On the use of Minimal Residual methods for solving indefinite symmetric structured linear systems

V. Simoncini

Dipartimento di Matematica, Università di Bologna

and CIRSA, Ravenna, Italy

valeria@dm.unibo.it

The problem

$$\mathcal{M}x = b$$

with $\mathcal{M}$ large, real indefinite symmetric matrix

A popular example:

$$\begin{bmatrix} A & B^T \\ B & -C \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix} \qquad A = A^T, \; C = C^T \geq 0$$

... Survey: Benzi, Golub and Liesen, Acta Num 2005

Iterative solver. Convergence considerations.

$$\mathcal{M}x = b$$

$\mathcal{M}$ is symmetric and indefinite $\quad\rightarrow\quad$ MINRES

$$x_k \in x_0 + K_k(\mathcal{M}, r_0), \quad \text{s.t.} \qquad \min \|b - \mathcal{M}x_k\|$$

$r_k = b - \mathcal{M}x_k$, $k = 0, 1, \ldots$, x_0 starting guess

Iterative solver. Convergence considerations.

$$\mathcal{M}x = b$$

$\mathcal{M}$ is symmetric and indefinite $\quad \rightarrow \quad$ MINRES

$$x_k \in x_0 + K_k(\mathcal{M}, r_0), \quad \text{s.t.} \qquad \min \|b - \mathcal{M}x_k\|$$

$r_k = b - \mathcal{M}x_k$, $k = 0, 1, \ldots$, x_0 starting guess

If $\mathrm{spec}(\mathcal{M}) \subset [-a, -b] \cup [c, d]$, with $|b - a| = |d - c|$, then

$$\|b - \mathcal{M}x_{2k}\| \leq 2 \left(\frac{\sqrt{ad} - \sqrt{bc}}{\sqrt{ad} + \sqrt{bc}} \right)^k \|b - \mathcal{M}x_0\|$$

Note: more general but less tractable bounds available

Features ...

... of MINRES

- Residual minimizing solver for *indefinite* linear systems
- Short-term recurrence (possibly with Lanczos recurrence)
- Positive definite preconditioner, but not necessarily factorized (not as LL^T)

Features ...

... of MINRES

- Residual minimizing solver for $indefinite$ linear systems
- Short-term recurrence (possibly with Lanczos recurrence)
- Positive definite preconditioner, but not necessarily factorized (not as LL^T)

... and of this talk

- Harmonic Ritz values and superlinear convergence
- Enhancing MINRES convergence
- Estimating the Saddle-point problem inf-sup constant

Harmonic Ritz values

$$K_k(\mathcal{M}, r_0) = \text{span}\{r_0, \mathcal{M}r_0, \dots, \mathcal{M}^{m-1}r_0\} \qquad (x_0 = 0)$$

$$x_m = \phi_{m-1}(\mathcal{M})r_0 \in K_k(\mathcal{M}, r_0), \quad \phi_{m-1}\text{polyn. of deg} \leq m-1$$

