On approximation classes of adaptive methods

Gantumur Tsogtgerel
McGill University

LMS–EPSRC Durham Symposium
Durham

Wednesday July 9, 2014
Plan of the talk

Background:
- approximation classes
- Besov spaces
- multilevel approximation

Ongoing work on
- approximation classes of adaptive finite element methods
### Basic setup

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Omega$</td>
<td>polyhedral Lipschitz domain in $\mathbb{R}^n$</td>
</tr>
<tr>
<td>$P_0$</td>
<td>triangulation of $\Omega$</td>
</tr>
<tr>
<td>$\mathcal{P}$</td>
<td>the family of all <strong>conforming</strong> triangulations obtained from $P_0$ by a sequence of newest vertex bisections</td>
</tr>
<tr>
<td>$S_P$</td>
<td>the Lagrange $C^0$ finite element space of piecewise polynomials of degree not exceeding $m$, subordinate to $P \in \mathcal{P}$</td>
</tr>
<tr>
<td>$X_0$</td>
<td>Examples: $X_0 = L^p(\Omega)$, $X_0 = H^1(\Omega)$</td>
</tr>
</tbody>
</table>

Let

$$E(u, P) = \min_{v \in S_P} \|u - v\|_{X_0}, \quad E_j(u) = \inf_{\{P \in \mathcal{P} : \#P \leq 2^j\}} E(u, P),$$

and define the approximation class $\mathcal{A}_s^\infty(X_0)$ for $s > 0$ by

$$u \in \mathcal{A}_s^\infty(X_0) \iff E_j(u) \lesssim 2^{-js} \iff \left[ 2^{js} E_j(u) \right]_{j \in \mathbb{N}} \in \ell^\infty.$$
Recall

\[ E(u, P) = \min_{v \in S_P} \| u - v \|_{X_0}, \quad E_j(u) = \inf_{\{P \in \mathcal{P} : \#P \leq 2^j\}} E(u, P), \]

and

\[ u \in A_s^{\infty}(X_0) \iff E_j(u) \lesssim 2^{-js} \iff \left[ 2^{js} E_j(u) \right]_{j \in \mathbb{N}} \in \ell^\infty. \]

We extend this definition by introducing \( A_q^s(X_0) \) for \( 0 < q \leq \infty \) by

\[ u \in A_q^s(X_0) \iff \left[ 2^{js} E_j(u) \right]_{j \in \mathbb{N}} \in \ell^q. \]

We have \( A_q^s(X_0) \subset A_r^s(X_0) \) for \( q \leq r \), and \( A_q^s(X_0) \subset A_r^\alpha(X_0) \) for \( s > \alpha \) and for any \( 0 < q, r \leq \infty \). In a typical situation, it is a quasi-Banach space.

We would like to compare, say, \( A_q^s(L^p(\Omega)) \) with known function spaces.
For best $N$-term approximations in a wavelet basis, we have

$$\mathcal{A}_q^s(L^p(\Omega)) = B_{q,q}^\alpha(\Omega), \quad \text{for} \quad s = \frac{\alpha}{n} = \frac{1}{q} - \frac{1}{p} > 0,$$

where $B_{q,r}^\alpha(\Omega)$ is a Besov space ($B_{p,p}^s \approx W^{s,p}$).

For $\frac{\alpha}{n} = \frac{1}{q} - \frac{1}{p}$ we have $B_{q,q}^\alpha(\Omega) \subset L^p(\Omega)$.

Less sharp characterizations are known for

- nonlinear spline approximations
- wavelet tree approximations
- adaptive finite element approximations
Direct and inverse embeddings

[Binev, Dahmen, DeVore, Petrushev '02], [Gaspoz, Morin '13]

\[ B^\alpha_{q,q}(\Omega) \subset \mathcal{A}^s(\mathcal{L}^p(\Omega)) \]

with \( s = \alpha/n \), if

\( \delta = \frac{\alpha}{n} + \frac{1}{p} - \frac{1}{q} > 0 \)

and \( 0 < \alpha < m + \max\{1, \frac{1}{q}\} \).

