Stability and Lebesgue constants in RBF interpolation

Stefano De Marchi[1]

[1]Dept. of Computer Science, University of Verona
http://www.sci.univr.it/~demarchi

Göttingen, 20 September 2008

Outline

Motivations

- Stability is very important in numerical analysis: desirable in numerical computations, it depends on the accuracy of algorithms [4, Higham's book].
- In polynomial interpolation, the stability of the process can be measured by the so-called Lebesgue constant, i.e the norm of the projection operator from $\mathcal{C}(K)$ (equipped with the uniform norm) to $\mathcal{P}_n(K)$ (or itselfs) ($K \subset \mathbb{R}^n$), which estimates also the interpolation error.
- The Lebesgue constant depends on the interpolation points via the fundamental Lagrange or cardinal polynomials.

Our approach

1. Good interpolation points [DeM. RSMT03; DeM. Schaback Wendland AiCM05].
2. Cardinal functions bounds [DeM. Schaback AiCM08] .
3. Lebesgue constants estimates and growth [DeM. Schaback AiCM08; Bos DeM. EJA08 (1d)].

Notations

- $X = \{x_1, \dots, x_N\} \subseteq \Omega \subseteq \mathbb{R}^d$, distinct; *data sites*;
- $\{f_1, \dots, f_N\}$, *data values*;
- $\Phi : \Omega \times \Omega \to \mathbb{R}$ *symmetric positive definite kernel*

the RBF interpolant

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

Letting $V_X = \mathrm{span}\{\Phi(\cdot, x) : x \in X\}$, $s_{f,X}$ can be written in terms of cardinal functions, $u_j \in V_X$, $u_j(x_k) = \delta_{jk}$, i.e.

$$s_{f,X} = \sum_{j=1}^{N} f(x_j) u_j\,. \tag{2}$$

Error estimates

- Take $V_\Omega := \text{span}\{\Phi(\cdot, x) : x \in \Omega\}$ on which Φ is the reproducing kernel. $clos(V_\Omega) = \mathcal{N}_\Phi(\Omega)$, the native Hilbert space to Φ.
- $f \in \mathcal{N}_\Phi(\Omega)$, using (2) and the reproducing kernel property of Φ on V_Ω, applying the Cauchy-Schwarz inequality, we get

$$|f(x) - s_{f,X}(x)| \le P_{\Phi,X}(x)\, \|f\|_\Phi \tag{3}$$

$P_{\Phi,X}$: power function.

A power function expression

Letting $\det\left(A_{\Phi,X}(y_1,...,y_N)\right) = \det\left(\Phi(y_i,x_j)\right)_{1\le i,j\le N}$, then

$$u_k(x) = \frac{\det_{\Phi,X}(x_1,\ldots,x_{k-1},x,x_{k+1},\ldots,x_N)}{\det_{\Phi,X}(x_1,\ldots,x_N)}\,, \qquad (4)$$

Letting $u_j(x),\ 0 \le j \le N$ with $u_0(x) := -1$ and $x_0 = x$, then

$$P^2_{\Phi,X}(x) = \mathbf{u}^T A_{\Phi,Y}\mathbf{u}\,, \qquad (5)$$

where $\mathbf{u}^T = (-1, u_1(x), \ldots, u_N(x)),\quad Y = X \cup \{x\}$.

The problem

Are there any good points for approximating all functions in the native space?

Our approach

1. Power function estimates.
2. Geometric arguments.

Literature

- A. Beyer: *Optimale Centerverteilung bei Interpolation mit radialen Basisfunktionen. Diplomarbeit*, Universität Göttingen, 1994.

  He considered numerical aspects of the problem.
- L. P. Bos and U. Maier: *On the asymptotics of points which maximize determinants of the form* $\det(g(|x_i - x_j|))$, in *Advances in Multivariate Approximation* (Berlin, 1999),

  They investigated on Fekete-type points for univariate RBFs, proving that if g is s.t. $g'(0) \neq 0$ then points that maximize the Vandermonde determinant are the ones asymptotically equidistributed.

Literature

- A. Iske:, *Optimal distribution of centers for radial basis function methods.* Tech. Rep. M0004, Technische Universität München, 2000.
  He studied admissible sets of points by varying the centers for stability and quality of approximation by RBF, proving that uniformly distributed points gives better results. He also provided a bound for the so-called *uniformity*: $\rho_{X,\Omega} \leq \sqrt{2(d+1)/d}$, $d=$ space dimension.
- R. Platte and T. A. Driscoll:, *Polynomials and potential theory for GRBF interpolation*, SINUM (2005),they used potential theory for finding near-optimal points for gaussians in 1d.

Main result

Idea: data set for good approximation for all $f \in \mathcal{N}_\Phi(\Omega)$ should have regions in Ω without large holes.
Assume Φ, translation invariant, integrable and its Fourier transform decays at infinity with $\beta > d/2$

Theorem

[DeM., Schaback&Wendland, AiCM 2005.] *For every $\alpha > \beta$ there exists a constant $M_\alpha > 0$ with the following property: if $\epsilon > 0$ and $X = \{x_1, \ldots, x_N\} \subseteq \Omega$ are given such that*

$$\|f - s_{f,X}\|_{L_\infty(\Omega)} \le \epsilon \|f\|_\Phi, \qquad \text{for all } f \in W_2^\beta(\mathbb{R}^d), \tag{6}$$

then the fill distance of X satisfies

$$h_{X,\Omega} \le M_\alpha \epsilon^{\frac{1}{\alpha - d/2}}. \tag{7}$$

Remarks

1. The interpolation error can be bounded by

$$\|f - s_{f,X}\|_{L_\infty(\Omega)} \le C\, h_{X,\Omega}^{\beta-d/2} \|f\|_{W_2^\beta(\mathbb{R}^d)}. \qquad (8)$$

2. $M_\alpha \to \infty$ when $\alpha \to \beta$, so from (8) we cannot get $h_{X,\Omega}^{\beta-d/2} \le C\,\epsilon$ but as close as possible.
3. The proof does not work for gaussians (no compactly supported functions in the native space of the gaussians).

To remedy, we made the additional assumption that

X is already quasi-uniform,i.e. $h_{X,\Omega} \approx q_{X,\Omega}$.

- As a consequence, $P_{\Phi,X}(x) \leq \epsilon$. The result follows from the lower bounds of $P_{\Phi,X}$ (cf. [Schaback AiCM95] where they are given in terms of q_X).
- Quasi-uniformity brings back to bounds in term of $h_{X,\Omega}$.

Observation: optimally distributed data sites are sets that cannot have a large region in Ω without centers, i.e. $h_{X,\Omega}$ is sufficiently small.

