Robust multilevel methods

Johannes Kraus

Johann Radon Institute for Computational and Applied Mathematics,
Austrian Academy of Sciences, Linz, Austria

SuperCA++ Workshop
Bansko, April 22-24, 2012

Outline

Outline

Outline

Outline

Scalar elliptic model problem

Let us consider the elliptic boundary value problem

$$-\nabla \cdot (\mathbf{a}(\mathbf{x})\nabla u(\mathbf{x})) = f(\mathbf{x}) \quad \text{in} \quad \Omega, \tag{1a}$$

$$u = 0 \quad \text{on} \quad \Gamma_D, \tag{1b}$$

$$(\mathbf{a}(\mathbf{x})\nabla u(\mathbf{x})) \cdot \mathbf{n} = 0 \quad \text{on} \quad \Gamma_N, \tag{1c}$$

for an unknown function $u(\mathbf{x})$ where Ω is a polygonal domain in $\mathbb{R}^d$, $d = 2, 3$, and $f(\mathbf{x})$ is a given squared Lebesgue integrable function, i.e.,

$$f \in L_2(\Omega) := \{v : v \text{ is defined on } \Omega \text{ and } \int_\Omega v^2 \, d\mathbf{x} < \infty\}.$$

We assume that $\mathbf{a}(\mathbf{x})$ in (1) is SPD and uniformly bounded in Ω, i.e.,

$$c_1 \|\mathbf{v}\|^2 \leq \mathbf{v}^T \mathbf{a}(\mathbf{x}) \, \mathbf{v} \leq c_2 \|\mathbf{v}\|^2 \quad \forall \mathbf{v} \in \mathbb{R}^d, \forall \mathbf{x} \in \Omega,$$

for some positive constants c_1 and c_2, and $\mathbf{n}$ is the outward unit vector normal to the boundary $\Gamma = \partial\Omega$ for which $\Gamma = \bar{\Gamma}_D \cup \bar{\Gamma}_N$ holds.

Conforming finite element (FE) methods

The numerical solution of (1) by a conforming finite element method uses a proper finite-dimensional subspace

$$\mathcal{V}_h := \{v \in C^0(\Omega) : v|_e \in P_r(e) \quad \forall e \in \mathcal{T}_h\} \quad \text{of}$$

$$H^1(\Omega) := \{v : v \text{ is defined on } \Omega \text{ and } \int_\Omega v^2 + \nabla v \cdot \nabla v \, d\mathbf{x} < \infty\}$$

defined for a shape-regular triangulation $\mathcal{T}_h$ of Ω into elements $e \in \mathcal{T}_h$ where $P_r(e) := \{w : w \text{ is a polynomial of degree } \leq r \text{ on } e\} \quad \forall e \in \mathcal{T}_h$.

Conforming FEM: Find $u_h \in \mathcal{V}_h$ where $\mathcal{V}_h \subset \mathcal{V}(= H^1(\Omega))$ such that

$$\mathcal{A}_h(u_h, v_h) = \mathcal{L}_h(v_h) \quad \forall v_h \in \mathcal{V}_h, \tag{2a}$$

$$\mathcal{A}_h(u_h, v_h) \equiv \mathcal{A}(u_h, v_h) := \int_\Omega \mathbf{a}(\mathbf{x}) \nabla u_h(\mathbf{x}) \cdot \nabla v_h(\mathbf{x}) \, d\mathbf{x}, \tag{2b}$$

$$\mathcal{L}_h(v_h) \equiv \mathcal{L}(v_h) := \int_\Omega f(\mathbf{x}) \, v_h(\mathbf{x}) \, d\mathbf{x}. \tag{2c}$$

Nonconforming finite element (FE) methods

Violating the conformity condition $\mathcal{V}_h \subset \mathcal{V}$ the $\mathcal{V}$-norm in general is no longer well-defined on $\mathcal{V}_h$. A remedy is to use mesh-dependent norms in the convergence analysis, e.g., given a partition $\mathcal{T}_h$ of Ω, we define

$$\|v\|_{m,h} := \sqrt{\sum_{e \in \mathcal{T}_h} \|v\|_{m,e}^2}$$

where $\|\cdot\|_{m,e}$ is the induced norm on the space $H^m(e)$.
The bilinear form $\mathcal{A}_h(\cdot,\cdot)$ in the the weak formulation (2) of problem (1) can be defined by

$$\mathcal{A}_h(u_h, v_h) := \sum_{e \in \mathcal{T}_h} \int_e \mathbf{a}(e)\, \nabla u_h(\mathbf{x}) \cdot \nabla v_h(\mathbf{x})\, d\mathbf{x}.$$

Here $\mathbf{a}(e)$ is a piecewise constant SPD matrix, defined by

$$\mathbf{a}(e) = \frac{1}{|e|} \int_e \mathbf{a}(\mathbf{x})\, d\mathbf{x} \quad \forall e \in \mathcal{T}_h.$$

The discrete problem

The general procedure is to consider a basis of the finite element space $\mathcal{V}_h$, which we denote by $\Phi = \{\phi_1, \phi_2, \dots, \phi_N\}$. Then

$$v_h = \sum_{i=1}^{N} v_i \phi_i$$

where the real numbers v_i are the expansion coefficients of v_h. Representing the solution u_h of (2) as $u_h = \sum_{i=1}^{N} u_i \phi_i$ it can easily be seen that (2) is equivalent to

$$\sum_{i=1}^{N} \mathcal{A}_h(\phi_i, \phi_j) u_i = \mathcal{L}_h(\phi_j), \quad j = 1, 2, \dots, N,$$

which in matrix form reads as

$$A\mathbf{u} = \mathbf{b}.$$

Here $\mathbf{u} = (u_i) \in \mathbb{R}^N$ is the vector of unknowns, and the right-hand side vector $\mathbf{b} = (b_j) \in \mathbb{R}^N$ is defined by $b_j = \mathcal{L}_h(\phi_j) \quad 1 \leq j \leq N$.

Why preconditioning?

PDE on Ω $\rightarrow$ PDE discretizations $\rightarrow$ $A\mathbf{u} = \mathbf{b}$

- A is large, sparse, positive definite, ill-conditioned ($\kappa(A) = O(h^{-2})$)

Why preconditioning?

PDE on Ω $\rightarrow$ PDE discretizations $\rightarrow$ $A\mathbf{u} = \mathbf{b}$

- A is large, sparse, positive definite, ill-conditioned ($\kappa(A) = O(h^{-2})$)

Solve algebraic linear systems $A\mathbf{u} = \mathbf{b}$:

- Direct methods (Gaussian elimination....LU, LDL^T):
  - ▷ Very robust but caution: Computational cost is $\mathcal{O}(N^\alpha)$, $\alpha > 1$, $N \approx 10^6$ to 10^9 in practice.
- Iterative methods $\checkmark$
- Multilevel preconditioning

Why preconditioning?

