Combination preconditioning and self-adjointness in non-standard inner products with application to saddle point problems

Martin Stoll & Andy Wathen

22nd BIENNIAL CONFERENCE ON NUMERICAL ANALYSIS
UNIVERSITY OF DUNDEE,SCOTLAND, UK

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 symmetric and positive definite and C symmetric positive semi-definite.

Motivating Example – 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

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

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 and positive definite under certain conditions imposed on A_0 in the inner product defined by

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

Original paper (Cited **181** times on June 6th 2007)!

So why the heck is this useful?

$\widehat{\mathcal{A}}$ is nonsymmetric and solvers would be GMRES QMR BICG ...

BUT

$\widehat{\mathcal{A}}$ is self-adjoint in $\mathcal{H}$ and we can use CG or MINRES

AND

in every step we minimize the error

$$\|e_i\|_{\mathcal{H}\widehat{\mathcal{A}}}$$

over

$$x_0 + \mathcal{K}_i(\mathcal{P}^{-1} r_0, \widehat{\mathcal{A}}).$$

Self-adjointness

We assume

$$\langle \cdot, \cdot \rangle_{\mathcal{H}} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$$

to be a symmetric bilinear form or an inner product where

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

A matrix $\mathcal{A} \in \mathbb{R}^{n \times n}$ is self-adjoint in $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ iff

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

Self-adjointness of the matrix $\mathcal{A}$ in $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ thus means that

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

for all x, y so that

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

is the basic relation for self-adjointness of $\mathcal{A}$ in $\langle \cdot, \cdot \rangle_{\mathcal{H}}$.

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–Bramble Pasciak CG

Introduced by Bramble and Pasciak (1988) it is a widely used CG technique with the preconditioner

$$\mathcal{P}^{-1} = \begin{bmatrix} A_0^{-1} & 0 \\ BA_0^{-1} & -I \end{bmatrix}$$

and inner product matrix

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

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) where the inner product matrix is changed to

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

Combination preconditioning

Lemma 4 shows that if we can find $\mathcal{P}_3$ and $\mathcal{H}_3$ such that

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

a new preconditioner and bilinear form are found. We want to combine the Bramble-Pasciak and the Benzi-Simoncini technique which gives

$$\alpha \mathcal{P}_1^{-T} \mathcal{H}_1 + \beta \mathcal{P}_2^{-T} \mathcal{H}_2 = \begin{bmatrix} (\alpha A_0^{-1} + \beta I) A - (\alpha + \beta \gamma) I & (\alpha A_0^{-1} + \beta I) B^T \\ -\beta B & -(\alpha + \beta \gamma) I \end{bmatrix}.$$

Combination preconditioning

One possibility for a splitting of $\alpha \mathcal{P}_1^{-T} \mathcal{H}_1 + \beta \mathcal{P}_2^{-T} \mathcal{H}_2$ is given by

$$\mathcal{P}_3^{-T} = \begin{bmatrix} \alpha A_0^{-1} + \beta I & 0 \\ 0 & -\beta I \end{bmatrix}$$

as the new preconditioner and by

$$\mathcal{H}_3 = \begin{bmatrix} A - (\alpha + \beta\gamma)(\alpha A_0^{-1} + \beta I)^{-1} & B^T \\ B & \frac{\alpha+\beta\gamma}{\beta} I \end{bmatrix}.$$

the symmetric matrix defining a bilinear form.

Numerical Experiments

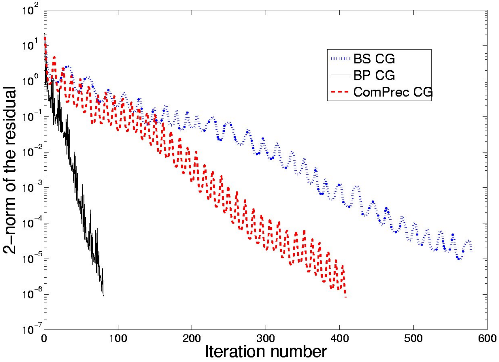

Figure: Combination preconditioning with $\alpha = 1/2$ and $\beta = 1/2$ compared to Bramble-Pasciak and Benzi-Simoncini CG .

Numerical Experiments

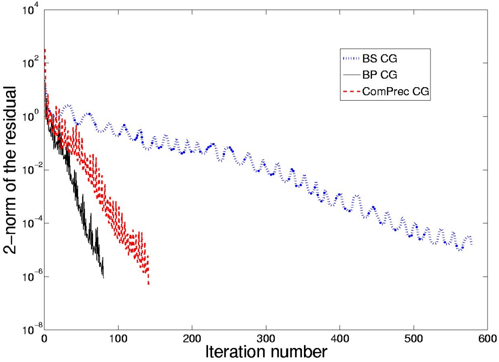

Figure: Combination preconditioning with $\alpha = 15$ and $\beta = 0.1$ compared to Bramble-Pasciak and Benzi-Simoncini CG .

Conclusions

- We provided insight in how preconditioners for saddle point problems can be combined.
- The basic analysis holds also for other classes of matrices.
- This is a more theoretical analysis but there might be applications where such techniques can be exploited.

Conclusions

- We provided insight in how preconditioners for saddle point problems can be combined.
- The basic analysis holds also for other classes of matrices.
- This is a more theoretical analysis but there might be applications where such techniques can be exploited.

More details in Andy's talk!