Adaptive tree approximation with finite elements

Andreas Veeser

Università degli Studi di Milano (Italy)

July 2015 / Cimpa School / Mumbai

Outline

1. Basic notions in constructive approximation
2. Tree approximation
3. Mesh refinement with tree structure for PDEs
4. Approximation of gradients
5. Convergence rates for tree approximation of gradients

Outline

Introduction

Replace a complicated function by a simple one, striving for a optimal trade-off.

An example: approximate

- a continuous function
- with piecewise constants over intervals of same length
- in the maximum norm

Applications

Such replacements are used in, e.g.,

- numerical integration
- lossy compression of an acoustic signal, image
- numerical solution of PDEs
- ...

Formalization – 1

- V: 'target functions', linear space
- $(S_n)_n$: with $S_n \subset S_{n+1}$, the functions in each S_n are determined by n parameters,
- $\|\cdot\|$: norm on all $V - S_n$

such that

$$\forall v \in V, \epsilon > 0 \, \exists n \in \mathbb{N}, s \in S_n \quad \|v - s\| \leq \epsilon$$

If all S_n are linear spaces, then

linear approximation, otherwise *nonlinear approximation*.

Formalization – 2

For any algorithm

$$A_n : V \to S_n, \qquad N \in \mathbb{N},$$

we have

$$\|v - A_n v\| \geq E(v, S_n) := \inf_{s \in S_n} \|v - s\| \qquad (\text{" best error "})$$

and

$$\#\text{ of operations for } A_n v \geq n$$

The practical problem

Find an algorithm with the opposite inequalities

$$\|v - A_n v\| \preccurlyeq E(v, S_n) \qquad (\text{"}A_n v \text{ is near best"})$$

and

$$\# \text{ of operations for } A_n v \preccurlyeq n,$$

up to constants.

Such algorithms are called *instance optimal*. Since the density assumption yields $\lim_{n\to\infty} E(v, S_n) = 0$, they are in particular convergent. This brings us to ...

The theoretical problem

Which properties of $v \in V$ determine the speed of

$$E(v, S_n) \to 0 \quad \text{as} \quad n \to \infty?$$

For example: under which conditions do we have

$$E(v, S_n) \le Cn^{-r}$$

with $r > 0$ and $C \ge 0$? What determines r and C?

Approximation spaces – linear case

Introduce

$$\mathcal{A}^r := \{v \in V \mid |v|_{\mathcal{A}^r} < \infty\}$$

with

$$|v|_{\mathcal{A}^r} := \sup_{n \in \mathbb{N}} n^r E(v, S_n),$$

which is a seminorm for linear approximation, since each $E(\cdot, S_n)$ is one.

Clearly,

- $|v|_{\mathcal{A}^r} < \infty$ is decided for big n
- $s \leq r \implies \mathcal{A}^r \subset \mathcal{A}^s$

Approximation spaces – nonlinear case

For nonlinear approximation, we assume that, for some $c \geq 1$,

$$S_n + S_n := \{v_1 + v_2 \mid v_1, v_2 \in S_n\} \subset S_{cn}$$

Then

$$|v|_{\mathcal{A}^r} := \sup_{n \in \mathbb{N}} n^r E(v, S_n),$$

is a *quasi-seminorm*, ie we have only

$$|v_1 + v_2|_{\mathcal{A}^r} \leq C \left(|v_1|_{\mathcal{A}^r} + |v_2|_{\mathcal{A}^r} \right)$$

with some $C > 1$.

Rate optimality

An algorithm $A_n : V \to S_n$, $n \in \mathbb{N}$, is r-rate-optimal whenever there is $C' \geq 0$ such that

$$E(v, S_n) \leq Cn^{-r} \implies \|v - A_n v\| \leq C' C n^{-r}$$

(and $\#$ of operations for $A_n \preccurlyeq n$)

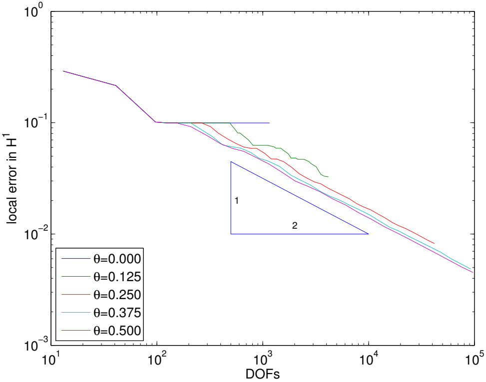
Notice:

- instance optimality implies rate optimality, but not vice versa.

The role of C'

Rate optimality is more than

$$E(v, S_n) = O(n^{-r}) \implies \|v - A_n v\| = O(n^{-r}) :$$



Outline

(Quasi-)norms

Given $\Omega \subset \mathbb{R}^d$ and $p \in (0, \infty]$, we define

$$\|v\|_{L^p(\Omega)} := \begin{cases} \left(\int_\Omega |v|^p\right)^{1/p} & p < \infty, \\ \sup_\Omega |v| & p = \infty, \end{cases}$$

which is a quasi-norm with constant $\max\{1, 2^{\frac{1}{p}-1}\}$.

