

Numerical behavior of saddle point solvers

Miro Rozložník

joint results with Pavel Jiránek and Valeria Simoncini

Institute of Computer Science, Czech Academy of Sciences,
Prague, Czech Republic

Seminar at the Hong Kong City University, August 27, 2009

Saddle point problems

We consider a saddle point problem with the symmetric 2×2 block form

$$\begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix}.$$

- ▶ A is a square $n \times n$ nonsingular (symmetric positive definite) matrix,
- ▶ B is a rectangular $n \times m$ matrix of (full column) rank m .

Applications: mixed finite element approximations, weighted least squares, constrained optimization, computational fluid dynamics, electromagnetism etc. [Benzi, Golub and Liesen, 2005]. For the updated list of applications leading to saddle point problems contact [Benzi, 2009].

PROBLEM

SOLUTION APPROACH

PRECONDITIONER

ITERATIVE SOLVER

Segregated or coupled solution approach

1. **Schur complement or null-space projection approach:** outer iteration for solving the reduced system;
2. **inexact solution of inner systems:** inner iteration loop with appropriate stopping criterion;

Numerous solution schemes: inexact Uzawa algorithms, inexact null-space methods, inner-outer iteration methods, two-stage iteration processes, multilevel or multigrid methods, domain decomposition methods

References: [Elman and Golub, 1994], [Bramble, Pasciak and Vassilev, 2000], [Zulehner, 2002], [Braess, Deufhard and Lipnikov, 2002], ...

Preconditioning and preconditioners

1. **preconditioning**: iteration scheme for solving the preconditioned system;
2. **approximate or incomplete factorization scheme**: structure-based or with appropriate dropping criterion;

Numerous techniques: block diagonal preconditioners, block triangular preconditioners, constraint preconditioning, Hermitian/skew-Hermitian preconditioning and other splittings, combination preconditioning

References: [Bramble and Pasciak, 1988], [Silvester and Wathen, Wathen and Silvester 1993, 1994], [Elman, Silvester and Wathen, 2002, 2005], [Kay, Loghin and Wathen, 2002], [Perugia, Simoncini, Arioli, 1999], [Keller, Gould and Wathen 2000], [Gould, Hribar and Nocedal, 2001], [Stoll, Wathen, 2008], ...

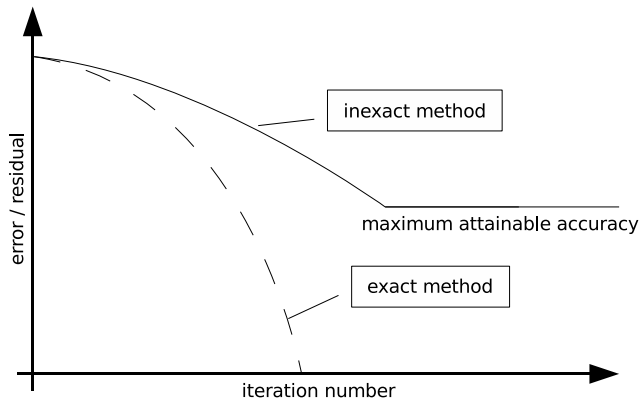
Exact and finite precision arithmetic

1. **iterative method**: finite termination property, theoretical rate of convergence;
2. **the rounding errors in floating point arithmetic**: numerical stability of the solver

Numerous iterative solvers: conjugate gradient (CG) method, MINRES, GMRES, flexible GMRES, GCR, BiCG, BiCGSTAB, ...

References: [Hestenes and Stiefel, 1952], [Paige and Saunders, 1975], [Saad and Schultz, 1986], [Elman, 1982], [Lanczos 1950], [Fletcher 1976], [van der Vorst 1992], [Paige, 1976], [Greenbaum and Strakoš 1991, 1992], [Greenbaum, Paige, R., Strakoš 1995, 1997, 2006], [Modersitzki, Sleijpen and van der Vorst, 1997], [Gutknecht, Jiránek, R, 2008], ...

Delay of convergence and limit on the final accuracy



Schur complement reduction method

- Compute y as a solution of the Schur complement system

$$B^T A^{-1} B y = B^T A^{-1} f,$$

- compute x as a solution of

$$A x = f - B y.$$

- Segregated vs. coupled approach: x_k and y_k approximate solutions to x and y , respectively.
- Inexact solution of systems with A : **every computed solution $\hat{u}$ of $Au = b$ is interpreted an exact solution of a perturbed system**

$$(A + \Delta A) \hat{u} = b + \Delta b, \quad \|\Delta A\| \leq \tau \|A\|, \quad \|\Delta b\| \leq \tau \|b\|, \quad \tau \kappa(A) \ll 1.$$

Iterative solution of the Schur complement system

choose y_0 , solve $Ax_0 = f - By_0$

compute α_k and $p_k^{(y)}$

$$y_{k+1} = y_k + \alpha_k p_k^{(y)}$$

$$\text{solve } Ap_k^{(x)} = -Bp_k^{(y)}$$

back-substitution:

$$\mathbf{A}: x_{k+1} = x_k + \alpha_k p_k^{(x)},$$

$$\mathbf{B}: \text{solve } Ax_{k+1} = f - By_{k+1},$$

$$\mathbf{C}: \text{solve } Au_k = f - Ax_k - By_{k+1},$$

$$x_{k+1} = x_k + u_k.$$

$$r_{k+1}^{(y)} = r_k^{(y)} - \alpha_k B^T p_k^{(x)}$$

inner
iteration

outer
iteration

Numerical experiments: a small model example

$$A = \text{tridiag}(1, 4, 1) \in \mathbb{R}^{100 \times 100}, \quad B = \text{rand}(100, 20), \quad f = \text{rand}(100, 1),$$

