Max-Planck-Institut
für Mathematik
in den Naturwissenschaften
Leipzig

Low-rank approximation of integral operators by interpolation

by

Steffen Börm and Lars Grasedyck

Preprint no.: 72 2002

Low-rank approximation of integral operators by interpolation

Steffen Börm and Lars Grasedyck

Max Planck Institute of Mathematics in the Sciences
Inselstr. 22–26, D-04103 Leipzig, Germany

August 27, 2002

A central component of the analysis of panel clustering techniques for the approximation of integral operators is the so-called *η-admissibility condition* "$\min\{\mathrm{diam}(\tau), \mathrm{diam}(\sigma)\} \leq 2\eta \mathrm{dist}(\tau, \sigma)$" that ensures that the kernel function is approximated only on those parts of the domain that are far from the singularity.

Typical techniques based on a Taylor expansion of the kernel function require the distance of such a subdomain to be "far enough" from the singularity such that the parameter η has to be smaller than a given constant depending on properties of the kernel function.

In this paper, we demonstrate that *any* η is sufficient if interpolation instead of Taylor expansion is used for the kernel approximation, which paves the way for grey-box panel clustering algorithms.

1 Introduction

1.1 Model problem

Let Ω be a subdomain or submanifold of $\mathbb{R}^d$. We consider a Fredholm integral operator of the form

$$\mathcal{G}[u](x) = \int_\Omega g(x, y) u(y) \mathrm{d}y$$

with an asymptotically smooth kernel function g, i.e., there exist constants C_{as1} and C_{as2} and a singularity degree $s \in \mathbb{N}$ such that for all $z \in \{x_j, y_j\}$ the inequality

$$|\partial_z^n g(x, y)| \leq C_{\mathrm{as1}} (C_{\mathrm{as2}} \|x - y\|)^{-n-s} n! \tag{1}$$

holds. This kind of operator occurs, e.g., in the integral equation formulation of the Poisson problem in $\mathbb{R}^3$, where g is the singularity function $g(x,y) = \|x-y\|^{-1}$. A standard Galerkin discretisation of $\mathcal{G}$ for a basis $(\varphi_i)_{i\in I}$, $V := \text{span}\{\varphi_i \,:\, i \in I\}$, yields a matrix G with entries

$$G_{i,j} := \int_\Omega \int_\Omega \varphi_i(x) g(x,y) \varphi_j(y) \, \mathrm{d}x \, \mathrm{d}y. \tag{2}$$

Since the support of the kernel g is in general not local, one expects a dense matrix G.

The algorithmic complexity for computing and storing a dense matrix is quadratic in the number of degrees of freedom, therefore different approaches have been introduced to avoid dense matrices: for translation-invariant kernel functions and simple geometries, the matrix G has Toeplitz structure, which can be exploited by algorithms based on the fast Fourier transformation. If the underlying geometry can be described by a small number of smooth maps, wavelet techniques can be used in order to compress the resulting dense matrix [3]. Our approach is a refined combination of the panel clustering method [6] and hierarchical matrices [1, 4, 5], which are based on the idea of replacing the kernel function locally by degenerate approximations.

1.2 Low rank approximation

Let $r \times s \subseteq I \times I$ be a sub-block of the product index set. We define the corresponding domains

$$\tau := \cup_{i\in r} \text{supp}(\varphi_i), \quad \sigma := \cup_{i\in s} \text{supp}(\varphi_i)$$

and (minimal) axially parallel boxes B_τ, B_σ containing τ, σ.

We assume that $\text{dist}(B_\tau, B_\sigma) > 0$ holds, which implies that $g|_{B_\tau \times B_\sigma}$ is smooth. For the corresponding sub-matrix $R := K|_{r\times s}$ we seek a low rank matrix $\tilde{R}$ such that the approximation error $\|R - \tilde{R}\|$ is of the same size as the discretisation error $\inf_{v\in V} \|u - v\|$ for the (unknown) solution u. The aim of this paper is to prove that the matrix $\tilde{R}$ of rank k can easily be constructed by interpolation of the kernel g in such a way that the approximation error behaves like

$$\|R - \tilde{R}\| = \mathcal{O}(C_{r,s}^{-k})$$

for a constant $C_{r,s} < 1$, i.e., exponential convergence with respect to the order k even if B_τ and B_σ are arbitrarily close to each other.

