A note on preconditioning for indefinite linear systems*

Malcolm F. Murphy[a]   Gene H. Golub[b]   Andrew J. Wathen[c]

[a]School of Mathematics, University of Bristol, University Walk, Bristol BS8 1TW, UK
[b]Department of Computer Science, Stanford University, California 94305, USA. email:golub@sccm.stanford.edu †
[c]Oxford University Computing Laboratory, Wolfson Building, Parks Road, Oxford OX1 3QD, UK. email: wathen@comlab.ox.ac.uk . Please address correspondence to this author.

Preconditioners are often conceived as approximate inverses. For non-singular indefinite matrices of saddle-point (or KKT) form, we show how preconditioners incorporating an exact Schur complement lead to preconditioned matrices with exactly two or exactly three distinct eigenvalues. Thus approximations of the Schur complement lead to preconditioners which can be very effective even though they are in no sense approximate inverses.

Oxford University Computing Laboratory
Numerical Analysis Group
Wolfson Building
Parks Road
Oxford, England   OX1 3QD   June, 1999

1 Introduction and Result

In many areas there arise matrix systems of the form

$$\mathcal{A}u = \left[\begin{array}{cc} A & B^T \\ C & 0 \end{array}\right]\left[\begin{array}{c} x \\ y \end{array}\right] = \left[\begin{array}{c} f \\ g \end{array}\right] \tag{1.1}$$

where $A \in \Re^{n \times n}$ and $B, C \in \Re^{m \times n}$ with $n \geq m$ (see for example [3]). When $\mathcal{A}$ arises from a constrained variational or optimisation problem it is usual that $B = C$. There are also many problems where additionally A is symmetric, so that $\mathcal{A}$ is symmetric (see for example [2], [4], [9], [11]). Whether symmetric or not, the matrix $\mathcal{A}$ is generally indefinite (i.e. it has eigenvalues with both positive and negative real parts).

Iterative solution of systems of the form (1.1) can be achieved by any of a number of methods: in particular Krylov subspace methods such as MINRES [10] or GMRES [12] are applicable in the symmetric and non-symmetric cases respectively. It is often advantageous (and in many situations necessary) to employ a preconditioner, $\mathcal{P}$, with such iterative methods. The role of $\mathcal{P}$ is to reduce the number of iterations required for convergence whilst not increasing significantly the amount of computation required at each iteration.

Intuitively if $\mathcal{P}$ can be chosen so that $\mathcal{P}^{-1}$ is an inexpensive approximate inverse of $\mathcal{A}$, then this might make a good preconditioner, however it is not necessary for a good preconditioner to have that $\mathcal{P}^{-1}$ be an approximate inverse of $\mathcal{A}$. A sufficient condition for a good preconditioner is that the preconditioned matrix $\mathcal{T} = \mathcal{P}^{-1}\mathcal{A}$ has a low degree minimum polynomial. This condition is more usually expressed in terms of $\mathcal{T}$ having only a few distinct eigenvalues: in this form we must insist that $\mathcal{T}$ is not degenerate (derogatory) or at least that it's Jordan canonical form has Jordan blocks of only small dimension.

In this note we show how preconditioners can be derived for systems of the form (1.1) based upon an 'exact' preconditioner which yields a preconditioned matrix with exactly 3 or exactly 2 distinct eigenvalues (in fact precisely with minimum polynomial of exact degree 3 or 2). This observation is based on the Schur complement:

Proposition 1. If

$$\mathcal{A} = \left[\begin{array}{cc} A & B^T \\ C & 0 \end{array}\right]$$

is preconditioned by

$$\mathcal{P} = \left[\begin{array}{cc} A & 0 \\ 0 & CA^{-1}B^T \end{array}\right] \tag{1.2}$$

then the preconditioned matrix $\mathcal{T} = \mathcal{P}^{-1}\mathcal{A}$ satisfies

$$\mathcal{T}(\mathcal{T} - I)(\mathcal{T}^2 - \mathcal{T} - I) = 0.$$

Proof. Simple calculation yields

$$\mathcal{T} = \mathcal{P}^{-1}\mathcal{A} = \begin{bmatrix} I & A^{-1}B^T \\ (CA^{-1}B^T)^{-1}C & 0 \end{bmatrix}$$

and

$$\left(\mathcal{T} - \frac{1}{2}I\right)^2 = \begin{bmatrix} \frac{1}{4}I + A^{-1}B^T(CA^{-1}B^T)^{-1}C & 0 \\ 0 & \frac{5}{4}I \end{bmatrix}.$$

But $A^{-1}B^T(CA^{-1}B^T)^{-1}C$ is a projection so that

$$\left[\left(\mathcal{T} - \frac{1}{2}I\right)^2 - \frac{1}{4}I\right]^2 = \left[\left(\mathcal{T} - \frac{1}{2}I\right)^2 - \frac{1}{4}I\right]$$

which simplifies to

$$\mathcal{T}(\mathcal{T} - I)(\mathcal{T}^2 - \mathcal{T} - I) = 0. \tag{1.3}$$

Remark 1. Since (1.3) can be factorised into distinct linear factors (over $\Re$) it follows that $\mathcal{T}$ is diagonalisable and has at most the four distinct eigenvalues

$$0, \quad 1, \quad \frac{1}{2} \pm \frac{\sqrt{5}}{2}. \tag{1.4}$$

If $\mathcal{T}$ is nonsingular then it has the 3 non-zero eigenvalues.

Remark 2. Proposition 1 specifically addresses *left* preconditioning, but by simple similarity transformations involving $\mathcal{P}$ or factorisations of $\mathcal{P}$ the result of the proposition applies equally for *right* preconditioning, $\mathcal{T} = \mathcal{A}\mathcal{P}^{-1}$, or in general any centred preconditioning $\mathcal{T} = \mathcal{P}_1^{-1}\mathcal{A}\mathcal{P}_2^{-1}$ where $\mathcal{P}_1\mathcal{P}_2 = \mathcal{P}$.

Remark 3. It directly follows from Proposition 1 that for any vector r, the Krylov subspace

$$\text{span}\{r, \mathcal{T}r, \mathcal{T}^2r, \mathcal{T}^3r, \ldots\}$$

is of dimension at most 3 if $\mathcal{T}$ is nonsingular (or 4 if $\mathcal{T}$ is singular). Thus in particular, any Krylov subspace iterative method with an optimality or Galerkin property (for example such as any minimum residual method - see [6]), will terminate in at most 3 iterations with the solution to a linear system of the form (1.1) if the preconditioner (1.2) is used.

Remark 4. If consideration of symmetry is not important then a similar argument to the above shows that the choice

$$\mathcal{P} = \begin{bmatrix} A & B^T \\ 0 & CA^{-1}B^T \end{bmatrix} \tag{1.5}$$

yields a preconditioned system with exactly the 2 eigenvalues ± 1. Use of (1.5) rather than (1.2) requires only one more multiplication of a vector by B^T per iteration. If

$CA^{-1}B^T$ is replaced by $-CA^{-1}B^T$ in (1.5), then the preconditioned matrix has only a single eigenvalue of 1, but it is *not* in this case diagonalisable: it is similarly elementary to show that the minimum polynomial is $(\mathcal{P}^{-1}\mathcal{A} - I)^2$. Since this is again quadratic, there seems nothing to gain by choice of sign.

We comment that in the context of general symmetric indefinite factorisation, Gill et al. [7] have proposed a preconditioner which also gives a preconditioned matrix with eigenvalues ± 1. In the notation of this paper (and in the non-symmetric case) their preconditioner is

$$\mathcal{P} = \left[\begin{array}{cc} A & B^T \\ C & 2CA^{-1}B^T \end{array} \right].$$

Remark 5. Proposition 1 is of practical use when inexpensive approximations of A and of the Schur complement $CA^{-1}B^T$ exist. Known examples arise in discretisations of the Stokes problem of incompressible fluid dynamics [13] and indeed in many saddle-point approximations where inf-sup or Babuška-Brezzi [1] stability holds. In such cases, the eigenvalues lie in three distinct regions - the precise clustering around the non-zero values in (1.4) depends on the quality of the approximations.

Remark 6. If A^{-1} is readily approximated then this approximation could be used in the Schur complement as well as in the (1,1) block of $\mathcal{P}$ (see for example ([8]) where semicirculant approximations which allow the use of Fast Fourier transforms are used). Also inner-outer iteration could be effective for this preconditioner (see for example [5]).

References

[1] F. BREZZI, *New applications of mixed finite element methods,* Proc. of the International Congress of Mathematicians 1986, pp. 1335-1347, 1987.

[2] P.T. BOGGS AND J.W.TOLLE, *Sequential Quadratic Programming,* Acta Numerica 1995, pp. 1-51, 1995.

[3] H.E. ELMAN AND D.J. SILVESTER, *Fast nonsymmetric iterations and preconditioning for Navier-Stokes equations,* SIAM J. Sci. Comput.**17**, pp. 33-46, 1996.

[4] H.E. ELMAN, D.J. SILVESTER AND A.J. WATHEN, *Iterative methods for problems in Computational Fluid Dynamics,* in 'Iterative Methods in Scientific Computing', Eds. R.H. Chan, T.F. Chan & G.H. Golub, Springer-Verlag, Singapore, pp. 271-327, 1997.

[5] H.E. ELMAN AND G.H. GOLUB, *Inexact and preconditioned Uzawa algorithms for saddle point problems,* SIAM J. Numer. Anal.**31**, pp. 1645-1661, 1994.

[6] R.FREUND, G.H. GOLUB AND N.NACHTIGAL, *Iterative solution of linear systems,* Acta Numerica 1992, pp. 57-100.

[7] P.E. GILL, W. MURRAY, D.B. PONCELEÓN AND M.A. SAUNDERS, *Preconditioners for indefinite systems arising in optimisation,* SIAM J. Matrix Anal. Appl.**13**, pp. 292-311.

[8] L. HEMMINGSSON AND A.J. WATHEN, *A nearly optimal preconditioner for the Navier-Stokes equations,* in preparation.

[9] J. MARYŠKA, M. ROZLOŽNÍK AND M. TUMA, *The potential fluid flow problem and the convergence rate of the minimum residual method,* Numer. Linear Alg. Appl.,**3**, pp. 525-542, 1996.

[10] C.C. PAIGE AND M.A. SAUNDERS, *Solution of Sparse Indefinite Systems of Linear Equations,* SIAM J. Numer. Anal.,**12**, pp. 617-629, 1975.

[11] I. PERUGIA, V. SIMONCINI AND M. ARIOLI, *Linear algebra methods in a mixed approximation of magnetostatics problems,* SIAM J. Sci. Stat. Comput., to appear.

[12] Y. SAAD AND M. SCHULTZ, *GMRES: A generalised minimum residual algorithm for solving non-symmetric linear systems,* SIAM J. Sci. Stat. Comput.,**7**, pp. 856-869, 1986.

[13] D.J. SILVESTER AND A.J. WATHEN, *Fast iterative solution of stabilised Stokes systems Part II: using general block preconditioners,* SIAM J. Numer. Anal. **31**, pp. 1352-1367, 1994.