Optimal Left and Right
Additive Schwarz Preconditioning for
Minimal Residual Methods with
Euclidean and Energy Norms

Marcus Sarkis
Worcester Polytechnic Inst., Mass. and IMPA, Rio de Janeiro

and

Daniel B. Szyld
Temple University, Philadelphia

NA Day, Bologna, 18 September 2006
• We want to solve certain PDEs (non-selfadjoint or indefinite elliptic) discretized by FEM (or divided differences)

• Use GMRES (or other Krylov subspace method)

• Precondition with Additive Schwarz (with coarse grid correction)

• Schwarz methods optimality (energy norm) and Minimal Residuals Methods (usually minimizing the 2-norm)

• Comments on left vs. right preconditioning
Examples

- Helmholtz equation \(-\Delta u + cu = f\)
- Advection diffusion equation \(-\Delta u + b.\nabla u + cu = f\)
- zero Dirichlet b.c.
General Problem Statement

Solve

\[ Bx = f \]

\( B \) non-Hermitian, discretization of \( b(u, v) = f(v) \)

\[ b(u, v) = a(u, v) + s(u, v) + c(u, v), \]

\[ a(u, v) = \int_{\Omega} \nabla u \cdot \nabla v \, dx, \]

\[ s(u, v) = \int_{\Omega} (b \cdot \nabla u)v + (\nabla \cdot bu)v \, dx, \quad b \in \mathbb{R}^d, \]

\[ c(u, v) = \int_{\Omega} cuv \, dx, \quad \text{and} \quad f(v) = \int_{\Omega} f v \, dx. \]
General Problem Statement (cont.)

\[ b(u, v) = a(u, v) + s(u, v) + c(u, v), \]
\[ a(u, v) = \int_\Omega \nabla u \cdot \nabla v \, dx, \]

- Let \( A \) be SPD, the discretization of \( a(\cdot, \cdot) \).

- So, we have \( B = A + R \).

- Here \( A = A^T \succ O \) is the discretization of \( \Delta u \).
Standard Finite Element Setting

Let $\Omega \subset \mathbb{R}^d$, with triangulation $\mathcal{T}_h(\Omega)$. Let $V$ be the traditional finite element space formed by piecewise linear and continuous functions vanishing on the boundary of $\Omega$. $V \subset \mathcal{H}^1_0(\Omega)$.

One-to-one correspondence between functions in finite element space and nodal values.

We abuse the notation and do not distinguish between them.

Let $\|v\|_a = a(v, v)$, and $\|v\|_A = (v^T A v)^{1/2}$ be the corresponding norms in $V$ and in $\mathbb{R}^n$, respectively.
Problem Statement (cont.)

• Use Krylov subspace iterative methods (e.g., GMRES)

• Left preconditioning: \( M^{-1}Bx = M^{-1}f \)

• Right preconditioning: \( BM^{-1}(Mu) = f \)
Krylov subspace methods: GMRES

\[ \mathcal{K}_m(B, r_0) = \text{span}\{r_0, Br_0, B^2r_0, \ldots, B^{m-1}r_0\}. \]

Given \( x_0, r_0 = f - Bx_0 \), find approximation

\[ x_m \in x_0 + \mathcal{K}_m(B, r_0), \quad x_m = x_0 + p(B)r_0 \]

(\( p(t) \) polynomial degree \( m - 1 \)) satisfying

\[ x_m = \arg \min \{\|f - Bx\|_2\}, \quad x \in x_0 + \mathcal{K}_m(B, r_0) \]
Krylov subspace methods: GMRES (cont.)
Preconditioned case, Implementation

Let $v_1, v_2, \ldots, v_m$ be an orthonormal basis of $\mathcal{K}_m(M^{-1}B, r_0) = \text{span}\{r_0, M^{-1}Br_0, (M^{-1}B)^2r_0, \ldots, (M^{-1}B)^{m-1}r_0\}$.

$$x_m = \arg \min \{\|f - M^{-1}Bx\|_2\}, \quad x \in x_0 + \mathcal{K}_m(M^{-1}B, r_0)$$