2 Interpolation scheme

2.1 Interpolation operators

We denote the space of k-th order polynomials in one spatial variable by $\mathcal{P}_k$ and fix a family $(\mathcal{I}_k)_{k\in\mathbb{N}_0}$ of interpolation operators

$$\mathcal{I}_k : C([-1,1]) \to \mathcal{P}_k$$

corresponding to interpolation points $(x_{i,k})_{i=0}^{k}$ and associated Lagrange polynomials $(\mathcal{L}_{i,k})_{i=0}^{k}$, such that for all $u \in C([-1,1])$

$$\mathcal{I}_k u = \sum_{i=0}^{k} u(x_{k,i})\mathcal{L}_{k,i}. \tag{3}$$

The operators are projections, i.e., for all $p \in \mathcal{P}_k$

$$\mathcal{I}_k p = p \tag{4}$$

holds. For each $k \in \mathbb{N}_0$, we introduce the Lebesgue constant $\Lambda_k \in \mathbb{R}_{\geq 1}$ by requiring that

$$\|\mathcal{I}_k u\|_{\infty,[-1,1]} \leq \Lambda_k \|u\|_{\infty,[-1,1]} \tag{5}$$

holds for all $u \in C([-1,1])$.

We assume that there are constants $C_\lambda, \lambda \in \mathbb{R}_{>0}$ satisfying

$$\Lambda_k \leq C_\lambda (k+1)^\lambda \tag{6}$$

for all $k \in \mathbb{N}$. For standard interpolation schemes, this estimate is fulfilled. E.g., for Chebyshev interpolation (cf. [7]), we even have $\Lambda_k \leq 2\log(k+1)/\pi + 1 \leq k+1$.

If $J = [a,b]$ is an arbitrary closed interval, the transformed interpolation operator is given by $\mathcal{I}_k^J := (\mathcal{I}_k(u \circ \Phi_J)) \circ \Phi_J^{-1}$, where $\Phi_J : [-1,1] \to J,\ x \mapsto ((b-a)x + (b+a))/2$ is the affine mapping from the reference interval to J. The properties (4) and (5) carry over to $\mathcal{I}_k^J$, the corresponding interpolation points and Lagrange polynomials are given by $x_{k,i}^J := \Phi_J(x_{k,i})$ and $\mathcal{L}_{k,i}^J := \mathcal{L}_{k,i} \circ \Phi_J^{-1}$.

2.2 Multidimensional interpolation operator

Let us fix an axially parallel box $B \subseteq \mathbb{R}^d$ with

$$B = J_1 \times \cdots \times J_d,$$

where $(J_j)_{j=1}^{d}$ are closed intervals.

The corresponding k-th order tensor product interpolation operator is given by

$$\mathcal{I}_k^B := \mathcal{I}_k^{J_1} \otimes \cdots \otimes \mathcal{I}_k^{J_d}. \tag{7}$$

This is a projection mapping from $C(B)$ to

$$\mathcal{Q}_k := \operatorname{span}\{p_1 \otimes \cdots \otimes p_d \ :\ p_i \in \mathcal{P}_k \text{ for all } i \in \{1,\dots,d\}\}$$

and the following stability result holds:

Lemma 2.1 (Stability) *For $k \in \mathbb{N}_0$, $m \in \{1,\dots,d\}$ and $u \in C(B)$, we have*

$$\left\| \left(\bigotimes_{i=1}^{m} \mathcal{I}_k^{J_i} \right) \otimes \left(\bigotimes_{i=m+1}^{d} I \right) u \right\|_{\infty,B} \leq \Lambda_k^m \|u\|_{\infty,B},$$

i.e., applying interpolation to the first m coordinate directions is stable. In the case $d = m$, this estimate takes the form

$$\|\mathcal{I}_k^B u\|_{\infty,B} \le \Lambda_k^d \|u\|_{\infty,B}.$$

Proof. We use the representation

$$\left(\bigotimes_{i=1}^m \mathcal{I}_k^{J_i}\right) \otimes \left(\bigotimes_{i=m+1}^d I\right) = \prod_{i=1}^m \left(\left(\bigotimes_{j=1}^{i-1} I\right) \otimes \mathcal{I}_k^{J_i} \otimes \left(\bigotimes_{j=i+1}^d I\right)\right)$$

and apply the one-dimensional estimate to each of the factors. ∎

3 Approximation

Our analysis is based on the following approximation result from [2, Lemma 3.12].