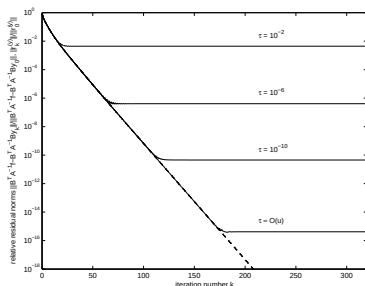
$$\kappa(A) = \|A\| \cdot \|A^{-1}\| = 7.1695 \cdot 0.4603 \approx 3.3001,$$

$$\kappa(B) = \|B\| \cdot \|B^\dagger\| = 5.9990 \cdot 0.4998 \approx 2.9983.$$

Accuracy in outer iteration

$$\| -B^T A^{-1} f + B^T A^{-1} B y_k - r_k^{(y)} \| \leq \frac{O(\tau) \kappa(A)}{1 - \tau \kappa(A)} \|A^{-1}\| \|B\| (\|f\| + \|B\| Y_k).$$

$$Y_k \equiv \max\{\|y_i\| \mid i = 0, 1, \dots, k\}.$$



$$B^T (A + \Delta A)^{-1} B \hat{y} = B^T (A + \Delta A)^{-1} f,$$

$$\|B^T A^{-1} f - B^T A^{-1} B \hat{y}\| \leq \frac{\tau \kappa(A)}{1 - \tau \kappa(A)} \|A^{-1}\| \|B\|^2 \|\hat{y}\|.$$

Does the final accuracy depend on the outer iteration method?

- ▶ Gap between the true and updated residual for any two-term recurrence method depends on the maximum norm of approximate solutions over the whole iteration process. [Greenbaum 1994, 1997]

$$\| -B^T A^{-1} f + B^T A^{-1} B y_k - r_k^{(y)} \| \leq \frac{O(\tau) \kappa(A)}{1 - \tau \kappa(A)} \| A^{-1} \| \| B \| (\| f \| + \| B \| Y_k).$$

$$Y_k \equiv \max\{\| y_i \| \mid i = 0, 1, \dots, k\}.$$

- ▶ Schur complement system is negative definite, some norm of the error or residual converges monotonically for almost all iterative methods. The quantity Y_k then does not play an important role and it can be replaced by $\| y_0 \|$ or a multiple of $\| y \|$.

Accuracy in the saddle point system

$$\|f - Ax_k - By_k\| \leq \frac{O(\alpha_1)\kappa(A)}{1 - \tau\kappa(A)} (\|f\| + \|B\|Y_k),$$

$$\| -B^T x_k - r_k^{(y)} \| \leq \frac{O(\alpha_2)\kappa(A)}{1 - \tau\kappa(A)} \|A^{-1}\| \|B\| (\|f\| + \|B\|Y_k),$$

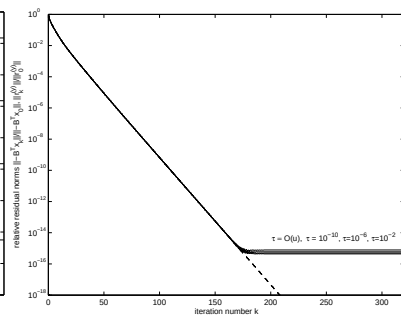
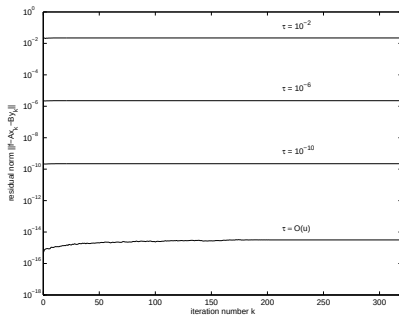
$$Y_k \equiv \max\{\|y_i\| \mid i = 0, 1, \dots, k\}.$$

Back-substitution scheme	α_1	α_2
A: Generic update $x_{k+1} = x_k + \alpha_k p_k^{(x)}$	τ	u
B: Direct substitution $x_{k+1} = A^{-1}(f - By_{k+1})$	τ	τ
C: Corrected dir. subst. $x_{k+1} = x_k + A^{-1}(f - Ax_k - By_{k+1})$	u	τ

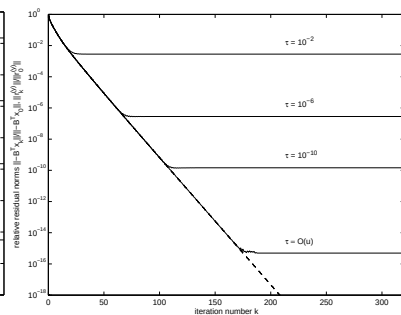
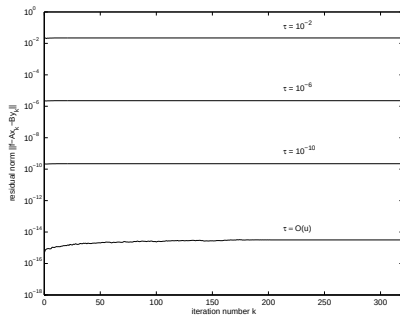
} additional system with A