On the other hand

\[ \mathcal{A}^s(\mathcal{L}^p(\Omega)) \subset B^\alpha_{q,q}(\Omega) \]

for \( s = \frac{\alpha}{n} = \frac{1}{q} - \frac{1}{p} > 0 \),

and \( \alpha < 1 + \frac{1}{q} \).
Main ingredients of the direct theorem

**Direct estimate** \[\text{[BDDP02,GM13]}\]
Let \( \delta = \frac{\alpha}{n} + \frac{1}{p} - \frac{1}{q} > 0 \) and \( 0 < \alpha < m + \max\{1, \frac{1}{q}\} \). Then for \( u \in B^\alpha_{q,q}(\Omega) \) and \( P \in \mathcal{P} \), there exists \( v \in S_P \) such that

\[
\| u - v \|_{L^p(\Omega)}^p \lesssim \sum_{\tau \in P} |\tau|^{p\delta} |u|^p_{B^\alpha_{q,q}(\hat{\tau})},
\]

where \( \hat{\tau} \) is the patch of triangles that touch \( \tau \).

**Proof:** Quasi-interpolator, Whitney estimates, Besov-Sobolev embedding.

**Mesh construction** \[\text{[BDDP02]}\]
For any \( u \in B^\alpha_{q,q}(\Omega) \) and \( N \), there exists \( P \in \mathcal{P} \) with \( \#P \leq N \) such that

\[
\sum_{\tau \in P} |\tau|^{p\delta} |u|^p_{B^\alpha_{q,q}(\hat{\tau})} \lesssim N^{-sp} \| u \|^p_{B^\alpha_{q,q}(\Omega)},
\]

where \( s = \frac{\alpha}{n} \).

**Proof:** Greedy algorithm to reduce \( e(\tau, P) = |\tau|^{\delta} |u|_{B^\alpha_{q,q}(\hat{\tau})} \).
Main ingredients of the inverse theorem

Inverse estimate [BDDP02]
Let $s = \frac{\alpha}{n} = \frac{1}{q} - \frac{1}{p} > 0$ and $\alpha < 1 + \frac{1}{q}$. Then we have

$$
\|v\|_{B_{q,q}^\alpha(\Omega)} \lesssim (\#P)^s \|v\|_{L^p(\Omega)}, \quad P \in \mathcal{P}, \quad v \in S_p.
$$

Proof: Multiscale decomposition of $v$.

Corollary [BDDP02]
For $s = \frac{\alpha}{n} = \frac{1}{q} - \frac{1}{p} > 0$ and $\alpha < 1 + \frac{1}{q}$ we have $\mathcal{A}^s(L^p(\Omega)) \subset B_{q,q}^\alpha(\Omega)$.

Proof: Real interpolation.

The embedding $\mathcal{A}^s_q(L^p(\Omega)) \subset B_{q,q}^\alpha(\Omega)$ cannot hold for $\alpha \geq 1 + \frac{1}{q}$ because in this range we have $S_p \subsetneq B_{q,q}^\alpha(\Omega)$.

This problem was dealt with in [GM13] by introducing generalized Besov spaces $A_{q,q}^\alpha(\Omega)$, and showing that $\mathcal{A}^s_q(L^p(\Omega)) \subset A_{q,q}^\alpha(\Omega)$ for all $\alpha > 0$. We call $A_{q,q}^\alpha(\Omega)$ multilevel approximation spaces.
For $j = 1, 2, \ldots$, let $P_j$ be the uniform refinement of $P_{j-1}$.

Let $G \subset \Omega$ be a domain consisting of elements from some $P_j$.

With $S_j = S_{P_j}$, and $0 < p < \infty$, we let

$$E(u, S_j, G)_p = \inf_{v \in S_j} \| u - v \|_{L^p(G)}, \quad u \in L^p(G).$$

Define the multilevel approximation spaces $A_{p,q}^{\alpha}(G) = A_{p,q}^{\alpha}({S_j}, G)$ by

$$u \in A_{p,q}^{\alpha}({S_j}, G) \iff \left( \lambda^j \alpha E(u, S_j, G)_p \right)_{j \geq 0} \in \ell^q,$$

where $\lambda = \sqrt{2}$.

Note that $u \in A_{p,q}^{\alpha}(G)$ implies $E(u, S_j, G)_p \lesssim 2^{-\alpha j/n} \sim h_j^{\alpha}$, with $h_j$ the typical meshwidth of $P_j$. 
We have $B_{q,q}^\alpha(\Omega) \subset A_{q,q}^\alpha(\Omega)$ for $0 < q < \infty$ and $0 < \alpha < m + \max\{1, \frac{1}{q}\}$.

