Multigrid absolute value preconditioning

Eugene Vecharynski[1] **Andrew Knyazev**[2] (speaker)

[1]Department of Computer Science and Engineering
University of Minnesota

[2]Department of Mathematical and Statistical Sciences
University of Colorado Denver

FIFTEENTH COPPER MOUNTAIN CONFERENCE
ON MULTIGRID METHODS



Acknowledgements

- Department of Mathematical and Statistical Sciences
  University of Colorado Denver
- Lynn Bateman Memorial Fellowship
- NSF DMS 0612751
- Copper Multigrid 2011 Organizing Committee

The results presented here are partially based on the PhD thesis of the first co-author, defended at the University of Colorado Denver, under the supervision of the second co-author, in December 2010.

Outline

- Brief intro: iterative methods, SPD and non-SPD preconditioning
- An ideal SPD preconditioner for a symmetric indefinite linear system
- **Absolute value preconditioning**. Definition
- Absolute value preconditioners for linear systems with strictly (block) diagonally dominant matrices
- MG absolute value preconditioner for a model problem. Numerical examples
- Conclusions

Symmetric indefinite linear systems

We consider a nonsingular linear system

$$Ax = b, \; A = A^* \in \mathbb{R}^{n \times n}.$$

Several origins of the problem

- Mixed finite element discretizations of PDEs in fluid and solid mechanics, acoustics
- Inner steps of interior point methods in linear and nonlinear optimization
- Solution of the correction equation in the Jacobi-Davidson method for a symmetric eigenvalue problem

General setting

- Large problem size, ill-conditioned
- Sparse matrices or matrix-free environment
- Direct methods are inapplicable. Iterate!
- Use of preconditioners to improve the convergence

How to solve a symmetric indefinite linear system?

What iterative method shall we use to approximate the solution of

$$Ax = b, \; A = A^*?$$

Let T be a *preconditioner*. Consider the *preconditioned* linear system

$$TAx = Tb.$$

- T is *symmetric indefinite* (e.g., $T \approx A^{-1}$) or *nonsymmetric*.
  $\Rightarrow$ TA is generally *nonsymmetric* in any inner product.
  The symmetry is lost: no short recurrence and complicated convergence properties. Possible solution techniques: **GMRES, BiCG, QMR**, etc.
- T is *symmetric positive definite (SPD)*.
  $\Rightarrow$ TA is *symmetric* in the T^{-1}-based inner product; $(x, y)_{T^{-1}} = (x, T^{-1}y)$.
  The symmetry is preserved:
  - **optimal short-term recurrent schemes, e.g., PMINRES**
  - PMINRES convergence speed is guaranteed by the positive and negative spectrum of TA

The Preconditioned Minimal Residual method

Let T be an SPD preconditioner.
The Preconditioned Minimal Residual method, at step i, constructs an approximation $x^{(i)}$ to the solution of system $Ax = b$ of the form

$$x^{(i)} \in x^{(0)} + \mathcal{K}_i\left(TA, Tr^{(0)}\right),$$

such that the residual vector $r^{(i)} = b - Ax^{(i)}$ satisfies the optimality condition

$$\|r^{(i)}\|_T = \min_{u \in A\mathcal{K}_i\left(TA, Tr^{(0)}\right)} \|r^{(0)} - u\|_T,$$

where $\mathcal{K}_i\left(TA, Tr^{(0)}\right) = \text{span}\left\{Tr^{(0)}, (TA)Tr^{(0)}, \ldots, (TA)^{i-1}Tr^{(0)}\right\}$ is the i-dimensional preconditioned Krylov subspace, $x^{(0)}$ is the initial guess.

Stable implementation: The PMINRES algorithm (Paige, Saunders, 1975).
Question: How do we define the SPD preconditioner T?

Matrix absolute value

For a given matrix A, its polar decomposition is $A = U|A|$, where $|A| = \sqrt{A^*A}$ and U is unitary. Let A be real indefinite symmetric and nonsingular, than the matrix absolute value $|A|$ is also nonsingular and U is the matrix sign of A, having only two distinct eigenvalues ± 1. Given the eigenvalue decomposition, $A = V\Lambda V^*$, where V is an orthogonal matrix of eigenvectors and $\Lambda = \mathrm{diag}\{\lambda_j\}$ is a diagonal matrix of the corresponding eigenvalues of A, we can compute

- the **matrix absolute value** of A as $|A| = V\,|\Lambda|\,V^*, \;\; |\Lambda| = \mathrm{diag}\{|\lambda_j|\}$.
- the **matrix sign** of A as

$$\mathrm{sign}(A) = V\mathrm{sign}(\Lambda)V^*, \; \mathrm{sign}(\Lambda) = \mathrm{diag}\left\{\mathrm{sign}(\lambda_j)\right\}.$$

The **polar decomposition** of a symmetric matrix can be written as

$$A = |A|\,\mathrm{sign}(A) = \mathrm{sign}(A)\,|A|\,.$$

Inverse of the matrix absolute value as an ideal SPD preconditioner

Let $T = |A|^{-1}$. The preconditioned linear system $TAx = Tb$ is

$$\text{sign}(A)x = |A|^{-1} b.$$

The matrix $TA = \text{sign}(A)$ has only two distinct eigenvalues: -1 and 1.