$$-B^T A^{-1} f + B^T A^{-1} B y_k = -B^T x_k - B^T A^{-1} (f - A x_k - B y_k)$$

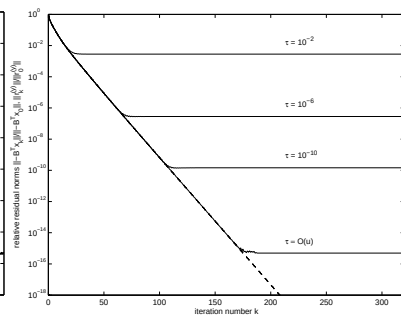
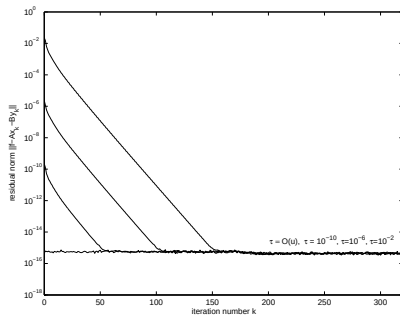
Generic update: $x_{k+1} = x_k + \alpha_k p_k^{(x)}$



Direct substitution: $x_{k+1} = A^{-1}(f - By_{k+1})$



Corrected direct substitution: $x_{k+1} = x_k + A^{-1}(f - Ax_k - By_{k+1})$

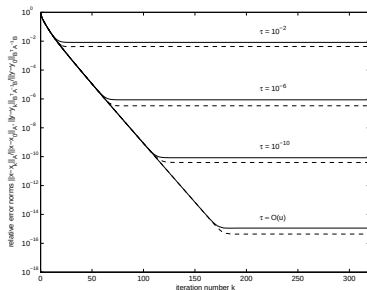


Forward error of computed approximate solution

$$\|x - x_k\| \leq \gamma_1 \|f - Ax_k - By_k\| + \gamma_2 \| -B^T x_k \|,$$

$$\|y - y_k\| \leq \gamma_2 \|f - Ax_k - By_k\| + \gamma_3 \| -B^T x_k \|,$$

$$\gamma_1 = \sigma_{\min}^{-1}(A), \quad \gamma_2 = \sigma_{\min}^{-1}(B), \quad \gamma_3 = \sigma_{\min}^{-1}(B^T A^{-1} B).$$



Null-space projection method

- ▶ compute $x \in N(B^T)$ as a solution of the projected system

$$(I - \Pi)A(I - \Pi)x = (I - \Pi)f,$$

- ▶ compute y as a solution of the least squares problem

$$By \approx f - Ax,$$

$\Pi = B(B^T B)^{-1} B^T$ is the orthogonal projector onto $R(B)$.

- ▶ Schemes with the inexact solution of least squares with B . Every computed approximate solution $\bar{v}$ of a least squares problem $Bv \approx c$ is interpreted as an exact solution of a perturbed least squares

$$(B + \Delta B)\bar{v} \approx c + \Delta c, \quad \|\Delta B\| \leq \tau \|B\|, \quad \|\Delta c\| \leq \tau \|c\|, \quad \tau \kappa(B) \ll 1.$$

Null-space projection method

choose x_0 , solve $By_0 \approx f - Ax_0$

compute α_k and $p_k^{(x)} \in N(B^T)$

$$x_{k+1} = x_k + \alpha_k p_k^{(x)}$$

$$\left| \begin{array}{l} \text{solve } Bp_k^{(y)} \approx r_k^{(x)} - \alpha_k Ap_k^{(x)} \end{array} \right.$$

back-substitution:

$$\mathbf{A: } y_{k+1} = y_k + p_k^{(y)},$$

$$\mathbf{B: } \text{solve } By_{k+1} \approx f - Ax_{k+1},$$

$$\mathbf{C: } \text{solve } Bv_k \approx f - Ax_{k+1} - By_k,$$

$$y_{k+1} = y_k + v_k.$$

$$r_{k+1}^{(x)} = r_k^{(x)} - \alpha_k Ap_k^{(x)} - Bp_k^{(y)}$$

} inner
iteration

} outer
iteration

Accuracy in the saddle point system

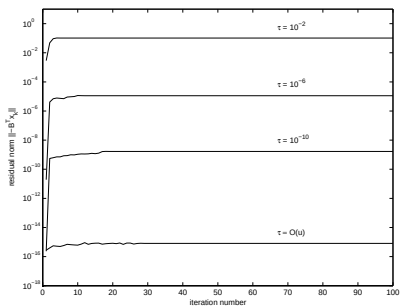
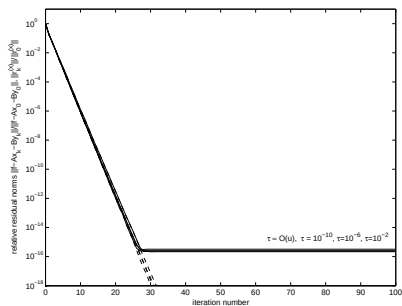
$$\|f - Ax_k - By_k - r_k^{(x)}\| \leq \frac{O(\alpha_3)\kappa(B)}{1 - \tau\kappa(B)} (\|f\| + \|A\|X_k),$$

$$\| -B^T x_k \| \leq \frac{O(\tau)\kappa(B)}{1 - \tau\kappa(B)} \|B\|X_k,$$

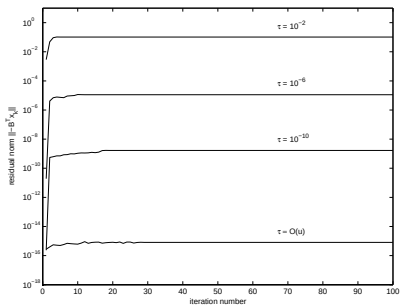
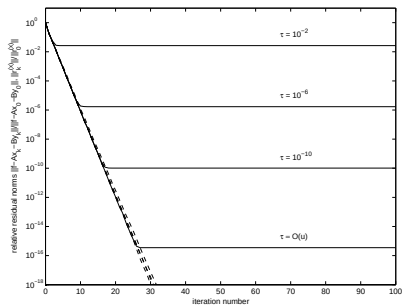
$$X_k \equiv \max\{\|x_i\| \mid i = 0, 1, \dots, k\}.$$

