The University of Reading
Department of Mathematics

Constraint-style preconditioners for regularized saddle point problems

H. S. Dollar

Numerical Analysis Report 3/06

March 27, 2006

Constraint-style preconditioners for regularized saddle point problems

H. S. Dollar[1 2]

Abstract

The problem of finding good preconditioners for the numerical solution of an important class of indefinite linear systems is considered. These systems are of a regularized saddle point structure

$$\begin{bmatrix} A & B^T \\ B & -C \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} c \\ d \end{bmatrix},$$

where $A \in \mathbb{R}^{n \times n}$, $C \in \mathbb{R}^{m \times m}$ are symmetric and $B \in \mathbb{R}^{m \times n}$.

In *Constraint preconditioning for indefinite linear systems,* SIAM J. Matrix Anal. Appl., 21 (2000), Keller, Gould and Wathen analyze the idea of using constraint preconditioners that have a specific 2 by 2 block structure for the case of C being zero. We shall extend this idea by allowing the (2,2) block to be symmetric and positive semi-definite. Results concerning the B

that quasidefinite matrices are strongly factorizable , i.e., a Cholesky-like factorization LDL^T exists for any symmetric row and column permutation of the quasidefinite matrix, [25]. The diagonal matrix has n positive and m negative pivots. However, we shall not confine ourselves to quasidefinite matrices.

It may be attractive to use iterative methods to solve systems such as (1), particularly for large m and n. In particular, Krylov subspace methods might be used. It is often advantageous to use a preconditioner, P, with such iterative methods. The preconditioner should reduce the number of iterations required for convergence but not significantly increase the amount of computation required at each iteration, [24, Chapter 13].

In Section 2 we shall firstly review the well known spectral properties of a technique commonly known as constraint preconditioning when $C = 0$ [15, 17]. For the case of $C = 0$; a constraint preconditioner exactly reproduces the (constraint) blocks B; B^T and the $C = 0$ block. It is restrictive to assume that the matrix C in the saddle point systems is always a zero matrix: a number of situations arise in which $C \neq 0$ [1, 16, 23]. In all these cases, C is positive semi-definite and, hence, we shall consider the idea of extending constraint preconditioners to the case of C being positive semi-definite. In particular, the preconditioner will exactly reproduce the B; B^T and C blocks, whilst the A block will be replaced by a symmetric block which we refer to as G; this is considered in Sections 3 and 4. Such a preconditioner has been considered before, for example, Perugia and Simoncini consider the case of $G = I$ [18], and Siefert and de Sturler assume that G is nonsingular [22], but we show that these assumptions can be relaxed. In Section 5 we shall report numerical results where our preconditioners have been used to solve various test problems.

2 Constraint preconditioners

Let us initially assume that $C = 0$: Keller, Gould and Wathen [15] investigated the spectral properties of the resulting preconditioned system when we use of a preconditioner of the form

$$P = \begin{bmatrix} G & B^T \\ B & 0 \end{bmatrix}; \qquad (2)$$

where G approximates but (in general) is not the same as A. They were able to prove various results about the eigenvalues and eigenvectors for the preconditioned systems $P^{-1}A$; where A and P are defined in (1) and (2), respectively. P is called a *constraint preconditioner*. Proof of the following theorem can be found in [15].

GA-10129B

Theorem 2.1. *Let* $A \in \mathbb{R}^{(n+m)\times}$ (DeA

[(P)]TJ/F216.97Tfo20TD[(n)]TJ/F66.97Tf4.9320TD[(+)]TJ/F206.97Tf6.1120TD[(m)]T

where $A \in \mathbb{R}^{n\times n}$ is symmetric and $B \in \mathbb{R}^{m\times n}$ is of full rank. Assume Z is an $n \times (n-m)$ basis for the nullspace of B. Preconditioning $\mathcal{A}$ by a matrix of the form

$$P = \begin{bmatrix} G & B^T \\ B & 0 \end{bmatrix},$$

where $G \in \mathbb{R}^{n\times n}$ is symmetric, and $B \in \mathbb{R}^{m\times n}$ is as above, implies that the matrix $P^{-1}\mathcal{A}$ has

1. *an eigenvalue at 1 with multiplicity $2m$;*
2. *$n-m$ eigenvalues λ which are defined by the generalized eigenvalue problem $Z^T A Z x_z = \lambda Z^T G Z x_z.$*

This accounts for all of the eigenvalues.

