Report no. 07/11

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

Martin Stoll
Oxford University Computing Laboratory, Numerical Analysis Group, Wolfson Building, Parks Road, Oxford OX1 3QD, U.K.,
martin.stoll@comlab.ox.ac.uk

Andy Wathen
Oxford University Computing Laboratory, Numerical Analysis Group, Wolfson Building, Parks Road, Oxford OX1 3QD, U.K.,
andy.wathen@comlab.ox.ac.uk

It is widely appreciated that the iterative solution of linear systems of equations with large sparse matrices is much easier when the matrix is symmetric. It is equally advantageous to employ symmetric iterative methods when a nonsymmetric matrix is self-adjoint in a non-standard inner product. Here, general conditions for such self-adjointness are considered. In particular, a number of known examples for saddle point systems are surveyed and combined to make new combination preconditioners which are self-adjoint in different inner products.

Key words and phrases: Linear systems, Krylov subspaces, Non-standard inner products

Oxford University Computing Laboratory
Numerical Analysis Group
Wolfson Building
Parks Road
Oxford, England OX1 3QD

April, 2007

1 Introduction

In 1988 Bramble and Pasciak [6] introduced a block triangular preconditioner for the discrete Stokes problem (matrix) which had the almost magical effect of turning the original indefinite symmetric matrix problem into a non-symmetric matrix which is both self-adjoint and, in certain practical circumstances, positive definite in a non-standard inner product; thus the conjugate gradient method could be used in the non-standard inner product.

Precisely, the symmetric saddle point problem

$$\underbrace{\begin{bmatrix} A & B^T \\ B & -C \end{bmatrix}}_{\mathcal{A}} x = b \tag{1.1}$$

with symmetric A and C, if preconditioned on the left by

$$\mathcal{P} = \begin{bmatrix} A_0 & 0 \\ B & -I \end{bmatrix} \text{ with } \quad \mathcal{P}^{-1} = \begin{bmatrix} A_0^{-1} & 0 \\ BA_0^{-1} & -I \end{bmatrix} \tag{1.2}$$

results in the non-symmetric matrix

$$\widehat{\mathcal{A}} = \mathcal{P}^{-1}\mathcal{A} = \begin{bmatrix} A_0^{-1}A & A_0^{-1}B^T \\ BA_0^{-1}A - B & BA_0^{-1}B^T + C \end{bmatrix} \tag{1.3}$$

which turns out to be self-adjoint (many would say 'symmetric') in the inner product $\langle\cdot,\cdot\rangle_{\mathcal{H}}$ defined by $\langle x, y\rangle_{\mathcal{H}}^2 := x^T\mathcal{H}y$ where

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

Moreover, $\langle x, \widehat{\mathcal{A}}x\rangle_{\mathcal{H}} > 0$ for all $x \neq 0$ so that $\widehat{\mathcal{A}}$ is also positive definite. For these results to hold, the matrix block A_0 has to be symmetric and positive definite and must be scaled in order that $A - A_0$ is also positive definite so that $\langle\cdot,\cdot\rangle_{\mathcal{H}}$ not only defines a symmetric bilinear form but also satisfies the positivity requirement $\langle x, x\rangle_{\mathcal{H}} > 0$ for $x \neq 0$ which ensures that it is an inner product.

The outcome is that the Conjugate Gradient (CG) method based on this inner product can be applied and the system can be solved efficiently. This Bramble and Pasciak CG method is a very powerful and widely used tool to solve saddle point systems. Further analysis and applications can be found in [1–3, 7, 9, 15, 19, 21, 25, 31, 32].

The Bramble and Pasciak Conjugate Gradient method is not the only solver of its kind where a matrix is preconditioned and then a non-standard inner product can be found such that efficient Krylov subspace solvers like CG can be applied. More examples are given in [5, 12, 19, 23, 27] and will be explained and used later in this paper.

We comment that alternative approaches for the Stokes and other saddle-point problems use symmetric preconditioners and iterative methods such as MINRES and

SQMR(see [11] Chapter 6, [4]). It is not our intention here to compare with such methods - see for example [9]. Rather, we explore some more abstract but elementary algebraic structures which enable some broadening of the set of available preconditioners for which saddle-point problems may be treated by symmetric iterative methods and in particular CG in non-standard inner-products. The outcome is some new preconditioning techniques which might be useful in practice. Our algebraic results apply generally but we have not considered other than saddle-point examples.

2 Background

For linear systems with large dimension, it is well-known that iterative methods are most often the only feasible solution approaches; direct methods for dense and sparse or structured matrices work well provided bandwidth/skyline is not too large, but if fill-in in the computed triangular factors is too great such methods are usually infeasible beyond a certain dimension. Amongst the available iterative methods, multigrid approaches are extremely attractive for certain classes of problems (see for example [11]), and in more generality Krylov subspace methods can be excellent solvers provided suitably fast convergence can be achieved: preconditioning is almost always required to achieve this.