Back-substitution scheme	α_3	} additional least square with B
A: Generic update $y_{k+1} = y_k + p_k^{(y)}$	u	
B: Direct substitution $y_{k+1} = B^\dagger(f - Ax_{k+1})$	τ	
C: Corrected dir. subst. $y_{k+1} = y_k + B^\dagger(f - Ax_{k+1} - By_k)$	u	

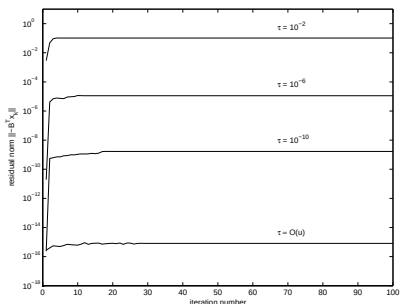
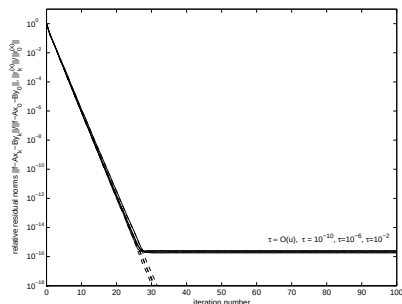
Generic update: $y_{k+1} = y_k + p_k^{(y)}$



Direct substitution: $y_{k+1} = B^\dagger(f - Ax_{k+1})$



Corrected direct substitution: $y_{k+1} = y_k + B^\dagger(f - Ax_{k+1} - By_k)$



$\mathcal{A}$ symmetric indefinite, $\mathcal{P}$ positive definite

$$\mathcal{A} = \begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix} \approx \mathcal{P} = \mathcal{R}^T \mathcal{R}$$

$$(\mathcal{R}^{-T} \mathcal{A} \mathcal{R}^{-1}) \mathcal{R} \begin{pmatrix} x \\ y \end{pmatrix} = \mathcal{R}^{-T} \begin{pmatrix} f \\ 0 \end{pmatrix}$$

$\mathcal{R}^{-T} \mathcal{A} \mathcal{R}^{-1}$ is symmetric indefinite!

$\mathcal{P}$ symmetric indefinite or nonsymmetric

$$\mathcal{P}^{-1} \mathcal{A} \begin{pmatrix} x \\ y \end{pmatrix} = \mathcal{P}^{-1} \begin{pmatrix} f \\ 0 \end{pmatrix}$$

$$(\mathcal{A} \mathcal{P}^{-1}) \mathcal{P} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix}$$

$\mathcal{P}^{-1} \mathcal{A}$ and $\mathcal{A} \mathcal{P}^{-1}$ are nonsymmetric!

Schur complement approach with indefinite preconditioner

$$\begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix}, \quad \mathcal{P} = \begin{pmatrix} A & B \\ B^T & B^T A^{-1} B - I \end{pmatrix}$$
$$\mathcal{A}\mathcal{P}^{-1} = \begin{pmatrix} I & 0 \\ (I - S)B^T A^{-1} & S \end{pmatrix}$$

$S = B^T A^{-1} B$, $\mathcal{A}\mathcal{P}^{-1}$ **nonsymmetric** but **diagonalizable** and it has a 'nice' spectrum!

$$\sigma(\mathcal{A}\mathcal{P}^{-1}) \subset \{1\} \cup \sigma(B^T A^{-1} B)$$

[Durazzi, Ruggiero 2003], [Fortin, El-Maliki, 2009?]

Krylov method with the preconditioner: basic properties

$$\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}, r_0 = \begin{pmatrix} 0 \\ s_0 \end{pmatrix}, e_{k+1} = \begin{pmatrix} x - x_{k+1} \\ y - y_{k+1} \end{pmatrix}$$

$$r_{k+1} = \begin{pmatrix} f \\ 0 \end{pmatrix} - \begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} x_{k+1} \\ y_{k+1} \end{pmatrix}$$

$$\begin{aligned} r_0 = \begin{pmatrix} 0 \\ s_0 \end{pmatrix} &\Rightarrow r_{k+1} = \begin{pmatrix} 0 \\ s_{k+1} \end{pmatrix} \\ &\Rightarrow Ax_{k+1} + By_{k+1} = f \end{aligned}$$

Preconditioned CG method: saddle point problem and indefinite preconditioner

$$r_{k+1}^T \mathcal{P}^{-1} r_j = 0, \quad j = 0, \dots, k$$

y_{k+1} is an iterate from CG applied to the Schur complement system

$$B^T A^{-1} B y = B^T A^{-1} f!$$

satisfying

$$\min_{u \in x_0 + K_{k+1}(B^T A^{-1} B, B^T A^{-1} f)} \|y - u\|_{B^T A^{-1} B} =$$