PDE on Ω $\rightarrow$ PDE discretizations $\rightarrow$ $A\mathbf{u} = \mathbf{b}$

- A is large, sparse, positive definite, ill-conditioned ($\kappa(A) = O(h^{-2})$)

Solve algebraic linear systems $A\mathbf{u} = \mathbf{b}$:

- Direct methods (Gaussian elimination....LU, LDL^T):
  - ▷ Very robust but caution: Computational cost is $\mathcal{O}(N^{\alpha})$, $\alpha > 1$, $N \approx 10^6$ to 10^9 in practice.
- Iterative methods $\checkmark$
- Multilevel preconditioning

Goal: Develop uniformly convergent iterative methods for $A\mathbf{u} = \mathbf{b}$

Algebraic multilevel iteration

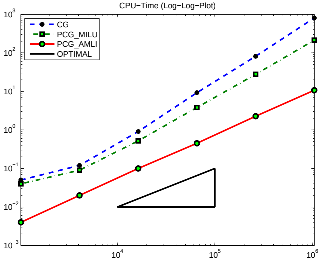
- Idea: Approximate block factorization $+$ stabilization
- Other related methods: Domain decomposition, (algebraic) multigrid, ...

Preconditioned conjugate gradient (PCG) method

Convergence rate estimate:

$$\|\mathbf{u} - \mathbf{u}_k\|_A \leq 2 \left(\frac{\sqrt{\kappa(BA)} - 1}{\sqrt{\kappa(BA)} + 1} \right)^k \|\mathbf{u} - \mathbf{u}_0\|_A$$

$$n_{\mathrm{it}} \leq \frac{1}{2}\sqrt{\kappa(BA)}\ln(2/\epsilon) + 1$$



Solution time (and n_{it}) for Poisson equation on a unit square

$h^{-2} \approx N$	DIRECT	CG	PCG-MILU	PCG-AMLI-V	PCG-AMLI-W
1024	0.02	0.05 (84)	0.04 (21)	$<$ 0.01 (16)	$<$ 0.01 (16)
4096	0.17	0.12 (163)	0.09 (30)	0.02 (18)	0.02 (17)
16384	2.21	0.91 (320)	0.52 (46)	0.09 (22)	0.09 (17)
65536	30.8	9.2 (630)	3.8 (68)	0.49 (25)	0.45 (17)
262144	–	81.6 (1256)	27.8 (102)	2.7 (28)	2.3 (17)
1048576	–	805 (2439)	214 (152)	13.3 (31)	10.5 (17)
		$\kappa = \mathcal{O}(h^{-2})$	$\kappa = \mathcal{O}(h^{-1})$		$\kappa = \mathcal{O}(1)$

Goals

There are many important classes of problems for which an efficient solution of the arising linear systems is a challenging task, e.g.,

- anisotropic problems
- linear systems obtained from nonconforming FE discretization
- problems with highly varying coefficients (multiscale problems)
- discontinuous Galerkin (DG) finite element systems
- linear elasticity problems for nearly incompressible materials
- problems with a large near null space, e.g., in $\mathcal{V} = H(\Omega, \mathrm{div})$
- etc.

We will discuss robust multilevel preconditioner for anisotropic problems, nonconforming FEM, and problems with highly varying coefficients.

Outline

Two-level hierarchical basis

Let $\mathcal{T}_H$ be a coarse mesh and $\mathcal{T}_h$ be a fine mesh obtained from regular refinement of $\mathcal{T}_H$. Consider the two-level hierarchical basis

$$\{\widetilde{\phi}_h^{(k)}, k = 1, 2, \cdots, N_h\} := \{\phi_H^{(l)} \text{ on } \mathcal{T}_H\} \cup \{\phi_h^{(m)} \text{ on } \mathcal{T}_h \backslash \mathcal{T}_H\} \tag{3}$$

and let us denote by $\widetilde{A}_h$ and by A_h the two-level hierarchical and nodal basis stiffness matrix, respectively. Under the splitting (3) we have

$$\begin{aligned} A_h &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{matrix} \}N_{h\backslash H} \\ \}N_H \end{matrix} \\ \widetilde{A}_h &= \begin{bmatrix} \widetilde{A}_{11} & \widetilde{A}_{12} \\ \widetilde{A}_{21} & \widetilde{A}_{22} \end{bmatrix} = \begin{bmatrix} A_{11} & \widetilde{A}_{12} \\ \widetilde{A}_{21} & A_H \end{bmatrix} \begin{matrix} \}N_{h\backslash H} \\ \}N_H \end{matrix} \quad . \end{aligned}$$

where the transformation which relates the nodal point vectors for the standard and for the hierarchical basis is given by

$$J_h = \begin{bmatrix} I & J_{12} \\ 0 & I \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix} = J_h \begin{bmatrix} \widetilde{\mathbf{v}}_1 \\ \widetilde{\mathbf{v}}_2 \end{bmatrix}, \qquad \begin{matrix} \mathbf{v}_1 = \widetilde{\mathbf{v}}_1 + J_{12}\widetilde{\mathbf{v}}_2 \\ \mathbf{v}_2 = \widetilde{\mathbf{v}}_2 \end{matrix} \quad .$$

Multiplicative and additive two-level preconditioners

Starting point is the exact two-by-two block factorization of $\widetilde{A}_h$ (or A):

$$A = \begin{bmatrix} I & \\ A_{21}A_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} A_{11} & \\ & S \end{bmatrix} \begin{bmatrix} I & A_{11}^{-1}A_{12} \\ & I \end{bmatrix}.$$

Then we define the two-level preconditioners B_{mul} and B_{add} by

$$B_{\mathrm{mul}}^{-1} := \begin{bmatrix} I & \\ A_{21}B_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} B_{11} & \\ & B_{22} \end{bmatrix} \begin{bmatrix} I & B_{11}^{-1}A_{12} \\ & I \end{bmatrix}$$

$$B_{\mathrm{add}}^{-1} := \begin{bmatrix} B_{11} & \\ & B_{22} \end{bmatrix}$$

where B_{11} and B_{22} approximate A_{11} and the exact Schur complement $S = A_{22} - A_{21}A_{11}^{-1}A_{12}$. We assume in an SPSD sense ($0 < \alpha_1, \ldots, \tilde{\beta}_2$):

$$\begin{array}{rclll} \alpha_1 A_{11} \le & B_{11} & \le \beta_1 A_{11} & & \text{for } B_{\mathrm{add}} \& B_{\mathrm{mul}} \\ \alpha_2 A_{22} \le & B_{22} & \le \beta_2 A_{22} & & \text{for } B_{\mathrm{add}} \\ \tilde{\alpha}_2 A_{22} \le & B_{22} + A_{21}B_{11}^{-1}A_{12} & \le \tilde{\beta}_2 A_{22} & & \text{for } B_{\mathrm{mul}} \end{array}$$