Lemma 3.1 (Melenk) *Let $J \subseteq \mathbb{R}$ be a closed finite interval. Let $u \in C^\infty(J)$ such that there are constants $C_u, \gamma_u \in \mathbb{R}_{\ge 0}$ satisfying*

$$\|u^{(n)}\|_{\infty,J} \le C_u \gamma_u^n n!$$

for all $n \in \mathbb{N}_0$. Then we have

$$\min_{v \in \mathcal{P}_k} \|u - v\|_{\infty,J} \le C_u 4e(1 + \gamma_u |J|)(k+1) \left(1 + \frac{2}{\gamma_u |J|}\right)^{-(k+1)}. \tag{8}$$

Theorem 3.2 (Interpolation error) *Let $u \in C^\infty(B)$ such that there are constants $C_u, \gamma_u \in \mathbb{R}_{\ge 0}$ satisfying*

$$\|\partial_j^n u\|_{\infty,B} \le C_u \gamma_u^n n! \tag{9}$$

for all $j \in \{1, \dots, d\}$ and $n \in \mathbb{N}_0$. Then we have

$$\|u - \mathcal{I}_k^B u\|_{\infty,B} \le 8e\Lambda_k^d C_u (1 + \gamma_u \operatorname{diam}(B))(k+1) \left(1 + \frac{2}{\gamma_u \operatorname{diam}(B)}\right)^{-(k+1)}. \tag{10}$$

Proof. Since $\mathcal{I}_k$ is a projection, we have for all $v \in \mathcal{P}_k$

$$\|u - \mathcal{I}_k^{J_i} u\|_{\infty,J_i} = \|(u - v) - \mathcal{I}_k^{J_i}(u - v)\|_{\infty,J_i} \le (1 + \Lambda_k)\|u - v\|_{\infty,J_i}.$$

Due to (9), we can combine this estimate with Lemma 3.1 and find

$$\left\| u - \left(\bigotimes_{j=1}^{i-1} I\right) \otimes \mathcal{I}_k^{J_i} \otimes \left(\bigotimes_{j=i+1}^d I\right) u \right\|_{\infty,B}$$

$$\leq C_u 4e(1+\Lambda_k)(1+\gamma_u|J_i|)(k+1)\left(1+\frac{2}{\gamma_u|J_i|}\right)^{-(k+1)}$$
$$\leq C_u 8e\Lambda_k(1+\gamma_u \mathrm{diam}(B))(k+1)\left(1+\frac{2}{\gamma_u \mathrm{diam}(B)}\right)^{-(k+1)}.$$

To conclude the proof, we apply this estimate to the following telescope sum:

$$\|u-\mathcal{I}_k^B u\|_{\infty,B} \leq \sum_{i=1}^{d}\left\|\left(\bigotimes_{j=1}^{i-1}\mathcal{I}_k^{J_j}\right)\otimes\left(\bigotimes_{j=i}^{d} I\right)u-\left(\bigotimes_{j=1}^{i}\mathcal{I}_k^{J_j}\right)\otimes\left(\bigotimes_{j=i+1}^{d} I\right)u\right\|$$
$$= \sum_{i=1}^{d}\left\|\left(\bigotimes_{j=1}^{i-1}\mathcal{I}_k^{J_j}\right)\otimes(I-\mathcal{I}_k^{J_i})\otimes\left(\bigotimes_{j=i+1}^{d} I\right)u\right\|$$
$$\stackrel{L.2.1}{\leq} \sum_{i=1}^{d}\Lambda_k^{i-1}\left\|\left(\bigotimes_{j=1}^{i-1} I\right)\otimes(I-\mathcal{I}_k^{J_i})\otimes\left(\bigotimes_{j=i+1}^{d} I\right)u\right\|$$
$$\leq 8e\Lambda_k^d d C_u(1+\gamma_u \mathrm{diam}(B))(k+1)\left(1+\frac{2}{\gamma_u \mathrm{diam}(B)}\right)^{-(k+1)}.$$

■

4 Application to the model problem

4.1 Approximation of the kernel

Let $r \times s \subseteq I \times I$ denote the index sub-set and $\tau \times \sigma$ the support of the corresponding basis functions from Section 1. For both τ and σ we fix axially parallel closed *bounding boxes* B_τ and B_σ satisfying

$$\tau \subseteq B_\tau, \quad \sigma \subseteq B_\sigma \quad \text{and} \quad \mathrm{dist}(B_\tau, B_\sigma) > 0.$$