On computing near-optimal points

We studied two algorithms.

1. Greedy Algorithm (GA)
2. Geometric Greedy Algorithm (GGA)

The Greedy Algorithm (GA)

At each step we determine a point where the power function attains its maxima w.r.t. the preceding set.

- **starting step:** $X_1 = \{x_1\}, \;\; x_1 \in \Omega$, *arbitrary*.
- **iteration step:** $X_j = X_{j-1} \cup \{x_j\}$ with $P_{\Phi, X_{j-1}}(x_j) = \|P_{\Phi, X_{j-1}}\|_{L_\infty(\Omega)}$.

The Geometric Greedy Algorithm (GGA)

This algorithm works quite well for subset Ω of cardinality n with small $h_{X,\Omega}$ and large q_X. The points are computed independently of the kernel Φ.

- **starting step:** $X_0 = \emptyset$ and define $\mathrm{dist}(x, \emptyset) := A,\ A > \mathrm{diam}(\Omega)$.
- **iteration step:** given $X_n \in \Omega, |X_n| = n$ pick $x_{n+1} \in \Omega \setminus X_n$ s.t. $x_{n+1} = \max_{x \in \Omega \setminus X_n} \mathrm{dist}(x, X_n)$. Then, form $X_{n+1} := X_n \cup \{x_{n+1}\}$.

Remarks on convergence

- Practical experiments show that the GA fills the currently largest hole in the data point close to the center of the hole and converges at least like

$$\|P_j\|_{L_\infty(\Omega)} \le C\, j^{-1/d}\,, \quad C > 0.$$

- Defining the *separation distance* for X_j as $q_j = \frac{1}{2}\min_{x \neq y \in X_j} \|x-y\|_2$ and the *fill distance* as $h_j = \max_{x \in \Omega} \min_{y \in X_j} \|x-y\|_2$ then, we can prove that

$$h_j \ge q_j \ge \frac{1}{2} h_{j-1} \ge \frac{1}{2} h_j, \quad j \ge 2$$

i.e. the GGA produces quasi-uniformly distributed points in the euclidean metric.

Connections with discrete Leja sequences

- Let Ω_N be a discretization of a compact domain of $\Omega \subset \mathbb{R}^2$ and let x_0 arbitrarily chosen in Ω. The points

$$x_n = \max_{x \in \Omega_N \setminus \{x_0,\dots,x_{n-1}\}} \min_{0 \le k \le n-1} \|x - x_k\|_2$$

are a Leja sequence on Ω.

- Hence, the construction technique of GGA is conceptually similar to finding Leja sequences : *both maximize a function of distances*.
- The construction of the GGA is independent of the Euclidean metric. If μ is any metric on Ω, the GGA algorithm produces points asymptotically equidistributed in that metric. In [Caliari,DeM.,Vianello AMC2005] the GGA was used with the Dubiner metric on the square.

How good are the point sets computed by GA and GAA?

We could check these quantities:

- Interpolation error
- Uniformity: in particular, the GGA maximizes

$$\rho_{X,\Omega} = \frac{q_X}{h_{X,\Omega}} ,$$

  since it works well with subset $\Omega_n \subset \Omega$ with large q_X and small $h_{X,\Omega}$.

- Lebesgue constant

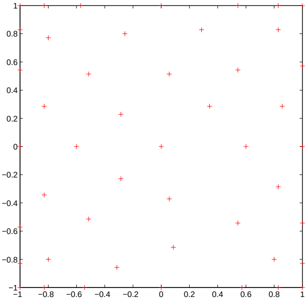
$$\Lambda_N := \max_{x \in \Omega} \lambda_N(x) = \max_{x \in \Omega} \sum_{k=1}^{N} |u_k(x)| .$$

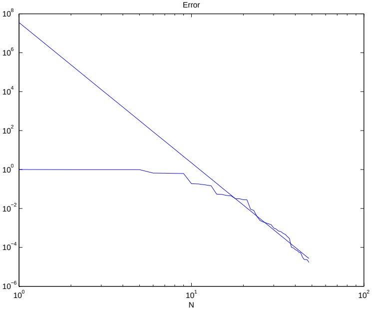
Numerical examples: details

1. We considered a discretization of $\Omega = [-1, 1]^2$ with 10000 random points.
2. The GA run until $\|P_{X,\Omega}\|_\infty \leq \eta$, η a chosen threshold.
3. The GGA, thanks to the connection with the Leja extremal sequences, run once and for all. We extracted 406 points from 406^3 random on $\Omega = [-1, 1]^2$, $406 = \dim(\Pi_{27}(\mathbb{R}^2))$.

GA: Gaussian

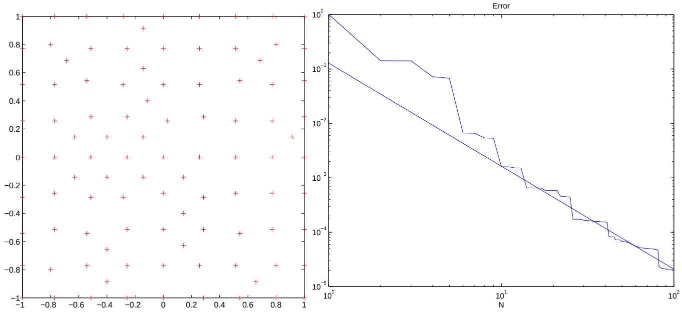
Gaussian with scale 1, 48 points, $\eta = 2 \cdot 10^{-5}$. The "error" in the right–hand figure is $\|P_N\|^2_{L_\infty(\Omega)}$ which has a decay as a function of the number N of data points. As determined by the regression line in the figure, the decay is like $N^{-7.2}$





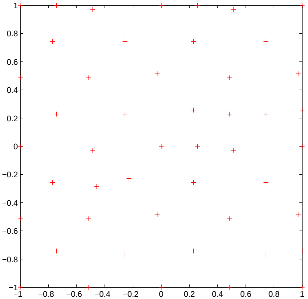
GA: Wendland

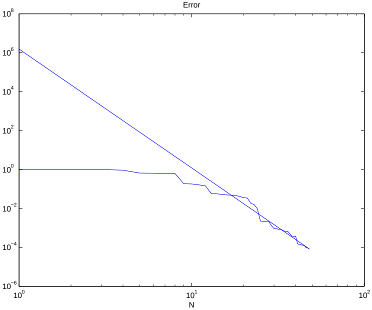
C^2 Wendland function scale 15, $N = 100$ points to depress the power function down to $2 \cdot 10^{-5}$. The error decays like $N^{-1.9}$ as determined by the regression line in the figure.