If either $Z^T AZ$ or $Z^T GZ$ are positive definite, then the indefinite preconditioner P applied to the indefinite saddle point matrix $\mathcal{A}$ with $C = 0$ yields a preconditioned matrix $P^{-1}\mathcal{A}$ which has real eigenvalues [15]. If both $Z^T AZ$ and $Z^T GZ$ are positive definite, then we can use a projected preconditioned conjugate gradient method to find x and y; see [12]. Results about the associated eigenvectors and the Krylov subspace dimension can also be found in [15].

3 Constraint preconditioners for the case of symmetric and positive definite C

In this section we shall assume that the matrix C is symmetric and positive definite. The term *constraint preconditioner* was introduced in [15] because the (1,2) and (2,1) matrix blocks of the preconditioner are exact representations of those in $\mathcal{A}$, where these blocks represent constraints. However, we also observe that the (2,2) matrix block is an exact representation when $C = 0$. This motivates the generalization of the constraint preconditioner to take the form

$$P = \begin{bmatrix} G & B^T \\ B & -C \end{bmatrix}, \tag{3}$$

where $G \in \mathbb{R}^{n\times n}$ approximates, but is, in general, not the same as A.

For symmetric matrix systems, the convergence of an applicable iterative method is determined by the distribution of the eigenvalues of the coefficient matrix. It is often desirable for the number of distinct eigenvalues to be small so that the rate of convergence is rapid. For non-normal systems the convergence is not so readily described, see [14, page 6].

We shall use the following assumptions in the theorems of this section:

A1 $C \in \mathbb{R}^{m\times m}$ is symmetric and positive definite,

A2 $A \in \mathbb{R}^{n\times n}$ is symmetric, A3 $B \in \mathbb{R}^{m \times n}$ $(m < n)$ has full rank,

A4 $G \in \mathbb{R}^{n \times n}$ is symmetric,

A5 $A \in \mathbb{R}^{(n+m)\times(n+m)}$ is as defined in (1),

A6 $P \in \mathbb{R}^{(n+m)\times(n+m)}$ is as defined in (3).

In the next section A1 will be relaxed.

Theorem 3.1. *Assume that A1-A6 hold, then the matrix*

If $A + B^T C^{-1} B$ or $G + B^T C^{-1} B$ are positive definite, then the preconditioned system has real eigenvalues. If both $A + B^T C^{-1} B$ and $G + B^T C^{-1} B$ are positive definite, then we can apply a projected preconditioned conjugate gradient method to find x and y [7, 11]. We also observe that if C has a small 2-norm, $\|A\|_2 = O(1)$ and $\|G\|_2 = O(1)$, then the $B^T C^{-1} B$ terms will dominate the generalized eigenvalue problem (8) for $Bx \neq 0$ and, hence, there will be at least a further m eigenvalues clustered about 1 for $\|C\|_2 \ll 1$. This additional clustering of part of the spectrum of $P^{-1} A$ will often translate into a speeding up of the convergence of a selected Krylov subspace method, [2, Section 1.3].

Theorem 3.2. *Assume that A1-A6 hold and $G + B^T C^{-1} B$ is positive definite, then the matrix $P^{-1} A$ has $n + m$ eigenvalues as defined in Theorem 3.1 and $m + i + j$ linearly independent eigenvectors. There are*

- *m eigenvectors of the form $\begin{bmatrix} 0^T & y^T \end{bmatrix}^T$ that correspond to the case $\lambda = 1$;*
- *i ($0 \leq i \leq n$) eigenvectors of the form $\begin{bmatrix} x^T & y^T \end{bmatrix}^T$ arising from $Ax = \sigma G x$ for which the i vectors x are linearly independent, $\sigma = 1$, and $\lambda = 1$; and*
- *j ($0 \leq j \leq n$) eigenvectors of the form $\begin{bmatrix} x^T & y^T \end{bmatrix}^T$ that correspond to the case $\lambda \neq 1$.*

