Preconditioner updates for sequences of sparse, large and nonsymmetric linear systems

Jurjen Duintjer Tebbens

Institute of Computer Science

Academy of Sciences of the Czech Republic

joint work with

Miroslav Tůma

Institute of Computer Science

Academy of Sciences of the Czech Republic

and Technical University in Liberec

GAMM–SIAM Conference on Applied Linear Algebra, Düsseldorf, July 27, 2006.

Outline

1. Motivation
2. The proposed preconditioner updates
3. Numerical experiments
4. Updates in matrix-free environment
5. Conclusions

1. Motivation / Newton's method

Solving systems of nonlinear equations (e.g. from computational fluid dynamics, structural mechanics, numerical optimization, etc)

$$F(x) = 0$$

$$\Downarrow$$

Sequences of linear algebraic systems of the form

$$J(x_k)\Delta x = -F(x_k),\ J(x_k) \approx F'(x_k)$$

solved until for some $k, k = 1, 2, \ldots$ $\quad \|F(x_k)\| < tol$

$J(x_k)$ may change both structurally and numerically

1. Motivation / Cost reduction

We focus on reduction of overall costs by sharing some of the computational effort among the subsequent linear systems. Some options include:

- Freezing approximate Jacobians over a couple of subsequent systems (MNK: Shamanskii, 1967; Brent, 1973).
- Freezing preconditioners over a couple of subsequent systems (MFNK: Knoll, McHugh, 1998).
- Preconditioned iterative Krylov-space solvers often restricted to simple preconditioners from stationary methods (e.g., Jacobi, Gauss-Seidel)
- Physics-based preconditioners: Preconditioning by discretized simpler operators like scaled diffusion operators for convection-diffusion equations; using other physics-based operator splittings (only a selection from huge bibliography: Concus, Golub, 1973; Elman, Schultz, 1986; Brown, Saad, 1990; Knoll, McHugh, 1995; Knoll, Keyes, 2004)

1. Motivation / Updates

2. Preconditioner updates

- Dense updates of ("exact") decompositions (Bartels, Golub, Saunders, 1970; Gill, Golub, Murray, Saunders, 1974), e.g. in the simplex method.
- Specific sparse updates of ("exact") decompositions (Hager, Davis, 1999–2004).
- Preconditioners from quasi-Newton updates, e.g. by low-rank updates (Morales, Nocedal, 2000), (Bergamaschi, Bru, Martínez, Putti, 2006)
- World of recycling of Krylov subspaces (e.g., Morgan 1995-2002); Baglama, Calvetti, Golub, Reichel, 1999; Carpentieri et. al., 2003; de Sturler, 1996; Erhel, Burrage, Pohl, 1996; Duff, Giraud, Langou, Martin, 2005; Giraud et. al. 2004–2005; Parks et al. 2004)
- Approximate diagonal updates of ("incomplete") decompositions (Benzi, Bertaccini, 2003; Bertaccini, 2004) for SPD systems.

2. The proposed preconditioner updates

- We focus on approximate preconditioner updates for nonsymmetric systems solved by arbitrary iterative methods.
- Updating frozen preconditioners for preconditioned iterative methods instead of their recomputation.
- Simple algebraic updates which may be potentionally considered in matrix-free computations.

Notation: Consider two systems

$$Ax = b, \quad A^{+}x = b^{+}; \text{ preconditioned by } M, M^{+}$$

We would like the update M^{+} to become as powerful as M.

2. The proposed preconditioner updates

Some information about the quality of M is given by

- a norm of $A - M$, which expresses accuracy of the preconditioner
- a norm of $I - M^{-1}A$ (or $I - AM^{-1}$), which expresses stability of the preconditioner
- our "ideal" preconditioner satisfies

$$||A - M|| = ||A^+ - M^+||$$

- Clearly,

$$M^+ \equiv M - (A - A^+)$$

  is ideal in this sense, but application of

$$(M^+)^{-1} = (M - (A - A^+))^{-1}$$

  is in general not feasible.

2. The proposed preconditioner updates

Factorized preconditioners for subsequent systems

$$\text{Let} \quad M = LDU \quad \text{where} \quad M \approx A \quad \text{or} \quad M \approx A^{-1}$$

$$\downarrow$$

Possible straightforward approximations of

$$(M^+)^{-1} = (M - (A - A^+))^{-1} \equiv (M - B)^{-1}$$

assuming L and U are not too far from identity:

- First choice
  $(M - B)^{-1} = U^{-1}(D - L^{-1}BU^{-1})^{-1}L^{-1} \approx U^{-1}(D - B)^{-1}L^{-1}$,
  $M^+ = L(\overline{D - B})U$ for $\overline{D - B} \approx D - B$ easily invertible
- Second choice
  $(M - B)^{-1} = (DU - L^{-1}B)^{-1}L^{-1} \approx (DU - B)^{-1}L^{-1}$,
  $M^+ = L(\overline{DU - B})$ for $\overline{DU - B} \approx DU - B$ easily invertible