GGA: Gaussian

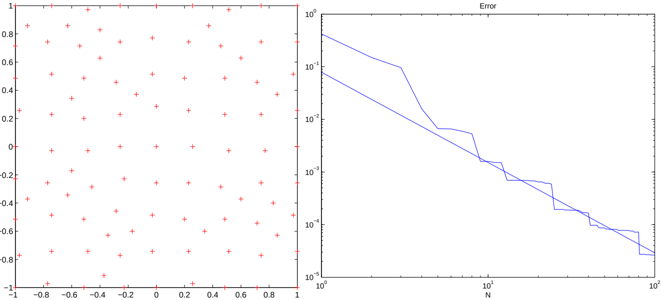
error decay when the Gaussian power function is evaluated on the data supplied by the geometric greedy method up to X_{48}. The final error is larger by a factor of 4, and the estimated decrease of the error is only like $N^{-6.1}$.





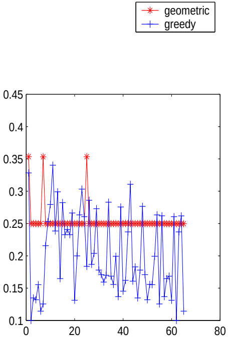
GGA: Wendland

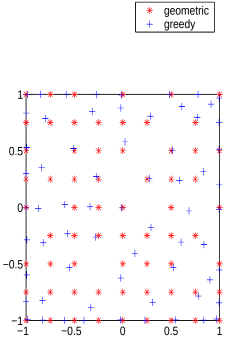
The error factor is only 1.4 bigger, while the estimated decay order is -1.72.



Gaussian

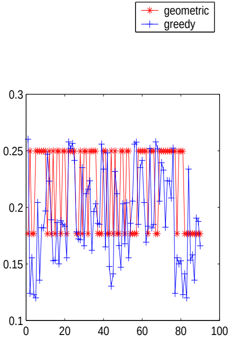
Below: 65 points for the gaussian with scale 1. Left: their separation distances; Right: the points $(+)$ are the one computed with the GA with $\eta = 2.0e - 7$, while the (*) the one computed with the GGA.

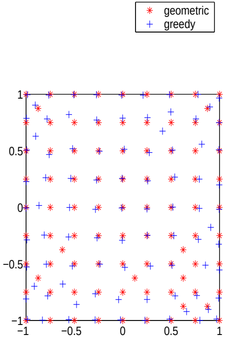




Inverse multiquadrics

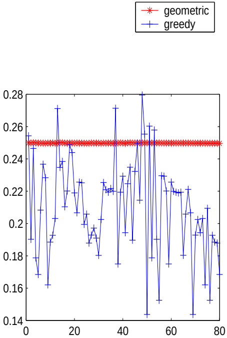
Below: 90 points for the IM with scale 1. Left: their separation distances; Right: the points (+) are the one computed with the GA with $\eta = 2.0e-5$, while the (*) the one computed with the GGA.

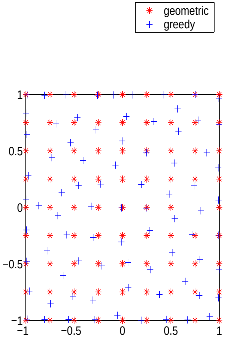




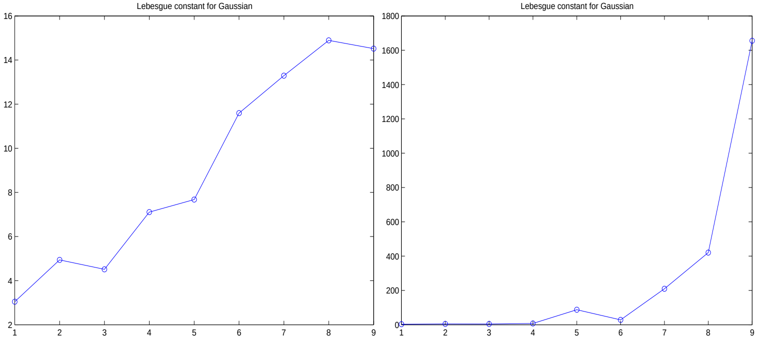
Wendland

Below: 80 points for the Wendland's RBF with scale 1. Left: their separation distances; Right: the points $(+)$ are the one computed with the GA with $\eta = 1.0e - 1$, while the (*) the one computed with the GGA.

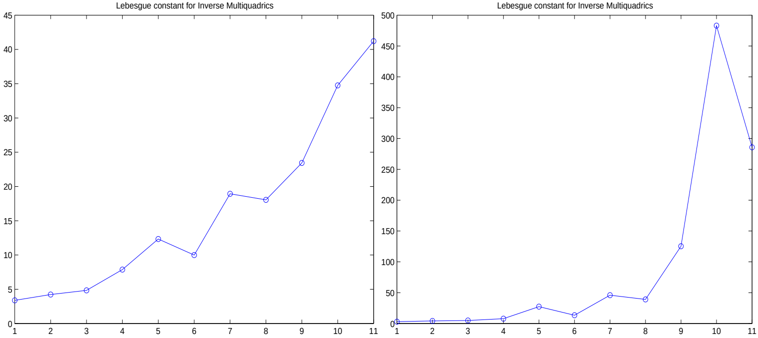




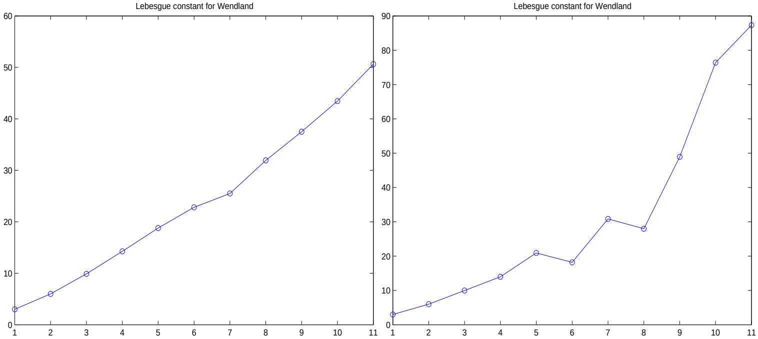
Lebesgue constants for the near-optimal points for the gaussian. Left: the growth of the data-dependent points. Right: the growth of the data-independent points.