Condition number estimates

In the simple case when $B_{11} = A_{11}$ and $B_{22} = A_{22}$ we have:

Theorem 1 ([Ax-94]):

Let $\mathbf{w} = \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix}$ be a block-vector which is consistent with the two-by-two representation of A. Further let $B_{ii} = A_{ii}$, $i = 1, 2$. Then

$$(1-\gamma)\mathbf{w}^T B_{\mathrm{add}}^{-1}\mathbf{w} \le \mathbf{w}^T A\mathbf{w} \le (1+\gamma)\mathbf{w}^T B_{\mathrm{add}}^{-1}\mathbf{w}.$$

The CBS constant can be defined as the minimal real number γ satisfying the strengthened Cauchy-Bunyakowski-Schwarz inequality

$$|\mathbf{v}_1^T A_{12}\mathbf{v}_2| \le \gamma \left\{\mathbf{v}_1^T A_{11}\mathbf{v}_1 \ \mathbf{v}_2^T A_{22}\mathbf{v}_2\right\}^{1/2}.$$

O. Axelsson: *Iterative Solution Methods*. Cambridge University Press, 1994.

Condition number estimates, cont.

Theorem 2 ([Ax-94]):

The following estimates hold for the multiplicative two-level preconditioner with $B_{ii} = A_{ii}$, $i = 1, 2$:

$$(1 - \gamma^2)\mathbf{w}^T B_{\mathrm{mul}}^{-1} \mathbf{w} \le \mathbf{w}^T A \mathbf{w} \le \mathbf{w}^T B_{\mathrm{mul}}^{-1} \mathbf{w}.$$

Corollary 1:

Let $B_{ii} = A_{ii}$, $i = 1, 2$. Then the following estimates for $\varkappa_{\mathrm{add}}$ and $\varkappa_{\mathrm{mul}}$ hold:

$$\begin{aligned} \varkappa_{\mathrm{add}} := \kappa\,(B_{\mathrm{add}} A) &\le \frac{1+\gamma}{1-\gamma} = \frac{(1+\gamma)^2}{1-\gamma^2}, \\ \varkappa_{\mathrm{mul}} := \kappa\,(B_{\mathrm{mul}} A) &\le \frac{1}{1-\gamma^2}. \end{aligned}$$

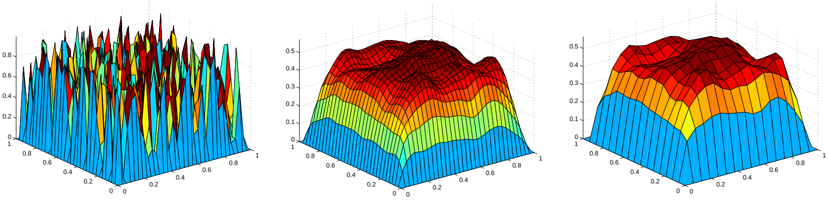
- In general, when $B_{ii} \approx A_{ii}$ the estimates of $\varkappa_{\mathrm{add}}$ and $\varkappa_{\mathrm{mul}}$ additionally depend on the spectral equivalence constants in $B_{ii} \approx A_{ii}$.

Outline

Motivation: The effect of smoothing

Classical stationary iterative methods, which are based on updating a current iterate at a node based on the values of the iterate at neighboring nodes, reduce the highly oscillatory error components fast.

The resulting smooth error can be represented accurately using fewer degrees of freedom, i.e., on a coarse grid.



Random initial error (left) after 5 Gauß-Seidel iterations (middle) represented on coarse mesh (right picture)

Construction of linear AMLI

Consider a sequence of meshes $\mathcal{T}_0, \mathcal{T}_1, \ldots, \mathcal{T}_\ell$ obtained by regular refinement of a coarsest mesh $\mathcal{T}_0$ and let

$$A^{(k)} = \begin{bmatrix} A_{11}^{(k)} & A_{12}^{(k)} \\ A_{21}^{(k)} & A_{22}^{(k)} \end{bmatrix}$$

denote the corresponding two-by-two block partitioned (hierarchical) matrix at any given level $k = 1, 2, \ldots, \ell$.

The AMLI-cycle multigrid preconditioner $B^{(k)}$ at level k is defined by

$$B^{(k)} := \tilde{R}^{(k)} + (I - (R^{(k)})^T A^{(k)}) P^{(k)} B_\nu^{(k-1)} (P^{(k)})^T (I - A^{(k)} R^{(k)})$$

where

- $\tilde{R}^{(k)} = R^{(k)}$ or $\tilde{R}^{(k)} = R^{(k)} + (R^{(k)})^T - (R^{(k)})^T A^{(k)} R^{(k)}$,
- $B_\nu^{(k-1)} = B^{(k-1)} q^{(k)}(A^{(k-1)} B^{(k-1)}) = q^{(k)}(B^{(k-1)} A^{(k-1)}) B^{(k-1)}$,

$R^{(k)}$ is a so-called smoothing iteration (e.g., Richardson, Gauss-Seidel, ILU), $P^{(k)}$ is an interpolation operator, and $q^{(k)}$ is a properly chosen polynomial.

Construction of linear AMLI, cont.

The (original) multiplicative (linear) AMLI preconditioner (as proposed in [AxVa-89]) satisfies the following recurrence relation:

- $B^{(0)} := (A^{(0)})^{-1}$,
- $B^{(k)} := {L^{(k)}}^T D^{(k)} L^{(k)}$,

$$L^{(k)} := \begin{bmatrix} I & 0 \\ -A_{21}^{(k)} {B_{11}^{(k)}}^{-1} & I \end{bmatrix}, \quad D^{(k)} := \begin{bmatrix} {B_{11}^{(k)}}^{-1} & 0 \\ 0 & B_\nu^{(k-1)} \end{bmatrix},$$

where

$$B_\nu^{(k-1)} := \left(I - p^{(k)}(B^{(k-1)} A^{(k-1)})\right) \left(A^{(k-1)}\right)^{-1}$$

and $p^{(k)}(t)$ is a properly scaled and shifted Chebyshev polynomial of degree ν_k satisfying $p^{(k)}(0) = 1$.

