The Bramble-Pasciak^{+} preconditioner for saddle point problems

Martin Stoll & Andy Wathen

Numerical Analysis Day
Oxford, April 13, 2007

The linear system

The Problem

We want to solve $\mathcal{A}x = b$ where

$$\underbrace{\begin{bmatrix} A & B^T \\ B & -C \end{bmatrix}}_{\mathcal{A}} \tag{1}$$

with $A \in \mathbb{R}^{n,n}$ symmetric and positive definite and $C \in \mathbb{R}^{m,m}$ symmetric negative semi-definite. $B \in \mathbb{R}^{m,n}$ is assumed to have full rank.

Saddle point problems

Saddle point problems arise in a variety of applications such as

- Mixed finite element methods for Fluid and Solid mechanics
- Interior point methods in optimisation

See Benzi, Golub, Liesen (2005), Elman, Silvester, Wathen (2005), Brezzi, Fortin (1991), Nocedal, Wright (1999)

Saddle point problems provide due to their indefiniteness and often poor spectral properties a challenge for people developing solvers.
Benzi, Golub, Liesen (2005)

Saddle point problems

Saddle point problems arise in a variety of applications such as

- Mixed finite element methods for Fluid and Solid mechanics
- Interior point methods in optimisation

See Benzi, Golub, Liesen (2005), Elman, Silvester, Wathen (2005), Brezzi, Fortin (1991), Nocedal, Wright (1999)

Saddle point problems provide due to their indefiniteness and often poor spectral properties a challenge for people developing solvers.
Benzi, Golub, Liesen (2005)

Some Background – Basic relations

We introduce the bilinear form induced by $\mathcal{H}$

$$\langle x, y \rangle_{\mathcal{H}} := x^T \mathcal{H} y$$

which is an inner product iff $\mathcal{H}$ is positive definite. A matrix $\mathcal{A} \in \mathbb{R}^{n \times n}$ is self-adjoint in $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ if and only if

$$\langle \mathcal{A}x, y \rangle_{\mathcal{H}} = \langle x, \mathcal{A}y \rangle_{\mathcal{H}} \qquad \text{for all } x, y$$

which can be reformulated to

$$\mathcal{A}^T \mathcal{H} = \mathcal{H} \mathcal{A}.$$

Some Background – Solvers

- CG needs symmetry in $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ **plus** positive definiteness in $\langle \cdot, \cdot \rangle_{\mathcal{H}}$
- MINRES needs the symmetry $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ but **no** definiteness in $\langle \cdot, \cdot \rangle_{\mathcal{H}}$

Spectral properties of $\mathcal{A}$ can be enhanced by preconditioning, ie. considering

$$\widehat{\mathcal{A}} = \mathcal{P}^{-1}\mathcal{A}$$

instead of $\mathcal{A}$.

Original matrix $\mathcal{A}$ is symmetric in $\langle \cdot, \cdot \rangle_I \Rightarrow$ MINRES can be used.

What about the symmetry of $\widehat{\mathcal{A}}$?

The Bramble-Pasciak CG

We consider saddle point problem

$$\mathcal{A} = \begin{bmatrix} A & B^T \\ B & -C \end{bmatrix}$$

with a block-triangular preconditioner

$$\mathcal{P} = \begin{bmatrix} A_0 & 0 \\ B & -I \end{bmatrix}$$

which results in

$$\widehat{\mathcal{A}} = \mathcal{P}^{-1}\mathcal{A} = \begin{bmatrix} A_0^{-1}A & A_0^{-1}B^T \\ BA_0^{-1}A - B & BA_0^{-1}B^T + C \end{bmatrix}.$$

The Bramble-Pasciak CG

The preconditioned matrix

$$\widehat{\mathcal{A}} = \mathcal{P}^{-1}\mathcal{A} = \begin{bmatrix} A_0^{-1}A & A_0^{-1}B^T \\ BA_0^{-1}A - B & BA_0^{-1}B^T + C \end{bmatrix}$$

is self-adjoint in the bilinear form defined by

$$\mathcal{H} = \begin{bmatrix} A - A_0 & 0 \\ 0 & I \end{bmatrix}.$$

Under certain conditions for A_0 $\mathcal{H}$ defines an inner product and $\widehat{\mathcal{A}}$ is also positive definite in this inner product, e.g. $A_0 = .5A$.