Preconditioned CG algorithm

$$\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}, r_0 = b - \mathcal{A} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} 0 \\ s_0 \end{pmatrix}$$

$$\begin{pmatrix} p_0^{(x)} \\ p_0^{(y)} \end{pmatrix} = \mathcal{P}^{-1} r_0 = \mathcal{P}^{-1} \begin{pmatrix} 0 \\ s_0 \end{pmatrix}$$

$$k = 0, 1, \dots$$

$$\alpha_k = \left(\begin{pmatrix} 0 \\ s_k \end{pmatrix}, \mathcal{P}^{-1} \begin{pmatrix} 0 \\ s_k \end{pmatrix} \right) / \left(\mathcal{A} \begin{pmatrix} p_k^{(x)} \\ p_k^{(y)} \end{pmatrix}, \begin{pmatrix} p_k^{(x)} \\ p_k^{(y)} \end{pmatrix} \right)$$

$$\alpha_k = \frac{(r_k, z_k)}{(\mathcal{A} p_k, p_k)}$$

$$\begin{pmatrix} x_{k+1} \\ y_{k+1} \end{pmatrix} = \begin{pmatrix} x_k \\ y_k \end{pmatrix} + \alpha_k \begin{pmatrix} p_k^{(x)} \\ p_k^{(y)} \end{pmatrix}$$

$$r_{k+1} = r_k - \alpha_k \mathcal{A} \begin{pmatrix} p_k^{(x)} \\ p_k^{(y)} \end{pmatrix} = \begin{pmatrix} 0 \\ s_{k+1} \end{pmatrix}$$

$$z_{k+1} = \mathcal{P}^{-1} r_{k+1}$$

$$\beta_k = \left(\begin{pmatrix} 0 \\ s_{k+1} \end{pmatrix}, \mathcal{P}^{-1} \begin{pmatrix} 0 \\ s_{k+1} \end{pmatrix} \right) / \left(\begin{pmatrix} 0 \\ s_k \end{pmatrix}, \mathcal{P}^{-1} \begin{pmatrix} 0 \\ s_k \end{pmatrix} \right)$$

$$\beta_k = \frac{(r_{k+1}, z_{k+1})}{(r_k, z_k)}$$

$$\begin{pmatrix} p_{k+1}^{(x)} \\ p_{k+1}^{(y)} \end{pmatrix} = \mathcal{P}^{-1} \begin{pmatrix} 0 \\ s_{k+1} \end{pmatrix} + \beta_k \begin{pmatrix} p_k^{(x)} \\ p_k^{(y)} \end{pmatrix} = \begin{pmatrix} -A^{-1} B p_{k+1}^{(y)} \\ p_{k+1}^{(y)} \end{pmatrix}$$

$$p_{k+1} = z_{k+1} + \beta_k p_k$$

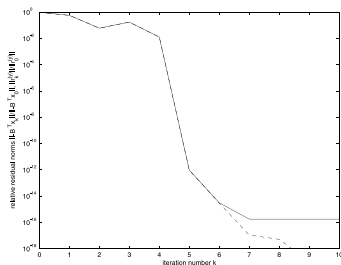
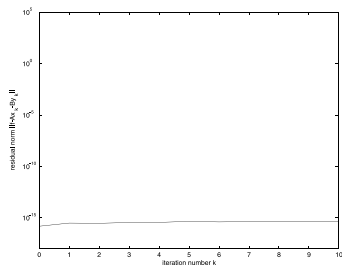
Numerical experiments: a small model example

$$A = \text{tridiag}(1, 4, 1) \in \mathbb{R}^{25 \times 25}, \quad B = \text{rand}(25, 5), \quad f = \text{rand}(25, 1),$$

$$\kappa(A) = \|A\| \cdot \|A^{-1}\| = 5.9854 \cdot 0.4963 \approx 2.9710,$$

$$\kappa(B) = \|B\| \cdot \|B^\dagger\| = 5.9990 \cdot 0.4998 \approx 2.9983.$$

Generic update: $x_{k+1} = x_k + \alpha_k p_k^{(x)}$ with $p_k^{(x)} = -A^{-1}Bp_k^{(y)}$



Saddle point problem and indefinite constraint preconditioner

$$\begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix}, \quad \mathcal{P} = \begin{pmatrix} I & B \\ B^T & 0 \end{pmatrix}$$

$$\mathcal{A}\mathcal{P}^{-1} = \begin{pmatrix} A(I - \Pi) + \Pi & (A - I)B(B^T B)^{-1} \\ 0 & I \end{pmatrix}$$

$\Pi = B(B^T B)^{-1} B^T$ - orth. projector onto $\text{span}(B)$

[Lukšan, Vlček, 1998], [Gould, Keller, Wathen 2000]
[Perugia, Simoncini, Arioli, 1999], [R, Simoncini, 2002]

Indefinite constraint preconditioner: spectral properties

$\mathcal{AP}^{-1}$ **nonsymmetric** and **non-diagonalizable!**
but it has a 'nice' spectrum:

$$\begin{aligned}\sigma(\mathcal{AP}^{-1}) &\subset \{1\} \cup \sigma(A(I - \Pi) + \Pi) \\ &\subset \{1\} \cup \sigma((I - \Pi)A(I - \Pi)) - \{0\}\end{aligned}$$

and only 2 by 2 Jordan blocks!