The k-th order *cluster interpolation operator* is defined in terms of the multidimensional interpolation operator (7) by $\mathcal{I}_k^\tau := \mathcal{I}_k^{B_\tau}$. We define the constants

$$C_g := \frac{C_{\mathrm{as1}}}{(C_{\mathrm{as2}}\,\mathrm{dist}(B_\tau, B_\sigma))^s} \quad \text{and} \quad \gamma_g := \frac{1}{C_{\mathrm{as2}}\,\mathrm{dist}(B_\tau, B_\sigma)}$$

The function $x \mapsto g(x,y)$ fulfils the assumption (9) due to (1). Theorem 3.2 yields

$$\|g(\cdot,y)-\mathcal{I}_k^\tau[g(\cdot,y)]\|_{\infty,B_\tau}$$
$$\leq 8ed\Lambda_k^d C_g(1+\gamma_g\,\mathrm{diam}(B_\tau))(k+1)\left(1+\frac{2}{\gamma_g\,\mathrm{diam}(B_\tau)}\right)^{-(k+1)}.$$

Analogously, we get for the interpolation operator $\mathcal{I}_k^\sigma := \mathcal{I}_k^{B_\sigma}$

$$\|g(x,\cdot)-\mathcal{I}_k^\sigma[g(x,\cdot)]\|_{\infty,B_\sigma}$$

$$\leq 8ed\Lambda_k^d C_g(1+\gamma_g \mathrm{diam}(B_\sigma))(k+1)\left(1+\frac{2}{\gamma_g \mathrm{diam}(B_\sigma)}\right)^{-(k+1)}.$$

Depending on the diameters of B_τ and B_σ we define the kernel approximation

$$\tilde{g}(x,y) := \begin{cases} \mathcal{I}_k^\tau[g(\cdot,y)](x) & \text{if } \mathrm{diam}(B_\tau) \leq \mathrm{diam}(B_\sigma) \\ \mathcal{I}_k^\sigma[g(x,\cdot)](y) & \text{otherwise} \end{cases}. \tag{11}$$

For $C_{\mathrm{diam}} := \min\{\mathrm{diam}(B_\tau), \mathrm{diam}(B_\sigma)\}$ we get the estimate

$$\|g-\tilde{g}\|_{\infty,B_\tau\times B_\sigma} \leq 8ed\Lambda_k^d C_g(1+\gamma_g C_{\mathrm{diam}})(k+1)\left(1+\frac{2}{\gamma_g C_{\mathrm{diam}}}\right)^{-(k+1)}. \tag{12}$$

4.2 Approximation of the matrix block

We define the entries of the matrix $\tilde{R}$ by

$$\tilde{R}_{ij} := \int_\tau \int_\sigma \phi_i(x)\tilde{g}(x,y)\phi_j(y)\,\mathrm{d}x\,\mathrm{d}y.$$

In the case $\mathrm{diam}(B_\tau) \leq \mathrm{diam}(B_\sigma)$, we have $\tilde{g}(x,y) = \mathcal{I}_k^\tau[g(\cdot,y)](x)$, i.e.,

$$\tilde{g}(x,y) = \sum_{\nu\in K} g(x_{K,\nu},y)\mathcal{L}_{K,\nu}(x)$$

with $K := \{0,\dots,k\}^d$ and

$$x_{K,\nu} := (x_{k,\nu_1},\dots,x_{k,\nu_d}) \quad \text{and} \quad \mathcal{L}_{K,\nu} := \mathcal{L}_{k,\nu_1}\otimes\cdots\otimes\mathcal{L}_{k,\nu_d}$$

due to (3). We have the representation $\tilde{R} = XY^\top$ with

$$X_{i\nu} = \int_\tau \phi_i(x)\mathcal{L}_{K,\nu}\,\mathrm{d}x \quad \text{and} \quad Y_{j\nu} = \int_\sigma \phi_j(y)g(x_{K,i},y)\,\mathrm{d}x$$

for $i\in r$, $j\in s$ and $\nu\in K$, which implies $\operatorname{rank}\tilde{R} \leq \#K = (k+1)^d$. By the same arguments, we can prove that $\operatorname{rank}\tilde{R} \leq (k+1)^d$ holds for the second case $\mathrm{diam}(B_\tau) \geq \mathrm{diam}(B_\sigma)$, too.