2. The proposed preconditioner updates

Factorized preconditioners for subsequent systems

Lemma: Let $||A - LDU|| = \varepsilon ||A|| < ||B||$. Then $M^+ = L(\overline{DU - B})$ satisfies

$$||A^+ - M^+|| \leq$$

$$\leq \frac{\|L\| \, \|DU - B - \overline{DU - B}\| + ||L - I|| \, \|B\| + \varepsilon ||A||}{||B|| - \varepsilon ||A||} \cdot ||A^+ - LDU||.$$

- $\overline{DU - B}$ should be close to $DU - B$
- $||L - I||$ should be small
- then we can get $||M^+ - A^+||$ even smaller than $||M - A||$.
- analogue results for norms of $I - A^+(M^+)^{-1}$ (as in Bertaccini, 2004).

2. The proposed preconditioner updates

1. Motivation

Benzi and Bertaccini (2003, 2004) use approximate diagonal updates.
This is motivated by solving equations with a parabolic term:

$$\frac{\partial u}{\partial t} - \Delta u = f,$$

e.g., 2D problem with 2^{nd} order centered differences in space and backward Euler time discretization for grid internal nodes (i, j) and time step $t+1$

$$h^2(u_{ij}^{t+1} - u_{ij}^{t}) + \tau(u_{i+1,j}^{t+1} + u_{i-1,j}^{t+1} + u_{i,j+1}^{t+1} + u_{i,j-1}^{t+1} - 4u_{ij}^{t+1}) = h^2 \tau f_{ij}^{t+1}$$

We get matrices with five diagonals where only diagonal entries may change with time steps $\Rightarrow \overline{DU - B} \equiv \mathrm{diag}(DU - B)$

2. The proposed preconditioner updates

2. Our motivation: Solving nonlinear convection-diffusion problems

$$-\Delta u + u\nabla u = f$$

$$\Downarrow$$

E.g., from the upwind discretization in 2D, with $u \geq 0$ we get for grid internal nodes (i, j)

$$u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1} - 4u_{ij} + hu_{ij}(2u_{ij} - u_{i-1,j} - u_{i,j-1}) = h^2 f_{ij}$$

It is a matrix with five diagonals

Entries of only three diagonals may change in subsequent linear systems

2. The proposed preconditioner updates

2. Solving nonlinear convection-diffusion problems (continued)

$$-\Delta u + u \nabla u = f$$

$$\begin{pmatrix}
* & * & \dots & * & & & & \\
* & * & * & \dots & * & & & \\
\vdots & * & * & * & \dots & * & & \\
* & \vdots & * & * & * & \dots & * & \\
 & * & \vdots & * & * & * & \dots & * \\
 & & * & \vdots & * & * & * & \dots \\
 & & & * & \vdots & * & * & * \\
 & & & & * & \vdots & * & *
\end{pmatrix}$$

2. Solving nonlinear convection-diffusion problems (continued)

$$-\Delta u + u\nabla u = f$$

$$\begin{pmatrix}
\spadesuit & \spadesuit & \dots & \spadesuit & & & & \\
* & \spadesuit & \spadesuit & \dots & \spadesuit & & & \\
\vdots & * & \spadesuit & \spadesuit & \dots & \spadesuit & & \\
* & \vdots & * & \spadesuit & \spadesuit & \dots & \spadesuit & \\
 & * & \vdots & * & \spadesuit & \spadesuit & \dots & \spadesuit \\
 & & * & \vdots & * & \spadesuit & \spadesuit & \dots \\
 & & & * & \vdots & * & \spadesuit & \spadesuit \\
 & & & & * & \vdots & * & \spadesuit
\end{pmatrix}$$

2. The proposed preconditioner updates

Proposed approximation of

$$DU - B:$$

2. The proposed preconditioner updates

Proposed approximation of

$$DU - B:$$

Define $\qquad B = -(L_B + D_B + U_B)$

and put $\qquad \overline{DU - B} \equiv DU - D_B - U_B.$

2. The proposed preconditioner updates

Proposed approximation of

$$DU - B:$$

Define $\qquad B = -(L_B + D_B + U_B)$

and put $\qquad \overline{DU - B} \equiv DU - D_B - U_B.$

Then

$$M^+ = L(DU - D_B - U_B)$$

is for free and its application asks for one forward and one backward solve.

2. The proposed preconditioner updates

Proposed approximation of

$$DU - B :$$

Define $\quad B = -(L_B + D_B + U_B)$

and put $\quad \overline{DU - B} \equiv DU - D_B - U_B.$

Then

$$M^+ = L(DU - D_B - U_B)$$

is for free and its application asks for one forward and one backward solve.