[Lukšan, Vlček 1998], [Gould, Wathen, Keller, 1999], [Perugia, Simoncini 1999]

Krylov method with the constraint preconditioner: basic properties

$$\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}, r_0 = \begin{pmatrix} s_0 \\ 0 \end{pmatrix}, e_{k+1} = \begin{pmatrix} x - x_{k+1} \\ y - y_{k+1} \end{pmatrix}$$

$$r_{k+1} = \begin{pmatrix} f \\ 0 \end{pmatrix} - \begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} x_{k+1} \\ y_{k+1} \end{pmatrix}$$

$$\begin{aligned} r_0 = \begin{pmatrix} s_0 \\ 0 \end{pmatrix} &\Rightarrow r_{k+1} = \begin{pmatrix} s_{k+1} \\ 0 \end{pmatrix} \\ &\Rightarrow B^T(x - x_{k+1}) = 0 \\ &\Rightarrow x_{k+1} \in \text{Null}(B^T)! \end{aligned}$$

Preconditioned CG algorithm

$$\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}, r_0 = b - \mathcal{A} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} s_0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} p_0^{(x)} \\ p_0^{(y)} \end{pmatrix} = \mathcal{P}^{-1} r_0 = \mathcal{P}^{-1} \begin{pmatrix} s_0 \\ 0 \end{pmatrix}$$

$$k = 0, 1, \dots$$

$$\alpha_k = \left(\begin{pmatrix} s_k \\ 0 \end{pmatrix}, \mathcal{P}^{-1} \begin{pmatrix} s_k \\ 0 \end{pmatrix} \right) / \left(\mathcal{A} \begin{pmatrix} p_k^{(x)} \\ p_k^{(y)} \end{pmatrix}, \begin{pmatrix} p_k^{(x)} \\ p_k^{(y)} \end{pmatrix} \right) \quad \alpha_k = (r_k, z_k) / (\mathcal{A} p_k, p_k)$$

$$\begin{pmatrix} x_{k+1} \\ y_{k+1} \end{pmatrix} = \begin{pmatrix} x_k \\ y_k \end{pmatrix} + \alpha_k \begin{pmatrix} p_k^{(x)} \\ p_k^{(y)} \end{pmatrix}$$

$$r_{k+1} = r_k - \alpha_k \mathcal{A} \begin{pmatrix} p_k^{(x)} \\ p_k^{(y)} \end{pmatrix} = \begin{pmatrix} s_{k+1} \\ 0 \end{pmatrix}$$

$$z_{k+1} = \mathcal{P}^{-1} r_{k+1}$$

$$\beta_k = \left(\begin{pmatrix} s_{k+1} \\ 0 \end{pmatrix}, \mathcal{P}^{-1} \begin{pmatrix} s_{k+1} \\ 0 \end{pmatrix} \right) / \left(\begin{pmatrix} s_k \\ 0 \end{pmatrix}, \mathcal{P}^{-1} \begin{pmatrix} s_k \\ 0 \end{pmatrix} \right) \quad \beta_k = (r_{k+1}, z_{k+1}) / (r_k, z_k)$$

$$\begin{pmatrix} p_{k+1}^{(x)} \\ p_{k+1}^{(y)} \end{pmatrix} = \mathcal{P}^{-1} \begin{pmatrix} s_{k+1} \\ 0 \end{pmatrix} + \beta_k \begin{pmatrix} p_k^{(x)} \\ p_k^{(y)} \end{pmatrix}$$

$$p_{k+1} = z_{k+1} + \beta_k p_k$$

Preconditioned CG method: error norm

$$r_{k+1}^T \mathcal{P}^{-1} r_j = 0, \quad j = 0, \dots, k$$

x_{k+1} is an iterate from CG applied to

$$(I - \Pi)A(I - \Pi)x = (I - \Pi)f!$$

satisfying

$$\|x - x_{k+1}\|_A = \min_{u \in x_0 + \text{span}\{(I - \Pi)s_j\}} \|x - u\|_A$$

[Lukšan, Vlček 1998], [Gould, Wathen, Keller, 1999]

Preconditioned CG method: residual norm

$$\|x_{k+1} - x\| \rightarrow 0$$

but in general

$$y_{k+1} \not\rightarrow y$$

which is reflected in

$$\|r_{k+1}\| = \left\| \begin{pmatrix} s_{k+1} \\ 0 \end{pmatrix} \right\| \not\rightarrow 0!$$

but under appropriate scaling yes!

Preconditioned CG method: residual norm

$$x_{k+1} \rightarrow x$$

$$x - x_{k+1} = \phi_{k+1}((I - \Pi)A(I - \Pi))(x - x_0)$$

$$s_{k+1} = \phi_{k+1}(A(I - \Pi) + \Pi)s_0$$

$$\sigma((I - \Pi)A(I - \Pi)) \sim \sigma(A(I - \Pi) + \Pi)?$$

$$\begin{aligned} \{1\} &\in \sigma((I - \Pi)A(I - \Pi)) - \{0\} \\ \Rightarrow \|r_{k+1}\| &= \left\| \begin{pmatrix} s_{k+1} \\ 0 \end{pmatrix} \right\| \rightarrow 0! \end{aligned}$$

How to avoid misconvergence?