Proof. The form of the eigenvectors follows directly from the proof of Theorem 3.1. It remains for us to show that the $m + i + j$ eigenvectors are linearly independent, that is, we need to show that

$$\begin{bmatrix} 0 & \cdots & 0 \\ y_1^{(1)} & \cdots & y_m^{(1)} \end{bmatrix} \begin{bmatrix} a_1^{(1)} \\ \vdots \\ a_m^{(1)} \end{bmatrix} + \begin{bmatrix} x_1^{(2)} & \cdots & x_i^{(2)} \\ y_1^{(2)} & \cdots & y_i^{(2)} \end{bmatrix} \begin{bmatrix} a_1^{(2)} \\ \vdots \\ a_i^{(2)} \end{bmatrix} + \begin{bmatrix} x_1^{(3)} & \cdots & x_j^{(3)} \\ y_1^{(3)} & \cdots & y_j^{(3)} \end{bmatrix} \begin{bmatrix} a_1^{(3)} \\ \vdots \\ a_j^{(3)} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \quad (9)$$

implies that the vectors $a^{(k)}$ ($k = 1, 2, 3$) are zero vectors. Multiplying (9) by $P^{-1} A$, and recalling that in the previous equation the first matrix arises from the case $\lambda_k = 1$ ($k = 1, \ldots, m$), the second matrix from the case $\lambda_k = 1$ and $\sigma_k = 1$ ($k = 1, \ldots, i$), and the last matrix from $\lambda_k \neq 1$ ($k = 1, \ldots, j$), gives

$$\begin{bmatrix} 0 & \cdots & 0 \\ y_1^{(1)} & \cdots & y_m^{(1)} \end{bmatrix} \begin{bmatrix} a_1^{(1)} \\ \vdots \\ a_m^{(1)} \end{bmatrix} + \begin{bmatrix} x_1^{(2)} & \cdots & x_i^{(2)} \\ y_1^{(2)} & \cdots & y_i^{(2)} \end{bmatrix} \begin{bmatrix} a_1^{(2)} \\ \vdots \\ a_i^{(2)} \end{bmatrix} + \begin{bmatrix} x_1^{(3)} & \cdots & x_j^{(3)} \\ y_1^{(3)} & \cdots & y_j^{(3)} \end{bmatrix} \begin{bmatrix} \lambda_1 a_1^{(3)} \\ \vdots \\ \lambda_j a_j^{(3)} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \quad (10)$$

Subtracting (9) from (10) we obtain

$$\begin{bmatrix} x_1^{(3)} & \cdots & x_j^{(3)} \\ y_1^{(3)} & \cdots & y_j^{(3)} \end{bmatrix} \begin{bmatrix} (\lambda_1 - 1)a_1^{(3)} \\ \vdots \\ (\lambda_j - 1)a_j^{(3)} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}.$$

The assumption that $G + B^T C^{-1} B$ is positive definite implies that $x_k^{(3)}$ $(k = 1, \ldots, j)$ are linearly independent and thus that $(\lambda_k - 1)a_1^{(3)} = 0,\ (k = 1, \ldots, j).$ The eigenvalues λ_k $(k = 1, \ldots, j)$ are non-unit which implies that $a_k^{(3)} = 0$ $(k = 1, \ldots, j).$ We also have linear independence of $x_k^{(2)}$ $(k = 1, \ldots, i)$ and thus $a_k^{(2)} = 0$ $(k = 1, \ldots, i).$ Equation (9) simplifies to

$$\begin{bmatrix} 0 & \cdots & 0 \\ y_1^{(1)} & \cdots & y_m^{(1)} \end{bmatrix} \begin{bmatrix} a_1^{(1)} \\ \vdots \\ a_m^{(1)} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}.$$

However, $y_k^{(1)}$ $(k = 1, \ldots, m)$ are linearly independent and thus $a_k^{(1)} = 0$ $(k = 1, \ldots, m).$ □

Krylov subspace theory states that iteration with any method with an optimality property, e.g. GMRES [21], will terminate when the degree of the minimum polynomial is attained. This is also true of some other (non-optimal) practical iterations such as BiCGTAB as long as failure does not occur. In particular, the degree of the minimum polynomial is equal to the dimension of the corresponding Krylov subspace $\mathcal{K}(P^{-1}A, b)$ (for general b), [20, Proposition 6.1].