Lemma 4.1 *The error in the Frobenius norm $\|M\|_F = \left(\sum_{i,j} M_{ij}^2\right)^{1/2}$ is bounded by*

$$\|R-\tilde{R}\|_F \leq \sqrt{\#r\#s}\,C_{g,r,s}\;\Lambda_k^d(k+1)\;\left(1+2C_{\mathrm{as2}}\frac{\mathrm{dist}(B_\tau,B_\sigma)}{\min\{\mathrm{diam}(B_\tau),\mathrm{diam}(B_\sigma)\}}\right)^{-(k+1)}, \tag{13}$$

where the constant $C_{g,r,s}$ is

$$C_{g,r,s} := 8ed\left(\max_i \|\phi_i\|_{L^1}\right)^2 C_{\mathrm{as1}}C_{\mathrm{as2}}^{-s}\mathrm{dist}(B_\tau,B_\sigma)^{-s}\left(1+\frac{\min\{\mathrm{diam}(B_\tau),\mathrm{diam}(B_\sigma)\}}{C_{\mathrm{as2}}\mathrm{dist}(B_\tau,B_\sigma)}\right)$$

Proof. The element-wise error is bounded by

$$|R_{i,j} - \tilde{R}_{i,j}| \overset{(12)}{\leq} \int_\Omega \int_\Omega |\phi_i(x)\phi_j(y)| \mathrm{d}x \, \mathrm{d}y$$

$$8ed\Lambda_k^d C_g (1 + \gamma_g C_{\mathrm{diam}})(k+1) \left(1 + \frac{2}{\gamma_g C_{\mathrm{diam}}}\right)^{-(k+1)}$$

$$\leq C_{g,r,s} \Lambda_k^d (k+1) \left(1 + 2C_{\mathrm{as2}} \frac{\mathrm{dist}(B_\tau, B_\sigma)}{\min\{\mathrm{diam}(B_\tau), \mathrm{diam}(B_\sigma)\}}\right)^{-(k+1)}$$

■

Since $\Lambda_k^d(k+1)$ is bounded by a polynomial and since C_{as2} and $\mathrm{dist}(B_\tau, B_\sigma)$ are positive, we get exponential convergence with respect to the order k. In order to find a uniform bound of the rate of the exponential convergence, one typically demands the *standard admissibility*

$$\min\{\mathrm{diam}(B_\tau), \mathrm{diam}(B_\sigma)\} \leq \mathrm{dist}(B_\tau, B_\sigma),$$

or η-admissibility

$$\min\{\mathrm{diam}(B_\tau), \mathrm{diam}(B_\sigma)\} \leq 2\eta \, \mathrm{dist}(B_\tau, B_\sigma)$$

for a fixed $\eta > 0$.

References

[1] Steffen Börm, Lars Grasedyck, and Wolfgang Hackbusch. Introduction to hierarchical matrices with applications. Technical Report 18, Max Planck Institute for Mathematics in the Sciences, 2002.

[2] Steffen Börm, Maike Löhndorf, and J. Markus Melenk. Approximation of integral operators by variable-order interpolation. Technical report, Max Planck Institute for Mathematics in the Sciences, 2002. to appear.

[3] Wolfgang Dahmen and Reinhold Schneider. Wavelets on manifolds I: Construction and domain decomposition. *SIAM J. of Math. Anal.*, 31:184–230, 1999.

[4] Wolfgang Hackbusch. A sparse matrix arithmetic based on $\mathcal{H}$-matrices. Part I: Introduction to $\mathcal{H}$-matrices. *Computing*, 62:89–108, 1999.

[5] Wolfgang Hackbusch and Boris Khoromskij. A sparse matrix arithmetic based on $\mathcal{H}$-matrices. Part II: Application to multi-dimensional problems. *Computing*, 64:21–47, 2000.

[6] Wolfgang Hackbusch and Zenon Paul Nowak. On the fast matrix multiplication in the boundary element method by panel clustering. *Numerische Mathematik*, 54:463–491, 1989.

[7] T.J. Rivlin. *The Chebyshev Polynomials.* Wiley-Interscience, 1984.