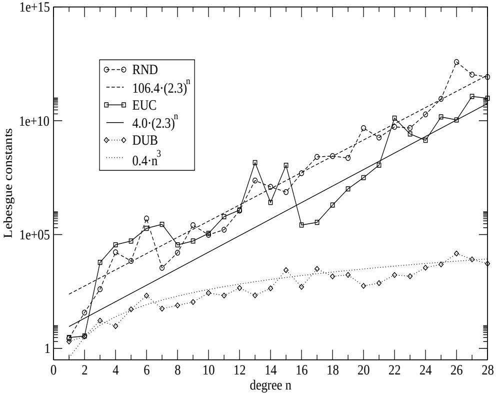
Lebesgue constants for the near-optimal points for the inverse multiquadrics. Left: the growth of the data-dependent points. Right: the growth of the data-independent points.



Lebesgue constants for the near-optimal points for the Wendland's rbf. Left: the growth of the data-dependent points. Right: the growth of the data-independent points.



A comparison of Lebesgue constants growth for points on the square: **RND** (random points), **EUC** (data-independent points), **DUB** (Du



$f_1(x,y) = exp(-8\,x^2 - 8\,y^2)$ and $f_2(x,y) = \sqrt{x^2+y^2-xy}$, on $\Omega = [-1,1]$.

	G-G65	GGA-G65	G-W80	GGA-W80	G-IMQ90	GGA-IMQ90
f_1	$5.5\,10^{-1}$	$**$	$5.6\,10^{-1}$	$**$	$4.9\,10^{-1}$	$**$
f_2	$7.3\,10^{-1}$	$**$	$**$	$**$	$**$	$**$

Table: Errors in L_2-norm for interpolation by the Gaussian. When errors are > 1.0 we put $**$.

	G-G65	GGA-G65	G-W80	GGA-W80	G-IMQ90	GGA-IMQ90
f_1	$2.1\,10^{-1}$	$1.6\,10^{-1}$	$1.3\,10^{-1}$	$1.1\,10^{-1}$	$1.4\,10^{-1}$	$1.0\,10^{-1}$
f_2	$6.1\,10^{-1}$	$8.7\,10^{-1}$	$6.1\,10^{-1}$	$9.7\,10^{-1}$	$4.6\,10^{-1}$	$5.8\,10^{-1}$

Table: Errors in L_2-norm for interpolation by the Wendland's function.

	G-G65	GGA-G65	G-W80	GGA-W80	G-IMQ90	GGA-IMQ90
f_1	$2.3\,10^{-1}$	$2.3\,10^{-1}$	$4.0\,10^{-2}$	$3.1\,10^{-2}$	$3.5\,10^{-2}$	$2.5\,10^{-2}$
f_2	$5.9\,10^{-1}$	$6.0\,10^{-1}$	$3.8\,10^{-1}$	$4.6\,10^{-1}$	$3.7\,10^{-1}$	$3.6\,10^{-1}$

Table: Errors in L_2-norm for interpolation by the inverse multiquadrics.

Remarks

1. The GGA is independent on the kernel and generates asymptotically equidistributed optimal sequences. It still inferior to the GA that considers the power function.
2. The points computed by the GGA is such that $h_{X_n,\Omega} = \max_{x\in\Omega} \min_{y\in X_n} \|x-y\|_2$. In [Caliari,DeM,Vianello2005], we proved that they are quasi-uniform in the Dubiner metric and connected to Leja sequences.
3. So far,we have no proof of the fact the GGA generates a sequence with $h_n \le Cn^{-1/d}$, as required by asymptotic optimality.
4. We could look for data-dependent adaptive strategies for reconstruction of functions from spans of translates of kernels using new techniques known from learning theory and algorithms, applying optimization techniques for data selection (proposed by Robert... not yet implemented!).

Initial ideas

Given the recovery process $f \to s_{f,X}$, where $s_{f,X} = \sum_{j=1}^{N} f(x_j)u_j$ for some $u_j : \Omega \subset \mathbb{R}^d \to \mathbb{R}$ we look for bounds of the form

$$\|s_{f,X}\|_{L_\infty(\Omega)} \leq C(X)\|f\|_{\ell_\infty(X)} \,. \tag{9}$$

$C(X)$, the stability constant, can be bounded below as

$$C(X) \geq \left\| \sum_{j=1}^{N} |u_j(x)| \right\|_{L_\infty(\Omega)} \tag{10}$$

i.e. by the Lebesgue constant $\Lambda_X := \max_{x\in\Omega} \sum_{j=1}^{N} |u_j(x)|$.

Remarks on Polynomial Interpolation

1. Looking for upper bounds for $C(X)$ and/or Λ_X is a classical problem.In recovering by **polynomials**, upper bounds for the Lebesgue constant exist, leading to the problem of finding near-optimal points.
2. For P.I., near-optimal points X of cardinality N, have Λ_X that, in 1D behaves like $\log(N)$ and in 2D on the square like $\log^2(N)$.An important set of near-optimal points in the square for P.I., are the Padua points [Bos,Caliari,DeM,Vianello,Xu JAT06, NM07], http://en.wikipedia.org/wiki/Padua_points.

Padua points

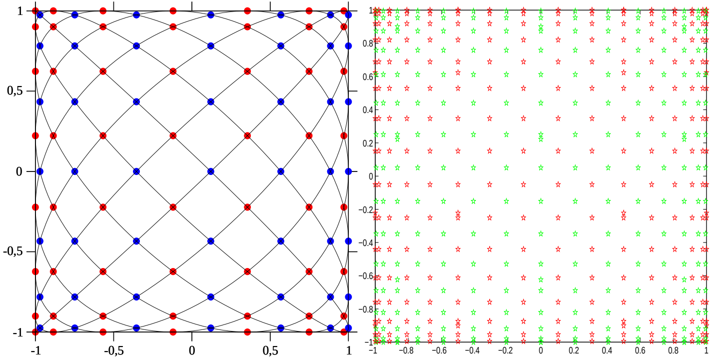


Figure: (Left) Padua points for $N = 13$ and the generating curve. (Right) Padua points for $N = 30$

Padua points

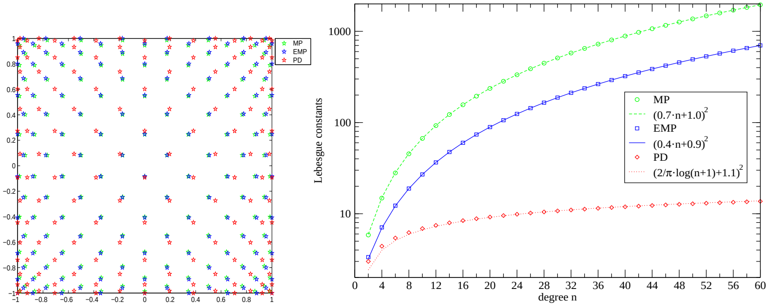


Figure: (Left): Morrow-Patterson, Extended Morrow-Patterson and Padua points for $N = 8$. (Right) Lebesgue constants growth

Motivations

Stability bounds for multivariate kernel–based recovery processes are missing.