The condition for A_0 usually involves the solution of an eigenvalue problem which can be expensive.

The Bramble-Pasciak^{+} CG

We always want an inner product for symmetric and positive definite A_0

$$\mathcal{H}^{+} = \begin{bmatrix} A + A_0 & \\ & I \end{bmatrix}.$$

Therefore, new preconditioner $\mathcal{P}^{+}$

$$\mathcal{P}^{+} = \begin{bmatrix} A_0 & 0 \\ -B & I \end{bmatrix}$$

is required. The preconditioned matrix

$$\widehat{\mathcal{A}} = \left(\mathcal{P}^{+}\right)^{-1}\mathcal{A} = \begin{bmatrix} A_0^{-1}A & A_0^{-1}B^T \\ BA_0^{-1}A + B & BA_0^{-1}B^T - C \end{bmatrix}$$

is self-adjoint in this inner product.

Definiteness in $\mathcal{H}^+$

If we split

$$\widehat{\mathcal{A}}^T\mathcal{H}^+ = \begin{bmatrix} AA_0^{-1}A + A & AA_0^{-1}B^T + B^T \\ BA_0^{-1}A + B & BA_0^{-1}B^T - C \end{bmatrix}$$

as

$$\begin{bmatrix} I & \\ BA^{-1} & I \end{bmatrix} \begin{bmatrix} AA_0^{-1}A + A & \\ & -BA_0^{-1}B^T - C \end{bmatrix} \begin{bmatrix} I & A^{-1}B^T \\ & I \end{bmatrix}$$

we see that since this is a congruence transformation the matrix is always indefinite. This means:

- No reliable CG can be applied
- In practice CG quite often works fine
- Augmented methods might be used.

An $\mathcal{H}^+$-inner product implementation of MINRES

Use that $\widehat{\mathcal{A}}$ symmetric in $\mathcal{H}$-inner product and therefore implement a version of **Lanczos process** with $\mathcal{H}$-inner product which gives

$$\widehat{\mathcal{A}} V_k = V_k T_k + \beta_k v_{k+1} e_k^T$$

with

$$T_k = \begin{bmatrix} \alpha_1 & \beta_1 & & \\ \beta_1 & \ddots & \ddots & \\ & \ddots & \ddots & \beta_{k-1} \\ & & \beta_{k-1} & \alpha_k \end{bmatrix} \text{ and } V_k = [v_1, v_2, \ldots, v_k]$$

as well as $V_k^T \mathcal{H}^+ V_k = I$.

An $\mathcal{H}^+$-inner product implementation of MINRES

The following condition holds for the residual

$$\begin{aligned}
\|r_k\|_{\mathcal{H}^+} &= \|b - Ax_k\|_{\mathcal{H}} &&= \|b - Ax_0 - AV_k y_k\|_{\mathcal{H}^+} \\
&= \|r_0 - V_{k+1} T_{k+1} y_k\|_{\mathcal{H}^+} &&= \left\| V_{k+1}(V_{k+1}^T \mathcal{H}^+ r_0 - T_{k+1} y_k) \right\|_{\mathcal{H}^+} \\
&= \left\| V_{k+1}^T \mathcal{H}^+ r_0 - T_{k+1} y_k \right\|_{\mathcal{H}^+} &&= \| \|r_0\| e_1 - T_{k+1} y_k \|_{\mathcal{H}^+} .
\end{aligned}$$

Minimizing $\| \|r_0\| e_1 - T_{k+1} y_k \|_{\mathcal{H}^+}$ can be done by the standard updated-QR factorization technique. Implementation details can be found in Greenbaum (1997).

The simplified Lanczos method

The non-symmetric Lanczos process generates two sequences of vectors where the following condition holds

$$v_j = \phi_j(\widehat{\mathcal{A}})v_1 \text{ and } w_j = \gamma_j \phi_j(\widehat{\mathcal{A}}^T)w_1$$

where ϕ is a polynomial of degree $j-1$ the so-called *Lanczos polynomial*. Setting $w_1 = \mathcal{H}v_1$ and using the self-adjointness of $\widehat{\mathcal{A}}$ in $\mathcal{H}^+$, ie. $\widehat{\mathcal{A}}^T\mathcal{H}^+ = \mathcal{H}^+\widehat{\mathcal{A}}$, gives

$$w_j = \gamma_j\phi_j(\widehat{\mathcal{A}}^T)w_1 = \gamma_j\phi_j(\widehat{\mathcal{A}}^T)\mathcal{H}^+v_1 = \gamma_j\mathcal{H}^+\phi_j(\widehat{\mathcal{A}})v_1 = \gamma_j\mathcal{H}^+v_j.$$

Therefore the non-symmetric Lanczos process can be simplified, ie. multiplications with $\widehat{\mathcal{A}}^T$ can be exchanged for multiplication by $\mathcal{H}^+$.

