The Bramble-Pasciak^{+} preconditioner and combination preconditioning

(Relaxing Conditions for the Bramble Pasciak CG)

Martin Stoll & Andy Wathen

Computational Methods with Applications,Harrachov, Czech Republic, August 19 - 25, 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 positive 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 $V_k^T \mathcal{H}^+ V_k = I$. The following condition holds for the residual

$$\|r_k\|_{\mathcal{H}^+} = \| \|r_0\| \, e_1 - T_{k+1} y_k \|_{\mathcal{H}^+} \, .$$

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

The ideal transpose-free QMR method (ITFQMR)

The matrix formulation of the non-symmetric Lanczos process

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

gives

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

Ignoring the term V_{k+1} gives QMR method.
Using $\widehat{\mathcal{A}}^T \mathcal{H}^+ = \mathcal{H}^+ \widehat{\mathcal{A}}$ a simplified Lanczos process can be implemented and we obtain

$$w_j = \gamma_j \mathcal{H}^+ v_j.$$

This is the basis for the ideal transpose-free QMR (ITFQMR) see Freund (1994).

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 $x = \frac{1}{\lambda - 1} A^{-1} B^T y$ 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 $x = \frac{1}{\lambda-1}A^{-1}B^T y$ 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.

Basic properties I

Lemma 1

If $\mathcal{A}_1$ and $\mathcal{A}_2$ are self-adjoint in $\langle\cdot,\cdot\rangle_{\mathcal{H}}$ then for any $\alpha,\beta\in\mathbb{R}$, $\alpha\mathcal{A}_1+\beta\mathcal{A}_2$ is self-adjoint in $\langle\cdot,\cdot\rangle_{\mathcal{H}}$.

Lemma 2

If $\mathcal{A}$ is self-adjoint in $\langle\cdot,\cdot\rangle_{\mathcal{H}_1}$ and in $\langle\cdot,\cdot\rangle_{\mathcal{H}_2}$ then $\mathcal{A}$ is self-adjoint in $\langle\cdot,\cdot\rangle_{\alpha\mathcal{H}_1+\beta\mathcal{H}_2}$ for every $\alpha,\beta\in\mathbb{R}$.

Lemma 3

For symmetric $\mathcal{A}$, $\widehat{\mathcal{A}}=\mathcal{P}^{-1}\mathcal{A}$ is self-adjoint in $\langle\cdot,\cdot\rangle_{\mathcal{H}}$ if and only if $\mathcal{P}^{-T}\mathcal{H}$ is self-adjoint in $\langle\cdot,\cdot\rangle_{\mathcal{A}}$.

Basic properties I

Lemma 1

If $\mathcal{A}_1$ and $\mathcal{A}_2$ are self-adjoint in $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ then for any $\alpha, \beta \in \mathbb{R}$, $\alpha\mathcal{A}_1 + \beta\mathcal{A}_2$ is self-adjoint in $\langle \cdot, \cdot \rangle_{\mathcal{H}}$.

Lemma 2

If $\mathcal{A}$ is self-adjoint in $\langle \cdot, \cdot \rangle_{\mathcal{H}_1}$ and in $\langle \cdot, \cdot \rangle_{\mathcal{H}_2}$ then $\mathcal{A}$ is self-adjoint in $\langle \cdot, \cdot \rangle_{\alpha\mathcal{H}_1 + \beta\mathcal{H}_2}$ for every $\alpha, \beta \in \mathbb{R}$.

Lemma 3

For symmetric $\mathcal{A}$, $\widehat{\mathcal{A}} = \mathcal{P}^{-1}\mathcal{A}$ is self-adjoint in $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ if and only if $\mathcal{P}^{-T}\mathcal{H}$ is self-adjoint in $\langle \cdot, \cdot \rangle_{\mathcal{A}}$.

Basic properties I

Lemma 1

If $\mathcal{A}_1$ and $\mathcal{A}_2$ are self-adjoint in $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ then for any $\alpha, \beta \in \mathbb{R}$, $\alpha\mathcal{A}_1 + \beta\mathcal{A}_2$ is self-adjoint in $\langle \cdot, \cdot \rangle_{\mathcal{H}}$.

