On Preconditioners for Generalized Saddle Point Problems with an Indefinite Block

Piotr Krzyżanowski

University of Warsaw, ul. Banacha 2, 02-097 Warszawa, Poland;
piotr.krzyzanowski@mimuw.edu.pl

1 Introduction

In many applications one needs to solve a discrete system of linear equations with a symmetric block matrix

$$\mathcal{M}\begin{pmatrix} u \\ p \end{pmatrix} \equiv \begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f \\ g, \end{pmatrix} \tag{1}$$

where the block $A = A^T$ is not necessarily positive definite and may even be singular. Such situation occurs, for example, after suitable finite element discretization of the generalized Stokes problem, cf. [5],

$$-\Delta u - \omega u + \nabla p = f, \tag{2}$$

$$\text{div } u = 0, \tag{3}$$

when for large enough ω one cannot preserve the ellipticity of $-\Delta - \omega$. Another example is the time-harmonic Maxwell equation, see [7],

$$\nabla \times \nabla \times u - \omega u + \nabla p = f, \tag{4}$$

$$\text{div } u = 0, \tag{5}$$

where large enough ω again results in an indefinite A. Although the whole system matrix (1) remains invertible when $\omega = 0$, the matrix A which then corresponds to the discrete curl-curl operator, has a large kernel.

In these examples, the discrete problem matrix (1) is ill conditioned with respect to the mesh parameter h. Our aim in this paper is to analyze block preconditioners for such systems, for which the preconditioned conjugate residuals (PCR) method, see [8], converges independently of h.

Block preconditioning allows for an efficient reuse of existing methods of preconditioning problems of simpler structure, such as symmetric positive definite systems. Actually, block diagonal or triangular preconditioners decompose in a natural way the large system (1) into several smaller and simpler problems. Since domain decomposition based preconditioners are very well developed for symmetric and positive definite problems, and high quality software, such as PETSc, see [1] contains implementations of very robust methods, the use of block preconditioners may be a reasonable solution method instead of more involved methods.

We present an analysis of some block preconditioning algorithms within a general framework, essentially assuming only that equation (1) is well posed and that $\mathcal{M}$ is symmetric. Our analysis is valid for inexact block solvers and shows that a successful preconditioner can be based on preconditioners for symmetric positive definite sub-problems. Let us note that probably the first observation that (diagonal) preconditioners based on positive definite blocks are applicable even in the case when A is not necessarily positive definite, was made in [9]. Here we generalize this observation to various kinds of block preconditioners.

2 General Assumptions

Let $\bar{V},\bar{W}$ be real Hilbert spaces with scalar products denoted by $((\cdot,\cdot))$ and $(\cdot,\cdot)$, respectively. The norms in these spaces, induced by the inner products, will be denoted by $\|\cdot\|$ and $|\cdot|$. We consider a family of finite dimensional subspaces indexed by the parameter $h \in (0,1)$: $V_h \subset \bar{V}, \quad W_h \subset \bar{W}$. If V_h, W_h come from finite element approximations, the dimension of these subspaces increases for decreasing h.

Following [4], let us introduce three continuous bilinear forms: $a : \bar{V} \times \bar{V} \to R$, $b : \bar{V} \times \bar{W} \to R$, $c : \bar{W} \times \bar{W} \to R$. We assume that $a(\cdot,\cdot)$ is symmetric and there exists a constant α, independent of h, such that

$$\exists\, \alpha > 0 \quad \forall h \in (0,1) \qquad \inf_{v \in V_h^0, v \neq 0} \sup_{u \in V_h^0, u \neq 0} \frac{a(u,v)}{\|u\|\,\|v\|} \geq \alpha, \tag{6}$$

where $V_h^0 = \{v \in V_h : \forall q \in W_h \quad b(v,q) = 0\}$. We shall also assume that the finite dimensional spaces V_h and W_h satisfy the uniform LBB condition,

$$\exists\, \beta > 0 \quad \forall h \in (0,1) \quad \forall p \in W_h \qquad \sup_{v \in V_h, v \neq 0} \frac{b(v,p)}{\|v\|} \geq \beta\, |p|. \tag{7}$$

Remark 1. Condition (6), when related to our motivating problems, generalized Stokes (2)–(3) or time-harmonic Maxwell equations (4)–(5), imposes some conditions on the values of ω, e.g., in the latter case, $\sqrt{\omega}$ has to be distinct from any Maxwell eigenvalue of the discrete problem.