- Ideal for upwind/downwind modifications
- Our experiments cover broader spectrum of problems

2. The proposed preconditioner updates

The triangular update

$$M^+ = L(DU - D_B - U_B)$$

combines an incomplete LU factorization with the structural, Gauss-Seidel type preconditioner

$$DU - B \approx triu(DU - B).$$

We also defined updates incorporating both triangles of $DU - B$ through

- a combination of forward and backward SOR approximation of $DU - B$,
- selection of dominating rows/columns based on Gauss-Jordan transformations,

see Duintjer Tebbens, Tůma, 200?.

3. Experimental results

- Preconditioned BiCGSTAB (but other Krylov space methods show similar behavior)
- Stopped after residual reduction by seven orders of magnitude (but close to linear convergence curves)
- Experiments in Fortran and Matlab
- Various preconditioners and various problems tested
- Black-box update procedure automatically choosing the dominating triangle

3. Experimental results: II.

Example: Numerical simulation of air flow at a low Mach number subject to the gravitational force.

- 2D longitudinal section of a tunnel
- the pressure and density varying only in the horizontal direction
- gravitational term balanced out by the pressure gradient
- von Neumann boundary conditions and Lax-Friedrichs fluxes
- first-order operator splitting
- the implicit Euler method combined with the first order discretization in space

kindly provided by Andreas Meister and Philipp Birken

3. Experimental results: III.

Table 1: *Air flow in a tunnel, n=4800, nnz=138024.*

ILUT$(0.001/5)$, timep=0.05, psize=135798						
A / M	Self prec		Freeze		Update	
	Its	Time	Its	Time	Its	Time
$A^{(10)}$	30	0.50	17	0.27	17	0.27
$A^{(20)}$	32	0.59	19	0.34	27	0.31
$A^{(30)}$	34	0.61	24	0.44	21	0.34
$A^{(40)}$	39	0.67	31	0.52	24	0.39
$A^{(50)}$	40	0.70	39	0.63	24	0.44
$A^{(60)}$	47	0.80	80	1.41	31	0.56
$A^{(65)}$	47	0.75	107	1.64	27	0.42
$A^{(70)}$	38	0.70	72	1.28	28	0.51
$A^{(75)}$	114	1.98	230	4.06	105	1.96
$A^{(80)}$	63	1.14	87	1.51	80	1.42

4. Updates in matrix-free environment: I

Assume we always update with upper triangles, i.e.
$M^+ = L(DU - D_B - U_B)$.

Cost comparison in case of explicitly given matrices:

	computational costs	storage costs
recomputation	factorization	A^+, current L, U
updating	0	A^+ + $triu(A)$, old L, U

Cost comparison in matrix-free environment:

	computational costs	storage costs
recomputation	$est(A^+)$ + factorization	current L, U
updating	$est(triu(A^+))$	$triu(A^+), triu(A)$, old L, U

4. Updates in matrix-free environment: II

- Typically, estimating A^+ with graph colouring techniques is about twice as expensive as estimating $triu(A^+)$
- In the previous example: $est(A^+) : \pm 16$ mvp, $est(triu(A^+)) : \pm 9$ mvp.
- This fact can only be exploited if we know *a priori* which triangle to compute. Therefore:
- Attempt to enhance dominance of one triangle by permutation P of the whole sequence:

$$PA^{(i)}P^Ty = Pb^{(i)}.$$

. . . work in progress, for the moment only weak but inexpensive permutations.

4. Updates in matrix-free environment: III

Example 2: Numerical simulation of a tunnel fire event at a low Mach number.

- 2D longitudinal section of a tunnel
- 10 Mega-Watt fire source
- von Neumann boundary conditions and Lax-Friedrichs fluxes
- first-order operator splitting
- the implicit Euler method combined with the first order discretization in space

4. Updates in matrix-free environment: IV

Table 2: *Tunnel fire event, n=59.392, nnz=1.119.944, ILU(0)*

A / M	Self prec	Freeze	Update	Permuted update
$A^{(5)}$	19	∞	35	35
$A^{(10)}$	31	∞	63	44
$A^{(15)}$	19	∞	37	37
$A^{(20)}$	76	∞	46	45
$A^{(25)}$	83	∞	42	39
$A^{(30)}$	374	∞	59	54
timing	3200 s	∞	465 s	444 s

5. Conclusions

- Nonsymmetric preconditioners in the form of decompositions can be successfully updated by algebraic techniques.

- The techniques seem to be efficient and robust, applied as a black-box.

- Combined with other update techniques for sequences (e.g. recycling of Krylov subspaces) they may lead to powerful solvers.

Thank you for your attention.

Supported by the project “Information Society” of the Academy of Sciences of the Czech Republic under No. 1ET400300415