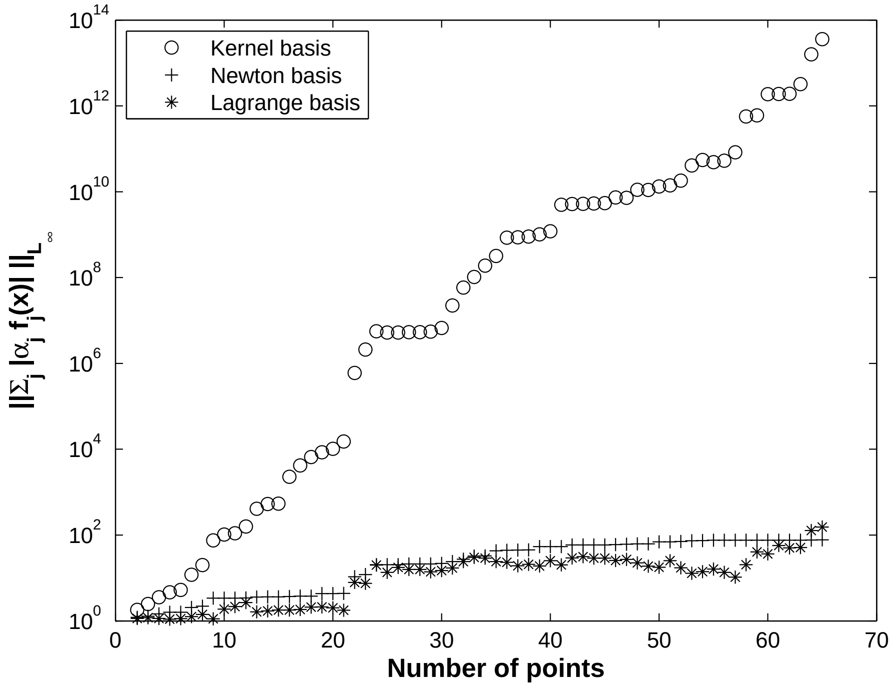

Fig. 3. Bound (2) for stability of evaluation