From now on, we drop the subscript h to simplify the notation. In what follows we consider preconditioners for a family of finite dimensional problems:

Problem 1. Find $(u,p) \in V \times W$ such that

$$\mathcal{M}\begin{pmatrix} u \\ p \end{pmatrix} \equiv \begin{pmatrix} A & B^* \\ B & 0 \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} F \\ G \end{pmatrix}. \tag{8}$$

The operators in (8) are defined by:

$$A : V \to V, \quad ((Au,v)) = a(u,v) \quad \forall u,v \in V,$$
$$B : V \to W, \quad (Bu,p) = b(u,p) \quad \forall u \in V,\, p \in W,$$

while the right hand side components $F \in V$, $G \in W$ satisfy $((F,v)) \equiv \langle\langle f,v\rangle\rangle$ and $(G,w) \equiv \langle g,w\rangle$, where f,g are given linear continuous functionals on $\bar{V}$, $\bar{W}$, and $\langle\langle\cdot,\cdot\rangle\rangle$, $\langle\cdot,\cdot\rangle$ denote the duality pairing in $\bar{V}$, $\bar{W}$, respectively. B^* denotes the formal adjoint operator to B, i.e. $(Bu,p) = ((u,B^*p))$ for all $u \in V$, $p \in W$. Let us recall the key theorem which we shall use throughout the paper. This is the classical result on the stability of (8):

Lemma 1. [4] *Under the above assumptions, there exits a unique pair* $(u,p) \in V \times W$ *which solves* (8). *Moreover,*

$$\|u\| + |p| \lesssim \|F\| + |G|. \tag{9}$$

Here, and it what follows, $x \lesssim y$ means that there exists a positive constant C, independent of x, y and h, such that $x \le Cy$. Similarly, $x \simeq y$ will denote that both $x \lesssim y$ and $y \lesssim x$ hold.

We introduce two more operators, $A_0 : V \to V$ and $J_0 : W \to W$. We assume that they are self-adjoint, their inverses are easy to apply, and that they define inner products spectrally equivalent to $((\cdot,\cdot))$ and $(\cdot,\cdot)$, respectively:

$$((A_0u,u)) \simeq ((u,u)) \quad \forall u \in V, \tag{10}$$
$$(J_0p,p) \simeq (p,p) \quad \forall p \in W. \tag{11}$$

In other words, we shall always assume that A_0 and J_0 define good preconditioners for the Grammian matrices for the chosen bases in V and W, respectively. For example, the A_0 preconditioner may be constructed using very efficient domain decomposition or multigrid techniques; for J_0, in some cases such as the generalized Stokes problem, one can also apply a very cheap diagonal scaling instead of domain decomposition.

With any X-elliptic, selfadjoint operator G we may associate an energy norm of $x \in X$, $\|f\|_G = ((Gf,f))^{1/2}$. From (10)–(11) it directly follows that the energy norms defined by A_0, J_0 and their inverses are equivalent, with constants independent of h, to the original norms in appropriate spaces:

Lemma 2. *For any* $f \in V$ *and* $g \in W$,

$$\|f\|_{A_0} \simeq \|f\| \simeq \|f\|_{A_0^{-1}}, \tag{12}$$
$$|g|_{J_0} \simeq |g| \simeq |g|_{J_0^{-1}}. \tag{13}$$

Lemma 3. *The norms of* A, B, A_0, J_0, $\mathcal{M}$ *in appropriate spaces are bounded independently of* h,

$$\|A\|_{V\to V}, \quad \|B\|_{V\to W}, \quad \|A_0\|_{V\to V}, \quad \|J_0\|_{W\to W}, \quad \|\mathcal{M}\|_{V\times W\to V\times W} \lesssim 1.$$