For symmetric matrix systems, the Conjugate Gradient (CG) method ([17]) for positive definite and minimum residual (MINRES) or SYMMLQ methods ([24]) for indefinite systems are based on short term recurrences and are the Krylov subspace methods of choice. The CG method is especially popular because of its efficiency. By contrast, for nonsymmetric matrix systems there is a large number of methods (GMRES,BICGSTAB, QMR, ... [26], [8], [14]) and various hybrid methods (GMRESR, BICGSTAB(ℓ),... [30], [29]) reflecting that convergence is less well understood than in the symmetric case and the best method for any particular problem is usually not clear, see [28] for a good overview and further references. This situation is rather unsatisfactory. Certainly if it were computationally possible to somehow convert a non-symmetric system to an equivalent one with a symmetric matrix so that CG or MINRES could be employed then this could be very attractive. Even in the case of a symmetric and indefinite matrix system, conversion to a symmetric and positive definite matrix system is attractive since it enables use of CG (the above case of the Bramble and Pasciak method is an example). Use of the normal equations approach — replacing $Ax = b$ with $A^T Ax = A^T b$ — is usually not so attractive because for other than a matrix which is very well-conditioned this usually leads to much poorer convergence of iterative methods.

One can broaden the possibilities by allowing non-standard inner-products — the important/desirable matrices are then the matrices which are self-adjoint with respect to such an inner product. The self-adjointness in a non-standard inner-product can enable the use of CG for positive definite systems and MINRES for self-adjoint but indefinite systems. A method like ITFQMR ([13]) can be used if only the self-adjointness in a symmetric bilinear form is given. For related considerations, in particular in connection with Krylov subspace methods, see [28], section 13.

Here we reserve the word symmetric when referring to matrices to mean those matrices which are self-adjoint in the usual Euclidean inner product defined by $\langle x, y\rangle := x^T y = \sum x_i y_i$ ie. those matrices whose entries satisfy $a_{i,j} = a_{j,i}$. Correspondingly, by $A^T(= B)$ we mean the matrix with entries $b_{i,j} = a_{j,i}$ ie. the adjoint matrix in the usual Euclidean inner product.

3 Basic properties

Firstly we review the basic mathematics. We consider here only real Euclidean vector spaces; we see no reason that our theory should not apply in the complex case or indeed for other vector spaces, but we have not done so.

We say that

$$\langle\cdot,\cdot\rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \tag{3.1}$$

is a symmetric bilinear form if

- $\langle x, y\rangle = \langle y, x\rangle$ for all $x, y \in \mathbb{R}^n$

- $\langle \alpha x + y, z\rangle = \alpha\langle x, z\rangle + \langle y, z\rangle$ for all $x, y, z \in \mathbb{R}^n$ and all $\alpha \in \mathbb{R}$.

With the addition of a non-degeneracy condition, Gohberg et al (cf . [16]) use the term 'indefinite inner product'; general properties of such forms can also be found here.

If additionally, the positivity conditions

$$\langle x, x\rangle > 0 \text{ for } x \neq 0 \quad \text{with} \quad \langle x, x\rangle = 0 \text{ if and only if } x = 0$$

are satisfied, then (3.1) defines and inner product on $\mathbb{R}^n$.

For any real symmetric matrix, $\mathcal{H}$, $\langle\cdot,\cdot\rangle_{\mathcal{H}}$ defined by

$$\langle x, y\rangle_{\mathcal{H}}^2 := x^T \mathcal{H} y \tag{3.2}$$

is easily seen to be a symmetric bilinear form which is an inner product if and only if $\mathcal{H}$ is positive definite.

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

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

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

$$\mathcal{A}^T \mathcal{H} = \mathcal{H}\mathcal{A} \tag{3.3}$$

is the basic relation for self-adjointness of $\mathcal{A}$ in $\langle\cdot,\cdot\rangle_{\mathcal{H}}$. We emphasize that $\langle\cdot,\cdot\rangle$, $\langle\cdot,\cdot\rangle_{\mathcal{H}}$ must be symmetric bilinear forms here, but we do not require them to be inner products in order that we use the self-adjointness concept. For practical reasons, we will consider positivity/non-positivity of symmetric bilinear forms and positive definiteness/indefiniteness of self-adjoint matrices separately from our considerations of symmetry and self-adjointness. Whenever we write $\langle\cdot,\cdot\rangle_{\mathcal{H}}$, $\mathcal{H}$ will be symmetric.

