An Updated Preconditioner for Sequences of General 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

XVII. Householder Symposium, Zeuthen, June 5, 2008.

1. Introduction

- The solution of *sequences* of linear systems arises in numerous applications
- Rather few work has been done in our community on efficient solution of *general* sequences of linear systems
- The central question of such work will be:

  How can we share part of the computational effort throughout the sequence ?

- Below we list some known strategies.

1. Introduction

- Very simple trick: Hot starts, i.e. use the solution of the previous system as initial guess.

- Sometimes exact updating of the factorizations for large problems is feasible: Rank-one updates of LU factorizations have been used since decades in the simplex method where the change of one system matrix to another is restricted to one column [Bartels, Golub, Saunders - 1970; Suhl, Suhl - 1993].

- General rank-one updates of an LU decomposition are discussed in [Stange, Griewank, Bollhoefer - 2005].

- If the linear solver is a Krylov subspace method, strategies to recycle information gained from previously generated Krylov subspaces have shown to be beneficial in many applications [Parks, de Sturler, Mackey, Johnson, Maiti - 2006], [Giraud, Gratton, Martin - 2007], [Frank, Vuik - 2001].

1. Introduction

- When shifted linear systems with identical right hand sides have to be solved, Krylov subspace methods allow advantageous implementations based on the fact that all systems generate the same subspace [Frommer, Glässner - 1998]

- In nonlinear systems solved with a Newton-type method one can skip evaluations of the (approximate) Jacobian during some iterations, leading to Shamanskii's combination of the chord and Newton method [Brent - 1973] $\Rightarrow$ linear solving techniques with multiple right hand sides can be exploited.

- Another option: Allow changing the system matrices but freeze the preconditioner [Knoll, Keyes - 2004].

1. Introduction

To enhance the power of a frozen preconditioner one may also use approximate updates:

- In [Meurant - 2001] we find approximate updates of incomplete Cholesky factorizations and
- in [Benzi, Bertaccini - 2003, 2004] banded updates were proposed for both symmetric positive definite approximate inverse and incomplete Cholesky preconditioners.
- In Quasi-Newton methods the difference between system matrices is of small rank and preconditioners may be efficiently adapted with approximate small-rank updates; this has been done in the symmetric positive definite case, see e.g. [Bergamaschi, Bru, Martínez, Putti - 2006, Nocedal, Morales - 2000].

2. The proposed preconditioner updates

- We focus on a black-box approximate preconditioner update for *general nonsymmetric* systems solved by arbitrary iterative methods.
- Updating frozen preconditioners for preconditioned iterative methods instead of their recomputation.
- Simple algebraic updates which can be considered in matrix-free computations.

Notation: Consider two systems

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

$$\text{let} \qquad B \equiv A - A^+.$$

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

2. The proposed preconditioner updates

If $||A - M||$ is the accuracy of the preconditioner M for A, we will try to find an updated M^+ for A^+ with comparable accuracy,

$$||A - M|| \approx ||A^+ - M^+||.$$

Let M be factorized as $M = LDU$, then the choice

$$M^+ = LDU - B$$

would give $||A - M|| = ||A^+ - M^+||$. We will approximate this ideal update $LDU - B$ in two steps, similarly to the techniques in [Benzi, Bertaccini - 2003, Bertaccini - 2004]. First we use

$$LDU - B = L(DU - L^{-1}B) \approx L(DU - B) \quad \text{or}$$

$$LDU - B = (LD - BU^{-1})U \approx (LD - B)U$$

depending on whether L is closer to identity or U.

2. The proposed preconditioner updates

Define the standard splitting

$$B = L_B + D_B + U_B.$$

Then the second approximation step is

$$LDU - B \approx L(DU - B) \approx L(DU - D_B - U_B) \equiv M^+$$

(upper triangular update) or

$$LDU - B \approx (LD - B)U \approx (LD - L_B - D_B)U \equiv M^+$$

(lower triangular update). Then M^+ 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

As an example consider a two-dimensional nonlinear convection-diffusion model problem: It has the form

$$-\Delta u + Ru\left(\frac{\partial u}{\partial x} + \frac{\partial u}{\partial y}\right) = 2000x(1-x)y(1-y), \tag{1}$$

on the unit square, discretized by 5-point finite differences on a uniform grid.