In the other direction, we have $A_{q,q}^\alpha(\Omega) \subset B_{q,q}^\alpha(\Omega)$ for $0 < q < \infty$ and $0 < \alpha < 1 + \frac{1}{q}$.

So in most interesting situations, we have $B_{q,q}^\alpha(\Omega) \subset A_{q,q}^\alpha(\Omega)$.

Gaspoz-Morin’s inverse theorem says that $A_q^s(L^p(\Omega)) \subset B_{q,q}^\alpha(\Omega)$ for $s = \frac{\alpha}{n} = \frac{1}{q} - \frac{1}{p} > 0$. Recall the inclusion $A_q^s(L^p(\Omega)) \subset B_{q,q}^\alpha(\Omega)$ cannot hold above the red line.

Their direct theorem says that $B_{q,q}^\alpha(\Omega) \subset A_q^s(L^p(\Omega))$ for $\frac{\alpha}{n} > \frac{1}{q} - \frac{1}{p}$ and $0 < \alpha < m + \max\{1, \frac{1}{q}\}$.

Question I: What is the difference between $A_{q,q}^\alpha$ and $B_{q,q}^\alpha$?

Question II: Do we have $A_{q,q}^\alpha(\Omega) \subset A_q^s(L^p(\Omega))$?
**Conjecture:** If \( u \in A_{p,q}^\alpha (\{ S_j \}, \Omega) \) for all possible initial triangulations \( P_0 \) of \( \Omega \), then \( u \in B_{p,q}^\alpha (\Omega) \).

**Lemma**

Let \( \phi \in S_k \) be such that \( \phi \notin C^1 (\Omega) \) for some \( k \). Then there exists an initial triangulation \( \overline{P}_0 \) of \( \Omega \), such that \( E(\phi, \overline{S}_j)_p \gtrsim \lambda^{-(1+\frac{1}{p})j} \) for \( 0 < p < \infty \), where \( \{ \overline{S}_j \} \) is the sequence analogous to \( \{ S_j \} \) with \( P_0 \) replaced by \( \overline{P}_0 \).

**Proof \(( n = 2)\):**

- There is an edge \( e \) of \( P_k \), such that \( |\phi(x,y)| \sim \max\{0,y\} \) under suitable transformation, where \( y \) is the coordinate normal to \( e \).
- We choose \( \overline{P}_0 \) so that \( e \) cuts through the “middle” of each triangle in any refinement of \( \overline{P}_0 \).
Proof \((n = 2)\):

- There is an edge \(e\) of \(P_k\), such that \(|\phi(x, y)| \sim \max\{0, y\}\) under a suitable transformation, where \(y\) is the coordinate normal to \(e\).
- We choose \(\overline{P}_0\) so that \(e\) cuts through the “middle” of each triangle in any refinement of \(\overline{P}_0\).

We have

\[
\|\phi\|_{L^p(V_j)}^p \sim \int_0^{h_j} y^p \, dy \sim h_j^{p+1} \sim \lambda^{-j(p+1)},
\]

where \(V_j\) is the shaded area, and

\[
E(\phi, \overline{S}_j)_p \gtrsim \|\phi\|_{L^p(V_j)} \sim \lambda^{-j(1 + \frac{1}{p})}.
\]
Theorem: We have $A^\alpha_{q,q}(\Omega) \subset \mathcal{A}_\infty^s(L^p(\Omega))$ for $s = \frac{\alpha}{n} > \frac{1}{q} - \frac{1}{p} \geq 0$.

Proof: The two ingredients are the same as before.

Mesh construction
For any $u \in A^\alpha_{q,q}(\Omega)$ and $N$, there exists $P \in \mathcal{P}$ with $\#P \leq N$ such that

$$\sum_{\tau \in P} |\tau|^p |u|^{p}_{A^\alpha_{q,q}(\hat{\tau})} \lesssim N^{-sp} \|u\|^{p}_{A^\alpha_{q,q}(\Omega)},$$

where $s = \frac{\alpha}{n}$.