O. Axelsson and P. Vassilevski: Algebraic multilevel preconditioning methods I. *Numer. Math.*, **56**, 157-177 (1989).

Analysis of linear AMLI

Let us now derive a bound for the condition number $\kappa(B^{(\ell)}A^{(\ell)})$. This estimate is based on the approximation property of the two-level preconditioner $\bar{B}^{(k)}$, i.e.,

$$\theta_0^{(k)} \mathbf{v}^T \bar{B}^{(k)} \mathbf{v} \le \mathbf{v}^T {A^{(k)}}^{-1} \mathbf{v} \le \theta_1^{(k)} \mathbf{v}^T \bar{B}^{(k)} \mathbf{v} \quad \forall \mathbf{v}, \quad k = 1, \ldots, \ell \qquad (4)$$

where $\quad \bar{B}^{(k)} = {L^{(k)}}^T \bar{D}^{(k)} L^{(k)}$,

$$L^{(k)} = \begin{bmatrix} I & \\ -A_{21}^{(k)} {B_{11}^{(k)}}^{-1} & I \end{bmatrix}, \quad \bar{D}^{(k)} = \begin{bmatrix} {B_{11}^{(k)}}^{-1} & \\ & {A^{(k-1)}}^{-1} \end{bmatrix}.$$

Let us further assume that

$$0 < \theta_0 < \theta_0^{(k)} \le \theta_1^{(k)} < \theta_1 < \infty \quad \text{for all } k = 1, \ldots, \ell. \qquad (5)$$

Then the following theorem provides an estimate of the (relative) condition number of the multilevel preconditioner $B^{(\ell)}$.

Analysis of linear AMLI, cont.

Theorem 3:

Assume that the approximation property (4) holds true for the two-level preconditioner on all levels $j = 1, 2, \ldots, \ell$, i.e., condition (5) is satisfied for some positive constants θ_0 and θ_1.
Further, let $0 < \rho_0 < \rho_1 < \infty$ be positive constants satsifying

$$\rho_0 \leq \frac{\theta_0}{\max\{1, r_1\}} \leq \frac{\theta_1}{\min\{1, r_0\}} \leq \rho_1 \tag{6}$$

where

$$\begin{aligned} r_0 &= \min_k \min_{x \in [\rho_1{}^{-1}, \rho_0{}^{-1}]} x\, q^{(k)}(x), \\ r_1 &= \max_k \max_{x \in [\rho_1{}^{-1}, \rho_0{}^{-1}]} x\, q^{(k)}(x). \end{aligned}$$

Then the estimate

$$\kappa(B^{(\ell)} A^{(\ell)}) \leq \frac{\rho_1}{\rho_0} \tag{7}$$

holds uniformly (in the number of levels ℓ).

Outline

Construction of nonlinear AMLI

Now we define the following nonlinear AMLI-cycle preconditioner: $B^{(k)}[\cdot] : \mathbb{R}^{N_k} \mapsto \mathbb{R}^{N_k}$ for $1 \leq k \leq \ell$ by

$$B^{(k)}[\mathbf{y}] := L^{(k)^T} D^{(k)}[L^{(k)}\mathbf{y}],$$

where

$$L^{(k)} := \begin{bmatrix} I & 0 \\ -A_{21}^{(k)} B_{11}^{(k)^{-1}} & I \end{bmatrix},$$

and

$$D^{(k)}[\mathbf{z}] = \begin{bmatrix} B_{11}^{(k)^{-1}} \mathbf{z}_1 \\ B_{\nu}^{(k-1)}[\mathbf{z}_2] \end{bmatrix}.$$

J. Kraus: An algebraic preconditioning method for M-matrices: Linear versus nonlinear multilevel iteration. *Numer. Lin. Alg. Appl.*, **9**, 599-618 (2002).

Construction of nonlinear AMLI, cont.

The (nonlinear) mapping $B_\nu^{(k-1)}[\cdot]$ is defined by

$$\begin{aligned} B_\nu^{(k-1)}[\cdot] &= {A^{(0)}}^{-1} \quad \text{if } k = 1, \\ B_\nu^{(k-1)}[\cdot] &:= B^{(k-1)}[\cdot] \quad \text{if } \nu = 1 \text{ and } k \geq 1, \end{aligned}$$

and for $k > 1$ and $\nu > 1$

$$B_\nu^{(k-1)}[\mathbf{d}] := \mathbf{x}_{(\nu)}$$

where $\mathbf{x}_{(\nu)}$ is the ν-th iterate obtained when applying the generalized conjugate gradient (GCG) algorithm to the linear system $A^{(k-1)}\mathbf{x} = \mathbf{d}$ thereby using $B^{(k-1)}[\cdot]$ as a preconditioner and starting with the initial guess $\mathbf{x}_{(0)} = \mathbf{0}$.

Outline

Optimality conditions

If $B_{11}^{(k)} = A_{11}^{(k)}$ at all levels $k = 1, 2, \ldots, \ell$ the linear AMLI W-cycle ($\nu = 2$) becomes a uniform preconditioner if

$$\sqrt{\vartheta} = \frac{1}{\sqrt{1-\gamma^2}} < 2 = \nu.$$

In fact it has been shown in [AxVa-90] under the milder assumption

$$A_{11}^{(k)} \leq B_{11}^{(k)} \leq (1+\delta_1) A_{11}^{(k)}$$

that the condition number of the linear AMLI preconditioner is uniformly bounded (independent of the mesh size h and the number of levels ℓ) if ν and γ satisfy the relation

$$\frac{1}{\sqrt{1-\gamma^2}} < \nu.$$

O. Axelsson and P. Vassilevski: Algebraic multilevel preconditioning methods II.
SIAM J. Numer. Anal., **27**, 1569-1590 (1990).

Optimality conditions, cont.

In order to keep the number of arithmetic operations of each PCG (or outer GCG) iteration at level ℓ at a total cost of order $O(N) = O(N_\ell)$ the polynomial degree (or equivalently the number of inner iterations) ν at the coarse levels has to be smaller than the coarsening factor ϱ, i.e.,

$$\nu < \varrho \approx \frac{N_k}{N_{k-1}}$$

which is typically $\varrho = 4$ for 2D problems and $\varrho = 8$ for 3D problems. This results in the following optimality conditions

$$\frac{1}{\sqrt{1-\gamma^2}} < \nu < \varrho$$

for the multiplicative AMLI and

$$\sqrt{\frac{1+\gamma}{1-\gamma}} < \nu < \varrho$$

for the additive AMLI.