- With $V_m = [v_1, v_2, \ldots, v_m]$, obtain Arnoldi relation:

$$M^{-1}BV_m = V_{m+1}H_{m+1,m}$$

$H_{m+1,m}$ is $(m + 1) \times m$ upper Hessenberg

- Element in $\mathcal{K}_m(M^{-1}B, v_1)$ is a linear combination of $v_1, v_2, \ldots, v_m$, i.e., of the form $V_my, \ y \in \mathbb{R}^m$
GMRES (cont.)
Preconditioned case, Implementation

- Find $y = y_m$ and we have $x_m = x_0 + V_m y_m$

$$\|M^{-1} f - M^{-1} B x\|_2 = \|M^{-1} r_0 - M^{-1} B V_m y\|_2 =$$

$$= \|V_{m+1} \beta e_1 - V_{m+1} \bar{H}_m y\|_2 = \|\beta e_1 - \bar{H}_m y\|_2$$

find $y$ using QR factorization of $\bar{H}_m$. 
One convergence bound for GMRES [Elman 1982]

(unpreconditioned version)

\[ \|r_m\| = \|f - Bx_m\| \leq \left( 1 - \frac{c^2}{C^2} \right)^{m/2} \|r_0\|, \]

where

\[ c = \min_{x \neq 0} \frac{(x, Bx)}{(x, x)} \quad \text{and} \quad C = \max_{x \neq 0} \frac{\|Bx\|}{\|x\|}. \]
Schwarz Preconditioning

Class of Preconditioners based on Domain Decomposition

Decomposition of $V$ into a sum of $N + 1$ subspaces $R^T_i V_i \subset V$, and

$$V = R^T_0 V_0 + R^T_1 V_1 + \cdots + R^T_N V_N.$$ 

$R^T_i : V_i \rightarrow V$ extension operator from $V_i$ to $V$. This decomposition usually NOT a direct sum.

Subspaces $R^T_i V_i$, $i = 1, \ldots, N$ are related to a decomposition of the domain $\Omega$ into overlapping subregions $\Omega^\delta_i$ of size $O(H)$ covering $\Omega$. The subspace $R^T_0 V_0$ is the coarse space.
Schwarz Preconditioning (cont.)

For \( u_i, v_i \in V_i \) define

\[
 b_i(u_i, v_i) = b(R_i^T u_i, R_i^T v_i), \quad a_i(u_i, v_i) = a(R_i^T u_i, R_i^T v_i).
\]

Let

\[
 B_i = R_i B R_i^T, \quad A_i = R_i A R_i^T
\]

be the matrix representations of these local bilinear forms, i.e., the local problems.

\( R_i \) is a restriction operator, \( R_i^T \) is a prolongation (embedding) operator.
Two versions of Additive Schwarz Preconditioning here

\[ M^{-1} = R_0^T B_0^{-1} R_0 + \sum_{i=1}^{p} R_i^T B_i^{-1} R_i, \]

or \[ M^{-1} = R_0^T B_0^{-1} R_0 + \sum_{i=1}^{p} R_i^T A_i^{-1} R_i, \]

where \( B_i = R_i B R_i^T \) and \( A_i = R_i A R_i^T \) (local problems)

\( R_i \) restriction, \( R_i^T \) prolongation with overlap \( \delta \)

\( B_0 \) coarse problem, size \( O(H) \), discretization \( O(h) \).
Let \( P = M^{-1}B \), be the preconditioned problem.

**Theorem.** [Cai and Widlund, 1993]

There exist constants \( H_0 > 0 \), \( c(H_0) > 0 \), and \( C(H_0) > 0 \), such that if \( H \leq H_0 \), then for \( i = 1, 2 \), and \( u \in V \),

\[
\frac{a(u, Pu)}{a(u, u)} \geq c_p,
\]

and

\[
\|Pu\|_a \leq C_p \|u\|_a,
\]

where \( C_p = C(H_0) \) and \( c_p = C_0^{-2} c(H_0) \).
Two-level Schwarz preconditioners are optimal in the sense that bounds for $M^{-1}B$ (or $BM^{-1}$) are independent of the mesh size and the number of subdomains, or slowly varying with them.