Proof: The same argument works basically because the spaces $A^\alpha_{q,q}(G)$ enjoy the locality property

$$\sum_{\tau \in P} |u|^{q}_{A^\alpha_{q,q}(\hat{\tau})} \lesssim \|u\|^{q}_{A^\alpha_{q,q}(\Omega)}.$$
Lemma: Let $\delta = \frac{\alpha}{n} + \frac{1}{p} - \frac{1}{q} > 0$. Then for $u \in A_{q,q}^{\alpha}(\Omega)$ and $P \in \mathcal{P}$ we have

$$\|u - Q_P u\|_{L^p(\Omega)} \lesssim \sum_{\tau \in P} |\tau|^{\delta} |u|_{A_{q,q}^{\alpha}(\hat{\tau})}^p,$$

where $Q_P$ is the quasi-interpolation operator from [GM13].

Proof ($q \leq 1$): We have

$$\|u - Q_P u\|_{L^p(\Omega)} = \sum_{\tau \in P} \|u - Q_P u\|_{L^p(\tau)} \lesssim \sum_{\tau \in P} \inf_{v \in S_P} \|u - v\|_{L^p(\hat{\tau})}^p.$$

Every triangle $\sigma \in P$ belongs to a unique $P_j$. Given $\tau \in P$ denote by $j(\tau)$ the highest index $j$ that occurs in the local patch surrounding $\tau$. We have

$$\inf_{v \in S_P} \|u - v\|_{L^p(\hat{\tau})} \leq \inf_{v \in S_{j(\tau)}} \|u - v\|_{L^p(\hat{\tau})},$$

because in $\hat{\tau}$, $P_{j(\tau)}$ is more refined than $P$. 
Proof of direct estimate continued

So far, we have

\[ \| u - Q_P u \|_{L^p(\Omega)} \lesssim \sum_{\tau \in P} \inf_{v \in S_{j(\tau)}} \| u - v \|_{L^p(\hat{\tau})}. \]

For each \( j \), let \( u_j \in S_j \) be such that \( \| u - u_j \|_{L^p(\hat{\tau})} = \inf_{v \in S_j} \| u - v \|_{L^p(\hat{\tau})} \). We have

\[ \| u - u_{j(\tau)} \|_{L^p(\hat{\tau})} \leq \sum_{j=j(\tau)}^{\infty} \| u_{j+1} - u_j \|_{L^p(\hat{\tau})} \lesssim \sum_{j=j(\tau)}^{\infty} \lambda^{(\frac{1}{q} - \frac{1}{p})jp^*} \| u_{j+1} - u_j \|_{L^q(\hat{\tau})}, \]

with \( p^* = \min\{1, p\} \). Putting \( \frac{1}{q} - \frac{1}{p} = \frac{\alpha}{n} - \delta \), we get

\[ \| u - u_{j(\tau)} \|_{L^p(\hat{\tau})} \lesssim \sum_{j=j(\tau)}^{\infty} \lambda^{-jn\delta p^*} \lambda^{j\alpha p^*} \| u - u_j \|_{L^q(\hat{\tau})} \]

\[ \leq \lambda^{-j(\tau)n\delta p^*} \sum_{j=j(\tau)}^{\infty} \lambda^{j\alpha p^*} \| u - u_j \|_{L^q(\hat{\tau})} \lesssim |\tau|^{\delta p^*} |u|_{A^{\alpha}_{p,p^*}}^p. \]
Consider the boundary value problem

\[ \Delta u = f \quad \text{in} \quad \Omega, \quad u = 0 \quad \text{on} \quad \partial \Omega. \]

A typical \textit{a posteriori} error estimate satisfies

\[ \left[ \eta(u, P) \right]^2 \sim \| u - u_P \|_{H^1(\Omega)}^2 + \sum_{\tau \in P} h_\tau^2 \| f - \Pi_\tau f \|_{L^2(\tau)}^2, \]

where \( u_P \in S_P \) is the Galerkin solution on \( P \), and \( \Pi_\tau : L^2(\tau) \to \mathbb{P}_d \) is the \( L^2(\tau) \)-orthogonal projection onto \( \mathbb{P}_d \), \( d \geq m - 2 \).