Outline

Model problem

Second-order elliptic boundary value problem:

$$
\begin{aligned}
-\nabla \cdot (a(x)\nabla u(x)) &= f(x) \quad \text{in } \Omega, && \text{(8a)}\\
u &= 0 \quad \text{on } \Gamma_D, && \text{(8b)}\\
(\mathbf{a}(x)\nabla u(x)) \cdot \mathbf{n} &= 0 \quad \text{on } \Gamma_N, && \text{(8c)}
\end{aligned}
$$

Ω ... polygonal domain in $\mathbb{R}^d$, $d = 2, 3$,

$f(x)$ ... source term in $L_2(\Omega)$,

$\Gamma = \partial\Omega$, $\Gamma = \Gamma_D \cup \Gamma_N$ ... boundary of Ω,

$\mathbf{n}$ ... outward normal unit vector on Γ,

$\mathbf{a}(x) = (a_{ij}(x))_{i,j=1}^2$ ... SPD coefficient matrix.

Model problem, cont.

Assumptions:

- quasi-uniform partition $\mathcal{T}_h$ of the domain Ω,
- over each element $e \in \mathcal{T}_h$ the functions $a_{ij}(x)$ are smooth, i.e., $\mathbf{a}(x) \approx \mathbf{a}(e) = \mathbf{a}_e$, where

$$\mathbf{a}_e = \begin{bmatrix} a_{e:11} & a_{e:12} \\ a_{e:21} & a_{e:22} \end{bmatrix} \quad \text{is SPD } \forall e \in \mathcal{T}_h.$$

In particular, we will consider Problem 8 with a diffusion tensor

$$\mathbf{a}(x) = \alpha(x) I = \alpha_e I \quad \forall e \in \mathcal{T}_h$$

where $\alpha_e > 0$ is a scalar quantity that may vary over several orders of magnitude across element interfaces.

After rescaling we may assume that the $\alpha_e \in (0, 1]$ for all $e \in \mathcal{T}_h$.

Outline

Preliminaries and notation

Consider a non-overlapping partition $\mathcal{T} = \mathcal{T}_h$ of Ω into elements $e \in \mathcal{T}$.

Definition of structures.

Let us denote by F a collection (union) of elements e from $\mathcal{T}$ which we shall call a *structure*. Further, let

$$\mathcal{F} = \mathcal{F}_h = \{F = F_i : i = 1, 2, \ldots, n_F\}$$

be a set of structures that covers $\mathcal{T}$, i.e., for all $e \in \mathcal{T}$ there exists a structure $F \in \mathcal{F}$ such that $e \subset F$.

Depending on whether the intersection of all mutually distinct (macro) structures is empty or not, we will refer to $\mathcal{F}$ (or $\mathcal{G}$) either as a *non-overlapping* or as an *overlapping covering* of $\mathcal{T}$ ($\mathcal{F}$).

Preliminaries and notation

Consider a non-overlapping partition $\mathcal{T} = \mathcal{T}_h$ of Ω into elements $e \in \mathcal{T}$.

Definition of macro structures.

Next, by G we denote a collection (union) of structures F from $\mathcal{F}$ which we shall call a *macro structure*. Further, let

$$\mathcal{G} = \mathcal{G}_h = \{G = G_i : \ i = 1, 2, \ldots, n_G\}$$

be a set of macro structures that covers $\mathcal{F}$ (and thus also $\mathcal{T}$). That is, for all $F \in \mathcal{F}$ there exists a macro structure $G \in \mathcal{G}$ such that $F \subset G$.

Depending on whether the intersection of all mutually distinct (macro) structures is empty or not, we will refer to $\mathcal{F}$ (or $\mathcal{G}$) either as a *non-overlapping* or as an *overlapping covering* of $\mathcal{T}$ ($\mathcal{F}$).

Preliminaries and notation, cont.

We can we can write the assembly of the global stiffness matrix $A = A_h$ as

$$A = \sum_{e \in \mathcal{T}_h} R_e^T A_e R_e,$$

where R_e is the restriction operator that restricts a global vector to the element e. Alternatively, we can assemble A from *local* matrices A_F or A_G associated with the coverings $\mathcal{F}$ or $\mathcal{G}$. That is,

$$\begin{aligned} A &= \sum_{F \in \mathcal{F}} R_F^T A_F R_F \ , \\ A &= \sum_{G \in \mathcal{G}} R_G^T A_G R_G \ , \\ A_F &= \sum_{e \subset F} \sigma_{e,F} R_{F \mapsto e}^T A_e R_{F \mapsto e} \ , \\ A_G &= \sum_{F \subset G} \sigma_{F,G} R_{G \mapsto F}^T A_F R_{G \mapsto F} \ . \end{aligned}$$

Preliminaries and notation, cont.

The non-negative scaling factors $\sigma_{e,F}$ and $\sigma_{F,G}$ have to be chosen in such a way that the above assembling properties are satisfied, which implies

$$\sum_{F \supset e} \sigma_{e,F} = 1 \quad \forall e \in \mathcal{T} \,, \qquad \sum_{G \supset F} \sigma_{F,G} = 1 \quad \forall F \in \mathcal{F} \,.$$

For instance, the scaling factors can be chosen as follows:

$$\sigma_{e,F} := \frac{1}{\sum_{F' \supset e} 1} \quad \forall e \subset F \,, \qquad \sigma_{F,G} := \frac{1}{\sum_{G' \supset F} 1} \quad \forall F \subset G \,.$$

Then the assembling property transfers from $\mathcal{A}_{\mathcal{F}} := \{A_F : \ F \in \mathcal{F}\}$ to $\mathcal{A}_{\mathcal{G}} := \{A_G : \ G \in \mathcal{G}\}$.

J. Kraus: Additive Schur complement approximation and application to multilevel preconditioning.
RICAM Report **2011-22**, Johann Radon Institute, Linz, 2011.

Algorithm

Assume that we are given:

- a set $\mathcal{D}$ of degrees of freedom (DOF) which is the union of a set of coarse degrees of freedom (CDOF) denoted by $\mathcal{D}_{\mathrm{c}}$ and its complement $\mathcal{D}_{\mathrm{f}} := \mathcal{D} \setminus \mathcal{D}_{\mathrm{c}}$ in $\mathcal{D}$,
- a non-overlapping or an overlapping covering $\mathcal{F}$ of $\mathcal{T}$,
- a set of structure matrices $\mathcal{A}_{\mathcal{F}} := \{A_F : F \in \mathcal{F}\}$ that satisfy the assembling property.