- The initial approximation is the discretization of $u_0(x, y) = 0$.
- We use here $R = 100$ and different grid sizes.
- We solve the resulting linear systems with BiCGSTAB with right preconditioning.
- Iterations were stopped when the Euclidean norm of the residual was decreased by seven orders.
- Matlab implementation.

2. The proposed preconditioner updates

BiCGSTAB iteration counts; reference factorization ILU(0)			
Matrix	Recomp	Freeze	Triangular update
$A^{(0)}$	40	40	40
$A^{(1)}$	25	37	37
$A^{(2)}$	24	41	27
$A^{(3)}$	20	48	26
$A^{(4)}$	17	56	30
$A^{(5)}$	16	85	32
$A^{(6)}$	15	97	35
$A^{(7)}$	14	106	43
$A^{(8)}$	13	97	44
$A^{(9)}$	13	108	45
$A^{(10)}$	13	94	50
$A^{(11)}$	15	104	45
$A^{(12)}$	13	156	49
overall time	13 s	13 s	7.5 s

2. The proposed preconditioner updates

The upper triangular update

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

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

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

However it does not take into account both triangular parts of $DU - B$. This can be simply corrected through the splitting

$$DU - B = \hat{L} + \hat{D} + \hat{U}$$

and defining the Gauss-Seidel update

$$M^+ = L\,(\hat{L} + \hat{D})\hat{D}^{-1}(\hat{D} + \hat{U})$$

and similarly for lower triangular update [Duintjer Tebbens, Tůma - 2007].

2. The proposed preconditioner updates

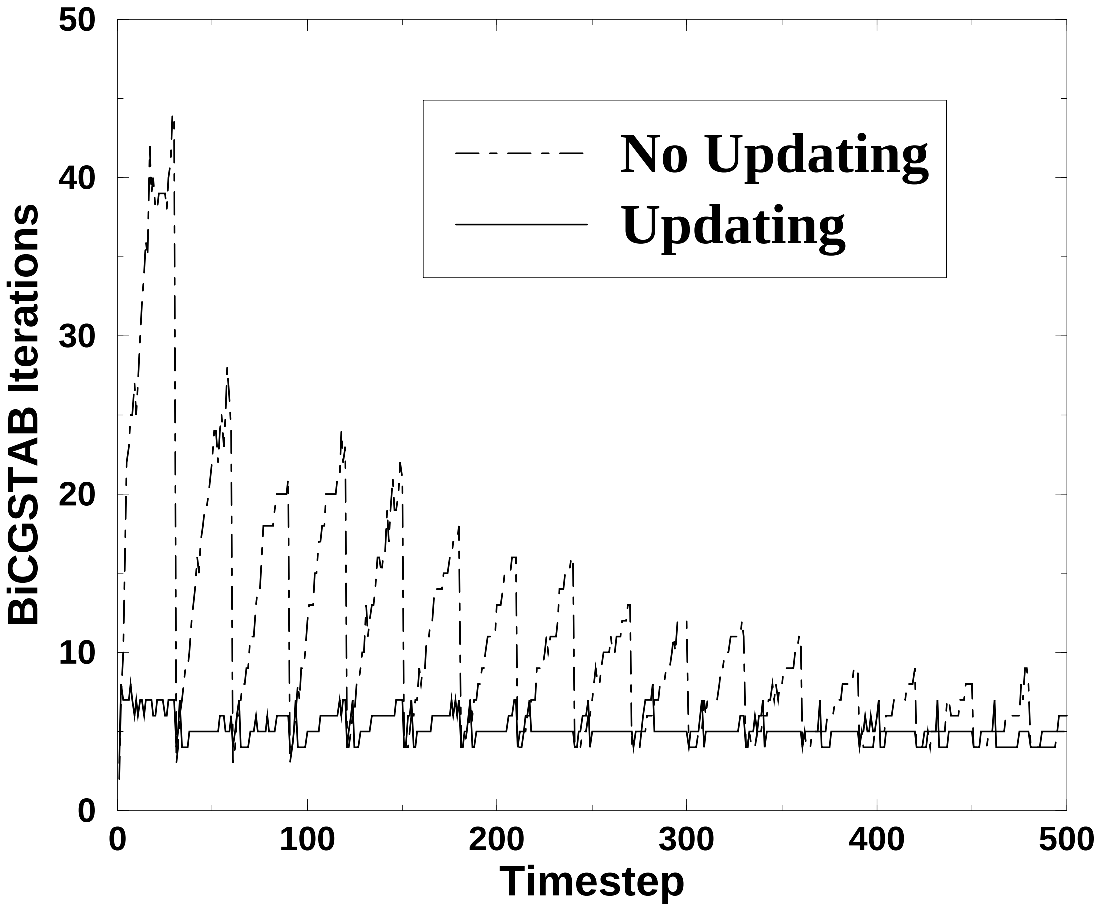
BiCGSTAB iteration counts; reference factorization ILU(0)				
Matrix	Recomp	Freeze	Triangular update	Gauss-Seidel update
$A^{(0)}$	40	40	40	40
$A^{(1)}$	25	37	37	27
$A^{(2)}$	24	41	27	27
$A^{(3)}$	20	48	26	19
$A^{(4)}$	17	56	30	21
$A^{(5)}$	16	85	32	25
$A^{(6)}$	15	97	35	29
$A^{(7)}$	14	106	43	31
$A^{(8)}$	13	97	44	40
$A^{(9)}$	13	108	45	38
$A^{(10)}$	13	94	50	44
$A^{(11)}$	15	104	45	35
$A^{(12)}$	13	156	49	42
overall time	13 s	13 s	7.5 s	6.5 s