In our PDEs, we have optimal bounds:

$$\frac{(x, M^{-1}Bx)_A}{(x, x)_A} \geq c_p \quad \text{and} \quad \|M^{-1}Bx\|_A \leq C_p\|x\|_A.$$ 

Cai and Zou [NLAA, 2002] observed:
Schwarz bounds use energy norms, while GMRES minimizes $l_2$ norms. Optimality may be lost!
In fact, Cai and Zou [NLAA, 2002] showed is that for Additive Schwarz $M^{-1}B$ is \textbf{NOT} positive real, i.e., there is no $c > 0$ for which

$$\frac{(x, M^{-1}Bx)}{(x, x)} \geq c.$$

Thus, this GMRES bound cannot be used in this case. We may not have the optimality.
Krylov Subspace Methods with Energy Norms

Proposed solution: Use GMRES minimizing the $A$-norm of the residual.

[Note: many authors mention this, e.g., Ashby-Manteuffel-Saylor, Essai, Greenbaum, Gutknecht, Weiss, ... ]

In this case, we have that $M^{-1}B$ is positive real with respect to the $A$-inner product since

$$
\frac{(x, M^{-1}Bx)_A}{(x, x)_A} \geq c_p \quad \text{and} \quad \|M^{-1}Bx\|_A \leq C_p \|x\|_A.
$$
Rework convergence bound for GMRES [Elman 1982] (preconditioned version)

\[ \| r_m \|_A = \| M^{-1} f - M^{-1} Bx_m \|_A \leq \left( 1 - \frac{c^2}{C^2} \right)^{m/2} \| M^{-1} r_0 \|_A , \]

where

\[ c = \min_{x \neq 0} \frac{(x, M^{-1} Bx)_A}{(x, x)_A} \quad \text{and} \quad C = \max_{x \neq 0} \frac{\| Bx \|_A}{\| x \|_A} . \]
Implementation:
Replace each inner product \((x, y)\) with \((x, y)_A = x^T A y\).
Only one matvec with \(A\) needed. Basis vectors are \(A\)-orthonormal.
Arnoldi relation: \(M^{-1} B V_m = V_{m+1} H_{m+1}\).

\[
\|M^{-1} r_0 - M^{-1} B V_m y\|_A = \|V_{m+1} \beta e_1 - V_{m+1} \bar{H}_m y\|_A = \|\beta e_1 - \bar{H}_m y\|_2
\]

Same QR factorization of \(\bar{H}_m\), same code for the minimization.

We use this for analysis, but sometimes also valid for computations.
Left vs. Right preconditioner

For right preconditioner $BM^{-1}u = f$, $M^{-1}u = x$.

$$(x, x)_A = (M^{-1}u, M^{-1}u)_A = (u, u)_G, \quad G = M^{-T}AM^{-1}.$$ 

- Every left preconditioned system $M^{-1}Bx = M^{-1}f$ with the $A$ norm is completely equivalent to a right preconditioned system with the $M^{-T}AM^{-1}$-norm.

$$\|r_0 - BM^{-1}Z_m y\|_{M^{-T}AM^{-1}} = \|M^{-1}r_0 - M^{-1}BM^{-1}Z_m y\|_A = \|\beta z_1 - M^{-1}BV m y\|_A = \|\beta e_1 - \bar{H}_m y\|_2.$$ 

$Z_m$ has the $G$-orthogonal basis of $\mathcal{K}_m(r_0, BM^{-1})$
Left vs. Right preconditioner

• Converse also holds: for every right preconditioner $M$ with $S$-norm, this is equivalent to left preconditioning with $M$ using the $M^TSM$-norm. (True in particular for $S = I$)

• When using the same inner product (norm), left and right preconditioning produce different upper Hessenberg matrices $H_m$.

• When using $A$-inner product for left preconditioning and $M^{-T}AM^{-1}$-inner product for right preconditioning, we have the same upper Hessenberg matrices $H_m$.

• The experiments we show with left preconditioning and $A$-norm minimization are the same as with right preconditioning with $G$-norm minimization, $G = M^{-T}AM^{-1}$.  