The ideal transpose-free QMR method (ITFQMR)

Based on the QMR method Freund (1994) a transpose-free QMR method with an implementation derived from the BICG procedure. Here, we use matrix formulation of the non-symmetric Lanczos process

$$\widehat{\mathcal{A}} V_k = V_{k+1} H_k$$

and

$$r_k = V_{k+1}(\|r_0\| \, e_1 - H_k y_k).$$

Ignoring the term V_{k+1} gives QMR method. Using simplification of the Lanczos process gives ITFQMR .

Eigenvalue analysis for $A_0 = A$

To get some insight into the convergence behaviour we the eigenvalues of

$$\widehat{\mathcal{A}} = \left(\mathcal{P}^{+}\right)^{-1} \mathcal{A} = \begin{bmatrix} I & A^{-1}B^T \\ 2B & BA^{-1}B^T \end{bmatrix}.$$

For the eigenpair $(\lambda, \begin{bmatrix} x \\ y \end{bmatrix})$ of $\widehat{\mathcal{A}}$ we know that

$$\begin{bmatrix} I & A^{-1}B^T \\ 2B & BA^{-1}B^T \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + A^{-1}B^T y \\ 2Bx + BA^{-1}B^T y \end{bmatrix} = \lambda \begin{bmatrix} x \\ y \end{bmatrix}$$

For $\lambda = 1$ we get

$$Ax + B^T y = Ax$$

which gives $B^T y = 0$ and $y = 0$ iff $Bx = 0$.

Since $\dim(\ker(B)) = n - m$ multiplicity of $\lambda = 1$ is $n - m$.

Eigenvalue analysis for $A_0 = A$

For $\lambda \neq 1$, we get that $x = \frac{1}{\lambda-1} A^{-1} B^T y$ which gives

$$BA^{-1}B^T y = \frac{\lambda(\lambda-1)}{\lambda+1} y.$$

For an eigenvalue σ of $BA^{-1}B^T$ we get

$$\sigma = \frac{\lambda(\lambda-1)}{\lambda+1}.$$

Eigenvalues of $\widehat{\mathcal{A}}$ become

$$\lambda_{1,2} = \frac{1+\sigma}{2} \pm \sqrt{\frac{(1+\sigma)^2}{4} + \sigma}.$$

Since $\sigma > 0$ we have m negative eigenvalues.

Eigenvalue analysis for $A_0 = A$

For $\lambda \neq 1$, we get that $x = \frac{1}{\lambda-1}A^{-1}B^Ty$ which gives

$$BA^{-1}B^Ty = \frac{\lambda(\lambda-1)}{\lambda+1}y.$$

For an eigenvalue σ of $BA^{-1}B^T$ we get

$$\sigma = \frac{\lambda(\lambda-1)}{\lambda+1}.$$

Eigenvalues of $\widehat{\mathcal{A}}$ become

$$\lambda_{1,2} = \frac{1+\sigma}{2} \pm \sqrt{\frac{(1+\sigma)^2}{4} + \sigma}.$$

Since $\sigma > 0$ we have m negative eigenvalues.

Eigenvalue analysis for $A_0 = A$

For $\lambda \neq 1$, we get that $x = \frac{1}{\lambda-1}A^{-1}B^Ty$ which gives

$$BA^{-1}B^Ty = \frac{\lambda(\lambda-1)}{\lambda+1}y.$$

For an eigenvalue σ of $BA^{-1}B^T$ we get

$$\sigma = \frac{\lambda(\lambda-1)}{\lambda+1}.$$

Eigenvalues of $\widehat{\mathcal{A}}$ become

$$\lambda_{1,2} = \frac{1+\sigma}{2} \pm \sqrt{\frac{(1+\sigma)^2}{4}+\sigma}.$$

Since $\sigma > 0$ we have m negative eigenvalues.