2. The proposed preconditioner updates

The following CFD problem (compressible supersonic flow) is a less academic example:

- Frontal flow with Mach-number 10 around a cylinder, which leads to a steady state.
- 3000 steps of the implicit Euler method are performed.
- The grid consists of 20994 points, we use Finite Volume discretization and system matrices are of dimension 83976. The number of nonzeroes is about 1.33$\cdot 10^6$ for all matrices of the sequence.
- In the beginning, a strong shock detaches from the cylinder, which then slowly moves backward through the domain until reaching the steady state position.
- The iterative solver is BiCGSTAB with stopping criterion 10^{-7}, the implementation is in C++.

2. The proposed preconditioner updates



BiCGSTAB iterations for the first 500 systems in the cylinder problem

2. The proposed preconditioner updates

Can we define a good indicator of the quality of the update compared to the frozen preconditioner ? For ILU(0) and lower triangular updates we can prove

Lemma 1 *Let*

$$\rho = \frac{\|(D_B + L_B)(I - U)\|_F \, (2\,\|E - U_B\|_F + \|(D_B + L_B)(I - U)\|_F)}{\|D_B + L_B\|_F^2} < 1,$$

where $E = A - LDU$. Then the accuracy $\|A^+ - (LD - D_B - L_B)U\|_F$ of the updated preconditioner is higher than the accuracy of the frozen preconditioner $\|A^+ - LDU\|_F^2$ with

$$\|A^+ - (LD - D_B - L_B)U\|_F \leq \sqrt{\|A^+ - LDU\|_F^2 - (1-\rho)\|D_B + L_B\|_F^2}. \qquad (2)$$

2. The proposed preconditioner updates

For the first systems of the CFD problem the previous bound has the values

i	$\Vert A^{(i)} - LDU \Vert_F$	$\Vert A^{(i)} - M^{(i)} \Vert_F$	Bound from (2)	ρ from (2)
2	37.454	34.277	36.172	0.571
3	37.815	34.475	36.411	0.551
4	42.096	34.959	36.938	0.245
5	50.965	35.517	37.557	0.104
6	55.902	36.118	38.308	0.083

Accuracy of the preconditioners and theoretical bounds

However, the indicator-value ρ

- describes only accuracy, not stability $\Vert I - A^{(i)}(M^{(i)})^{-1} \Vert$ of the update
- is not fully for free because of the matrix product $(D_B + L_B)(I - U)$.