It is easy to check that

Lemma 3.1 *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 $\langle\cdot,\cdot\rangle_{\mathcal{H}}$.*

Also

Lemma 3.2 *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 $\alpha,\beta\in\mathbb{R}$.*

Now if $\mathcal{A}$ is preconditioned on the left by $\mathcal{P}$, then from (3.3), $\widehat{\mathcal{A}}=\mathcal{P}^{-1}\mathcal{A}$ is self-adjoint in $\langle\cdot,\cdot\rangle_{\mathcal{H}}$ if and only if

$$(\mathcal{P}^{-1}\mathcal{A})^T\mathcal{H}=\mathcal{H}\mathcal{P}^{-1}\mathcal{A} \tag{3.4}$$

which is

$$\mathcal{A}^T\mathcal{P}^{-T}\mathcal{H}=\mathcal{H}\mathcal{P}^{-1}\mathcal{A}$$

or

$$\mathcal{A}^T(\mathcal{P}^{-T}\mathcal{H})=(\mathcal{P}^{-T}\mathcal{H})^T\mathcal{A}$$

since $\mathcal{H}$ is symmetric. Thus if $\mathcal{A}$ is also symmetric we get

$$(\mathcal{P}^{-T}\mathcal{H})^T\mathcal{A}=\mathcal{A}(\mathcal{P}^{-T}\mathcal{H}) \tag{3.5}$$

and so

Lemma 3.3 *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 $\langle\cdot,\cdot\rangle_{\mathcal{A}}$.*

Proof Follows directly from the above and (3.3).

□

Remark 3.4 *Lemma 3.3 includes the even more simple situations that $\mathcal{P}^{-1}\mathcal{A}$ is self-adjoint in $\langle\cdot,\cdot\rangle_{\mathcal{P}}$ and $\mathcal{A}\mathcal{P}^{-1}$ is self-adjoint in $\langle\cdot,\cdot\rangle_{\mathcal{A}^{-1}}$ when both $\mathcal{A}$ and $\mathcal{P}$ are symmetric since I is trivially self-adjoint in any symmetric bilinear form. Clearly invertibility of $\mathcal{P}$ and $\mathcal{A}$ respectively are needed in these two cases.*

Now for symmetric $\mathcal{A}$, if $\mathcal{P}_1$ and $\mathcal{P}_2$ are such that $\mathcal{P}_i^{-1}\mathcal{A}$ is self-adjoint in $\langle\cdot,\cdot\rangle_{\mathcal{H}_i}$, $i=1,2$ for symmetric matrices $\mathcal{H}_1$, $\mathcal{H}_2$, then

$$(\mathcal{P}_1^{-1}\mathcal{A})^T\mathcal{H}_1=\mathcal{H}_1(\mathcal{P}_1^{-1}\mathcal{A}) \quad \text{and} \quad (\mathcal{P}_2^{-1}\mathcal{A})^T\mathcal{H}_2=\mathcal{H}_2(\mathcal{P}_2^{-1}\mathcal{A}). \tag{3.6}$$

Using Lemma 3.3, $\mathcal{P}_i^{-T}\mathcal{H}_i$ is self-adjoint in $\langle\cdot,\cdot\rangle_{\mathcal{A}}$ for $i=1,2$ and thus by Lemma 3.1

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

is also self-adjoint in $\langle\cdot,\cdot\rangle_{\mathcal{A}}$ for any $\alpha,\beta\in\mathbb{R}$. Now, if for some α,β we are able to multiplicatively split the matrix $(\alpha\mathcal{P}_1^{-T}\mathcal{H}_1+\beta\mathcal{P}_2^{-T}\mathcal{H}_2)=\mathcal{P}_3^{-T}\mathcal{H}_3$ for some symmetric matrix $\mathcal{H}_3$, then $\mathcal{P}_3^{-T}\mathcal{H}_3$ is self-adjoint in $\langle\cdot,\cdot\rangle_{\mathcal{A}}$ and a further application of Lemma 3.3 yields that $\mathcal{P}_3^{-1}\mathcal{A}$ is self-adjoint in $\langle\cdot,\cdot\rangle_{\mathcal{H}_3}$. We have proved

Lemma 3.5 *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 $\langle\cdot,\cdot\rangle_{\mathcal{H}_3}$.

Lemma 3.5 shows a possible way to generate new preconditioners for $\mathcal{A}$. In Section 5 we show a practical example of its use.

The construction of $\mathcal{P}_3$, $\mathcal{H}_3$ in Lemma 3.5 also allows straightforward inheritance of positive definiteness — for this to be a useful property it is essential that $\langle\cdot,\cdot\rangle_{\mathcal{H}}$ defines an inner product ie. that $\mathcal{H}$ is positive definite. It is trivial to construct examples of indefinite diagonal matrices $\mathcal{A}$ and $\mathcal{H}$ for which $\langle\mathcal{A}x, x\rangle_{\mathcal{H}} > 0$ for all non-zero x, but in order to be able to take advantage of positive definiteness, for example by employing Conjugate Gradients, it is important that $\langle x, x\rangle_{\mathcal{H}} = x^T\mathcal{H}x > 0$ for all non-zero x.