However, if $p \in (0, 1)$, then

$$\|v_1 + v_2\|^p_{L^p(\Omega)} \leq \|v_1\|^p_{L^p(\Omega)} + \|v_2\|^p_{L^p(\Omega)}$$

often helps.

One-dimensional meshes

Let $\Omega \subset \mathbb{R}$ be an interval. A mesh $\mathcal{M}$ of Ω, is a partition $\{K\}_{K\in\mathcal{M}}$ into intervals such that

$$\overline{\Omega} = \cup_{K\in\mathcal{M}} K \quad \text{and} \quad \mathring{K} \cap \mathring{K}' = \emptyset \text{ whenever } K \neq K'$$

We denote by $\mathbb{M}(\Omega)$ the set of all meshes of Ω.

$\mathcal{M}$ is a uniform mesh whenever

$$|K| = h_{\mathcal{M}} := \frac{|\Omega|}{\#\mathcal{M}}.$$

Linear approximation with piecewise polynomials

Fix a maximal polynomial degree $\ell \in \mathbb{N}$ and set

$$\mathbb{P}^\ell(\mathcal{M}_n) := \{v \mid \forall K \in \mathcal{M}_n \; v_{|K} \in \mathbb{P}^\ell\}.$$

Let $\mathcal{M}_n$ be the equidistant mesh of Ω with n intervals; doubling n halves the meshsize.

Consider the linear spaces

$$S_n = \mathbb{P}^\ell(\mathcal{M}_n), \qquad n \in \mathbb{N}.$$

Nonlinear approximation with free breakpoints

Given $n \in \mathbb{N}$, set

$$\mathbb{M}_n(\Omega) := \{\mathcal{M} \in \mathbb{M}(\Omega) \mid \#\mathcal{M} \leq n\}$$

and consider

$$S_n = \mathbb{P}^\ell\big(\mathbb{M}_n(\Omega)\big) := \bigcup_{\mathcal{M} \in \mathbb{M}_n(\Omega)} \mathbb{P}^\ell(\mathcal{M}).$$

Observe

- the functions in S_n can be described with the help of $O(n)$ parameters,
- we have

$$S_n + S_n \subset S_{2n}$$

Nonlinear approximation with bisection

A bisection of an interval $[a, b]$ is its replacement with the two subintervals $[a, m]$, $[m, b]$ where $m = \frac{1}{2}(a + b)$ is the midpoint.

Given $n \in \mathbb{N}$ and starting form the initial mesh $\{\Omega\}$, set

$$\mathbb{M}_n^{\text{bisect}}(\Omega) := \{\mathcal{M} \in \mathbb{M}(\Omega) \mid \mathcal{M} \text{ is obtained by } n-1 \text{ bisections}\}$$

and consider

$$S_n = \mathbb{P}^\ell\left(\mathbb{M}_n^{\text{bisect}}(\Omega)\right) := \cup_{\mathcal{M}\in\mathbb{M}_n^{\text{bisect}}(\Omega)} \mathbb{P}^\ell(\mathcal{M}).$$

Also here

$$S_n + S_n \subset S_{2n}.$$

Comparison – uniform versus free breakpoints

Best errors in the maximum norm for various functions:

functions on $(0,1)$	$E(\cdot, \mathbb{P}^\ell(\mathcal{M}_n))$	$E(\cdot, \mathbb{P}^\ell(\mathbb{M}_n))$
$\sin(2\pi kx)$	$= 1$ for $n \leq k$	$= 1$ for $n \leq k$
$x^\rho, \rho \geq 1$	$\leq \frac{\rho}{2} n^{-1}$	$\leq \frac{1}{2} n^{-1}$
$x^\rho, \rho \in (0,1)$	$\leq \frac{1}{2} n^{-\rho}$	$\leq \frac{1}{2} n^{-1}$

Notice

- they may be no difference,
- the bigger n, the bigger may be the difference

.

Comparison – bisection

For $n = 2^k$, we have

$$\mathcal{M}_n \in \mathbb{M}_n^{\text{bisect}}(\Omega) \subset \mathbb{M}_n(\Omega)$$

and so

$$\mathbb{P}^\ell(\mathcal{M}_n) \subset \mathbb{P}^\ell\big(\mathbb{M}_n^{\text{bisect}}(\Omega)\big) \subset \mathbb{P}^\ell\big(\mathbb{M}_n(\Omega)\big)$$

In $\mathbb{M}_n^{\text{bisect}}(\Omega)$, the positions of the breakpoints are chosen from a priori fixed set, which becomes dense as $n \to \infty$.