Theorem 3.3. *Assume that A1-A6 hold and $G + B^T C^{-1} B$ is positive definite, then the dimension of the Krylov subspace $\mathcal{K}(P^{-1}A, b)$ is at most $\min\{n+2, n+m\}.$*

Proof. As in the proof of Theorem 3.1, the generalized eigenvalue problem is

$$\begin{bmatrix} A & B^T \\ B & -C \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \lambda \begin{bmatrix} G & B^T \\ B & -C \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}. \tag{11}$$

Suppose that the preconditioned matrix $P^{-1}A$ takes the form

$$P^{-1}A = \begin{bmatrix} \Theta_1 & \Theta_3 \\ \Theta_2 & \Theta_4 \end{bmatrix}, \tag{12}$$

where $\Theta_1 \in \mathbb{R}^{n \times n}$, $\Theta_2 \in \mathbb{R}^{m \times n}$, $\Theta_3 \in \mathbb{R}^{n \times m}$, and $\Theta_4 \in \mathbb{R}^{m \times m}.$ Using the facts that $P(P^{-1}A) = A$ and B has full row rank, we obtain $\Theta_3 = 0$ and $\Theta_4 = I.$ The precise forms of Θ_1 and Θ_2 are irrelevant for the argument that follows.

From the earlier eigenvalue derivation, it is evident that the characteristic polynomial of the preconditioned linear system (12) is

$$\left(P^{-1}A - I\right)^m \prod_{i=1}^{n} \left(P^{-1}A - \lambda_i I\right).$$

In order to prove the upper bound on the Krylov subspace dimension, we need to show that the order of the minimum polynomial is less than or equal to $\min\{n+2, n+m\}.$ Expanding the polynomial $\left(P^{-1}A - I\right)\prod_{i=1}^{n}\left(P^{-1}A - \lambda_i I\right)$ of degree $n+1,$ we obtain

$$\begin{bmatrix} (\Theta_1 - I)\prod_{i=1}^{n}(\Theta_1 - \lambda_i I) & 0 \\ \Theta_2 \prod_{i=1}^{n}(\Theta_1 - \lambda_i I) & 0 \end{bmatrix}.$$

Since Θ_1 has a full set of linearly independent eigenvectors, Θ_1 is diagonalizable. Hence,

$$(\Theta_1 - I)\prod_{i=1}^{n}(\Theta_1 - \lambda_i I) = 0.$$

We therefore obtain

$$\left(P^{-1}A - I\right)\prod_{i=1}^{n}\left(P^{-1}A - \lambda_i I\right) = \begin{bmatrix} 0 & 0 \\ \Theta_2 \prod_{i=1}^{n}(\Theta_1 - \lambda_i I) & 0 \end{bmatrix}. \qquad (13)$$

If $\Theta_2 \prod_{i=1}^{n}(\Theta_1 - \lambda_i I) = 0,$ then the order of the minimum polynomial of $P^{-1}A$ is less than or equal to $\min\{n+1, n+m\}.$ If $\Theta_2 \prod_{i=1}^{n}(\Theta_1 - \lambda_i I) \neq 0,$ then the

B1 $C \in \mathbb{R}^{m\times m}$ is symmetric and positive semi-definite, and has rank p where $0 < p < m$;

B2 C is factored as $C = EDE^T$; where $E \in \mathbb{R}^{m\times p}$; and $D \in \mathbb{R}^{p\times p}$ is non-singular,

B3 The matrix $F \in \mathbb{R}^{m\times(m-p)}$ is such that its columns span the nullspace of C;

B4 $\begin{bmatrix} E & F \end{bmatrix} \in \mathbb{R}^{m\times m}$ is orthogonal,

B5 The columns of $N \in \mathbb{R}^{n\times(n-m+p)}$ span the nullspace of F^TB.

The exact form of the factorization of C in **B2** is clearly not relevant and, also, clearly not unique – a spectral decomposition is a possibility.

Theorem 4.1. *Assume that* **A2-A6** *and* **B1-B5** *hold, then the matrix* $P^{-1}A$ *has*

- *an eigenvalue at 1 with multiplicity* $2m-p$; *and*
- $n-m+p$ *eigenvalues which are defined by the generalized eigenvalue problem*

$$N^T\left(A + B^TED^{-1}E^TB\right)Nz = \lambda N\left(G + B^TED^{-1}E^TB\right)Nz.$$

This accounts for all of the eigenvalues.