Lemma 3.6 *If the conditions of Lemma 3.5 are satisfied and additionally if $\mathcal{P}_i^{-1}\mathcal{A}$ is positive definite in $\langle\cdot,\cdot\rangle_{\mathcal{H}_i}$, $i = 1, 2$ then $\mathcal{P}_3^{-1}\mathcal{A}$ is positive definite in $\langle\cdot,\cdot\rangle_{\mathcal{H}_3}$ at least for positive values of α and β.*

Proof Positive definiteness of $\mathcal{P}^{-1}\mathcal{A}$ in $\langle\cdot,\cdot\rangle_{\mathcal{H}}$ means that

$$\langle\mathcal{P}^{-1}\mathcal{A}x, x\rangle_{\mathcal{H}} > 0, \quad \text{for } x \neq 0$$

ie. that $x^T\mathcal{A}\mathcal{P}^{-T}\mathcal{H}x > 0$ so that $\mathcal{A}\mathcal{P}^{-T}\mathcal{H}$ is a symmetric matrix with all eigenvalues positive. Thus each of $\mathcal{A}\mathcal{P}_1^{-T}\mathcal{H}_1$ and $\mathcal{A}\mathcal{P}_2^{-T}\mathcal{H}_2$ is symmetric and positive definite and it follows that

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

must also be symmetric and positive definite at least for positive values of α and β.

$\square$

We comment that there will in general be some negative values of α or β for which $\mathcal{P}_3^{-1}\mathcal{A}$ remains positive definite but at least one of α and β needs to be positive in this case. The precise limits on the values that α and β can take whilst positive definiteness is preserved depend on the extreme eigenvalues of $\mathcal{A}\mathcal{P}_1^{-T}\mathcal{H}_1$ and $\mathcal{A}\mathcal{P}_2^{-T}\mathcal{H}_2$. Unfortunately, even if $\mathcal{H}_1$ and $\mathcal{H}_2$ are positive definite there is no guarentee that $\mathcal{H}_3$ will be also.

We can also consider right preconditioning: if $\widehat{\mathcal{A}} = \mathcal{A}\mathcal{P}^{-1}$ is self-adjoint in $\langle\cdot,\cdot\rangle_{\mathcal{H}}$ then

$$\mathcal{P}^{-T}\mathcal{A}^T\mathcal{H} = \mathcal{H}\mathcal{A}\mathcal{P}^{-1} \tag{3.7}$$

which is

$$(\mathcal{P}^{-1})^T(\mathcal{A}^T\mathcal{H}) = (\mathcal{A}^T\mathcal{H})^T\mathcal{P}^{-1}. \tag{3.8}$$

Thus

Lemma 3.7 *If the right preconditioner $\mathcal{P}$ is symmetric and $\widehat{\mathcal{A}} = \mathcal{A}\mathcal{P}^{-1}$ is self-adjoint in $\langle\cdot,\cdot\rangle_{\mathcal{H}}$ for some symmetric matrix $\mathcal{H}$, then $\mathcal{A}^T\mathcal{H}$ is self-adjoint in $\langle\cdot,\cdot\rangle_{\mathcal{P}^{-1}}$.*

Lemma 3.7 shows that we could combine problem matrices and symmetric bilinear forms for the same preconditioner. This is obviously more a theoretical than a practical result compared to obtaining new preconditioners for a given problem as in the case of left preconditioning above. The splitting in $\mathcal{P}_3^{-T}\mathcal{H}_3$ introduced in Section 5 will provide not only a symmetric inner product matrix but also a symmetric preconditioner and therefore fulfills the conditions of Lemma 3.7.
We now want to discuss very briefly the eigenvalues of matrices which are self-adjoint according to our definition which allows indefinite symmetric bilinear forms. Assume that $\mathcal{A}^T\mathcal{H} = \mathcal{H}\mathcal{A}$ holds and that (λ, x) is a given eigenpair of $\mathcal{A}$. Thus,

$$\mathcal{A}x = \lambda x, \qquad x \neq 0. \tag{3.9}$$

Multiplying (3.9) from the left by $x^T\mathcal{H}$ gives

$$x^T\mathcal{H}\mathcal{A}x = \lambda x^T\mathcal{H}x \tag{3.10}$$

and using (3.3) we obtain

$$\frac{1}{2}x^T(\mathcal{H}\mathcal{A} + \mathcal{A}^T\mathcal{H})x = x^T\mathcal{H}\mathcal{A}x = \lambda x^T\mathcal{H}x. \tag{3.11}$$

Notice that the left hand side of (3.11) is the field of values of a symmetric matrix which is obviously real, see Chapter 8.1 in [20]. The right hand side is given by $\lambda x^T\mathcal{H}x$ where $x^T\mathcal{H}x$ is also real since $\mathcal{H}$ is symmetric. Therefore, the eigenvalue λ must be real. Furthermore, if the matrix $\mathcal{H}$ is definite the eigenvalues of $\mathcal{A}$ are bounded. Note that, a matrix $\mathcal{H}$ always exists such that $\mathcal{A}^T\mathcal{H} = \mathcal{H}\mathcal{A}$ since any matrix is similar to its transpose, see for example Chapter 3.2 in [18]. In the context of the above theory, the interesting candidates for $\mathcal{H}$ are the real symmetric matrices.

Note that the above arguments establish that there is no symmetric bilinear form in which $\mathcal{A}$ is self-adjoint unless $\mathcal{A}$ has real eigenvalues.