Lemma 2

If $\mathcal{A}$ is self-adjoint in $\langle \cdot, \cdot \rangle_{\mathcal{H}_1}$ and in $\langle \cdot, \cdot \rangle_{\mathcal{H}_2}$ then $\mathcal{A}$ is self-adjoint in $\langle \cdot, \cdot \rangle_{\alpha\mathcal{H}_1 + \beta\mathcal{H}_2}$ for every $\alpha, \beta \in \mathbb{R}$.

Lemma 3

For symmetric $\mathcal{A}$, $\widehat{\mathcal{A}} = \mathcal{P}^{-1}\mathcal{A}$ is self-adjoint in $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ if and only if $\mathcal{P}^{-T}\mathcal{H}$ is self-adjoint in $\langle \cdot, \cdot \rangle_{\mathcal{A}}$.

Basic properties II

Lemma 4

If $\mathcal{P}_1$ and $\mathcal{P}_2$ are left preconditioners for the symmetric matrix $\mathcal{A}$ for which symmetric matrices $\mathcal{H}_1$ and $\mathcal{H}_2$ exist with $\mathcal{P}_1^{-1}\mathcal{A}$ self-adjoint in $\langle\cdot,\cdot\rangle_{\mathcal{H}_1}$ and $\mathcal{P}_2^{-1}\mathcal{A}$ self-adjoint in $\langle\cdot,\cdot\rangle_{\mathcal{H}_2}$ and if

$$\alpha\mathcal{P}_1^{-T}\mathcal{H}_1 + \beta\mathcal{P}_2^{-T}\mathcal{H}_2 = \mathcal{P}_3^{-T}\mathcal{H}_3$$

for some matrix $\mathcal{P}_3$ and some symmetric matrix $\mathcal{H}_3$ then $\mathcal{P}_3^{-1}\mathcal{A}$ is self-adjoint in $\langle\cdot,\cdot\rangle_{\mathcal{H}_3}$.

This Lemma shows that if we can find such a splitting we have found a new preconditioner and a bilinear form in which the matrix is self-adjoint.

(St. & Wathen 2007, Oxford preprint).

Some examples–BP with Schur complement preconditioner

For the Bramble-Pasciak technique an extensions, see Klawonn (1998), Meyer et al. (2001), Simoncini (2001) include the preconditioner

$$\mathcal{P}^{-1} = \begin{bmatrix} A_0^{-1} & 0 \\ S_0^{-1} B A_0^{-1} & -S_0^{-1} \end{bmatrix}$$

where S_0 is a Schur complement preconditioner. The inner product then becomes

$$\mathcal{H} = \begin{bmatrix} A - A_0 & 0 \\ 0 & S_0 \end{bmatrix}.$$

Some examples–Benzi-Simoncini CG ($C = 0$)

Introduced by Benzi and Simoncini (2006) it is an extension to the CG method of Fischer et. al. (1998) with the preconditioner

$$\mathcal{P}^{-1} = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}$$

and inner product matrix

$$\mathcal{H} = \begin{bmatrix} A - \gamma I & B^T \\ B & \gamma I \end{bmatrix}$$

Some examples–Extensions for $C \neq 0$

The Benzi and Simoncini technique was extended by Liesen (2006) or Liesen and Parlett (2007) where the inner product matrix is changed to

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

Combining BP$^+$ and BP$^-$

We want to combine the classical method (BP$^-$) and the new method (BP$^+$) with

$$\mathcal{P}_- = \begin{bmatrix} A_0 & 0 \\ B & -I \end{bmatrix} \text{ and } \quad \mathcal{P}_+ = \begin{bmatrix} A_0 & 0 \\ -B & I \end{bmatrix}$$

as preconditioners and

$$\mathcal{H}_- = \begin{bmatrix} A - A_0 & 0 \\ 0 & I \end{bmatrix} \text{ and } \quad \mathcal{H}_+ = \begin{bmatrix} A + A_0 & 0 \\ 0 & I \end{bmatrix}.$$

for the inner products.