How can we proceed to derive them?

Recovering by kernels

Given a kernel $\Phi : \Omega \times \Omega \to \mathbb{R}$ (positive definite), construct

$$s_{f,X} := \sum_{j=1}^{N} \alpha_j \, \Phi(\cdot, x_j) \qquad (11)$$

from $V_X := \text{span}\, \{\Phi(\cdot, x) : x \in X\}$ of translates of Φ so that

$$f(x_k) = s_{f,X}(x_k), \;\; 1 \le k \le N \qquad (12)$$

with matrix $A_{\Phi,X} = (\Phi(x_k, x_j)),\; 1 \le j, k \le N$.

The matrix $A_{\Phi,X}$

1. Unfortunately the kernel matrix has bad condition number if the data locations come close, i.e. if q_X is small.
2. Then, the coefficients of the representation (11) get very large even if the data values $f(x_k)$ are small, and simple linear solvers will fail. Users often report that the final function $s_{f,X}$ of (11) behaves nicely in spite of the large coefficients, and using stable solvers (for instance Riley's algorithm) lead to useful results even in case of unreasonably large condition numbers [Fasshauer's talk]
3. The interpolant can be stably calculated (in the sense of (9)), while the coefficients in the basis supplied by the $\Phi(x, x_j)$ are unstable.

Error estimates and (in)stability

1. $h_{X,\Omega}$ and q_X are used for standard error and stability estimates for multivariate interpolants. The inequality

$$q_X \leq h_{X,\Omega}$$

holds in most cases.

2. 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 holds the inequality

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

Kernels and Fourier transforms

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

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

which are reproducing in their "native" Hilbert space $\mathcal{N}_\Phi$ 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

$$0 < c(1 + \|\omega\|_2^2)^{-\tau} \le \hat{K}(\omega) \le C(1 + \|\omega\|_2^2)^{-\tau} \qquad (13)$$

at infinity. This includes polyharmonic splines, thin-plate splines, the Sobolev/Matérn kernel, and Wendland's compactly supported kernels.

Theorem 1

Theorem

The classical Lebesgue constant for interpolation with Φ on $N = |X|$ data locations in a bounded $\Omega \subseteq \mathbb{R}^d$ has a bound of the form

$$\Lambda_X \le C\sqrt{N}\left(\frac{h_{X,\Omega}}{q_X}\right)^{\tau-d/2}. \qquad (14)$$

For quasi-uniform sets, with uniformity bounded by $\gamma < 1$, this simplifies to $\Lambda_X \le C\sqrt{N}$.
Each single cardinal function is bounded by

$$\|u_j\|_{L_\infty(\Omega)} \le C\left(\frac{h_{X,\Omega}}{q_X}\right)^{\tau-d/2}, \qquad (15)$$

which, in the quasi-uniform case, simplifies to $\|u_j\|_{L_\infty(\Omega)} \le C$.

Corollary

Corollary

Interpolation on sufficiently many quasi–uniformly distributed data is stable in the sense of

$$\|s_{f,X}\|_{L_\infty(\Omega)} \leq C \left(\|f\|_{\ell_\infty(X)} + \|f\|_{\ell_2(X)}\right) \tag{16}$$

and

$$\|s_{f,X}\|_{L_2(\Omega)} \leq C h_{X,\Omega}^{d/2} \|f\|_{\ell_2(X)} \tag{17}$$

with a constant C independent of X.

- In the right-hand side of (17), ℓ_2 is a properly scaled discrete version of the L_2 norm.
- Proofs have been done by resorting to classical error estimates. An alternative proof based on sampling inequality [Rieger, Wendland NM05], has been proposed in [Schaback, DeM. RR59-08,UniVR].

Proof sketch

1. Bound u_j. Using standard error estimates ([Corol. 11.33,Wendland's book]), we get

$$\|u_j\|_{L_\infty(\Omega)} \le 1+\left\| I_X \Psi\left(\frac{\cdot - x_j}{q_X}\right) - \Psi\left(\frac{\cdot - x_j}{q_X}\right)\right\|_{L_\infty(\Omega)} \le 1+C\, h_{X,\Omega}^{\tau-d/2}\left\|\Psi\left(\frac{\cdot}{q_X}\right)\right\|_{\mathcal{N}}. \quad (18)$$

$\Psi \in \mathcal{C}^\infty$, having support in the unit ball and such that $\Psi(0) = 1$, $\|\Psi\|_{L_\infty(\Omega)} = 1$ (i.e. a "bump" function).

2. Estimate the native space norm of $\Psi(\frac{\cdot}{q_X})$ getting

$$\left\|\Psi\left(\frac{\cdot}{q_X}\right)\right\|_{\mathcal{N}}^2 \le C_1\, q_X^{d-\tau/2}\|\Psi\|_{L_2}^2 .$$

Thus, the estimates easily follow.

Proof sketch

Finally, for the Lebesgue constant, observe that $p_{f,X}(x) = \sum_{j=1}^{N} f(x_j)\Psi\left(\frac{x-x_j}{q_X}\right)$ Then

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

- $\|p_{f,X}\|_{L_\infty(\Omega)} \le \|f\|_{\ell_\infty(X)}$, since $p_{f,X}$ is a sum of functions with nonoverlapping supports.
- $$\|I_X p_{f,X} - p_{f,X}\|_{L_\infty(\Omega)} \le C h_{X,\Omega}^{\tau-d/2} \|p_{f,X}\|_{\mathcal{N}} \,.$$

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

$$\|p_{f,X}\|_{\mathcal{N}} \le C q_X^{d-2\tau} \|\Psi\|_{W_2^\tau} \left(\sum_{i=1}^{N} |f(x_j)|^2\right)^{1/2} \le C q_X^{d-2\tau} \|\Psi\|_{W_2^\tau} \sqrt{N} \|f\|_{\ell_\infty(X)} \,.$$

□

Kernels

1. **Matérn/Sobolev kernel** (finite smoothness, definite positive)

$$\Phi(r) = (r/c)^{\nu} K_{\nu}(r/c), \text{ of order } \nu\,.$$

K_{ν} is the modified Bessel function of second kind. Examples were done with $\nu = 1.5$ at scale $c = 20, 320$. Schaback call them *Sobolev splines*.

2. **Gauss kernel** (infinite smoothness, definite positive)

$$\Phi(r) = \mathrm{e}^{-\nu r},\ \nu > 0\,.$$

Examples with $\nu = 1$ at scale $c = 0.1,\ 0.2,\ 0.4$.