It is also known that for a real diagonalizable matrix $\mathcal{A}$ which has only real eigenvalues there always do exist inner products in which $\mathcal{A}$ is self-adjoint.

Lemma 3.8 *If $\mathcal{A} = R^{-1}\Lambda R$ is a diagonalization of $\mathcal{A}$ with the diagonal matrix Λ of eigenvalues being real, then $\mathcal{A}$ is self-adjoint in $\langle\cdot,\cdot\rangle_{R^T\Theta R}$ for any real diagonal matrix Θ.*

Proof The conditions (3.3) for self-adjointness of $\mathcal{A}$ in $\langle\cdot,\cdot\rangle_{\mathcal{H}}$ are

$$R^T\Lambda R^{-T}\mathcal{H} = \mathcal{H}R^{-1}\Lambda R$$

which are clearly satisfied for $\mathcal{H} = R^T\Theta R$ whenever Θ is diagonal because then Θ and Λ commute.

$\square$

We remark that this result is not of great use in practice since knowledge of the complete eigensystem of $\mathcal{A}$ is somewhat prohibitive.

4 Self-adjointness for saddle point systems

For the remainder of this paper we consider only saddle-point matrices $\mathcal{A}$ of the form

$$\mathcal{A} = \begin{bmatrix} A & B^T \\ B & 0 \end{bmatrix}, \tag{4.1}$$

that is as given in (1.1) with $A = A^T$ and $C = 0$. Using the same block structure, we consider the preconditioner

$$\mathcal{P}^{-1} = \begin{bmatrix} X & Y^T \\ Z & W \end{bmatrix} \tag{4.2}$$

so that

$$\widehat{\mathcal{A}} = \mathcal{P}^{-1}\mathcal{A} = \begin{bmatrix} XA + Y^T B & XB^T \\ ZA + WB & ZB^T \end{bmatrix}. \tag{4.3}$$

To define the symmetric bilinear form $\langle\cdot,\cdot\rangle_{\mathcal{H}}$ we take the symmetric matrix

$$\mathcal{H} = \begin{bmatrix} E & F^T \\ F & G \end{bmatrix}. \tag{4.4}$$

The main intention is now to identify suitable blocks of $\mathcal{P}$ and $\mathcal{H}$ such that $\widehat{\mathcal{A}}$ is self-adjoint in $\langle\cdot,\cdot\rangle_{\mathcal{H}}$. Using (3.3) this reduces to

$$\begin{bmatrix} AX^T + BY^T & AZ^T + B^T W^T \\ BX^T & BZ^T \end{bmatrix} \begin{bmatrix} E & F^T \\ F & G \end{bmatrix} = \begin{bmatrix} E & F^T \\ F & G \end{bmatrix} \begin{bmatrix} XA + Y^T B & XB^T \\ ZA + WB & ZB^T \end{bmatrix}. \tag{4.5}$$

Expanding we get

$$\begin{aligned} AX^T E + B^T Y E + AZ^T F + B^T W^T F &= EXA + EY^T B + F^T ZA + F^T WB && (4.6)\\ AX^T F^T + B^T Y F^T + AZ^T G + B^T W^T G &= EXB^T + F^T ZB^T && (4.7)\\ BX^T E + BZ^T F &= FXA + FY^T B + GZA + GWB && (4.8)\\ BX^T F^T + BZ^T G &= FXB^T + GZB^T. && (4.9) \end{aligned}$$

in which the assumed symmetry of $\mathcal{H}$ and therefore of E and G ensure that (4.7) and (4.8) are transposes of each other and are therefore the same. Under the assumption that $\mathcal{H}$ is block-diagonal ($F = 0$) this simplifies to

$$\begin{aligned} A^T X^T E + B^T Y E &= EXA + EY^T B && (4.10)\\ A^T Z^T G + B^T W^T G &= EXB^T && (4.11)\\ BX^T E &= GZA + GWB && (4.12)\\ BZ^T G &= GZB^T && (4.13) \end{aligned}$$

where again (4.11) and (4.12) are the same.

We will now describe some examples where it can be checked that the above conditions are satisfied. The descriptions will be brief and we refer the reader to the original papers [5, 6, 19, 22] for more details.

The first example is the classical method by Bramble and Pasciak already mentioned in the Introduction. The preconditioner is given by

$$\mathcal{P} = \begin{bmatrix} A_0 & 0 \\ B & -I \end{bmatrix} \quad \text{and} \quad \mathcal{P}^{-1} = \begin{bmatrix} A_0^{-1} & 0 \\ BA_0^{-1} & -I \end{bmatrix} \tag{4.14}$$

and the symmetric bilinear form defined by the matrix

$$\mathcal{H} = \begin{bmatrix} A - A_0 & 0 \\ 0 & I \end{bmatrix}. \tag{4.15}$$

In 2006 Benzi and Simoncini gave a further example, see [5]. Namely,

$$\mathcal{P} = \mathcal{P}^{-1} = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \tag{4.16}$$

and

$$\mathcal{H} = \begin{bmatrix} A - \gamma I & B^T \\ B & \gamma I \end{bmatrix}. \tag{4.17}$$