- Scaling by a constant $\alpha > 0$ such that

$$\{1\} \in \text{conv}(\sigma((I - \Pi)\alpha A(I - \Pi)) - \{0\})$$

$$\begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix} \iff \begin{pmatrix} \alpha A & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} x \\ \alpha y \end{pmatrix} = \begin{pmatrix} \alpha f \\ 0 \end{pmatrix}$$

$$v : \quad \|(I - \Pi)v\| \neq 0, \quad \alpha = \frac{1}{((I - \Pi)v, A(I - \Pi)v)}!$$

- Scaling by a diagonal $A \rightarrow (\text{diag}(A))^{-1/2} A (\text{diag}(A))^{-1/2}$ often gives what we want!
- Different direction vector so that $\|r_{k+1}\| = \|s_{k+1}\|$ is locally minimized!

$$y_{k+1} = y_k + (B^T B)^{-1} B^T s_k$$

[Braess, Deuffhard, Lipikov 1999], [Hribar, Gould, Nocedal, 1999], [Jiránek, R, 2008]

Numerical experiments: a small model example

$$A = \text{tridiag}(1, 4, 1) \in \mathbb{R}^{25,25}, \quad B = \text{rand}(25, 5) \in \mathbb{R}^{25,5} \\ f = \text{rand}(25, 1) \in \mathbb{R}^{25}$$

$$\sigma(A) \subset [2.0146, 5.9854]$$

$$\alpha = 1/\tau \quad \sigma\left(\begin{pmatrix} \alpha A & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} I & B \\ B^T & 0 \end{pmatrix}^{-1}\right)$$

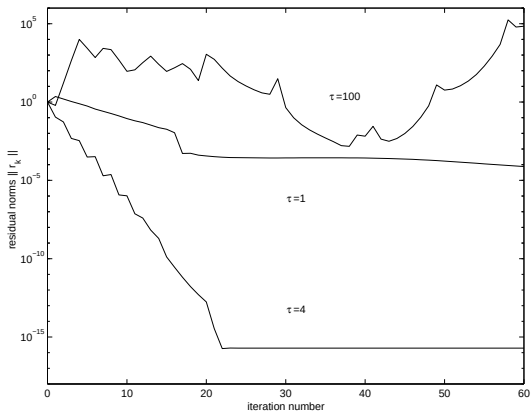
$$1/100 \quad [0.0207, 0.0586] \cup \{1\}$$

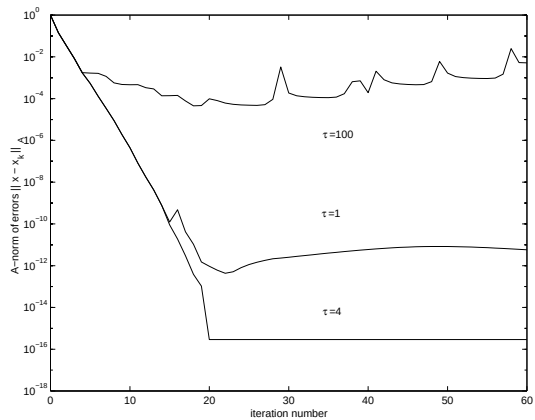
$$1/10 \quad [0.2067, 0.5856] \cup \{1\}$$

$$\mathbf{1/4} \quad [\mathbf{0.5170}, \mathbf{1.4641}]$$

$$1 \quad \{1\} \cup [2.0678, 5.8563]$$

$$4 \quad \{1\} \cup [8.2712, 23.4252]$$





Error norm of the computed approximate solution

Finite precision arithmetic:

$$\begin{pmatrix} \bar{x}_{k+1} \\ \bar{y}_{k+1} \end{pmatrix}, \quad \bar{r}_{k+1} = \begin{pmatrix} \bar{s}_{k+1}^{(1)} \\ \bar{s}_{k+1}^{(2)} \end{pmatrix} \rightarrow 0$$

$$\|x - \bar{x}_{k+1}\|_A^2 = (\Pi A(x - \bar{x}_{k+1}), \Pi(x - \bar{x}_{k+1})) + ((I - \Pi)A(x - \bar{x}_{k+1}), (I - \Pi)(x - \bar{x}_{k+1}))$$

$$\|x - \bar{x}_{k+1}\|_A \leq \gamma_1 \|\Pi(x - \bar{x}_{k+1})\| + \gamma_2 \|(I - \Pi)A(I - \Pi)(x - \bar{x}_{k+1})\|$$

Exact arithmetic:

$$\|\Pi(x - x_{k+1})\| = 0$$

$$\|(I - \Pi)A(I - \Pi)(x - x_{k+1})\| \rightarrow 0$$

Error norm of the computed approximate solution

departure from the null-space of B^T + projection of the residual onto it

$$\|x - \bar{x}_{k+1}\|_A \leq \gamma_3 \|B^T(x - \bar{x}_{k+1})\| + \gamma_2 \|(I - \Pi)(f - A\bar{x}_{k+1} - B\bar{y}_{k+1})\|$$

can be monitored by easily computable quantities:

$$B^T(x - \bar{x}_{k+1}) \sim \bar{s}_{k+1}^{(2)}$$

$$(I - \Pi)(f - A\bar{x}_{k+1} - B\bar{y}_{k+1}) \sim (I - \Pi)\bar{s}_{k+1}^{(1)}$$

Residuals: maximum attainable accuracy

$$\|(f - A\bar{x}_{k+1} - B\bar{y}_{k+1}) - \bar{s}_{k+1}^{(1)}\|, \|B^T(x - \bar{x}_{k+1}) - \bar{s}_{k+1}^{(2)}\| \leq$$

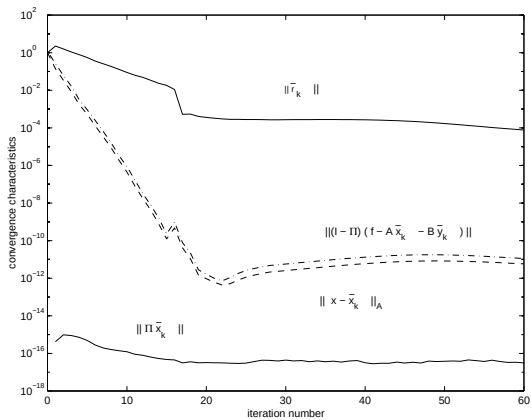
$$\leq \left\| \begin{pmatrix} f \\ 0 \end{pmatrix} - \begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} \bar{x}_{k+1} \\ \bar{y}_{k+1} \end{pmatrix} - \begin{pmatrix} \bar{s}_{k+1}^{(1)} \\ \bar{s}_{k+1}^{(2)} \end{pmatrix} \right\|$$