22
Energy Norms vs. $\ell_2$ Norm

Now, we “have” the optimality with energy norms. What can we say about the $\ell_2$ norm? Use equivalence of norms:

$$\|x\|_2 \leq \frac{1}{\sqrt{\lambda_{\text{min}}(A)}} \|x\|_A, \quad \|x\|_A \leq \sqrt{\lambda_{\text{max}}(A)} \|x\|_2$$

$$\|M^{-1}r^L_m\|_2 \leq \|M^{-1}r^A_m\|_2 \leq \frac{1}{\sqrt{\lambda_{\text{min}}(A)}} \|M^{-1}r^A_m\|_A$$

$$\leq \frac{1}{\sqrt{\lambda_{\text{min}}(A)}} \left(1 - \frac{c^2}{C^2}\right)^{m/2} \|M^{-1}r_0\|_A$$

$$\leq \frac{\sqrt{\lambda_{\text{max}}(A)}}{\sqrt{\lambda_{\text{min}}(A)}} \left(1 - \frac{c^2}{C^2}\right)^{m/2} \|M^{-1}r_0\|_2$$

$$= \sqrt{\kappa(A)} \left(1 - \frac{c^2}{C^2}\right)^{m/2} \|M^{-1}r_0\|_2$$
“Asymptotic” Optimality of $\ell_2$ Norm

$$\|M^{-1}r_m^L\|_2 \leq \sqrt{\kappa(A)} \left(1 - \frac{C^2}{C^2}\right)^{m/2} \|M^{-1}r_0\|_2$$

For a fixed mesh size $h$, Additive Schwarz preconditioned GMRES (2-norm) has a bound that goes to zero at the same speed as the optimal bound (energy norm), except for a factor $\sqrt{\kappa(A)}$ (which of course depends on $h$)
Numerical Experiments

• Helmholtz equation \(-\Delta u + cu = f, c = -5\ or \ c = -120.\)

• Advection diffusion equation \(-\Delta u + b.\nabla u + cu = f\)
  \(b^T = [10, 20], c = 1,\) upwind finite differences

• both on unit square, zero Dirichlet b.c., \(f \equiv 1\)

• Discretization: \(64 \times 64 \ (n = 3969),\)
  \(128 \times 128 \ (n = 16129),\) or \(256 \times 256 \ (n = 65025)\) nodes
  \(p = 4 \times 4\ or \ p = 8 \times 8\ subdomains\)
  Overlap: \(0, 1, 2 \ (1,3\ or 5\ lines\ of\ nodes)\)

• Tolerance \(\varepsilon = 10^{-8}\)
Figure 1: Helmholtz equation $c = -5$. GMRES minimizing the $\ell_2$ norm (o), and the $G$-norm (*). $64 \times 64$ grids, $4 \times 4$ subdomains. $\delta = 0$ Left: $G$-norm of both residuals. Right: $\ell_2$ norm of both residuals.
Figure 2: Helmholtz equation $c = -120$. GMRES minimizing the $\ell_2$ norm (o), and the $G$-norm (*). 128 $\times$ 128 grids, 8 $\times$ 8 subdomains. $\delta = 1$ Left: $G$-norm of both residuals. Right: $\ell_2$ norm of both residuals.
Figure 3: Helmholtz equation $c = -120$. GMRES minimizing the $\ell_2$ norm (o), and the $G$-norm (*). $256 \times 256$ grids, $8 \times 8$ subdomains. $\delta = 0$ Left: $G$-norm of both residuals. Right: $\ell_2$ norm of both residuals.
Figure 4: Advection-diffusion equation. GMRES minimizing the $\ell_2$ norm (o), and the $G$-norm (*). 128 × 128 grids, 4 × 4 subdomains. $\delta = 2$

Left: $G$-norm of both residuals. Right: $\ell_2$ norm of both residuals.
Figure 5: Advection-diffusion equation. GMRES minimizing the $\ell_2$ norm (o), and the $G$-norm (*). 256 × 256 grids, 8 × 8 subdomains. $\delta = 0$
Left: $G$-norm of both residuals. Right: $\ell_2$ norm of both residuals.
Conclusions

- GMRES in energy norm maintains optimality
- GMRES in $\ell_2$ norm achieves “asymptotic” optimality
- Observations on left vs. right preconditioning
- Numerical experiments illustrate this
Paper to appear in *CMAME*

available at  [http://www.math.temple.edu/szyld](http://www.math.temple.edu/szyld)