Recently, Liesen made an extension to this method taking a non-zero matrix C in (1.1) into account, see [22]. The preconditioner is again

$$\mathcal{P} = \mathcal{P}^{-1} = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \tag{4.18}$$

but the symmetric bilinear form is now defined by

$$\mathcal{H} = \begin{bmatrix} A - \gamma I & B^T \\ B & \gamma I - C \end{bmatrix}. \tag{4.19}$$

There are certain conditions which must be satisfied by the parameter γ in order to guarantee positive definiteness of $\mathcal{H}$ so that CG in the inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ can be reliably employed – see [5], [22].

There are also extensions to the classical Bramble-Pasciak case which are not restricted to the case $C = 0$ in (1.1), see [19, 23, 27]. In [23] for example, a Schur complement preconditioner S_0 is introduced into $\mathcal{P}$ giving

$$\mathcal{P} = \begin{bmatrix} A_0 & 0 \\ B & -S_0 \end{bmatrix} \quad \text{and} \quad \mathcal{P}^{-1} = \begin{bmatrix} A_0^{-1} & 0 \\ S_0^{-1}BA_0^{-1} & -S_0^{-1} \end{bmatrix}; \tag{4.20}$$

under certain conditions positive definiteness of the preconditioned saddle-point system can still be guaranteed in a non-standard inner product similar to (4.15).

5 Combination preconditioning

Here we demonstrate just one example of how Lemma 3.5 can be used to combine preconditioners in order to create new ones. In the process, new symmetric bilinear forms are created which sometimes have the desirable positivity property making them inner products in which for example CG can be employed.

We consider $\mathcal{P}_1, \mathcal{H}_1$ defined by the classical Bramble-Pasciak method (4.14),(4.15) and $\mathcal{P}_2, \mathcal{H}_2$ defined by the Benzi-Simoncini approach (4.16),(4.17). From Lemma 3.5 we get

$$(\alpha \mathcal{P}_1^{-T}\mathcal{H}_1 + \beta \mathcal{P}_2^{-T}\mathcal{H}_2) = \begin{bmatrix} (\alpha A_0^{-1} + \beta I)A - (\alpha + \beta\gamma)I & (\alpha A_0^{-1} + \beta I)B^T \\ -\beta B & -(\alpha + \beta\gamma)I \end{bmatrix} \tag{5.1}$$

which is self-adjoint $\forall \alpha, \beta \in \mathbb{R}$ in $\langle \cdot, \cdot \rangle_{\mathcal{A}}$. If we are able to multiplicatively split this into a new preconditioner $\mathcal{P}_3$ and a symmetric matrix $\mathcal{H}_3$, Lemma 3.5 guarentees that $\mathcal{P}_3^{-1}\mathcal{A}$ will be self-adjoint in $\langle \cdot, \cdot \rangle_{\mathcal{H}_3}$.

One possibility is

$$\mathcal{P}_3^{-T} = \begin{bmatrix} \alpha A_0^{-1} + \beta I & 0 \\ 0 & -\beta I \end{bmatrix} \quad \text{and } \mathcal{H}_3 = \begin{bmatrix} A - (\alpha + \beta\gamma)(\alpha A_0^{-1} + \beta I)^{-1} & B^T \\ B & \frac{\alpha+\beta\gamma}{\beta} I \end{bmatrix}. \tag{5.2}$$

Numerical results for the CG solution of $\mathcal{A}x = b$ with preconditioner $\mathcal{P}_3$ and symmetric bilinear form $\langle \cdot, \cdot \rangle_{\mathcal{H}_3}$ are presented here for various choices of α,β and γ. The test problem is a Stokes flow problem generated using the IFISS software in MATLAB ([10]). We assume that the conditions for γ in the Benzi-Simoncini case and for A_0 in the Bramble-Pasciak case are fulfilled. The matrix generated by IFISS is of dimension 659 and represents the flow over a channel domain. Figure 6 shows the comparison of the Bramble-Pasciak CG method (solid black line) compared with the Benzi-Simoncini CG (dotted blue line) and also the combination of the two methods with combination parameters $\alpha = 1/2$ and $\beta = 1/2$ (dashed red line). The slow convergence of the Benzi and Simoncini method can be explained by the lack of preconditioning in this method whereas the results for the Bramble-Pasciak CG justify its popularity. Figure 6 shows the same constellation with only a change in the combination parameters $\alpha = 15$ and $\beta = 0.1$.

6 Conclusions

We have explained the general concept of self-adjointness in non-standard inner products or symmetric bilinear forms and in the specific case of saddle point problems have shown how a number of known examples fit into this paradigm. We have indicated how self-adjointness may be taken advantage of in the choice of iterative solution method of Krylov subspace type—in general it is more desirable to be able to work with iterative methods for self-adjoint matrices rather than general nonsymmetric matrices because of the greater efficiency in general of symmetric iterative methods and certainly because the understanding of the convergence of symmetric iterative methods is much more secure and descriptive than for nonsymmetric methods.