3. Updates in matrix-free environment

In many applications the linear solver is chosen such that it is not necessary to store system matrices; a matrix-vector product subroutine suffices. Can the updates be used in matrix-free environment ?

First note that to compute an incomplete factorization like ILU in matrix-free environment at all, the system matrix has to be *estimated*. This can be done with a graph coloring algorithm that tries to minimize the number of matvecs for a good estimate [Cullum, Tůma - 2006].

Recall the upper triangular update

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

is based on the splitting

$$L_B + D_B + U_B = B = A - A^+.$$

Thus the update needs information on A and A^+.

3. Updates in matrix-free environment

We propose two strategies:

1. Estimation of *only the upper triangular part* of A^+ (the matrix A has been estimated anyway). Depending on the sparsity structure this might cost significantly less matvecs than estimation of the whole matrix A^+.

2. The second strategy circumvents any estimation of A^+. Let the matvec be replaced with a function evaluation

$$A^+ \cdot v \quad \rightarrow \quad F^+(v), \quad F^+ : \mathbb{R}^n \rightarrow \mathbb{R}^n.$$

We assume it is possible to compute the components $F_i^+ : \mathbb{R}^n \rightarrow \mathbb{R}$ of F^+ individually,

$$e_i^T \cdot A^+ \cdot v \quad \rightarrow \quad F_i^+(v), \quad F_i^+ : \mathbb{R}^n \rightarrow \mathbb{R}.$$

- The forward solves with L in $M^+ = L(DU - D_B - U_B)$ are trivial.

3. Updates in matrix-free environment

- For the backward solves, use a mixed explicit-implicit strategy: Split $DU - D_B - U_B = DU - triu(A) + triu(A^+)$ in the explicitly given part $X \equiv DU - triu(A)$ and the implicit part $triu(A^+)$. We then have to solve the triangular systems

$\left(X + triu(A^+)\right) z = y,$ yielding the standard backward substitution cycle

$$z_i = \frac{y_i - \sum_{j>i} x_{ij} z_j - \sum_{j>i} a^+_{ij} z_j}{x_{ii} + a^+_{ii}}, \qquad i = n, n-1, \ldots, 1. \qquad (3)$$

The sum $\sum_{j>i} a^+_{ij} z_j$ can be computed by the function evaluation

$$\sum_{j>i} a^+_{ij} z_j = F^+_i\left((0, \ldots, 0, z_{i+1}, \ldots, z_n)^T\right). \qquad (4)$$

The diagonal $\{a^+_{11}, \ldots, a^+_{nn}\}$ can be found by computing

$$a^+_{ii} = F^+_i(e_i), \qquad 1 \leq i \leq n.$$

4. Conclusions, future work

- We described a black-box preconditioner update for general nonsymmetric sequences of linear systems
- It can be combined with other techniques for solving specific types of sequences
- It can be applied in matrix-free environment
- Future work includes permutations to enhance triangular dominance, incorporating a priori estimators of the quality of the update and formulation for approximate inverse factorizations

For more details see:

- Birken Ph, Duintjer Tebbens J, Meister A, Tůma M: *Preconditioner Updates Applied to CFD Model Problems*, published online in Applied Numerical Mathematics in October 2007.
- Duintjer Tebbens J, Tůma M: *Improving Triangular Preconditioner Updates for Nonsymmetric Linear Systems*, LNCS vol. 4818, pp. 737–744, 2007 (proceedings of the 6th International Conference on Large-Scale Scientific Computations).
- Duintjer Tebbens J, Tůma M: *Efficient Preconditioning of Sequences of Nonsymmetric Linear Systems*, SIAM J. Sci. Comput., vol. 29, no. 5, pp. 1918–1941, 2007.

Thank you for your attention.

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