It is known that certain practical adaptive FEM converges optimally w.r.t. approximation classes associated to

\[ E(u, P) = \left( \min_{\nu \in S_P} \| u - \nu \|_{H^1(\Omega)}^2 + \sum_{\tau \in P} h_\tau^2 \| f - \Pi_\tau f \|_{L^2(\tau)}^2 \right)^{\frac{1}{2}}. \]
Generalized approximation classes

Let \( \rho(u, v, P) = \left( \| u - v \|^2_{H^1(\Omega)} + \sum_{\tau \in P} h^2_{\tau} \| f - \Pi_{\tau} f \|^2_{L^2(\tau)} \right)^{\frac{1}{2}} \), and define

\[
E(u, P) = \min_{v \in S_P} \rho(u, v, P), \quad E_j(u) = \inf_{\{P \in \mathcal{P} : \#P \leq 2^j\}} E(u, P).
\]

We introduce the approximation class \( \mathcal{A}^s_q(\rho) \) given by

\[
u \in \mathcal{A}^s_q(\rho) \iff \left[ 2^{js} E_j(u) \right]_{j \in \mathbb{N}} \in \ell^q.
\]

Also, define the oscillation class \( \mathcal{O}^s_q \) by

\[
f \in \mathcal{O}^s_q \iff \inf_{\{P \in \mathcal{P} : \#P \leq 2^j\}} \sum_{\tau \in P} h^2_{\tau} \| f - \Pi_{\tau} f \|^2_{L^2(\tau)} \lesssim 2^{-2js}.
\]

**Lemma:** If \( u \in \mathcal{A}^\infty_s(H^1_0(\Omega)) \) and \( f \in \mathcal{O}^s \) then \( u \in \mathcal{A}^\infty_s(\rho) \).

**Proof:** Overlay of meshes.

**Example:**
\( H^\alpha(\Omega) \subset \mathcal{O}^{1+\alpha} \) for \( \alpha \geq 0 \), so \( \mathcal{A}^\infty_s(H^1_0(\Omega)) \cap \Delta^{-1}(H^{s-1}(\Omega)) \subset \mathcal{A}^\infty_s(\rho) \) for \( s \geq 1 \).
Morally, $\mathcal{O}^s \approx \mathcal{A}^s(H^{-1}(\Omega))$, so we expect $B^\alpha_{q,q}(\Omega) \subset \mathcal{O}^{1+\alpha}$.

**Theorem:** We have $B^\alpha_{q,q}(\Omega) \subset \mathcal{O}^{1+\alpha}$ for $\frac{\alpha}{n} \geq \frac{1}{q} - \frac{1}{2}$, hence $\mathcal{A}^s(H^1_0(\Omega)) \cap \Delta^{-1}(B^{s-1}_{q,q}(\Omega)) \subset \mathcal{A}^s(\rho)$ for $\frac{s-1}{n} \geq \frac{1}{q} - \frac{1}{2}$.
Theorem: We have $B^\alpha_{q,q}(\Omega) \subset \mathcal{O}^{1+\alpha}$ for $\frac{\alpha}{n} \geq \frac{1}{q} - \frac{1}{2}$, hence $\mathcal{A}^s_\infty(H^1_0(\Omega)) \cap \Delta^{-1}(B^{s-1}_{q,q}(\Omega)) \subset \mathcal{A}^s_\infty(\rho)$ for $\frac{s-1}{n} \geq \frac{1}{q} - \frac{1}{2}$.

Proof: The mesh construction part works the same as before. For the direct estimate, with $\delta = \frac{\alpha}{n} - \frac{1}{q} + \frac{1}{2} \geq 0$, we have

$$\|f - \Pi_\tau f\|_{L^2(\tau)} \leq \|f - p\|_{L^2(\tau)} \lesssim |\tau|^\delta \|f - p\|_{L^q(\tau)} + |\tau|^\delta \|f\|_{B^\alpha_{q,q}(\tau)},$$

for any $p \in \mathbb{P}_d$, and

$$\min_{p \in \mathbb{P}_d} \|f - p\|_{L^q(\tau)} \lesssim \omega_{d+1}(f, \tau)_q \lesssim |f|_{B^\alpha_{q,q}(\tau)},$$

which gives

$$\sum_{\tau \in P} h_T^2 \|f - \Pi_\tau f\|_{L^2(\tau)}^2 \lesssim \sum_{\tau \in P} |\tau|^{2\delta + 2/n} |f|_{B^\alpha_{q,q}(\tau)}^2.$$
The arguments can be adapted to
- red refinements,
- splines,
- higher order problems,
- Stokes equations, etc.
- For variable coefficient equations, one loses “an epsilon” in the convergence rate.

Plans:
- inverse theorems for adaptive FEM
- boundary elements
- finite element exterior calculus