The combination

$$\alpha \mathcal{P}_{-}^{-T}\mathcal{H}_{-} + (1-\alpha)\mathcal{P}_{+}^{-T}\mathcal{H}_{+} = \begin{bmatrix} A_0^{-1}A + (1-2\alpha)I & A_0^{-1}B^T \\ 0 & (1-2\alpha)I \end{bmatrix}$$

gives the splitting

$$\mathcal{P}^{-T} = \begin{bmatrix} A_0^{-1} & A_0^{-1}B^T \\ 0 & (1-2\alpha)I \end{bmatrix} \Longrightarrow \mathcal{P} = \begin{bmatrix} A_0 & 0 \\ \frac{1}{(2\alpha-1)}B & \frac{1}{1-2\alpha}I \end{bmatrix}$$

as the new preconditioner and the bilinear form is then defined by

$$\mathcal{H} = \begin{bmatrix} A + (1-2\alpha)A_0 & 0 \\ 0 & I \end{bmatrix}.$$

Numerical Experiments – Stokes problem

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

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

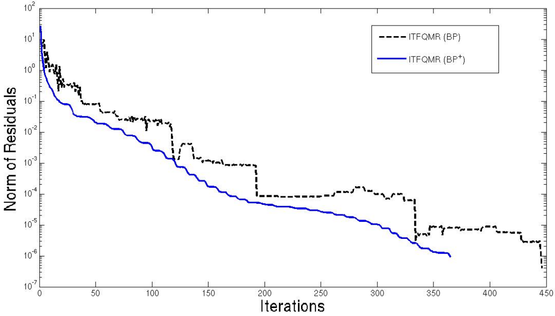
where $C = 0$.

All examples come from the IFISS package.

We compare our method to the block-diagonal preconditioned MINRES .

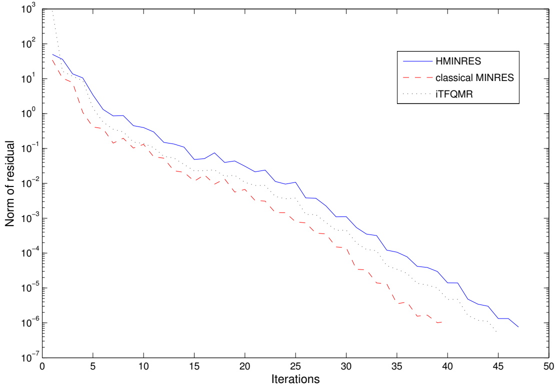
Example 1 – Stokes problem Channel domain

Results for classical BP and BP$^+$ for matrix of size 2467. Preconditioner A_0 incomplete Cholesky.



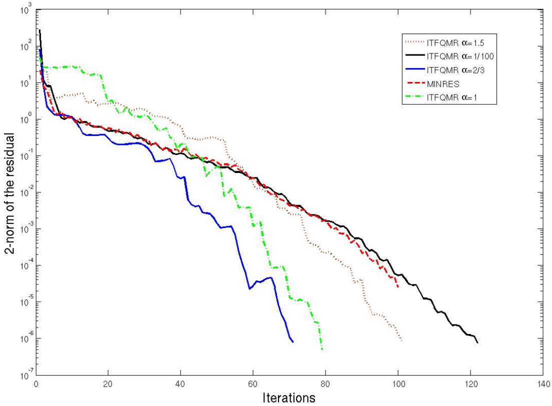
Example 2 – Stokes problem Channel domain

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 3 – Stokes problem Channel domain

Results for combination preconditioning and classical Bramble-Pasciak for problem of dimension 2467. Preconditioner A_0 being the Incomplete Cholseky decomposition and S_0 the Gramian.



Conclusions

- An alternative approach for saddle point problems
- A general concept how to combine preconditioners
- Some encouraging numerical results (more in the future)