Then we permute the rows and columns of A according to a two-level partitioning of $\mathcal{D}$, i.e.,

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = A_h = \begin{bmatrix} A_{\mathrm{ff}} & A_{\mathrm{fc}} \\ A_{\mathrm{cf}} & A_{\mathrm{cc}} \end{bmatrix} \begin{matrix} \}\ \mathcal{D}_{\mathrm{f}} \\ \}\ \mathcal{D}_{\mathrm{c}} \end{matrix}$$

The corresponding Schur complement we denote by

$$S = S_{\mathrm{c}} = S_H = A_{\mathrm{cc}} - A_{\mathrm{cf}} A_{\mathrm{ff}}^{-1} A_{\mathrm{fc}}.$$

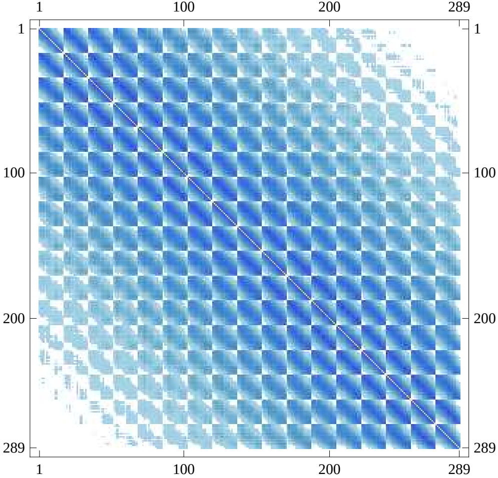
Algorithm, cont.

Additive Schur complement approximation ([Kr-11]):

1. Determine a global two-level numbering according to $\mathcal{D} = \mathcal{D}_{\mathrm{f}} \oplus \mathcal{D}_{\mathrm{c}}$.
2. Determine a covering $\mathcal{G}$ of $\mathcal{F}$ and a set of scaling factors $\{\sigma_{F,G}\}$.
3. For all $G \in \mathcal{G}$ perform the following steps:
   (a) Determine a "local" two-level numbering of the DOF of G.
   (b) Compute
   $$A_G = \begin{array}{cc} \begin{bmatrix} A_{G:\mathrm{ff}} & A_{G:\mathrm{fc}} \\ A_{G:\mathrm{cf}} & A_{G:\mathrm{cc}} \end{bmatrix} & \begin{array}{l} \} \, \mathcal{D}_{G:\mathrm{f}} \\ \} \, \mathcal{D}_{G:\mathrm{c}} \end{array} \end{array}$$
   (c) Compute the "local" Schur complement
   $$S_G = A_{G:\mathrm{cc}} - A_{G:\mathrm{cf}} A_{G:\mathrm{ff}}^{-1} A_{G:\mathrm{fc}}.$$
   (d) Determine the "local-to-global" mapping for the CDOF in $\mathcal{D}_{G:\mathrm{c}}$.
4. Assemble the global Schur complement approximation Q, i.e.,
$$Q = \sum_{G \in \mathcal{G}} R_{G:c}^T S_G R_{G:c}$$

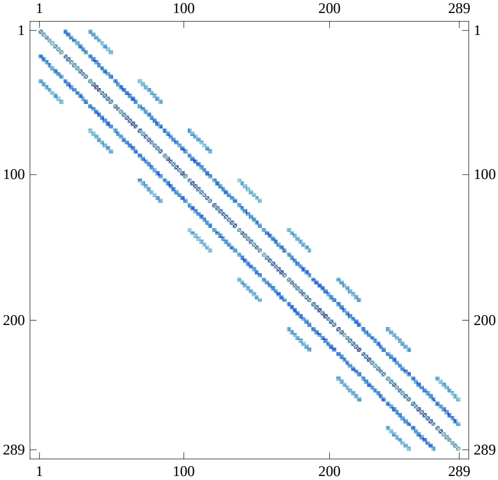
Outline

Nonzero pattern



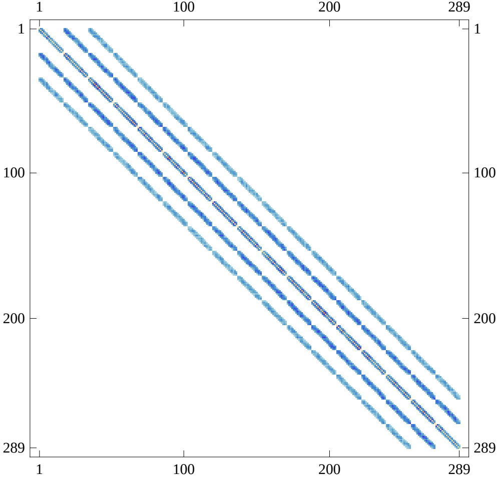
Exact Schur complement

Nonzero pattern



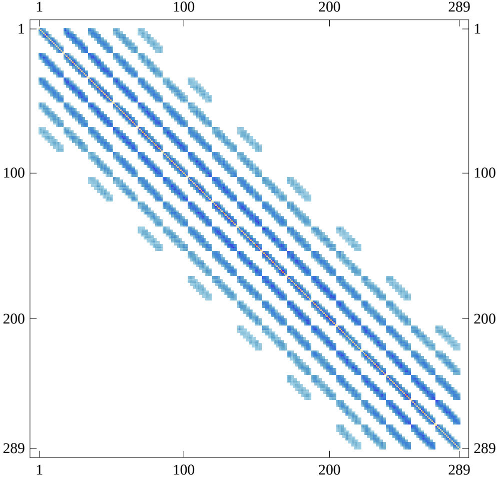
ASCA: Example 1

Nonzero pattern



ASCA: Example 2

Nonzero pattern

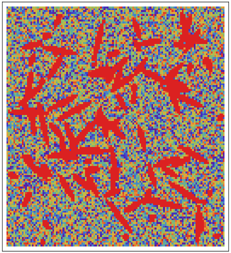


ASCA: Example 3

Multilevel convergence

In this experiment we have islands with a high diffusion coefficient $\alpha_e = \alpha_{\max}$ on a macro scale and inbetween a highly oscillatory coefficient $\alpha_e = 10^{p\ \mathrm{rand}}$ with $p \in \{1, 2, \ldots, q\}$, i.e., $\alpha_{\max}/\alpha_{\min} \approx 10^q$, see Figure.