The possibility of combination preconditioning by exploiting self-adjointness in different non-standard inner products or symmetric bilinear forms has been analysed and an example given of how two methods can be combined to obtain a new preconditioner and a different symmetric bilinear form.

Our analysis may provide the basis for the discovery of further useful examples where self-adjointness may hold in non-standard inner products and also shows how preconditioning can usefully be employed to create rather than destroy symmetry of matrices.

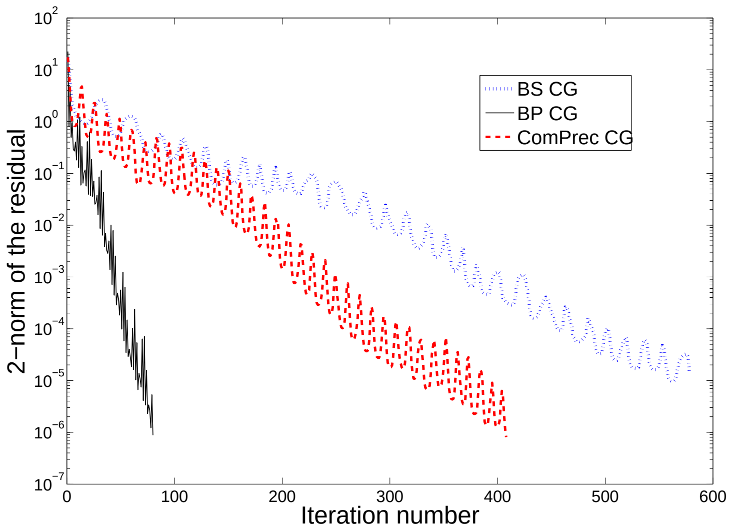

Figure 1: Combination preconditioning with $\alpha = 1/2$ and $\beta = 1/2$ compared to Bramble-Pasciak and Benzi-Simoncini CG .

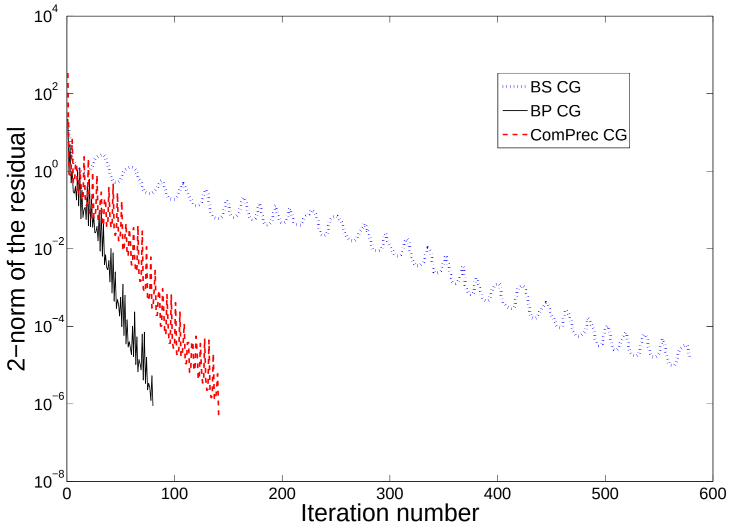

Figure 2: Combination preconditioning with $\alpha = 15$ and $\beta = 0.1$ compared to Bramble-Pasciak and Benzi-Simoncini CG .

References

[1] Mark Ainsworth and Spencer Sherwin, 1999. Domain decomposition preconditioners for p and hp finite element approximation of Stokes equations. *Comput. Methods Appl. Mech. Engrg.*, **175**(3-4): 243–266.

[2] Owe Axelsson and Maya Neytcheva, 2003. Preconditioning methods for linear systems arising in constrained optimization problems. *Numer. Linear Algebra Appl.*, **10**(1-2): 3–31. Dedicated to the 60th birthday of Raytcho Lazarov.

[3] Owe Axelsson and Maya Neytcheva, 2003. Robust preconditioners for saddle point problems. In *Numerical methods and applications*, Volume 2542 of *Lecture Notes in Comput. Sci.*, pages 158–166. Springer, Berlin.

[4] Michele Benzi, Gene H. Golub and Jörg Liesen, 2005. Numerical solution of saddle point problems. *Acta Numer.*, **14**: 1–137.

[5] Michele Benzi and Valeria Simoncini, 2006. On the eigenvalues of a class of saddle point matrices. *Numer. Math.*, **103**(2): 173–196.

[6] James H. Bramble and Joseph E. Pasciak, 1988. A preconditioning technique for indefinite systems resulting from mixed approximations of elliptic problems. *Math. Comp.*, **50**(181): 1–17.

[7] C. Carstensen, M. Kuhn and U. Langer, 1998. Fast parallel solvers for symmetric boundary element domain decomposition equations. *Numer. Math.*, **79**(3): 321–347.