$\Rightarrow$ The Preconditioned Minimal Residual method converges to the exact solution in at most two steps (cannot go any quicker!).

- $T = |A|^{-1}$ is an *ideal* SPD preconditioner for a symmetric indefinite linear system
- Construction of the *exact* $|A|^{-1}$ is generally prohibitively expensive
- Construct T to attain a relatively small norm $\left\| T - |A|^{-1} \right\|$. Can, in principle, be done, by approximating the action of a *matrix function* $f(A) = |A|^{-1}$ on a vector using A-based Krylov subspaces. Typically still too costly.

Absolute value preconditioning

Our idea: Construct a *practical* SPD preconditioner T as *spectrally equivalent* to the ideal preconditioner $|A|^{-1}$. Let us define $\delta_1 \geq \delta_0 > 0$ as

$$\delta_0(v, T^{-1}v) \leq (v, |A|\, v) \leq \delta_1(v, T^{-1}v), \ \forall v \in \mathbb{R}^n,$$

where A is the nonsingular symmetric indefinite coefficient matrix for a linear system $Ax = b$, we want to solve. We call T an **absolute value preconditioner** if the ratio $\delta_1/\delta_0 \geq 1$, which bounds the spectral condition number of the matrix $T\,|A|$, is **reasonably small**. For *mesh problems*, the ratio is independent of the mesh size. It does not have to be close to one!

The ratio $\delta_1/\delta_0 \geq 1$ measures the quality of the **absolute value preconditioner** T in terms of the convergence speed of the Preconditioned Minimal Residual method. At the same time, the costs of the construction and application of T should preferably be **similar** to the costs of the matrix-vector multiplication of the coefficient matrix A.

Spectrally equivalent absolute value preconditioning

Theorem

Let us be given a symmetric indefinite $A \in \mathbb{R}^{n \times n}$, an SPD $T \in \mathbb{R}^{n \times n}$, and constants $\delta_1 \geq \delta_0 > 0$, such that

$$\delta_0 (v, T^{-1} v) \leq (v, |A| \, v) \leq \delta_1 (v, T^{-1} v), \ \forall v \in \mathbb{R}^n.$$

Then all the eigenvalues of TA are located in the union of two intervals

$$[-\delta_1, -\delta_0] \bigcup [\delta_0, \delta_1] \, .$$

Interestingly, the converse does not hold!

Is the idea of absolute value preconditioning crazy enough to be practical?

- Remember, that neither $|A|^{-1}$, nor $|A|$ are available to us.
- How do we construct efficient absolute value preconditioning?

Absolute value preconditioners for strictly (block) diagonally dominant matrices

Matrix $A = \{A_{ij}\} \in \mathbb{R}^{n \times n}$, $i, j = 1, \ldots, s$, is strictly block diagonally dominant if

$$\left(\|A_{ii}^{-1}\|\right)^{-1} > \sum_{\substack{j=1 \\ j \neq i}}^{s} \|A_{ij}\|, \; i = 1, \ldots, s.$$

Theorem

Let A be a strictly block diagonally dominant symmetric indefinite matrix, such that

$$\delta \left(\|A_{ii}^{-1}\|\right)^{-1} \geq \sum_{\substack{j=1 \\ j \neq i}}^{s} \|A_{ij}\|, \; i = 1, \ldots, s,$$

for a fixed $\delta \in [0, 1)$. Let $T = diag\left\{|A_{11}|^{-1}, |A_{22}|^{-1}, \ldots, |A_{ss}|^{-1}\right\}$. Then all the eigenvalues of the matrix TA are located in the union of intervals

$$\{y \in \mathbb{R} : |y + 1| \leq \delta\} \bigcup \{y \in \mathbb{R} : |y - 1| \leq \delta\}.$$

Absolute value preconditioning for a model problem

Consider the **"shifted Laplacian" equation on a unit square** with Dirichlet boundary conditions and a relatively small shift value,

$$-\Delta u(x, y) - c^2 u(x, y) = f(x, y), \quad (x, y) \in \Omega = (0, 1) \times (0, 1)$$
$$u|_{\Gamma} = 0.$$

The discretization of the boundary value problem using a standard 5-point FD stencil on a uniform mesh leads to the linear system

$$(L - c^2 I)x = b.$$

- The shifted negative discrete Laplace operator $L - c^2 I$ is symmetric and indefinite. We assume it to be nonsingular
- The preconditioner T is intended to be spectrally equivalent to $\left|L - c^2 I\right|^{-1}$
- Use a geometric MG approach to construct $w = Tr$
- We have no proof, only numerical results

Two-grid absolute value preconditioner for the model problem

Algorithm (Two-grid absolute value preconditioner)

Input r, output $w = Tr$.