Proof. Any $y \in \mathbb{R}^m$ can be written as $y = Ey_e + Fy_f$. Substituting this into the generalized eigenvalue problem (4) and premultiplying by $\begin{bmatrix} I & 0 \\ 0 & E^T \\ 0 & F^T \end{bmatrix}$ we obtain

$$\left[\begin{array}{cc|c} A & B^TE & B^TF \\ E^TB & -D & 0 \\ \hline F^TB & 0 & 0 \end{array}\right]\begin{bmatrix} x \\ y_e \\ y_f \end{bmatrix} = \lambda \left[\begin{array}{cc|c} G & B^TE & B^TF \\ E^TB & -D & 0 \\ \hline F^TB & 0 & 0 \end{array}\right]\begin{bmatrix} x \\ y_e \\ y_f \end{bmatrix}. \quad (14)$$

Noting that the (3,3) block has dimension $(m-p)\times(m-p)$ and is a zero matrix in both coefficient matrices, we can apply Theorem 2.1 from [15] to obtain that $P^{-1}A$ has

- an eigenvalue at 1 with multiplicity $2(m-p)$; and
- $n-m+2p$ eigenvalues which are defined by the generalized eigenvalue problem

$$\overline{N}^T\begin{bmatrix} A & B^TE \\ E^TB & -D \end{bmatrix}\overline{N}w_n = \lambda\overline{N}^T\begin{bmatrix} G & B^TE \\ E^TB & -D \end{bmatrix}\overline{N}w_n, \quad (15)$$

where $\overline{N}$ is an $(n+p)\times(n-m+2p)$ basis for the nullspace of $\begin{bmatrix} F^TB & 0 \end{bmatrix} \in \mathbb{R}^{(m-p)\times(n+p)}$, and

$$\begin{bmatrix} x \\ y_e \end{bmatrix}^T = \overline{N}w_n + \begin{bmatrix} B^TF \\ 0 \end{bmatrix} w_b.$$

Letting $\overline{N} = \begin{bmatrix} N & 0 \\ 0 & I \end{bmatrix}$, then (15) becomes

$$\begin{bmatrix} N^TAN & N^TB^TE \\ E^TBN & -D \end{bmatrix} \begin{bmatrix} w_{n1} \\ w_{n2} \end{bmatrix} = \lambda \begin{bmatrix} N^TGN & N^TB^TE \\ E^TBN & -D \end{bmatrix} \begin{bmatrix} w_{n1} \\ w_{n2} \end{bmatrix}. \quad (16)$$

- *j $(0 \le j \le n)$ eigenvectors of the form $\begin{bmatrix} x^T & y^T \end{bmatrix}^T$ that correspond to the case $\lambda \neq 1$.*

Proof. Proof of the form and linear independence of the $m + i + j$ eigenvalues obtained in a similar manner to the proof Theorem 3.2. □

To show that both the lower and upper bounds on the number of linearly independent eigenvectors can be attained we need only consider variations on Examples 2.5 and 2.6 from [15].

Example 4.1 (minimum bound). *Consider the matrices*

$$A = \begin{bmatrix} 1 & 2 & 1 & 0 \\ 2 & 2 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad P = \begin{bmatrix} 1 & 3 & 1 & 0 \\ 3 & 4 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},$$

such that $m = 2$, $n = 2$ and $p = 1$. The preconditioned matrix $P^{-1}A$ has an eigenvalue at 1 with multiplicity 4, but only two linearly independent eigenvectors which arise from case (1) of Theorem 4.2. These eigenvectors may be taken to be $\begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix}^T$ and $\begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}^T$.

Example 4.2 (maximum bound). *Let $A \in \mathbb{R}^{4\times 4}$ be as defined in Example 4.1, but assume that $G = A$. The preconditioned matrix $P^{-1}A$ has an eigenvalue at 1 with multiplicity 4 and clearly a complete set of eigenvectors. These may be taken to be*

From the earlier eigenvalue derivation, it is evident that the characteristic polynomial of the preconditioned linear system (18) is

$$\left(P^{-1}A - I\right)^{2m-p} \prod_{i=1}^{n-m+p} \left(P^{-1}A - \lambda_i I\right).$$