Number of iterations k_0 for residual reduction by 10^8



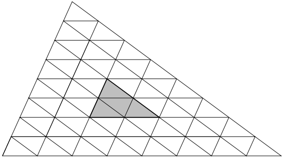
W-cycle	AMLI convergence: Example 3				
$B_{11} = B_{11}^{\mathrm{MILUE}}$	$\ell = 2$	$\ell = 3$	$\ell = 4$	$\ell = 5$	$\ell = 6$
$q = 1$	6	6	6	6	6
$q = 2$	7	7	7	7	7
$q = 3$	7	7	7	7	7
$q = 4$	8	7	8	8	8
$q = 5$	8	8	8	8	8
$q = 6$	8	8	8	8	8

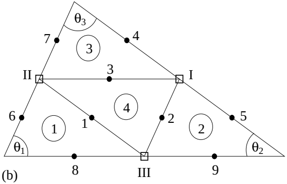
Outline

Crouzeix-Raviart (CR) finite elements:

For the nonconforming CR finite element, where the nodal basis functions are defined at the face centers, the natural vector spaces $\mathcal{V}_H(E) := \operatorname{span}\{\phi_I, \phi_{II}, \phi_{III}\}$ and $\mathcal{V}_h(E) := \operatorname{span}\{\phi_i\}_{i=1}^9$ are no longer nested, i.e. $\mathcal{V}_H(E) \nsubseteq \mathcal{V}_h(E)$.

This makes the direct construction with $\mathcal{V}_2(E) := \mathcal{V}_H(E)$ impossible.


(a)



(b)

The hierarchical basis functions have to be chosen s.t. the resulting subspaces $\mathcal{V}_1(E)$ and $\mathcal{V}_2(E)$ satisfy the direct sum condition

$$\mathcal{V} = \mathcal{V}_1 \oplus \mathcal{V}_2$$

Two-level splitting: General form

Let $\Phi_E := \{\phi^{(i)}\}_{i=1}^{9}$ denotes the set of the "midpoint" basis functions of the four congruent elements in the macro-element E. Then we define

$$\mathcal{V}_1(E) := \mathrm{span}\,\{\phi_1,\, \phi_2,\, \phi_3,\, \phi_4 - \phi_5,\, \phi_6 - \phi_7,\, \phi_8 - \phi_9\}\,,$$
$$\mathcal{V}_2(E) := \mathrm{span}\,\{\phi_1^C + \phi_4 + \phi_5,\, \phi_2^C + \phi_6 + \phi_7,\, \phi_3^C + \phi_8 + \phi_9\}\,,$$

where $\phi_i^C := \sum_k c_{ik}\phi_k$ with $i, k \in \{1, 2, 3\}$.
The transformation matrix corresponding to this general splitting is given by

$$J_E = J_E(C) = \begin{bmatrix} I_3 & 0 & C \\ 0 & J_- & J_+ \end{bmatrix} \quad (\in \mathbb{R}^{9\times 9}),$$

$$J_- := \frac{1}{2}\begin{bmatrix} 1 & -1 & & & & \\ & & 1 & -1 & & \\ & & & & 1 & -1 \end{bmatrix}^T, \quad J_+ := \frac{1}{2}\begin{bmatrix} 1 & 1 & & & & \\ & & 1 & 1 & & \\ & & & & 1 & 1 \end{bmatrix}^T$$

$$\tilde{\phi}_E := (\tilde{\phi}^{(i)})_{i=1}^{9} = J_E^T\,\phi_E, \quad \tilde{A}_E = J_E^T A_E J_E = \begin{bmatrix} \tilde{A}_{E:11} & \tilde{A}_{E:12} \\ \tilde{A}_{E:12}^T & \tilde{A}_{E:22} \end{bmatrix} \begin{matrix} \}\in \mathcal{V}_1(E) \\ \}\in \mathcal{V}_2(E) \end{matrix}\,,$$

$$\tilde{A}_h := \sum_{E\in\mathcal{T}_H} R_E^T \tilde{A}_E R_E\,.$$

Specific two-level splittings

- The transformation for the standard DA splitting, cf. [BlMaNe-04], is obtained for the choice $C = \frac{1}{2}I$.
- A generalized DA splitting (GDA), cf. [KrMaSy-08], is obtained for $C = \frac{1}{2}I + \mu(\mathbb{1} - 3I)$ where the optimal choice of $\mu \in [0, \frac{1}{4}]$ depends on the minimum angle in the triangular mesh. Here $\mathbb{1}$ denotes the 3×3 matrix of all ones.

R. Blaheta, S. Margenov, and M. Neytcheva: Uniform estimate of the constant in the strengthened CBS inequality for anisotropic non-conforming FEM systems.

Numer. Lin. Alg. Appl., **11**(4), 309-326, (2004).

J. Kraus, J. Synka, and S. Margenov: On the multilevel preconditioning of Crouzeix-Raviart elliptic problems.

Numer. Lin. Alg. Appl., **15**, 395-416, (2008).

First reduce (FR) splitting

- The first-reduce splitting (FR) is obtained for $C = -A_{E:11}^{-1}\bar{A}_{E:13}$ where

$$\bar{A}_E := J_\pm^T A_E J_\pm = \begin{bmatrix} \bar{A}_{E:11} & \bar{A}_{E:12} & \bar{A}_{E:13} \\ \bar{A}_{E:12}^T & \bar{A}_{E:22} & \bar{A}_{E:23} \\ \bar{A}_{E:13}^T & \bar{A}_{E:23}^T & \bar{A}_{E:33} \end{bmatrix} \begin{array}{l} \left.\vphantom{\begin{matrix} a \\ a \end{matrix}}\right\} \mathcal{V}_1(E) \\ \left.\vphantom{a}\right\} \mathcal{V}_2(E) \end{array}$$

$$\text{and} \quad J_\pm := \begin{bmatrix} I & 0 & 0 \\ 0 & J_- & J_+ \end{bmatrix}.$$

The name “First Reduce” (FR) expresses that the (global) stiffness matrix is first reduced (via static condensation) to a system with its Schur complement

$$B = \begin{bmatrix} \bar{A}_{22} & \bar{A}_{23} \\ \bar{A}_{23}^T & \bar{A}_{33} \end{bmatrix} - \begin{bmatrix} \bar{A}_{12}^T \\ \bar{A}_{13}^T \end{bmatrix} A_{11}^{-1} \begin{bmatrix} \bar{A}_{12}, & \bar{A}_{13} \end{bmatrix}.$$

This can be written equivalently as a new (combined) transformation with

$$J_{\mathrm{FR}} := J_\pm J_B = \begin{bmatrix} I & -A_{11}^{-1}\bar{A}_{12} & -A_{11}^{-1}\bar{A}_{13} \\ 0 & J_- & J_+ \end{bmatrix} \quad \Rightarrow \quad C = -A_{11}^{-1}\bar{A}_{13}.$$

Outline

Local analysis of the CBS constant

Theorem 5 ([KrMaSy-08]):

Let $\tilde{A}_{E:33}(C)$ be the macro-element submatrix associated with the coarse grid. Then, the minimum value for γ_E is attained for $C := -A_{E:11}^{-1}\bar{A}_{E:13}$.