1. **Pre-smoothing**. *Apply ν pre-smoothing steps,*
$$w^{(i+1)} = w^{(i)} + M^{-1}(r - Lw^{(i)}),\ i = 0, \ldots, \nu - 1,\ w^{(0)} = 0.$$
This step results in the pre-smoothed vector $w^{pre} = w^{(\nu)}$, $\nu \geq 1$.
2. **Coarse grid correction**. *Restrict $r - Lw^{pre}$ to the coarse grid (R), multiply it by $\left|L_H - c^2 I_H\right|^{-1}$, prolongate to the fine grid (P), and add to w^{pre},*
$$w^{cgc} = w^{pre} + P \left|L_H - c^2 I_H\right|^{-1} R \left(r - Lw^{pre}\right).$$
3. **Post-smoothing**. *Apply ν post-smoothing steps,*
$$w^{(i+1)} = w^{(i)} + M^{-*}(r - Lw^{(i)}),\ i = 0, \ldots, \nu - 1,\ w^{(0)} = w^{cgc}.$$
Return $w = w^{(\nu)}$.

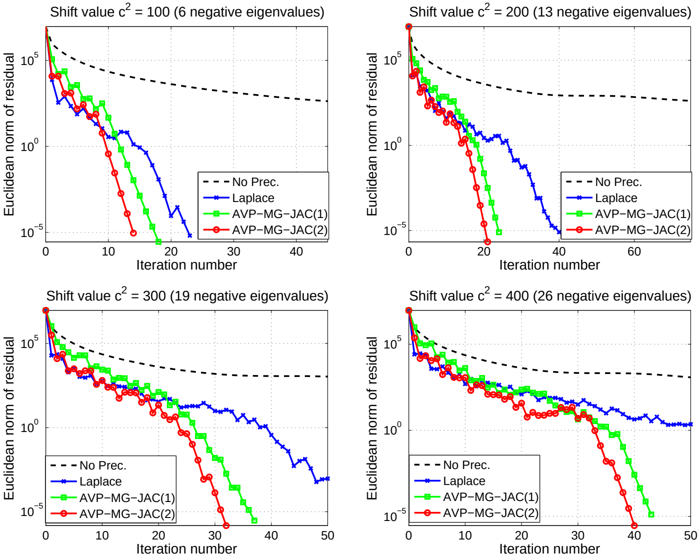
MG absolute value preconditioner for the model problem

- The resulting preconditioner T is SPD under mild assumptions on the smoother, restriction and prolongation
- **In practice, we use the MG extension of the two-grid algorithm ("V-cycle")**
- Compare our MG absolute value preconditioner with the preconditioner based on the inverse of the Laplacian (A. Bayliss, C. Goldstein, E. Turkel, 1983)

In our numerical tests

- Standard coarsening scheme. The coarsest grid is of the mesh size 2^{-4}, and the finest grid is of the mesh size 2^{-7}
- Full weighting for the restriction and piecewise multilinear interpolation for the prolongation
- The smoother: ω-damped Jacobi, $\omega = 4/5$

MG absolute value preconditioner for the model problem: a fixed mesh, different shifts



MG absolute value preconditioner for the model problem: mesh-independent convergence

Number of steps performed to achieve the decrease by the factor 10^{-8} in the error norm.

	$h = 2^{-7}$	$h = 2^{-8}$	$h = 2^{-9}$	$h = 2^{-10}$
$c^2 = 100$	15	14	14	14
$c^2 = 200$	21	21	21	21
$c^2 = 300$	31	32	32	30
$c^2 = 400$	40	39	40	40

Table: Mesh-independent convergence of PMINRES with the MG absolute value preconditioner

MG absolute value preconditioner for the model problem: the influence of the course-grid size

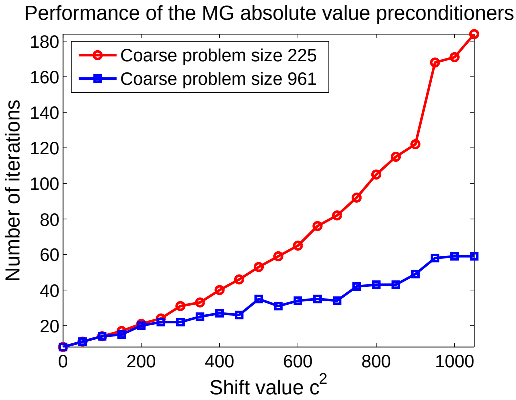


Figure: Performance of the MG absolute value preconditioners for the model problem with different shift values. The problem size $n = (2^7 - 1)^2 \approx 1.6 \times 10^4$. The number of negative eigenvalues varies from 0 to 75. The initial error norm is decreased by 10^{-8}.

Conclusions, current and future work

Conclusions

1. Introduced (SPD) absolute value preconditioning for symmetric indefinite linear systems
2. Constructed several examples of efficient absolute value preconditioning
3. Good results for the "shifted Laplacian" with a relatively small shift value

Current and future work

- Absolute value preconditioning for nonsymmetric linear systems, eigenvalue and singular value problems
- Algebraic multilevel absolute value preconditioning

We are looking for collaborators — experts in practical parallel preconditioning, to try our absolute value preconditioning ideas in established preconditioning software packages.

Thank you!