Lebesgue constants

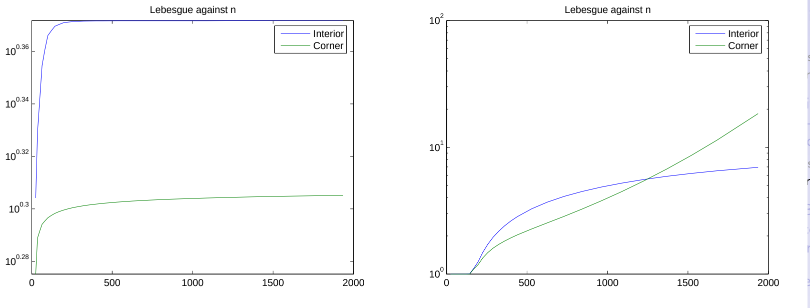
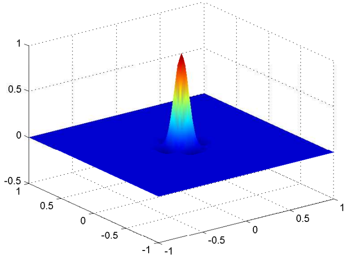


Figure: Lebesgue constants for the Matérn/Sobolev kernel (left) and Gauss kernel (right)

Lebesgue functions



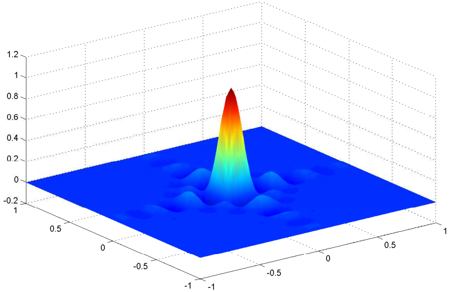


Figure: Lagrange basis function on 225 data points, Gaussian kernel with scale 0.1 (left) and scale 0.2 (right). See how scaling influences the Lagrange basis.

Lebesgue functions

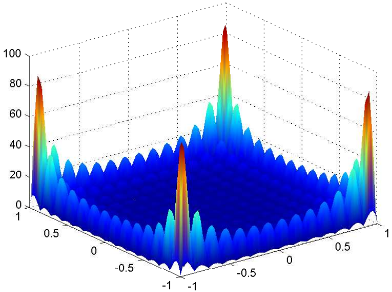


Figure: Gauss kernel with scale 0.4: Lebesgue function on 225 regular. The maximum of the Lebesgue function is attained near the corners for large scales, while the behavior in the interior is as stable as for kernels with limited smoothness.

Lebesgue functions

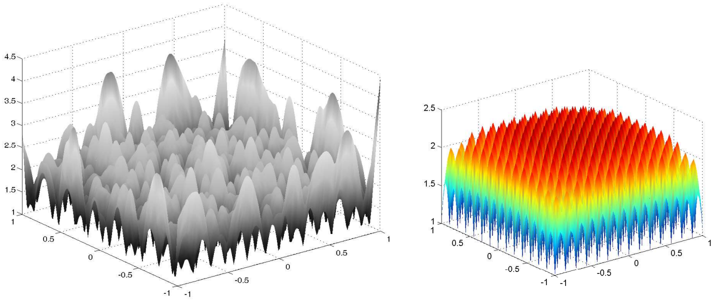
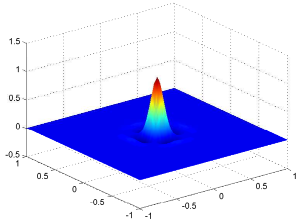


Figure: Matérn/Sobolev kernel with scale 320. Lebesgue function on 225 scattered points (left) and on 225 equidistributed points (right).

Lagrange basis functions



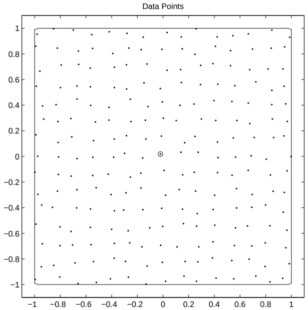
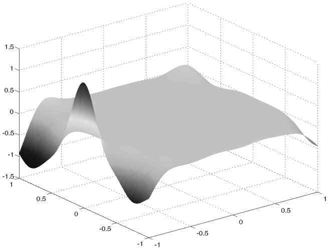


Figure: Matern/Sobolove kernel with scale 320: Lagrange basis (left) on 225 random points (right)

Lagrange basis functions



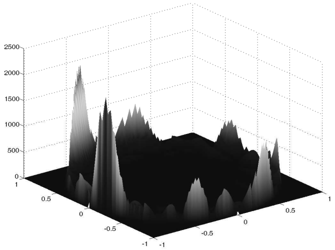


Figure: Lagrange basis (left) and Lebesgue function (right) for 168 scattered data points on the circle, Gaussian kernel with scale 0.4

Lagrange basis functions

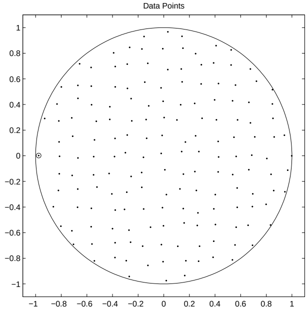


Figure: Data points for the previous figure

Lebesgue constants

Here we collect some computed Lebesgue constants on a grid of centers consisting of 225 pts on $[-1,1]^2$. The constants were computed on a finer grid made of 7225 pts. Matérn and Wendland had scaled by 10, IMQ and GA scaled by 0.2.

Matern	W2	IMQ	GA
2.3	2.3	2.7	4.3
1.3	1.3	1.3	1.7

First line contains the max of Lebesgue functions. The second are the estimated constants, by the Lebesgue function computed by the formula [Wendland's book, p. 208]

$$1+\sum_{i=1}^{N}(u_j^*(x))^2 \leq \frac{P^2_{\Phi,X}(x)}{\lambda_{min}(A_{\Phi,X\cup\{x\}})}, \quad x \notin X\,.$$

in a neighborhood of the point that maximizes the "classical" Lebesgue constant.

Remarks on the finite smooth case

1. In all examples, our bounds on the Lebesgue constants, are confirmed.
2. In all experiments, the Lebesgue constants seem to be uniformly bounded.
3. The maximum of the Lebesgue function is attained in the interior points.

Remarks on the infinite smoothness

... things are moreless specular ...

1. The Lebesgue constants do not seem to be uniformly bounded.
2. In all experiments, the Lebesgue function attains its maximum near the corners (for large scales).
3. The limit for large scales is called flat limit which corresponds to the Lagrange basis function for polynomial interpolation (see Larsson and Fornberg talks, [Driscoll, Fornberg 2002], [Schaback 2005],...).