Therefore

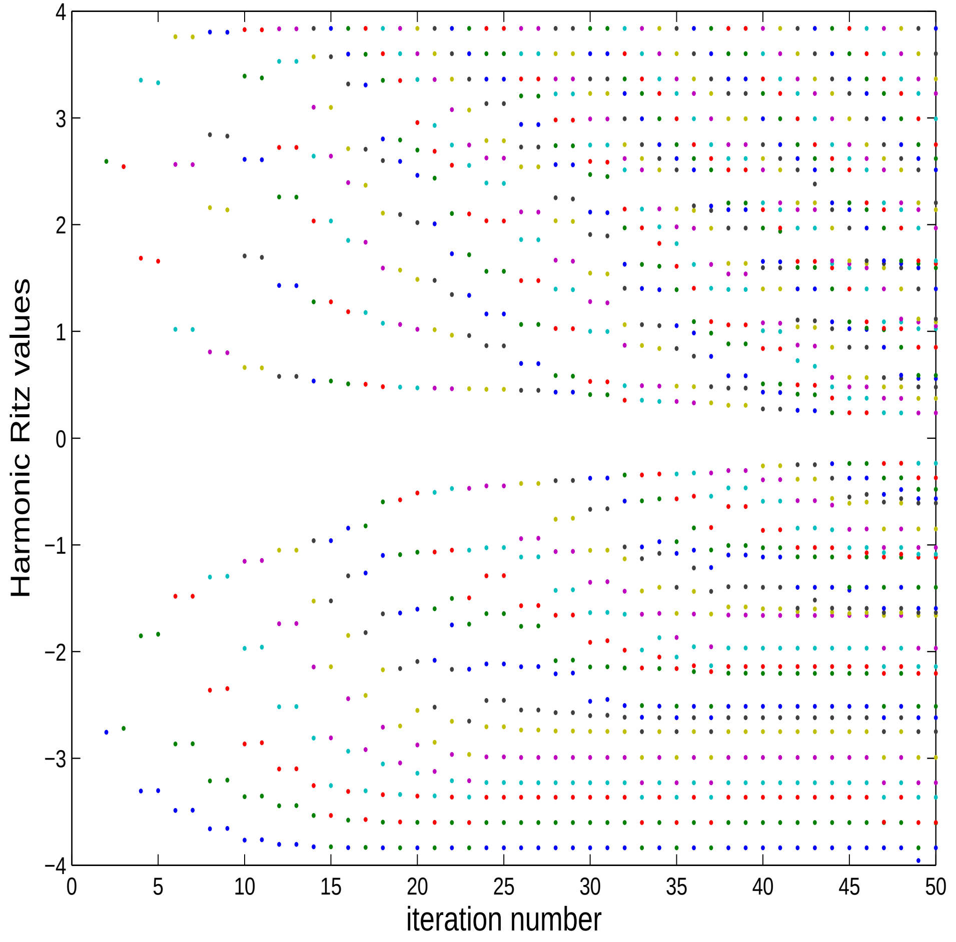
$$r_m = r_0 - \mathcal{M}x_m = \varphi_m(\mathcal{M})r_0, \quad \varphi_m\text{polyn. of deg} \leq m, \varphi_m(0) = 1$$

Harmonic Ritz values: roots of φ_m (residual polynomial)

Remark: Harmonic Ritz values approximate eigenvalues of $\mathcal{M}$

(Paige, Parlett & van der Vorst, '95)

Typical convergence pattern



Harmonic Ritz values as iterations proceed

Superlinear convergence

Generalization of CG well-known result (van der Sluis & van der Vorst '86)

MINRES: (van der Vorst and Vuik '93, van der Vorst '03)

(λ_i, z_i) eigenpairs of $\mathcal{M}$

Assume $r_m = \bar{r}_0 + s$ with $\bar{r}_0 \perp z_1$

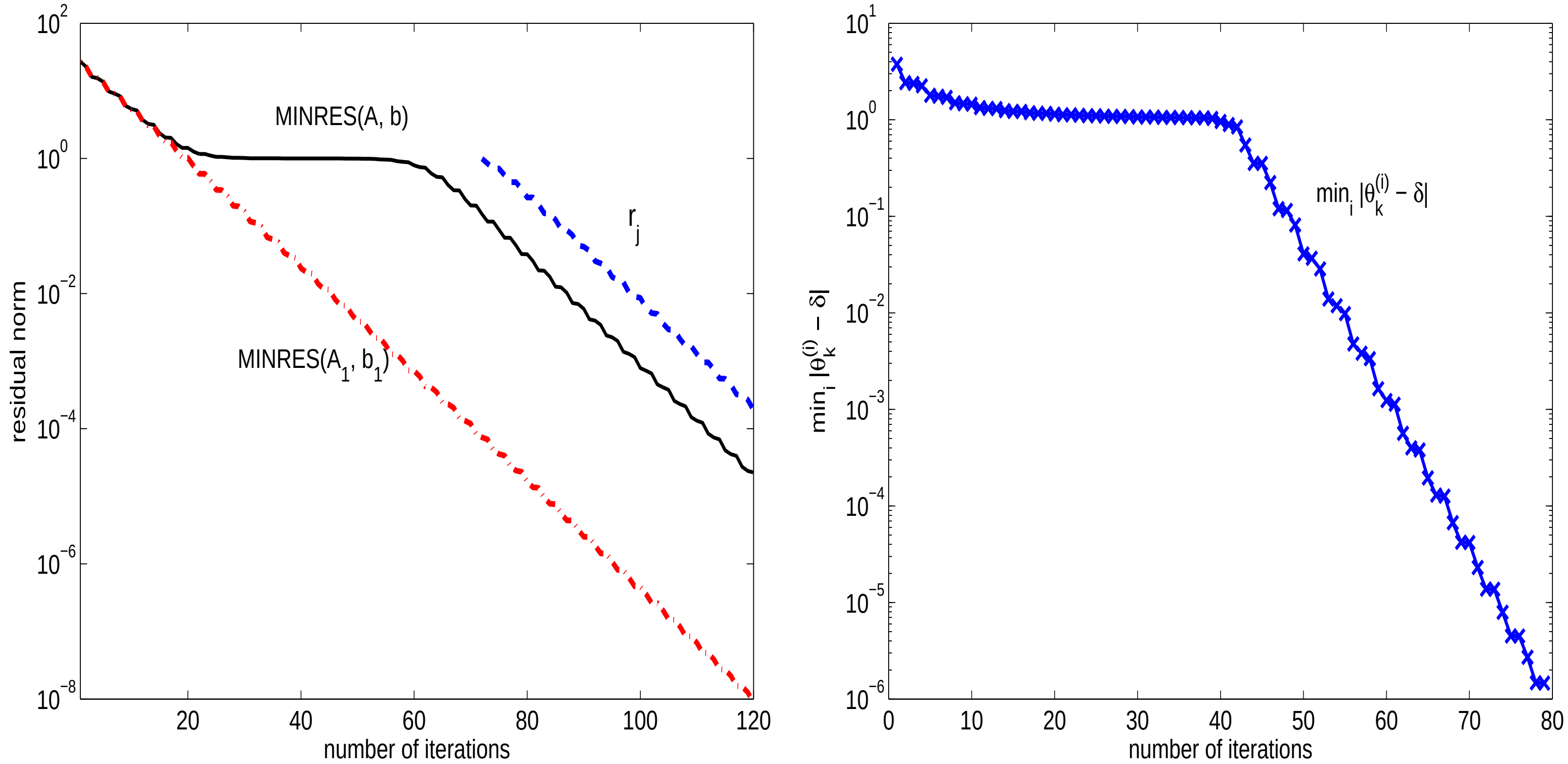
Let $\bar{r}_j$ be the GMRES residual after j iterations with $\bar{r}_0$. Then

$$\|r_{m+j}\| \le F_m \|\bar{r}_j\|, \quad \text{where} \quad F_m = \max_{k \ge 2} \frac{|\theta_1^{(m)}|}{|\lambda_1|} \frac{|\theta_1^{(m)} - \lambda_k|}{|\lambda_1 - \lambda_k|}$$

and $\theta_1^{(m)}$ is the harmonic Ritz value closest to λ_1 in $K_m(A, r_0)$.

(for a proof, Simoncini & Szyld '11, unpublished)

Superlinear convergence. An experiment.



A_1, b_1 data with negative eigenvalue closest to zero removed

Enhancing MINRES convergence for Saddle point problems

Spectral properties

$$\mathcal{M} = \begin{bmatrix} A & B^T \\ B & O \end{bmatrix} \qquad \begin{array}{ll} 0 < \lambda_n \leq \cdots \leq \lambda_1 & \text{eigs of } A \\ 0 < \sigma_m \leq \cdots \leq \sigma_1 & \text{sing. vals of } B \end{array}$$

$\mathrm{spec}(\mathcal{M})$ subset of (Rusten & Winther 1992)

$$\left[\frac{1}{2}(\lambda_n - \sqrt{\lambda_n^2 + 4\sigma_1^2}), \frac{1}{2}(\lambda_1 - \sqrt{\lambda_1^2 + 4\sigma_m^2})\right] \quad \cup \quad \left[\lambda_n, \frac{1}{2}(\lambda_1 + \sqrt{\lambda_1^2 + 4\sigma_1^2})\right]$$

More results under different hypotheses

Block diagonal Preconditioner

⋆ $\mathcal{P}_{\text{ideal}} = \begin{bmatrix} A & 0 \\ 0 & BA^{-1}B^T \end{bmatrix}$ MINRES converges in at most 3 its.

Block diagonal Preconditioner

$\star$ $\mathcal{P}_{\text{ideal}} = \begin{bmatrix} A & 0 \\ 0 & BA^{-1}B^T \end{bmatrix}$ MINRES converges in at most 3 its.

A more practical choice:

$$\mathcal{P} = \begin{bmatrix} \widetilde{A} & 0 \\ 0 & \widetilde{S} \end{bmatrix} \qquad \text{spd.} \quad \widetilde{A} \approx A \qquad \widetilde{S} \approx BA^{-1}B^T$$

spectrum in $\quad [-a, -b] \cup [c, d], \qquad a, b, c, d > 0$

A quasi-optimal approximate Schur complement

$$\widetilde{S} \approx BA^{-1}B^T$$

For certain operators, $\widetilde{S}$ is quasi-optimal:

spec$(BA^{-1}B^T\widetilde{S}^{-1})$ well clustered except for few eigenvalues

A quasi-optimal approximate Schur complement

$$\widetilde{S} \approx BA^{-1}B^T$$

For certain operators, $\widetilde{S}$ is quasi-optimal:

spec$(BA^{-1}B^T\widetilde{S}^{-1})$ well clustered except for few eigenvalues

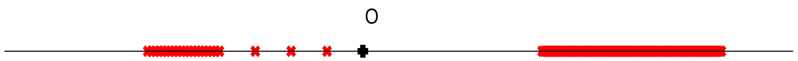
Possibly: well clustered eigs also mesh-independent

The role of $\widetilde{S}$

Claim:

The presence of outliers in $BA^{-1}B^T\widetilde{S}^{-1}$ is accurately inherited by the preconditioned matrix $\mathcal{M}\mathcal{P}^{-1}$ so that $\kappa(\mathcal{M}\mathcal{P}^{-1}) \gg 1$

spec($BA^{-1}B^T\widetilde{S}^{-1}$)



spec($\mathcal{M}\mathcal{P}^{-1}$)

(for a proof, see Olshanskii & Simoncini, SIMAX '10)

Eliminating the stagnation phase: “Deflated” MINRES

$Y = [y_1, \ldots, y_s]$: *approximate* eigenbasis of $\mathcal{M}$

* Approximation space: Augmented Lanczos sequence

$$v_{j+1} \perp \mathrm{span}\{Y, v_1, v_2, ..., v_j\}, \qquad \|v_{j+1}\| = 1$$

obtained by standard Lanczos method with coeff.matrix

$$\mathcal{G} := \mathcal{M} - \mathcal{M}Y(Y^T\mathcal{M}Y)^{-1}Y^T\mathcal{M}$$

* MINRES method:

$$r_j = \hat{b} - \mathcal{M}\hat{u}_j \;\perp\; \mathcal{G}K_j(\mathcal{G}, v_1)$$

$\Rightarrow$ $\hat{u}_j$ obtained with a short-term recurrence

Stokes type problem with variable viscosity in $\Omega \subset \mathbb{R}^d$

$$
\begin{aligned}
-\mathbf{div}\nu(\mathbf{x})\mathbf{Du} + \nabla p = \mathbf{f} \quad & \text{in} \quad \Omega, \\
-\operatorname{div} \mathbf{u} = 0 \quad & \text{in} \quad \Omega, \\
\mathbf{u} = 0 \quad & \text{on} \quad \partial\Omega,
\end{aligned}
$$

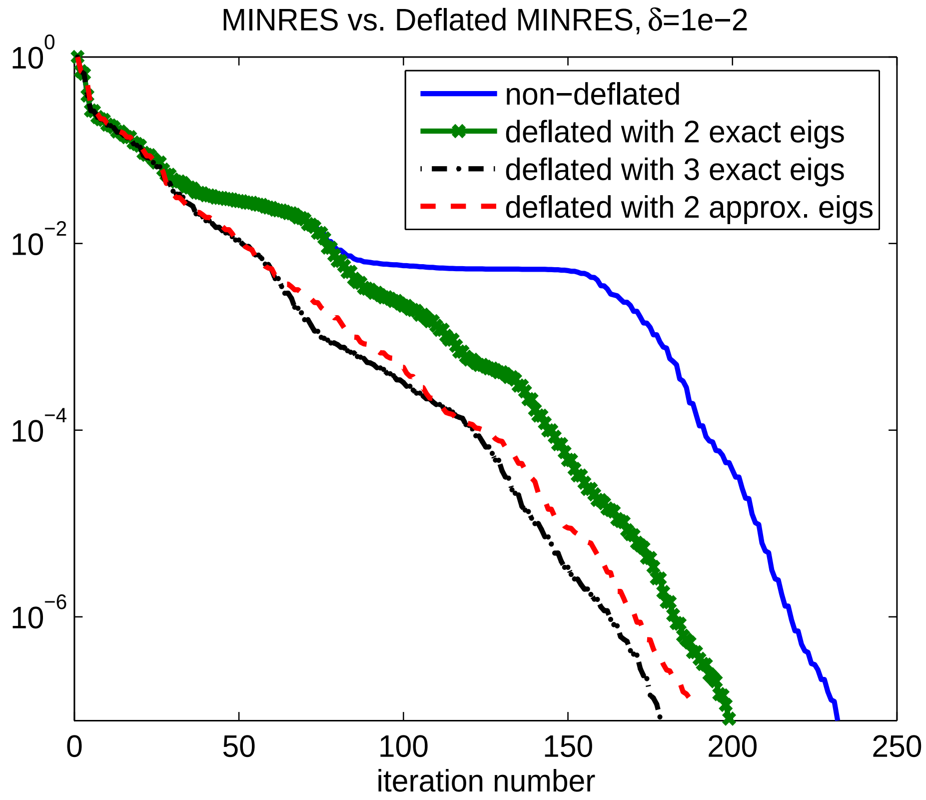
with $0 < \nu_{\min} \leq \nu(\mathbf{x}) \leq \nu_{\max} < \infty$. (Here, $\nu(\mathbf{x}) = 2\mu + \frac{\tau_s}{\sqrt{\varepsilon^2 + |\mathbf{Du(x)}|^2}}$)

$\mathbf{u}$: velocity vector field $\qquad p$: pressure

$\mathbf{Du} = \frac{1}{2}(\nabla\mathbf{u} + \nabla^T\mathbf{u})$ rate of deformation tensor;

Prec. S: pressure mass matrix wrto weighted product $(\nu^{-1}\cdot, \cdot)_{L^2(\Omega)}$

Exact and approximate eigenvectors



$\widetilde{A} = \mathrm{IC}(A, \delta), \delta = 10^{-2}$ poor approximation $\Rightarrow$ one small positive eig

Bercovier-Engelman model of the Bingham viscoplastic fluid

A stopping criterion for Stokes mixed approximation

$$\mathcal{M} = \begin{bmatrix} A & B^T \\ B & 0 \end{bmatrix}, \quad \mathcal{P} = \begin{bmatrix} \widetilde{A} & 0 \\ 0 & \widetilde{S} \end{bmatrix} \qquad \widetilde{A} \sim A, \quad \widetilde{S} \sim BA^{-1}B^T$$

For stable discretization, heuristic relation between error and residual:

$$\|x - x_j\|_{\mathcal{P}_{\text{ideal}}} \le \frac{\sqrt{2}}{\gamma^2}\|b - \mathcal{M}x_j\|_{\mathcal{P}^{-1}_{\text{ideal}}} \sim \frac{\sqrt{2}}{\gamma^2}\|b - \mathcal{M}x_j\|_{\mathcal{P}^{-1}} < tol$$

γ inf-sup constant

Estimating the inf-sup constant

For the preconditioned problem (Elman etal, '05):

$$\lambda_- \le \frac{1}{2}(\delta - \sqrt{\delta^2 + 4\delta\gamma^2}) \qquad \delta \le \lambda_+$$

with $\delta = \lambda_{\min}(A\widetilde{A}^{-1})$

If these bounds are tight (equalities), then

$$\gamma^2 = \frac{\lambda_-^2 - \lambda_-\lambda_+}{\lambda_+}$$

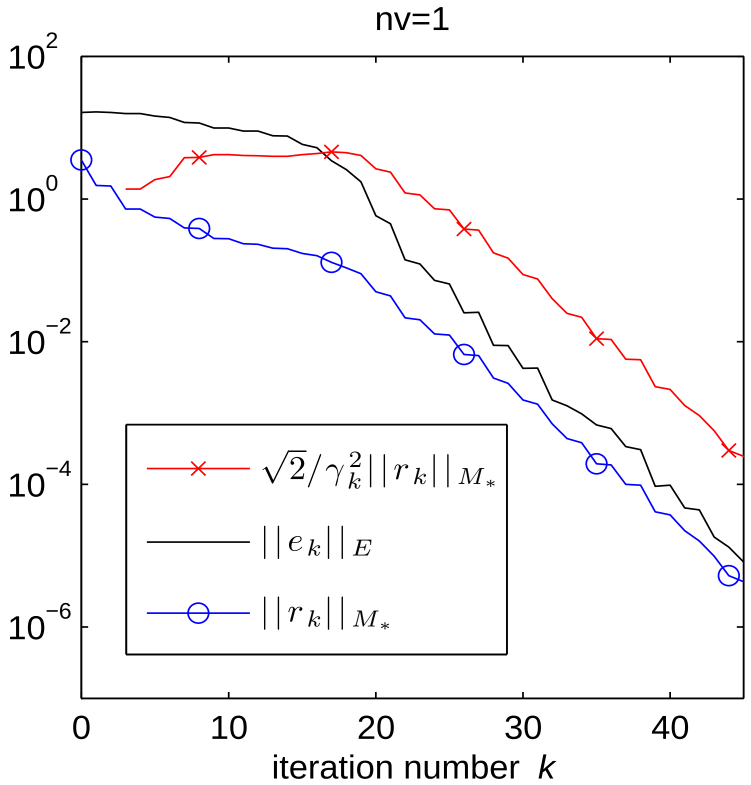
In practice, adaptive estimate with Harmonic Ritz values:

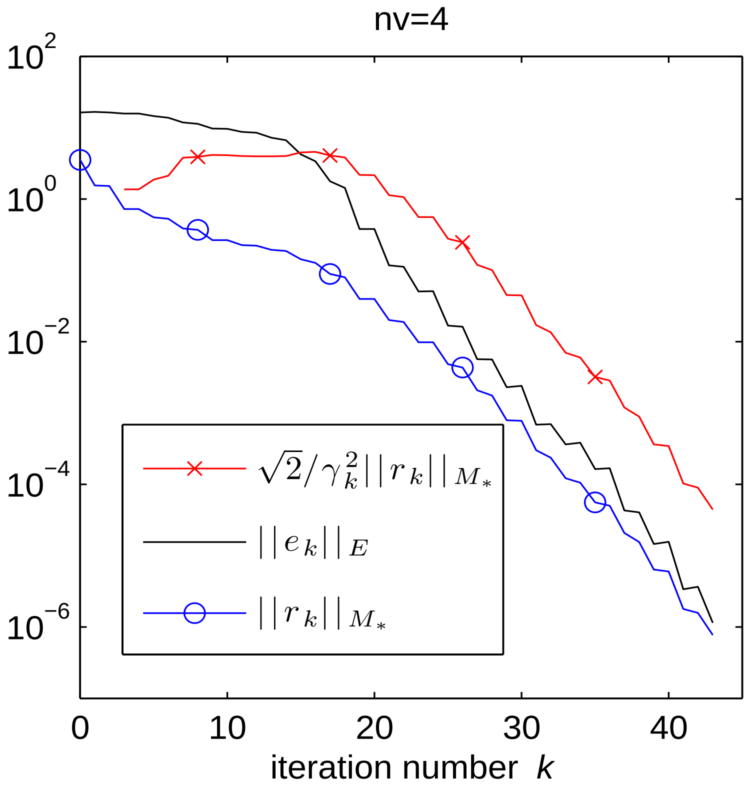
$$\gamma \approx \gamma_j^2 = \frac{(\theta_-^{(j)})^2 - \theta_-^{(j)}\theta_+^{(j)}}{\theta_+^{(j)}}, \qquad jth \text{ MINRES iteration}$$

(Silvester & Simoncini, '11)

An example

e_k: error at iteration k $\qquad$ r_k: residual at iteration k





$$E = \mathcal{P}_{\text{ideal}}, \qquad M_* = \mathcal{P}^{-1}$$

Conclusions

- MINRES effective for preconditioned sym indefinite problems
- Rich in information to be exploited
- Adaptive problem-related stopping criteria available

References:

* Maxim A. Olshanskii and V. Simoncini *Acquired clustering properties and solution of certain saddle point systems*. SIMAX, 2010.

* David J. Silvester and V. Simoncini *An Optimal Iterative Solver for Symmetric Indefinite Systems stemming from Mixed Approximation*. ACM TOMS, 2011.

* V. Simoncini and Daniel B. Szyld , unpublished, 2011.