[8] Van der Vorst, 1992. BiCGSTAB: A fast and smoothly converging variant of biCG for the solution of nonsymmetric linear systems. *SIAM J. Sci. Stat. Comput*, **13**: 631–644.

[9] H. C. Elman, 1996. Multigrid and Krylov subspace methods for the discrete Stokes equations. *In* N. D. Melson, T. A. Manteuffel, S. F. McCormick and C. C. Douglas, editors, *Seventh Copper Mountain Conference on Multigrid Methods*, Volume CP 3339, pages 283–299, Hampton, VA. NASA.

[10] Howard C. Elman, Alsison Ramage and David J. Silvester, June 2007. Algorithm xxx: IFISS, a Matlab toolbox for modelling incompressible flow. *ACM Transactions on Mathematical Software*, **33**(2).

[11] Howard C. Elman, David J. Silvester and Andrew J. Wathen, 2005. *Finite elements and fast iterative solvers: with applications in incompressible fluid dynamics.* Numerical Mathematics and Scientific Computation. Oxford University Press, New York.

[12] B. Fischer, A. Ramage, D. J. Silvester and A. J. Wathen, 1998. Minimum residual methods for augmented systems. *BIT*, **38**(3): 527–543.

[13] Roland W. Freund, 1994. Transpose-free quasi-minimal residual methods for non-Hermitian linear systems. In *Recent advances in iterative methods*, Volume 60 of *IMA Vol. Math. Appl.*, pages 69–94. Springer, New York.

[14] Roland W. Freund and Noël M. Nachtigal, 1991. QMR: a quasi-minimal residual method for non-Hermitian linear systems. *Numer. Math.*, **60**(3): 315–339.

[15] Gabriel N. Gatica and Norbert Heuer, 2002. Conjugate gradient method for dual-dual mixed formulations. *Math. Comp.*, **71**(240): 1455–1472.

[16] Israel Gohberg, Peter Lancaster and Leiba Rodman, 2005. *Indefinite linear algebra and applications.* Birkhäuser Verlag, Basel.

[17] Magnus R. Hestenes and Eduard Stiefel, 1952. Methods of conjugate gradients for solving linear systems. *J. Research Nat. Bur. Standards*, **49**: 409–436 (1953).

[18] Roger A. Horn and Charles R. Johnson, 1990. *Matrix analysis.* Cambridge University Press, Cambridge.

[19] Axel Klawonn, 1998. Block-triangular preconditioners for saddle point problems with a penalty term. *SIAM J. Sci. Comput.*, **19**(1): 172–184. Special issue on iterative methods (Copper Mountain, CO, 1996).

[20] Peter Lancaster and Miron Tismenetsky, 1985. *The theory of matrices.* Computer Science and Applied Mathematics. Academic Press Inc., Orlando, FL, second edition.

[21] U. Langer and M. Kuhn. Parallel solvers for large-scale, coupled finite and boundary element equations.

[22] J. Liesen, 2006. A note on the eigenvalues of saddle point matrices. Technical Report 10-2006, TU Berlin.

[23] Arnd Meyer and Torsten Steidten, 2001. Improvements and experiments on the Bramble-Pasciak Type CG for mixed problems in elasticity. Technical report, TU Chemnitz.

[24] C. C. Paige and M. A. Saunders, 1975. Solutions of sparse indefinite systems of linear equations. *SIAM J. Numer. Anal.*, **12**(4): 617–629.

[25] Jörg Peters, Volker Reichelt and Arnold Reusken, 2005. Fast iterative solvers for discrete Stokes equations. *SIAM J. Sci. Comput.*, **27**(2): 646–666.

[26] Youcef Saad and Martin H. Schultz, 1986. GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM J. Sci. Statist. Comput.*, **7**(3): 856–869.

[27] V. Simoncini, 2004. Block triangular preconditioners for symmetric saddle-point problems. *Appl. Numer. Math.*, **49**(1): 63–80.

[28] V. Simoncini and D.B. Szyld, February 2007. Recent computational developments in Krylov subspace methods for linear systems. *Numer. Linear Algebra Appl.*, **14**(1): 1–61.

[29] Gèrard L. G. Sleijpen and Diederik R. Fokkema, 1993. BiCGstab(l) for linear equations involving unsymmetric matrices with complex spectrum. *Electron. Trans. Numer. Anal.*, **1**(Sept.): 11–32 (electronic only).

[30] H. A. van der Vorst and C. Vuik, 1994. GMRESR: a family of nested GMRES methods. *Numer. Linear Algebra Appl.*, **1**(4): 369–386.

[31] Andrew Wathen and David Silvester, 1993. Fast iterative solution of stabilised Stokes systems. I. Using simple diagonal preconditioners. *SIAM J. Numer. Anal.*, **30**(3): 630–649.

[32] Walter Zulehner, 2002. Analysis of iterative methods for saddle point problems: a unified approach. *Math. Comp.*, **71**(238): 479–505.