A possible solution

Schaback, in a recent paper with S. Müller [Müeller, Scahaback JAT08], studied a Newton's basis for overcoming the ill-conditioning of linear systems in RBF interpolation. The basis is orthogonal in the native space in which the kernel is reproducing and more stable.

The case $\phi(x) = x$

This is based on the work [Bos,DeM. EJA2008].

- Sites $x_1 < x_2 < \cdots < x_n$ belong to some interval $[a, b]$
- Interpolation problem (correct): find coefficients $a_j \in \mathbb{R}$

$$\sum_{j=1}^{n} a_j \, |x - x_j| \; = \; y_j, \tag{19}$$

for function values y_j in two ways

1. solve the linear system with Vandermonde matrix;
2. give formulas for the associated cardinal functions u_j.

Formulas for the cardinal functions

The cardinal functions are the classical hat functions given as follows. When x_j is an interior point

$$u_j(x) = \begin{cases} 0 & \text{if } x \le x_{j-1} \\ \frac{x-x_{j-1}}{x_j-x_{j-1}} & \text{if } x_{j-1} < x \le x_j \\ \frac{x_{j+1}-x}{x_{j+1}-x_j} & \text{if } x_j < x \le x_{j+1} \\ 0 & \text{if } x > x_{j+1} \end{cases}, \quad 2 \le j \le n-1, \tag{20}$$

while for the boundary points x_1 and x_n

$$u_1(x) = \begin{cases} \frac{x_2-x}{x_2-x_1} & \text{if } x_1 \le x \le x_2 \\ 0 & \text{if } x > x_2; \end{cases} \tag{21}$$

$$u_n(x) = \begin{cases} 0 & \text{if } x \le x_{n-1} \\ \frac{x-x_{n-1}}{x_n-x_{n-1}} & \text{if } x_{n-1} < x \le x_n. \end{cases} \tag{22}$$

Formulas for the cardinal functions

These “hat” functions has another interesting property:

u_j is a combination of just 3 translates, $|x - x_{j-1}|$, $|x - x_j|$ and $|x - x_{j+1}|$. This also holds for u_1 and u_n identifying $x_0 = x_n$ and $x_{n+1} = x_1$.

For instance, for x_j an interior point

$$u_j(x) = \frac{1}{2(x_j - x_{j-1})}|x - x_{j-1}| - \frac{x_{j+1} - x_{j-1}}{2(x_{j+1} - x_j)(x_j - x_{j-1})}|x - x_j| + \frac{1}{2(x_{j+1} - x_j)}|x - x_{j+1}| \quad 2 \le j \le n-1, \tag{23}$$

Remark: (23) is defined for all $x \in \mathbb{R}$, but is identically zero outside $[x_{j-1}, x_{j+1}]$. The boundary points x_1 and x_n are again slightly different.

The case $\phi''(x) = \lambda^2 \phi(x)$

Assume $\lambda \in \mathbb{C}$.
We proved

1. u_j still a combination of 3 consecutive translates of $\phi(|x|)$ and support $[x_{j-1}, x_{j+1}]$.
2. uniqueness of this class of functions.

$\lambda = 0$ is essentially $\phi(x) = x$, then assume $\lambda \neq 0$. Hence,

$$\phi(x) = ae^{\lambda x} + be^{-\lambda x} \tag{24}$$

for some $a, b \in \mathbb{C}$.

Formulas for the cardinal functions

Observe that the interpolation problem for functions of the form

$$s(x) = \sum_{j=1}^{n} a_j \phi(|x - x_j|) \qquad (25)$$

is correct provided $b \neq a$ and $ae^{\lambda x_n} \neq \pm b e^{\lambda x_1}$.

Theorem

For $\phi(x)$ of the form (24) we have

$$det\left([\phi(|x_i - x_j|)]_{1 \leq i,j \leq n}\right) =$$

$$(b-a)^{n-2} e^{-2\lambda \sum_{j=1}^{n} x_j} \left(\prod_{j=1}^{n-1} (e^{2\lambda x_{j+1}} - e^{2\lambda x_j}) \right) \left(b^2 e^{2\lambda x_1} - a^2 e^{2\lambda x_n} \right).$$

Formulas for the cardinal functions

Proposition

For $\phi(x)$ of the form (24) with a, b so that the interpolation problem is correct, we have for $2 \le j \le n-1$,

$$u_j(x) = A_1 \phi(|x - x_{j-1}|) + A_2 \phi(|x - x_j|) + A_3 \phi(|x - x_{j+1}|)$$

where

$$A_1 = -\frac{e^{\lambda x_{j-1}} e^{\lambda x_j}}{(e^{2\lambda x_j} - e^{2\lambda x_{j-1}})(b-a)},$$

$$A_2 = \frac{(e^{2\lambda x_{j+1}} - e^{2\lambda x_{j-1}})e^{2\lambda x_j}}{(e^{2\lambda x_{j+1}} - e^{2\lambda x_j})(e^{2\lambda x_j} - e^{2\lambda x_{j-1}})(b-a)},$$

$$A_3 = -\frac{e^{\lambda x_j} e^{\lambda x_{j+1}}}{(e^{2\lambda x_{j+1}} - e^{2\lambda x_j})(b-a)}.$$

Formulas for the cardinal functions

- These u_j are *identically* zero outside the interval $[x_{j-1}, x_j]$.
- Similar formulas hold for u_1 and u_n.

Cardinal functions

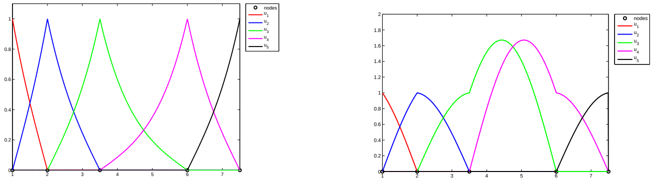


Figure: Cardinal functions for the nodes [1, 2, 3.5, 6, 7.5], $a = 2$, $b = 3$, and $\lambda = 1$ (left), $\lambda = i$ (right)

Uniqueness of the class

Theorem

Suppose that $\phi : \mathbb{R}^+ \to \mathbb{R}$ *is analytic. Suppose further that for any* $x_1 < x_2 < \cdots < x_n$, *the cardinal functions for interpolation of the form (25) can be given as a linear combination of three consecutive translates, i.e., there exist constants* α_j, β_j *and* γ_j *such that*

$$u_j(x) = \alpha_j \phi(|x - x_{j-1}|) + \beta_j \phi(|x - x_j|) + \gamma_j \phi(|x - x_{j+1}|),$$