In order to prove the upper bound on the Krylov subspace dimension, we need to show that the order of the minimum polynomial is less than or equal to $\min\{n - m+p+2, n+m\}.$ Expanding the polynomial $\left(P^{-1}A - I\right)\prod_{i=1}^{n-m+p}\left(P^{-1}A - \lambda_i I\right)$ of degree $n + 1,$ we obtain

$$\begin{bmatrix} (\Theta_1 - I)\prod_{i=1}^{n-m+p}(\Theta_1 - \lambda_i I) & 0 \\ \Theta_2 \prod_{i=1}^{n-m+p}(\Theta_1 - \lambda_i I) & 0 \end{bmatrix}.$$

Since $G + B^T E D^{-1} E^T B$ is positive definite, Θ_1 has a full set of linearly independent eigenvectors and is diagonalizable. Hence, $(\Theta_1 - I)\prod_{i=1}^{n-m+p}(\Theta_1 - \lambda_i I) = 0.$ We therefore obtain

$\left(P^{-1}\right.$

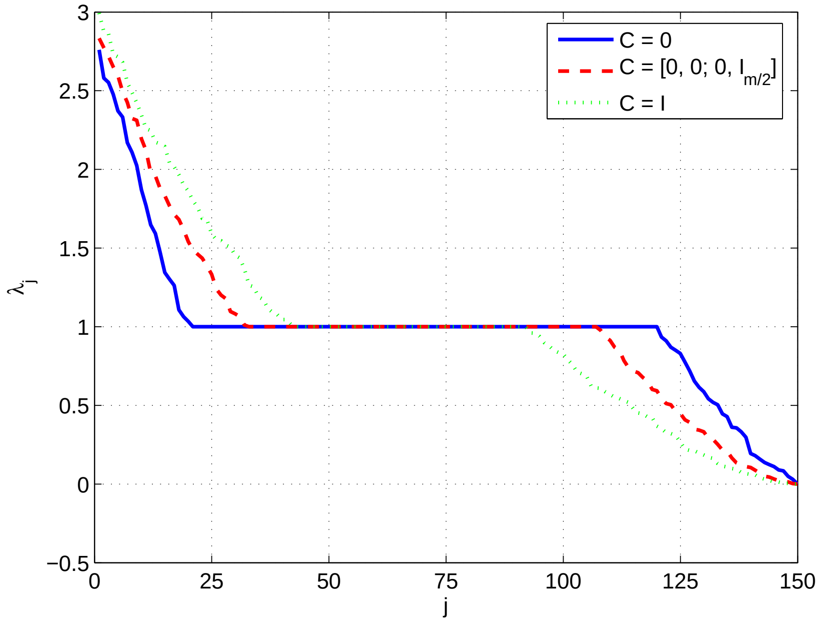

Figure 1: Distribution of the eigenvalues of $P^{-1}A$ for various choices of C.

interior-point method for such problems. We shall set $G = diag(A)$, $C = diag(0, \ldots, 0, 1, \ldots, 1)$ and vary the number of zeros on the diagonal of C so as to change its rank.

In Figure 1, we illustrate the change in the eigenvalues of the preconditioned system $P^{-1}A$ for three different choices of C. The eigenvalues are sorted so that

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{n+m}.$$

When $C = 0$, we expect there to be at least $2m$ unit eigenvalues [15]. We observe that our example has exactly $2m$ eigenvalues at 1. From Theorem 3.1, when $C = I$ there will be at least m unit eigenvalues. Our example has exactly m unit eigenvalues, Figure 1.

When C has rank $\frac{m}{2}$, then the preconditioned system $P^{-1}A$ has at least $\frac{3m}{2}$ unit eigenvalues, Theorem 4.1. Once again the number of unit eigenvalues for our example is exactly the lower bound given by the theorem.

Now suppose that we use (full) GMRES preconditioned by our extended constraint preconditioner with $G = diag(A)$ and vary the rank of C by changing the number of 1s along the diagonal of C (all other entries are zero). Figure 2 shows that with this choice of G there is a strong correlation between the upper bound on the Krylov subspace dimension and the number of iterations required to reduce the residual by at least a factor of 10^{-12}. This is an extreme example