Moreover,

$$\|A_0^{-1}\|_{V\to V}, \quad \|J_0^{-1}\|_{W\to W}, \quad \|\mathcal{M}^{-1}\|_{V\times W\to V\times W} \lesssim 1.$$

In the rest of the paper, we shall analyze preconditioners for the Preconditioned Conjugate Residual method (PCR), which is known to be applicable to indefinite symmetric systems, provided the preconditioner is a symmetric, positive definite operator. Other methods may also be applicable, such as QMR, BiCGStab, GMRES, etc. When applied to $\mathcal{M}$ with a preconditioner $\mathcal{P}$, its convergence rate, according to [8], depends on the quantity $\kappa(\mathcal{P}^{-1}\mathcal{M}) = \rho(\mathcal{P}^{-1}\mathcal{M})\rho(\mathcal{M}^{-1}\mathcal{P})$, where ρ denotes the spectral radius of a matrix.

3 Block Diagonal Preconditioner

In the section, we recall a result regarding the block diagonal preconditioner,

$$\mathcal{M}_D = \begin{pmatrix} A_0 & 0 \\ 0 & J_0 \end{pmatrix}.$$

This preconditioner has been thoroughly analyzed for symmetric saddle point problems, assuming either V-ellipticity, see [11], or only V^0-ellipticity of A, see e.g. [9, Sec. 3.2]. These results directly apply to the more general case, when A only satisfies (6). Actually, the only non-trivial property of $\mathcal{M}$ which is required in the proof is the stability result of Lemma 1.

Lemma 4 ([9]). *The preconditioned operator* $\mathcal{P}_D = \mathcal{M}_D^{-1}\mathcal{M}$ *satisfies*

$$\kappa(\mathcal{P}_D) \lesssim 1.$$

4 Block Upper Triangular Preconditioner

Another preconditioner for the operator $\mathcal{M}$ is based on a block upper triangular matrix

$$\mathcal{M}_U = \begin{pmatrix} A_0 & B^* \\ 0 & J_0 \end{pmatrix}. \tag{14}$$

The preconditioned operator $\mathcal{P}_U = \mathcal{M}_U^{-1} M$ is equal to

$$\mathcal{P}_U = \begin{pmatrix} A_0 & \\ & J_0 \end{pmatrix}^{-1} \begin{pmatrix} A - B^* J_0^{-1} B & B^* \\ B & 0 \end{pmatrix}, \tag{15}$$

so the triangular preconditioner acts as the diagonal preconditioner $\mathcal{M}_D$ applied to an *augmented* matrix

$$\tilde{\mathcal{M}}_U = \begin{pmatrix} A - B^* J_0^{-1} B & B^* \\ B & 0 \end{pmatrix}.$$

Remark 2. Usually, block systems (1) are augmented by *adding* a non-negative matrix to A, see e.g. [2] or [6]. Klawonn's preconditioner also results in a positively augmented matrix, cf. [10]. Here, we end up with a negatively augmented matrix, that is, we *subtract* a non-negative definite matrix from A. Numerical results provided in the final section, as well as some theoretical considerations, indicate that this approach improves the overall convergence of the iterative solver.

Due to the decomposition (15), it is still possible to use a PCR method to solve the preconditioned problem. The analysis of the upper triangular preconditioner reduces to the previous case of block diagonal preconditioning.

Lemma 5. *Lemma* 3 *holds for the augmented matrix* $\tilde{\mathcal{M}}_U$.

Applying the estimates from the block-diagonal case and using this lemma, we conclude that

Theorem 1. $\kappa(\mathcal{P}_U) \lesssim 1$.

5 Lower Block Triangular Preconditioner

It is also possible, with some additional assumptions, to analyze, in the same framework, the lower triangular block preconditioner

$$\mathcal{M}_L = \begin{pmatrix} A_0 & 0 \\ B & J_0 \end{pmatrix}. \tag{16}$$

The preconditioned operator $\mathcal{P}_L = \mathcal{M}_L^{-1} M$ then equals

$$\mathcal{P}_L = \begin{pmatrix} A_0 - A & \\ & J_0 \end{pmatrix}^{-1} \begin{pmatrix} A - AA_0^{-1}A & (A_0 - A)A_0^{-1}B^* \\ BA_0^{-1}(A_0 - A) & -BA_0^{-1}B^* \end{pmatrix}. \tag{17}$$

so the upper triangular preconditioner acts as a diagonal preconditioner

$$\mathcal{M}_{DL} = \begin{pmatrix} A_0 - A & \\ & J_0 \end{pmatrix}$$

applied to some symmetric matrix

$$\tilde{\mathcal{M}}_L = \begin{pmatrix} A - AA_0^{-1}A & (A_0 - A)A_0^{-1}B^* \\ BA_0^{-1}(A_0 - A) & -BA_0^{-1}B^* \end{pmatrix}.$$

See [3] for an analysis of the case when A is positive definite. In order to use the PCR framework, which requires the preconditioner to be positive definite, we have to assume some scaling of A_0; see [3].

Theorem 2. *If there exists a constant* $m > 0$*, independent of* h*, such that*

$$(((A_0 - A)u, u)) > m((u, u)) \quad \forall u \in V, \tag{18}$$

then

$$\kappa(\mathcal{P}_L) \lesssim 1.$$

6 Numerical Experiments

The numerical experiments were performed using a MATLAB implementation of a Taylor-Hood finite element discretization of the generalized Stokes problem (2)–(3) on a unit square, with homogeneous boundary condition for the velocity. The discretization resulted in a matrix $A = D - \omega M$, where D is the discrete Laplacian and M corresponds to the velocity mass matrix. We conducted two kinds of tests. First, we experimented with $A_0 = D + M$ and $J_0 = M$ (the pressure mass matrix), calling this preconditioner as the "exact" preconditioner. Then, in order to show a more realistic application, we used "inexact" preconditioners with A_0^{-1} defined as the incomplete Cholesky solve of $D + M$, with drop tolerance 10^{-3}.

In both cases, we investigated the convergence rate of the block diagonal, upper triangular and lower triangular preconditioners discussed above, for several values of ω and varying mesh size h. The stopping criterion was the reduction of the residual norm by a factor of 10^6. To provide sufficient scaling for the A_0 block in the lower triangular preconditioner, we have set $A_0 = 2D + M$ in $\mathcal{M}_L$ in the "exact" case. For comparison with the upper triangular solver, we also included a diagonally preconditioned positively augmented system, see Remark 2,

$$\mathcal{P}_{aug} = \mathcal{P}_D^{-1} \cdot \begin{pmatrix} A + B^T J_0^{-1} B & B^T \\ B & 0 \end{pmatrix}.$$

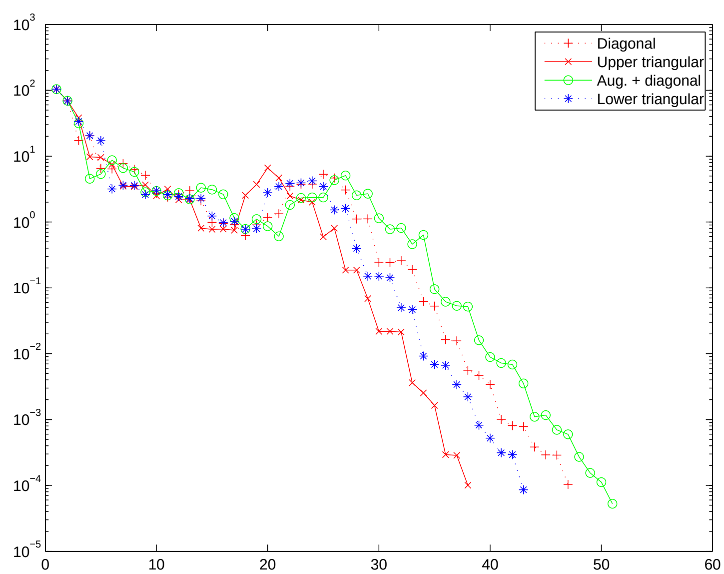

Fig. 1. A comparison of convergence histories of the PCR using four preconditioners for discretized generalized Stokes problem with $\omega = 10$. Exact A_0 solver (see details in the text).

Table 1. Iteration counts for various parameters and preconditioners; left panel: "exact" case, right panel: "inexact" case.

ω	h	$\mathcal{P}_D$	$\mathcal{P}_U$	$\mathcal{P}_{aug}$	$\mathcal{P}_L$	$\mathcal{P}_D$	$\mathcal{P}_U$	$\mathcal{P}_{aug}$	$\mathcal{P}_L$
10	1/4	39	32	39	33	56	44	57	47
10	1/64	45	38	50	43	116	113	130	101
100	1/4	73	66	86	84	103	92	126	109
100	1/64	133	114	150	135	144	128	171	119

As expected, good preconditioners such as those used in the "exact" case, led to iteration counts virtually independent of h. On the other hand, the number of iterations seems to grow sublinearly with the increase of ω, cf. Table 1.

7 Conclusions

Block preconditioning using optimal preconditioners for simple symmetric positive definite operators leads to optimal results with respect to the mesh size h under very mild assumptions on the A block in (1). There is a connection between the (left-) upper triangular preconditioning and the augmented Lagrangian method, with a prospective advantage of the former over the latter.

A general drawback of these preconditioners is that, in some situations, for example, when $A = D - \omega M$ with both D and M positive semidefinite (the case of time-harmonic Maxwell's equations), our bounds also depend on ω. Clearly, in such a case, if ω is very large, one should rather, instead of A, treat M as the dominant term in this block. How to choose the inexact preconditioning blocks in a robust way so that the block preconditioners would perform well independently of ω remains an open problem.

References

[1] Balay, S., Buschelman, K., Gropp, W.D., Kaushik, D., Knepley, M.G., McInnes, L.C., Smith, B.F., Zhang, H.: PETSc Web page, 2001. `http://www.mcs.anl.gov/petsc`.

[2] Benzi, M., Liu, J.: Block preconditioning for saddle point systems with indefinite $(1,1)$ block. *Internat. J. Comput. Math.*, 84(8):1117–1129, 2007.

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

[4] Brezzi, F., Fortin, M.: *Mixed and hybrid finite element methods*, vol. 15 of *Springer Series in Computational Mathematics*. Springer, New York, 1991. ISBN 0-387-97582-9.

[5] Cliffe, K.A., Garratt, T.J., Spence, A.: Eigenvalues of block matrices arising from problems in fluid mechanics. *SIAM J. Matrix Anal. Appl.*, 15(4):1310–1318, 1994.
[6] Golub, G.H., Greif, C., Varah, J.M.: An algebraic analysis of block diagonal preconditioner for saddle point systems. *SIAM J. Matrix Anal. Appl.*, 27(3): 779–792 (electronic), 2005.
[7] Greif, C., Schötzau, D.: Preconditioners for the discretized time-harmonic Maxwell equations in mixed form. *Numer. Linear Algebra Appl.*, 14(4):281–297, 2007.
[8] Hackbusch, W.: *Iterative solution of large sparse systems of equations*, vol. 95 of *Applied Mathematical Sciences*. Springer, New York, 1994. Translated and revised from the 1991 German original.
[9] Klawonn, A.: *Preconditioners for Indefinite Problems*. PhD thesis, Universität Münster, Germany, 1996.
[10] Klawonn, A.: Block-triangular preconditioners for saddle point problems with a penalty term. *SIAM J. Sci. Comput.*, 19(1):172–184 (electronic), 1998. Special issue on iterative methods (Copper Mountain, CO, 1996).
[11] Klawonn, A.: An optimal preconditioner for a class of saddle point problems with a penalty term. *SIAM J. Sci. Comput.*, 19(2):540–552 (electronic), 1998.