$2 \le j \le n-1$. *Suppose further that* u_j *has support in the interval* $[x_{j-1}, x_{j+1}]$. *Then there exists a* $\lambda \in \mathbb{C}$ *such that*

$$\phi''(x) = \lambda^2 \phi(x).$$

Formulas for the cardinal functions

Theorem

Suppose that $x_1 < x_2 < \cdots x_n$ and that $\phi(x) = ae^{\lambda x} + be^{-\lambda x}$ is such that the interpolation problem is correct. Then, independently of the values of a and b,

$$u_j(x) = e^{\lambda(x_j - x)} \begin{cases} \frac{e^{2\lambda x} - e^{2\lambda x_{j-1}}}{e^{2\lambda x_j} - e^{2\lambda x_{j-1}}} & \text{if } x \in [x_{j-1}, x_j] \\ \frac{e^{2\lambda x} - e^{2\lambda x_{j+1}}}{e^{2\lambda x_j} - e^{2\lambda x_{j+1}}} & \text{if } x \in [x_j, x_{j+1}] \\ 0 & \text{otherwise} \end{cases} \quad 2 \le j \le n-1\,,$$

and similarly for $u_1,\ u_n$.

Since u_j are positive functions

Proposition

Suppose that $x_1 < x_2 < \cdots x_n$ and that $\phi(x) = ae^{\lambda x} + be^{-\lambda x}$ for $\lambda \in \mathbb{R}$, is such that the interpolation problem is correct. Then, independently of the values of a and b,

$$\sum_{j=1}^{n} |u_j(x)| = \frac{e^{\lambda x} + e^{\lambda(x_j + x_{j+1} - x)}}{e^{\lambda x_j} + e^{\lambda x_{j+1}}}, \quad x \in [x_j, x_{j+1}].$$

In particular,

$$\max_{x_1 \le x \le x_n} \sum_{j=1}^{n} |u_j(x)| = 1.$$

The Case of λ Complex

Consider $\lambda = i$ with $a = -i/2$ and $b = i/2$ so that $g(x) = \sin(x)$. If we make the restriction that $x_n - x_1 < \pi$, one can prove that interpolation problem is correct.It follows

- $u_j(x) \geq 0$ on $[x_1, x_n]$ with $x_n - x_1 < \pi$.
- $$\sum_{j=1}^{n} |u_j(x)| = \frac{\cos(x - \frac{x_j + x_{j+1}}{2})}{\cos(\frac{x_{j+1} - x_j}{2})}, \quad x \in [x_j, x_{j+1}].$$

  The maximum is clearly attained at the midpoint $x = (x_j + x_{j+1})/2$ at which

$$\sum_{j=1}^{n} |u_j(x)| = \frac{1}{\cos(\frac{x_{j+1} - x_j}{2})}.$$

The Case of λ Complex

Hence

$$\Lambda_n := \max_{x_1 \le x \le x_n} \sum_{j=1}^{n} |u_j(x)|$$
$$= \max_{1 \le j \le n-1} \frac{1}{\cos(\frac{x_{j+1}-x_j}{2})}$$
$$= \frac{1}{\cos(\max_{1 \le j \le n-1} \frac{x_{j+1}-x_j}{2})}. \qquad (26)$$

The Case of λ Complex

Theorem

Suppose that $\phi(x) = \sin(x)$. Then, among all distributions of points $a = x_1 < x_2 < \cdots < x_n = b$ in the interval $[a, b]$ with $b - a < \pi$, the one for which Λ_n is uniquely minimized is the equally spaced one, i.e, for

$$x_j = a + \frac{(j-1)(b-a)}{(n-1)}, \;\; 1 \leq j \leq n.$$

The Case of λ Complex

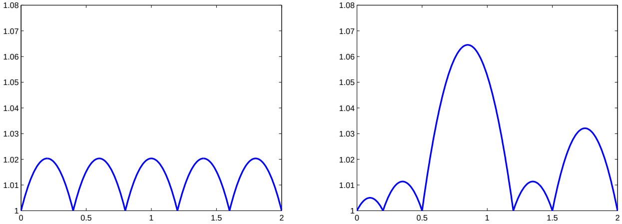


Figure: Lebesgue functions for $\lambda = i$ and equally spaced points (Left) and non-equally spaced points [0 0.2 0.5 1.2 1.5 2] (Right)

Work to do

1. In 1d, we studied the case $\phi(x) = x^3$ and discovered, for nearly equidistributed point set, a behavior similar to that of periodic cubic splines(?) [F. Schurer, Indag. Math.30 (1968)] giving $\Lambda_n < \frac{1}{4}(1+3\sqrt{3})$. More investigations are then necessary!
2. Study better the behavior of the cardinal functions u_j: why do they concentrate around x_j and "decay" at infinity?
3. Efficient computations (for overtaking ill-conditioning and instability) using Nick's Trefethen definition 10 digits, 5 sec. and 1 page!).

Important references

Bos L. and S. De Marchi, *Univariate Radial Basis Functions with Compact Support Cardinal Functions*, East J. Approx., Vol. 14(1) 2008, 69-80.

De Boor, C. and Ron A., *The least solution for the polynomial interpolation problem*, Math. Z., Vol. 210 1992, 347-378.

Driscoll T. and Fornberg B, *Interpolation in the limit of increasingly flat radial basis functions*, Computers Math. Appl. 43 2002, 413-422.

Nicholas J. Higham, *Accuracy and Stability of Numerical Algorithms*, SIAM, Philadelphia, 1996

S. Jokar and B. Meheri,*Lebesgue function for multivariate interpolation by RBF*, Appl. Math. Comp. 187(1) 2007, 306–314.

S. Müller and R. Schaback, *A Newton basis for Kernel Spaces*, To appear on J. Approx. Theory 2008, available at R. Schaback's home page.

H. Wendland, Scattered Data Approximation, *Cambridge Monographs on Applied and Computational Mathematics*, Vol. 17, 2005.

H. Wendland, C. Rieger, *Approximate interpolation with applications to selecting smoothing parameters,*, Num. Math. **101** 2005, 643-662.

DWCAA09

First announcement

2nd Dolomites Workshop on Constructive Approximation and Applications

Alba di Canazei, 3-9 Sept. 2009.

- Keynote speakers (confirmed so far!): Carl de Boor, Robert Schaback, Nick Trefethen, Holger Wendland, Yuan Xu
- Sessions on: Polynomial and rational approximation (Org.: J. Carnicer, A. Cuyt), Approximation by radial bases (Org.: A. Iske, J. Levesley), Quadrature and cubature (Org. B. Bojanov, E. Venturino, Approximation in linear algebra (Org. C. Brezinski, M. Eiermann).

Thank you for your attention!