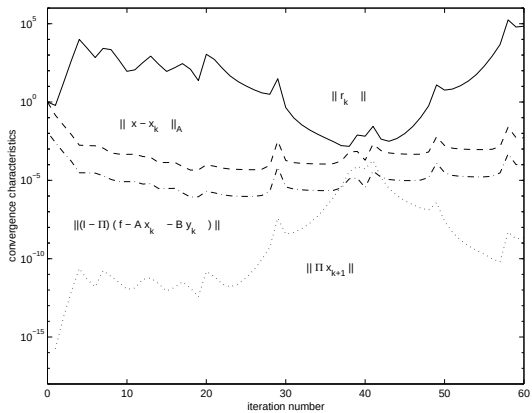
$$\leq c_1 \varepsilon \kappa(\mathcal{A}) \max_{j=0, \dots, k+1} \|\bar{r}_j\|$$

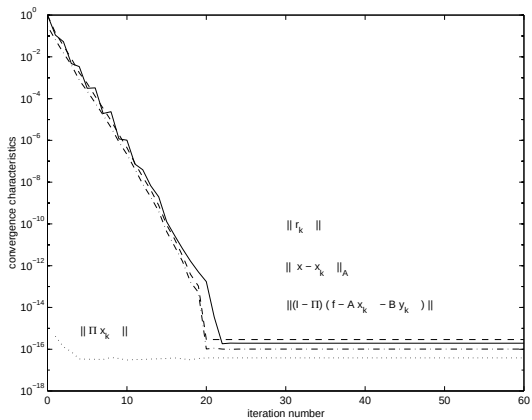
[Greenbaum 1994,1997], [Sleijpen, et al. 1994]

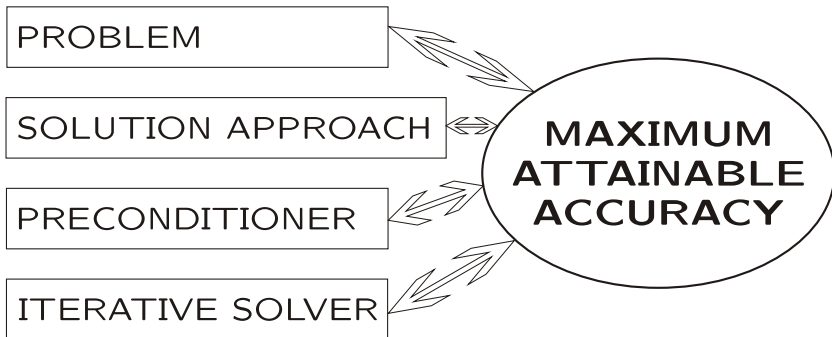
good scaling: $\|\bar{r}_j\| \rightarrow 0$ nearly monotonically

$$\|\bar{r}_0\| \sim \max_{j=0, \dots, k+1} \|\bar{r}_j\|$$



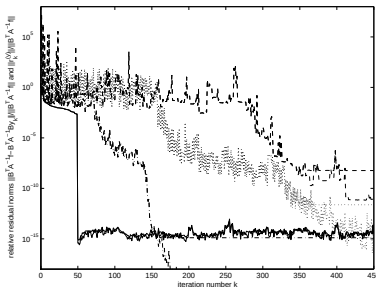






Conclusions: segregated solution approach

- ▶ The accuracy measured by the residuals of the saddle point problem depends on the choice of the back-substitution scheme [Jiránek, R, 2008]. The schemes with (generic or corrected substitution) updates deliver approximate solutions which satisfy either the first or second block equation to working accuracy.
- ▶ Care must be taken when solving nonsymmetric systems [Jiránek, R, 2008], all bounds of the limiting accuracy depend on the maximum norm of computed iterates, cf. [Greenbaum 1994,1997], [Sleijpen, et al. 1994].



Conclusions: coupled approach with indefinite preconditioner

- ▶ Short-term recurrence methods are applicable for saddle point problems with indefinite preconditioning at a cost comparable to that of symmetric solvers.
- ▶ The convergence of CG applied to saddle point problem with indefinite preconditioner for all right-hand side vectors is not guaranteed. For a particular set of right-hand sides the convergence can be achieved by the appropriate scaling of the saddle point problem.
- ▶ Since the maximum attainable accuracy depends heavily on the size of computed residuals, a good scaling of the problems leads to approximate solutions satisfying both two block equations to the working accuracy.

Thank you for your attention.

<http://www.cs.cas.cz/~miro>

P. Jiránek and M. Rozložník. Maximum attainable accuracy of inexact saddle point solvers. *SIAM J. Matrix Anal. Appl.*, 29(4):1297–1321, 2008.

P. Jiránek and M. Rozložník. Limiting accuracy of segregated solution methods for nonsymmetric saddle point problems. *J. Comput. Appl. Math.* 215 (2008), pp. 28-37.

M. Rozložník and V. Simoncini, Krylov subspace methods for saddle point problems with indefinite preconditioning, *SIAM J. Matrix Anal. Appl.*, 24 (2002), pp. 368–391.