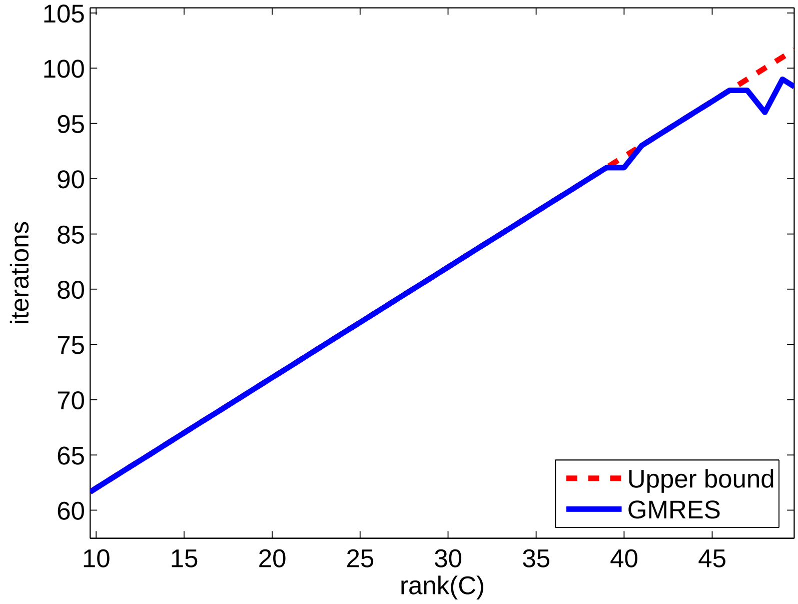

Figure 2: Comparison of upper bound on the Krylov subspace dimension and the number of iterations required to reduce the residual by $10^{-12}.$

and, as we will see in the following results, the number of iterations is often a lot lower than the upper bound on the Krylov subspace dimension.

Let us compare five different approaches for solving problems of the form (1). The matrix C is set to have rank $\lceil m/2 \rceil$ and to be diagonal with just entries of 0 and 1, as above. The indefiniteness of the matrix suggests the use of MINRES; we shall use the unpreconditioned version, although positive definite preconditioning could be employed, see [23]. We note that unpreconditioned MINRES is equivalent (in exact arithmetic) to unpreconditioned GMRES for these examples because A

Problem	m	n	MINRES	GMRES(I)	GMRES(D)	GMRES($\tilde{A}$)	PPCG(D)	PPCG($\tilde{A}$)
CVXQP1_M	500	1000	—	547	251	396	95	90
CVXQP2_M	750	1000	—	623	240	192	82	31
GOULDQP2_S	349	699	108	23	20	76	10	1
KSIP	1001	1021	41	9	1	13	1	1
MOSARQP1	700	3200	147	57	10	30	8	3

Table 1: Comparison of different Krylov subspace methods and preconditioners for some of the CUTEr test problems

(20) which we denote by PPCG($\tilde{A}$). Dollar, Gould, Schilders and Wathen show that the PPCG($\tilde{A}$) method will terminate (with exact arithmetic) in at most $\min\{2m, n - m + \lceil \frac{m}{2} \rceil\} + 1$ iterations.

The saddle point systems are all preprocessed such that the first

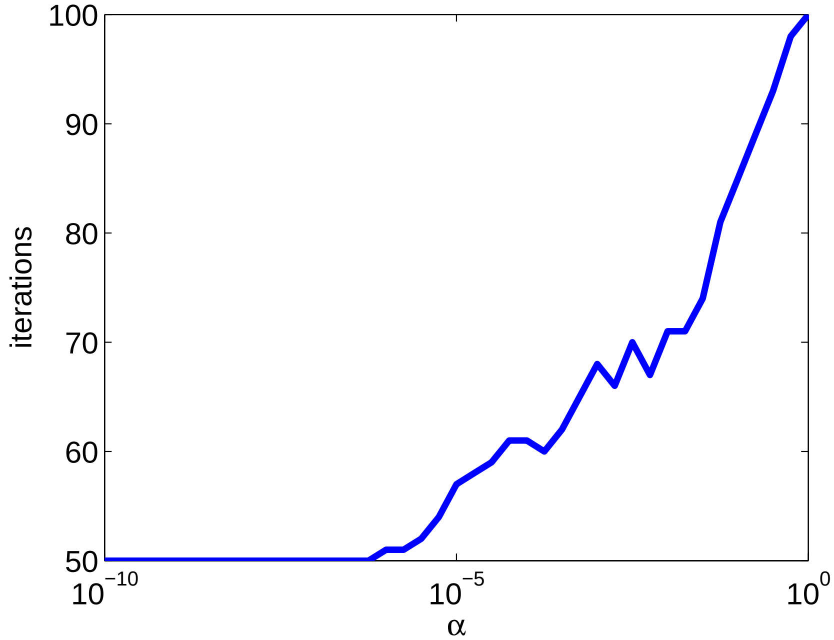

Figure 3: Comparison of the number of PPCG iterations for $C = \alpha I$ and varying $\alpha.$ The right hand sides have been set to be equal to the sum of the columns of A.

the possible use of implicit-factorization constraint preconditioners which only require small factorizations to be carried out [7, 8, 9]. Dollar, Gould, Schilders and Wathen also consider how G might be chosen to further increase the number of eigenvalues at 1 [7].

