

Acquired clustering properties of certain saddle point systems and applications

V. Simoncini

Dipartimento di Matematica, Università di Bologna and CIRSA, Ravenna, Italy valeria@dm.unibo.it

Mostly joint work with Maxim A. Olshanskii, Lomonosov Moscow State University

The problem

$$\begin{bmatrix} A & B^T \\ B & -C \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix} \qquad \mathcal{M}x = b$$

- Computational Fluid Dynamics (Elman, Silvester, Wathen 2005)
- Elasticity problems
- Mixed (FE) formulations of II and IV order elliptic PDEs
- Linearly Constrained Programs
- Linear Regression in Statistics
- Image restoration and registration
- ... Survey: Benzi, Golub and Liesen, Acta Num 2005

The problem

$$\begin{bmatrix} A & B^T \\ B & -C \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix} \qquad \mathcal{M}x = b$$

Hypotheses:

- $\star A \in \mathbb{R}^{n \times n}$ symmetric
- $\star~B^T \in \mathbb{R}^{n \times m}$ tall, $m \leq n$
- \star C symmetric positive (semi)definite

More hypotheses later on specific problems...

Outline

- Approximating the Schur complement
- Facing a spectral difficulty
- Fix during the iterative solve
- Application to
 - A Stokes type problem
 - A PDE-constrained problem

Iterative solver. Convergence considerations.

 $\mathcal{M}x = b$

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

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

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

Iterative solver. Convergence considerations.

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

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

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

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

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

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

Note: more general but less tractable bounds available







Block diagonal Preconditioner
* A nonsing.,
$$C = 0$$
:

$$\mathcal{P}_0 = \begin{bmatrix} A & 0 \\ 0 & BA^{-1}B^T \end{bmatrix}$$

$$\Rightarrow \quad \mathcal{P}_0^{-\frac{1}{2}} \mathcal{M} \mathcal{P}_0^{-\frac{1}{2}} = \begin{bmatrix} I & A^{-\frac{1}{2}}B^T (BA^{-1}B^T)^{-\frac{1}{2}} \\ (BA^{-1}B^T)^{-\frac{1}{2}}BA^{-\frac{1}{2}} & 0 \end{bmatrix}$$

MINRES converges in at most 3 iterations. $\operatorname{spec}(\mathcal{P}_0^{-\frac{1}{2}}\mathcal{MP}_0^{-\frac{1}{2}}) = \{1, \frac{1}{2}(1 \pm \sqrt{5})\}$

A more practical choice:

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

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

Approximating the Schur complement

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

 \widetilde{S} optimal^a approximation when spec $(BA^{-1}B^T\widetilde{S}^{-1})$ well clustered

Typical choices for \widetilde{S} :

- Incomplete Cholesky fact. of $BA^{-1}B^T$
- Algebraic/Geometric Multigrid method
- Operator-based approximation
-

^aWith some abuse of language

A quasi-optimal approximate Schur complement

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

For certain operators, the approximate Schur complement \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

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

For certain operators, the approximate Schur complement \widetilde{S} is quasi-optimal:

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



Possibly: well clustered eigs are mesh-independent

Questions

• Do these spectral peculiarities have an effect on the preconditioned problem \mathcal{MP}^{-1} ?

Questions

- Do these spectral peculiarities have an effect on the preconditioned problem \mathcal{MP}^{-1} ?
- Do these spectral properties influence the convergence of MINRES on \mathcal{MP}^{-1} ?

Questions

- Do these spectral peculiarities have an effect on the preconditioned problem \mathcal{MP}^{-1} ?
- Do these spectral properties influence the convergence of MINRES on \mathcal{MP}^{-1} ?
- Can we eliminate this influence?

All answers in the affirmative...

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{MP}^{-1}



Spectral intervals of preconditioned problem

 $\begin{aligned} & \operatorname{spec}(A\widetilde{A}^{-1}): & 0 < \gamma_1 \leq \cdots \leq \gamma_n \\ & \operatorname{spec}(BA^{-1}B^T\widetilde{S}^{-1}): & 0 < \mu_1 \leq \cdots \leq \mu_m \end{aligned}$ $\begin{aligned} & \operatorname{Assume that for some } \ell \ll m: \\ & 0 < \mu_1 \leq \cdots \leq \mu_\ell \leq \varepsilon_0 \ll c_0 < \mu_{\ell+1} \leq \cdots \leq \mu_m \end{aligned}$ $\begin{aligned} & \text{for some } \varepsilon_0, c_0. \end{aligned}$

Spectral intervals of preconditioned problem

spec $(A\widetilde{A}^{-1})$: $0 < \gamma_1 \leq \cdots \leq \gamma_n$ $\operatorname{spec}(BA^{-1}B^T\widetilde{S}^{-1}): \quad 0 < \mu_1 \leq \cdots \leq \mu_m$ Assume that for some $\ell \ll m$: $0 < \mu_1 \leq \cdots \leq \mu_\ell \leq \varepsilon_0 \ll c_0 < \mu_{\ell+1} \leq \cdots \leq \mu_m$ for some ε_0, c_0 . Then spec (\mathcal{MP}^{-1}) is contained in: $\left[\frac{1}{2}\left(\gamma_1 - \sqrt{\gamma_1^2 + 4\gamma_n\mu_m}\right), \frac{1}{2}\left(\gamma_n - \sqrt{\gamma_n^2 + 4\gamma_1\mu_{\ell+1}}\right) + \sqrt{\varepsilon_0\gamma_n}\right]$ $\bigcup \left[\gamma_1, \frac{1}{2} \left(\gamma_n + \sqrt{\gamma_n^2 + 4\gamma_n \mu_m} \right) \right] \bigcup \left[-\sqrt{\varepsilon_0 \gamma_n}, \ \gamma_n - \sqrt{\gamma_n^2 + 4\gamma_1 \mu_1} \right]$ IE

For sufficiently small ε_0 at most ℓ eigenvalues of \mathcal{MP}^{-1} are in I_{ε}



Eliminating the stagnation phase: "Deflated" MINRES Consider $\underbrace{L^{-1}\mathcal{M}L^{-T}}_{\widehat{\mathcal{M}}}\hat{u} = \hat{b}, \quad \mathcal{P} = LL^{T}$

Let Y be an approximate eigenbasis of $\widehat{\mathcal{M}}$.

* Approximation space: Augmented Lanczos sequence

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

obtained by standard Lanczos method with coeff.matrix

$$\mathcal{G} := \widehat{\mathcal{M}} - \widehat{\mathcal{M}} Y (Y^T \widehat{\mathcal{M}} Y)^{-1} Y^T \widehat{\mathcal{M}}, \qquad \text{giving} \quad K_j(\widehat{\mathcal{M}}, Y, v_1)$$

* MINimal RESidual method:

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

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

Augmented MINRES ("Stanford" style) Given \mathcal{M}, b , maxit, tol, \mathcal{P} , and Y with orthonormal columns $\mathbf{u} = \mathbf{Y}(\mathbf{Y}^{T}\mathcal{M}\mathbf{Y})^{-1}\mathbf{Y}^{T}\mathbf{b}$ starting approx, $r = b - \mathcal{M}u$, $r_{1} = r$, $y = \mathcal{P}^{-1}r$, etc while $(i < \max i)$ i = i + 1 $v = y/\beta;$ $\mathbf{y} = \mathcal{M}\mathbf{v} - \mathcal{M}\mathbf{Y}(\mathbf{Y}^{T}\mathcal{M}\mathbf{Y})^{-1}\mathbf{Y}^{T}\mathcal{M}\mathbf{v}$ if $i \ge 2$, $y = y - (\beta/\beta_0)r_1$ $\alpha = v^T y, \qquad y = y - r_2 \alpha / \beta$ $r_1 = r_2, \qquad r_2 = y$ $y = \mathcal{P}^{-1} r_2$ $\beta_0 = \beta, \qquad \beta = \sqrt{r_2^T y}$ $e_0 = e, \qquad \delta = c\dot{\bar{d}} + s\alpha \qquad \bar{g} = s\bar{d} - c\alpha \qquad e = s\beta \qquad \bar{d} = -c\beta$ $\gamma = \|[\bar{g}, \beta]\|$ $c = \bar{g}/\gamma$, $s = \beta/\gamma$, $\phi = c\bar{\phi}$, $\bar{\phi} = s\bar{\phi}$ $w_1 = w_2, \qquad w_2 = w$ $w = (v - e_0 w_1 - \delta w_2) \gamma^{-1}$ $\mathbf{g} = \mathbf{Y} (\mathbf{Y}^{\mathrm{T}} \mathcal{M} \mathbf{Y})^{-1} \mathbf{Y}^{\mathrm{T}} \mathcal{M} \mathbf{w} \boldsymbol{\phi}$ $u = u - \mathbf{g} + \phi w$ $\zeta = \chi_1 / \gamma, \qquad \chi_1 = \chi_2 - \delta z, \qquad \chi_2 = -e\zeta$ Check preconditioned residual norm $(\bar{\phi})$ for convergence

end

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

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

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

u : velocity vector field p : pressure $\mathbf{D}\mathbf{u} = \frac{1}{2}(\nabla \mathbf{u} + \nabla^T \mathbf{u})$ rate of deformation tensor;

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

Experiments with Bercovier-Engelman model of the Bingham viscoplastic fluid

- * One zero pressure mode (eigvec easy to approx)
- * One small eigenvalue of precon'd Schur Complement
- \Rightarrow rough eigvec approximation : $\{\widetilde{u}_2, \widetilde{p}_2\}^T \approx \{u_2, p_2\}^T$

$$\widetilde{p}_{2} = \begin{cases} 0 & \text{if } \frac{1}{2} - \tau_{s} \leq y \leq \frac{1}{2} + \tau_{s}, \\ 1 & \text{if } 0 \leq y < \frac{1}{2} - \tau_{s}, \\ -1 & \text{if } 1 \geq y > \frac{1}{2} + \tau_{s}, \end{cases} \quad \text{and} \quad \widetilde{u}_{2} = -\widetilde{A}^{-1}B^{T}\widetilde{p}_{2}$$

Bercovier-Engelman model of the Bingham viscoplastic fluid $\widetilde{A} = \mathrm{IC}(A, \delta)$







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

A parameter identification problem

$$\min_{q} \frac{1}{2} ||F(q) - z||^2 + \alpha \mathcal{J}_{reg}(q)$$
 (1)

$$\mathcal{B}(q)u \equiv -\nabla \cdot (\sigma \nabla u) = f \tag{2}$$

(2): model for groundflow

u: fluid pressure $\sigma(x)$: (spatially dep.) hydraulic conductivity f(x): in/out-going fluid (incompressible flows) Parameter of interest: $q(x) = \log(\sigma(x))$, obs from soln: $u_{obs} = Cu$ (1): $F(q) = C\mathcal{B}(q)^{-1}f$ (parameter-to-obs map, non-linear function) \mathcal{J}_{reg} regularization functional (e.g. total variation) PDE-constrained formulation

$$\min_{q,u} \frac{1}{2} ||Cu - z||^2 + \alpha \mathcal{J}_{reg}(q)$$
$$\mathcal{B}(q)u - f = 0.$$

Space discretization + inexact Newton method provide linear systems:

$$\begin{bmatrix} H & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} v \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} -L_v \\ -L_{\boldsymbol{\lambda}} \end{bmatrix}$$

- H: Hessian of the operator
- $B{:}\ {\sf Jacobian}\ {\sf of}\ {\sf the}\ {\sf constraint}$
- v = [q, u] variables, λ : Lagrange multipliers

A similar setting: Monge-Kantorovich mass transfer problem

Pb: Given two density functions u_0 and u_T on the set Ω , find an "optimal" mapping from u_0 to u_T

Formulation (time in [0, T]):

$$\min_{u,m} \frac{1}{2} \|u(T, \mathbf{x}) - u_T(\mathbf{x})\|^2 + \frac{1}{2} \alpha T \int_{\Omega} \int_0^T u \|m\|^2 dt d\mathbf{x}$$

s.t. $u_t + \nabla \cdot (um) = 0, \quad u(0, \mathbf{x}) = u_0$

 $u(t, \mathbf{x})$: density field $m(t, \mathbf{x})$: velocity field

A preconditioning technique for a class of PDE-constrained optimization problems M. BENZI, E. HABER AND L. TARALLI, Adv. in Comput. Math. '10 A similar setting: Monge-Kantorovich mass transfer problem

Time and space discretization + Gauss-Newton approximation on the Lagrangian

Sequence of "Newton step-depending" linear systems:

$$\begin{bmatrix} Q^T Q & 0 & B_1^T \\ 0 & L & B_2^T \\ B_1 & B_2 & 0 \end{bmatrix} \begin{bmatrix} \widetilde{u}_k \\ \widetilde{m}_k \\ \widetilde{p}_k \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

with $Q^T Q$ diagonal and highly singular, L > 0 diagonal B_2 rank deficient

Reduced order problem

$$\begin{bmatrix} Q^T Q & 0 & B_1^T \\ 0 & L & B_2^T \\ B_1 & B_2 & 0 \end{bmatrix} \begin{bmatrix} \widetilde{u}_k \\ \widetilde{m}_k \\ \widetilde{p}_k \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

into

$$\begin{bmatrix} Q^T Q & B_1^T \\ B_1 & -B_2 L^{-1} B_2^T \end{bmatrix} \begin{bmatrix} \widetilde{u}_k \\ \widetilde{p}_k \end{bmatrix} = \begin{bmatrix} b_2 \\ \widetilde{b}_3 \end{bmatrix}$$

Both $Q^T Q$, $B_2 L^{-1} B_2^T$ pos. semi-definite, B_1 square nonsing.

Augmented Block Diagonal Preconditioning

$$\mathcal{M} = \begin{bmatrix} Q^T Q & B_1^T \\ B_1 & -B_2 L^{-1} B_2^T \end{bmatrix}$$

Both diagonal blocks are singular. Augmented preconditioning:

$$\mathcal{P}_{ad} = \begin{bmatrix} D & 0 \\ 0 & C(D) \end{bmatrix}, \qquad \begin{array}{l} D > 0 \\ C(D) \approx B_2 L^{-1} B_2^T + B_1 D^{-1} B_1^T \end{array}$$

Work in Progress

Exact preconditioner

$$\mathcal{P}_{ad} = \begin{bmatrix} D & 0 \\ 0 & B_2 L^{-1} B_2^T + B_1 D^{-1} B_1^T \end{bmatrix} \qquad \begin{array}{l} D = Q^T Q + \gamma \mathcal{N}(Q^T Q) \\ \gamma = \|L\| \\ \end{array}$$

n_x	n_t	n	#	CPU time	n_x	n_t	n	#	CPU time
20	20	8000	14	8.17	30	10	9000	21	34.16
25	25	15625	15	33.31	40	20	32000	21	376.54
30	30	27000	16	109.25	50	20	50000	-	-
35	35	42875	16	309.97					
40	40	64000	-	-					

Practical preconditioner

$$\mathcal{P}_{ad} = \begin{bmatrix} D & 0\\ 0 & C(D) \end{bmatrix}$$

$$D: \quad D = Q^T Q + \gamma \mathcal{N}(Q^T Q), \qquad \gamma = \|L\|$$

C(D): Algebraic Multigrid of $B_2L^{-1}B_2^T + B_1D^{-1}B_1^T$ (routine HSL_MI20)



Approximate smallest eigenvalues

C(D): Algebraic Multigrid of $B_2L^{-1}B_2^T + B_1D^{-1}B_1^T$ Eigs of "preconditioned augmented Schur complement"

 $\operatorname{spec}((B_2L^{-1}B_2^T + B_1D^{-1}B_1^T)C(D)^{-1})$

approx λ_i	$n_x = 10$	$n_x = 20$	$n_x = 30$	
i=1	9.4632 e-04	1.7676e-05	1.3247 e-05	
i=2	9.4999e-04	1.8274e-05	1.3937 e- 05	
i=3	8.5302e-01	$1.5449\mathrm{e}\text{-}04$	4.6116e-05	
i=4	8.5400e-01	$1.5484\mathrm{e} extsf{-}04$	4.6200 e- 05	
i=5	8.5722e-01	6.2973e-01	3.7707e-01	
i=6	8.5752e-01	6.3079e-01	3.7791e-01	
i=7	8.6410e-01	6.6089e-01	3.8408e-01	
i=8	8.6825e-01	6.6091e-01	3.8409e-01	
i=9	8.7875e-01	6.6557e-01	4.0425e-01	
i=10	8.8097e-01	6.6674e-01	4.0514e-01	



(at first Newton step)

Complete timings

n	\mathcal{P}_{ad}		\mathcal{P}_{ad} w/AMG		\mathcal{P}_{ad} w/AMG+DEFL			
	# its	time	# its	time	# its	time		
1000	12	0.19	24	0.29	17	0.26	+ 0.52, 2 eigs	
3375	13	1.28	52	1.43	25	0.80	+ 1.53, 5 eigs	
8000	14	8.17	57	4.64	21	1.95	+ 4.51, 4 eigs	
15625	15	33.31	85	14.64	36	6.68	+ 9.57, 5 eigs	
27000	16	109.25	81	28.37	28	10.61	+19.03, 4 eigs	
42875	16	309.97	122	65.82	47	27.33	+29.82, 5 eigs	
64000	-		90	88.63	37	37.82	+52.23, 4 eigs	

Conclusions

- Construction of Schur complement type preconditioner is spectrally crucial
- An effective way to overcome an "isolated" difficulty

Bibliography:

Acquired clustering properties and solution of certain saddle point systems,

Maxim A. Olshanskii and V. Simoncini

To appear in SIAM J. Matrix Analysis and Appl.

see also www.dm.unibo.it://~ simoncin