Numerical Experiments – Stokes problem

We are going to solve saddle point systems coming from the finite element method for the Stokes problem

$$\begin{aligned} -\nabla^2 u + \nabla p &= 0 \\ \nabla \cdot u &= 0 \end{aligned}$$

The linear system governing the finite element method for the Stokes problem is a saddle point problem

$$\begin{bmatrix} A & B^T \\ B & -C \end{bmatrix}$$

where $C \neq 0$ for stabilized systems. In our examples $C = 0$.

All examples come from the IFISS package.

Block diagonal preconditioning

Silvester and Wathen (1993,1994) use preconditioner

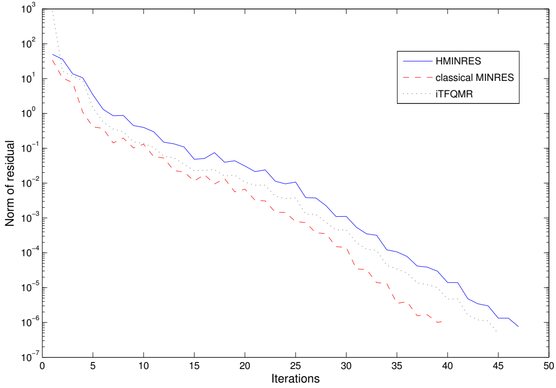
$$\mathcal{P} = \begin{bmatrix} A_0 & 0 \\ 0 & S_0 \end{bmatrix}$$

which is symmetric positive definite.

$\Longrightarrow$ Preconditioned MINRES can be applied.

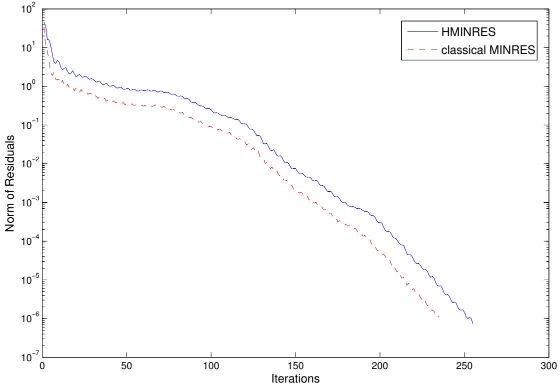
Example 1 – Stokes problem Channel domain

Results for $\mathcal{H}$-MINRES and classical Preconditioned MINRES with problem dimension 9539. Preconditioner $A_0 = A$ and S_0 being the Gramian (Mass matrix).



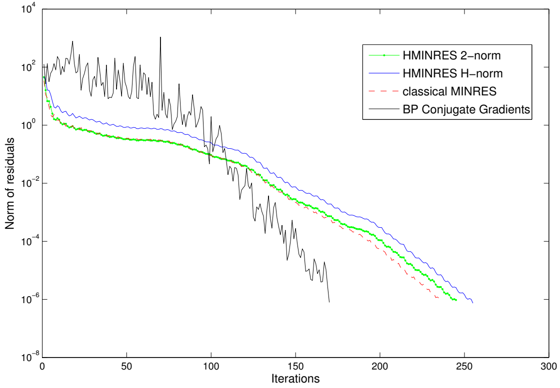
Example 2 – Stokes problem Channel domain

Results for $\mathcal{H}$-MINRES and classical Preconditioned MINRES with problem dimension 9539. Preconditioner A_0 is Incomplete Cholesky of A and S_0 being the Gramian (Mass matrix).



Example 3 – Stokes problem Channel domain

Again $\mathcal{H}$-MINRES and classical Preconditioned MINRES for problem dimension 9539. Preconditioner A_0 is Incomplete Cholesky of A and S_0 being the Gramian (Mass matrix). Additionally, $\mathcal{H}$-MINRES residual in the 2-norm and the classical Bramble-Pasciak CG.



Conclusions

- We presented a alternative approach
- Method could be used with augmented techniques to become competitive
- Presented algorithm could be used for combination preconditioning

Conclusions

- We presented a alternative approach
- Method could be used with augmented techniques to become competitive
- Presented algorithm could be used for combination preconditioning

Thank you for your attention!

Difficult questions can be discussed in 20 minutes in the pub!