6 Conclusions

In this paper, we investigated a class of preconditioners for regularized saddle point matrix systems that incorporate the (1,2), (2,1) and (2,2) blocks of the original matrix. We showed that the inclusion of these blocks in the preconditioner clusters at least $2m - p$ eigenvalues at 1, regardless of the structure of $G.$ However, the standard convergence theory for Krylov subspace methods is not readily applicable because, in general, $P^{-1}A$ does not have a complete set of linearly independent eigenvectors. Using a minimum polynomial argument, we found a general (sharp) upper bound on the number of iterations required to solve linear systems of the form (1).

To confirm the analytical results of this paper we used a subset of problems

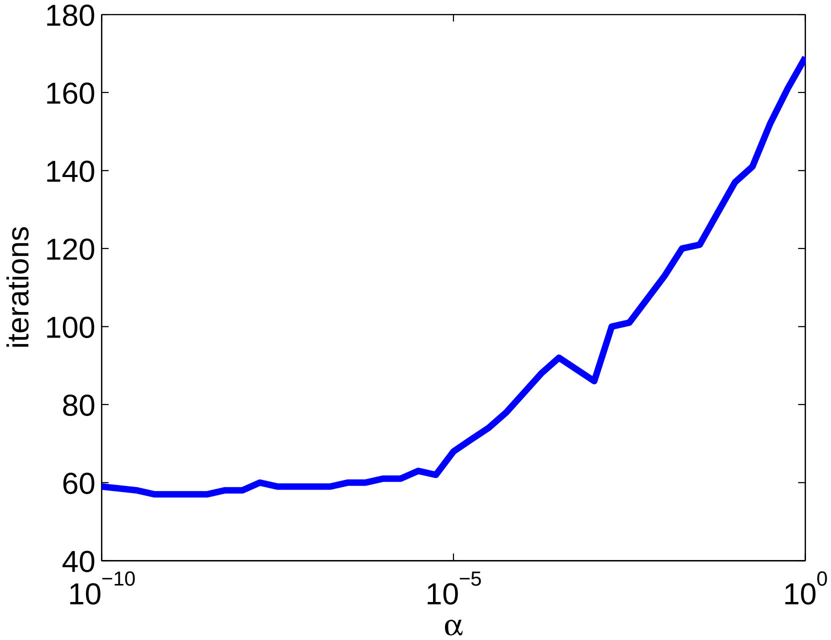

Figure 4: Comparison of the number of PPCG iterations for $C = \alpha I$ and varying α. The right hand side is a random vector.

from the CUTEr test set. We firstly used the CVXQP1_S problem and varied the rank of C to confirm the lower bound on the number of unit eigenvalues and the upper bound on the Krylov subspace dimension. We also compared MINRES for the unpreconditioned matrix system with the GMRES and PPCG methods where the preconditioner incorporate the (1,2), (2,1) and (2,2) blocks of the original matrix. We observed that the preconditioned methods resulted in a considerable reduction in the number of iterations required to reach our desired tolerance. Since GMRES and PPCG minimize different quantities, the number of iterations required may vary although the same preconditioner is used; indeed, we observe this in our results. We also confirmed that as the entries of C approach zero the number of PPCG iterations will decrease because of the additional clustering of eigenvalues around.

We have assumed that the sub-matrices B, B^T and $-C$ in (1) are exactly reproduced in the preconditioner. For truly large-scale problems this will be unrealistic [4, 5] but the theorems in this paper may still be of some interest in the inexact setting as a guide for choosing preconditioners. We wish to investigate this possibility in our future work.

,-521(S.)3740(Dol3sic]TJ/F4)37400-n3sicAS.then

Acknowledgements

The author would like to thank Nick Gould, Wil Schilders, