The Bramble-Pasciak\textsuperscript{+} 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 $Ax = b$ where

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

(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 arise in a variety of applications such as

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


*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


_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 $A \in \mathbb{R}^{n \times n}$ is self-adjoint in $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ if and only if

$$\langle Ax, y \rangle_{\mathcal{H}} = \langle x, Ay \rangle_{\mathcal{H}} \quad \text{for all } x, y$$

which can be reformulated to

$$A^T \mathcal{H} = \mathcal{H} A.$$
Some Background – Solvers

- CG needs symmetry in $\langle \cdot, \cdot \rangle_H$ plus positive definiteness in $\langle \cdot, \cdot \rangle_H$
- MINRES needs the symmetry $\langle \cdot, \cdot \rangle_H$ but no definiteness in $\langle \cdot, \cdot \rangle_H$

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

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

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 $\tilde{\mathcal{A}}$?
The Bramble-Pasciak CG

We consider saddle point problem

\[ A = \begin{bmatrix} A & B^T \\ B & -C \end{bmatrix} \]

with a block-triangular preconditioner

\[ P = \begin{bmatrix} A_0 & 0 \\ B & -I \end{bmatrix} \]

which results in

\[ \hat{A} = \hat{P}^{-1}A = \begin{bmatrix} A_0^{-1}A & A_0^{-1}B^T \\ BA_0^{-1}A - B & BA_0^{-1}B^T + C \end{bmatrix}. \]
The Bramble-Pasciak CG

The preconditioned matrix

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

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 $\hat{A}$ 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 $P^+$

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

is required. The preconditioned matrix

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

is self-adjoint in this inner product.
Definiteness in $\mathcal{H}^+$

If we split

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

as

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

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

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

with

$$T_k = \begin{bmatrix}
\alpha_1 & \beta_1 \\
\beta_1 & \ddots & \ddots \\
& \ddots & \ddots & \beta_{k-1} \\
& & \beta_{k-1} & \alpha_k 
\end{bmatrix}$$

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 \textsc{minres}

The following condition holds for the residual

\[
\|r_k\|_{\mathcal{H}^+} = \|b - Ax_k\|_{\mathcal{H}} = \|b - A\mathbf{x}_0 - AV_ky_k\|_{\mathcal{H}^+}
\]
\[
= \|r_0 - V_{k+1}T_{k+1}y_k\|_{\mathcal{H}^+} = \|V_{k+1}(V_{k+1}^T\mathcal{H}^+r_0 - T_{k+1}y_k)\|_{\mathcal{H}^+}
\]
\[
= \|V_{k+1}^T\mathcal{H}^+r_0 - T_{k+1}y_k\|_{\mathcal{H}^+} = \|\|r_0\|e_1 - T_{k+1}y_k\|_{\mathcal{H}^+}.
\]

Minimizing $\|\|r_0\|e_1 - T_{k+1}y_k\|_{\mathcal{H}^+}$ can be done by the standard updated-\textsc{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

\[ v_j = \phi_j(\hat{A})v_1 \]  \[ w_j = \gamma_j\phi_j(\hat{A}^T)w_1 \]

where \( \phi \) 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 \( \hat{A} \) in \( \mathcal{H}^+ \), ie. \( \hat{A}^T\mathcal{H}^+ = \mathcal{H}^+\hat{A} \), gives

\[ w_j = \gamma_j\phi_j(\hat{A}^T)w_1 = \gamma_j\phi_j(\hat{A}^T)\mathcal{H}^+v_1 = \gamma_j\mathcal{H}^+\phi_j(\hat{A})v_1 = \gamma_j\mathcal{H}^+v_j. \]

Therefore the non-symmetric Lanczos process can be simplified, ie. multiplications with \( \hat{A}^T \) 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

\[ \hat{A}V_k = V_{k+1}H_k \]

and

\[ r_k = V_{k+1}(\|r_0\|e_1 - H_ky_k). \]

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

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

For the eigenpair $(\lambda, \begin{bmatrix} x \\ y \end{bmatrix})$ of $\hat{A}$ 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

$$Ax + B^Ty = Ax$$

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

Since $\text{dim}(\text{ker}(B)) = n - m$ 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 $\sigma$ of $BA^{-1}B^T$ we get

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

Eigenvalues of $\hat{A}$ become

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

Since $\sigma > 0$ we have $m$ negative eigenvalues.
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 $\sigma$ of $BA^{-1} B^T$ we get

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

Eigenvalues of $\hat{A}$ become

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

Since $\sigma > 0$ we have $m$ negative eigenvalues.
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 $\sigma$ of $BA^{-1} B^T$ we get

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

Eigenvalues of $\hat{A}$ 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

\[-\nabla^2 u + \nabla p = 0\]
\[\nabla \cdot u = 0\]

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

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

which is symmetric positive definite.

$\implies$ 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 $H$-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, $H$-MINRES residual in the 2-norm and the classical Bramble-Pasciak CG.
Conclusions

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

- We presented an alternative approach.
- The method could be used with augmented techniques to become competitive.
- The presented algorithm could be used for combination preconditioning.

Thank you for your attention!

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