Theorem 6 ([KrMaSy-08]):

Consider the Crouzeix-Raviart finite element discretization of the elliptic model problem.
Then, comparing the DA, the GDA, and the FR splitting we have the relation

$$\gamma_{\mathrm{FR}}^2 \leq \gamma_{\mathrm{GDA}}^2 \leq \gamma_{\mathrm{DA}}^2 \leq 3/4. \tag{9}$$

Outline

Additive and multiplicative preconditioners

Theorem 7 ([BlMaNe-05]):

The following statements hold for any element size and shape and any coefficient anisotropy:

(a) If C_{11} is the additive preconditioner to B_{11} then

$$\kappa\left(C_{11}^{-1}B_{11}\right) < \frac{1}{4}(11+\sqrt{105}). \tag{10}$$

(b) If C_{11} is the multiplicative preconditioner to B_{11} then

$$\kappa\left(C_{11}^{-1}B_{11}\right) < \frac{15}{8}. \tag{11}$$

(c) The cost of the application of the preconditioner in both cases is proportional to the number of unknowns.

R. Blaheta, S. Margenov, and M. Neytcheva: Robust optimal multilevel preconditioners for non-conforming finite element systems.
Numer. Lin. Alg. Appl., **12**, 495-514, (2005).

Outline

Nonconforming linear finite elements

We consider the elliptic model problem on the unit square $\Omega = (0,1)\times(0,1)$ for the coefficient matrix

$$\mathbf{a}(\mathbf{x}) = \alpha_e \begin{bmatrix} \varepsilon & -\delta \\ -\delta & 1 \end{bmatrix}, \quad \forall e \in \mathcal{T}$$

Setting of the experiment:

- homogeneous Dirichlet boundary conditions,
- ratio $1:10$ of anisotropy, i.e., $\varepsilon = 10^{-1}$,
- δ is varied from 0 to $1/4$,
- $\bar{\Omega} = \bar{\Omega}_1 \cup \ldots \cup \bar{\Omega}_4$, where $\bar{\Omega}_1 = [0,1/2]^2$, $\bar{\Omega}_2 = [1/2,1] \times [0,1/2]$, $\bar{\Omega}_3 = [0,1/2] \times [1/2,1]$, $\bar{\Omega}_4 = [1/2,1]^2$
- jump of two orders of magnitude in the coefficient α_e at interfaces of subdomains $\bar{\Omega}_i$,
- uniform mesh with mesh size $h \in \{1/64, \ldots, 1/1024\}$,
- stopping criterion $\|\mathbf{r}^{(n_{\mathrm{it}})}\|/\|\mathbf{r}^{(0)}\| \leq 10^{-6}$.

Nonconforming linear finite elements, cont.

Nonlinear AMLI W-cycle: Number of iterations and CPU-time t in sec.

$1/h$	64		128		256		512		1024	
	n_{it}	t	n_{it}	t	n_{it}	t	n_{it}	t	n_{it}	t
	$\delta = 0$									
DA	17	0.12	18	0.71	19	3.58	19	16.5	19	75.6
GDA	13	0.13	13	0.70	13	3.49	13	17.0	13	80.2
FR	10	0.10	10	0.51	10	2.57	10	12.5	10	59.0
	$\delta = 1/8$									
DA	20	0.16	21	0.94	22	4.86	21	22.5	22	106
GDA	13	0.13	13	0.74	13	3.72	13	17.9	13	82.3
FR	12	0.11	12	0.63	12	3.17	11	14.3	11	66.2
	$\delta = 1/4$									
DA	19	0.15	20	0.91	20	4.50	20	21.4	20	98.5
GDA	16	0.15	16	0.85	16	4.31	16	20.9	16	97.1
FR	14	0.12	14	0.69	14	3.47	14	17.0	14	78.7

Nonconforming linear finite elements, cont.

Comparison of multiplicative and additive AMLI W-cycle (FR)

$\frac{1}{h}$	128	256	512	1024	2048
	multiplicative (additive) preconditioner				
	$\varepsilon = 1$, $\delta = 0$, no jump				
n_{it}	9 (17)	9 (17)	9 (17)	9 (17)	9 (17)
ρ	0.18 (0.43)	0.18 (0.44)	0.18 (0.44)	0.18 (0.44)	0.18 (0.44)
t	0.09 (0.13)	0.37 (0.56)	1.75 (2.63)	8.54 (12.5)	40.2 (58.1)
	$\varepsilon = 0.001$, $\delta = 0.01$, jump: three orders of magnitude				
n_{it}	6 (14)	6 (16)	6 (16)	6 (17)	6 (17)
ρ	0.09 (0.36)	0.10 (0.40)	0.10 (0.41)	0.10 (0.43)	0.10 (0.43)
t	0.08 (0.12)	0.35 (0.59)	1.62 (2.86)	7.77 (14.2)	36.6 (66.4)

Conclusions:

We

- described (optimal order) linear and nonlinear AMLI methods
- summarized convergence results
- introduced techniques for sparse Schur complement approximations
- showed extensions of the AMLI framework to nonconforming FEM
- discussed numerical results for elliptic problems
  - with highly oscillatory coefficients
  - with anisotropic coefficients

References:

J. Kraus and S. Margenov: Robust Algebraic Multilevel Methods and Algorithms. *Radon Series Comp. Appl. Math., vol. 5, Walter de Gruyter*, 2009.

P. Vassilevski: Multilevel Block Factorization Preconditioners. *Springer-Verlag*, 2008.

Recent and ongoing work:

Recent and ongoing collaborations on related topics:

- J. Kraus: Additive Schur complement approximation and application to multilevel preconditioning. *RICAM Report* **2011-22**, Johann Radon Institute, Linz, 2011.
- J. Kraus, M. Lymbery, S. Margenov: Robust multilevel methods for quadratic finite element anisotropic elliptic problems. Submitted, 2012.
- J. Brannick, Y. Chen, J. Kraus, L. Zikatanov: Algebraic multilevel preconditioners for the graph Laplacian based on matching in graphs. Submitted, 2012.
- J. Kraus, P. Vassilevski, L. Zikatanov: On using the polynomial of best uniform approximation to $1/x$ in two-level methods. To be submitted, 2012.
- J. Kraus, M. Lymbery, S. Margenov: Multilevel preconditioning of bicubic FEM elasticity systems. Work in progress, 2012.

THANK YOU FOR YOUR ATTENTION!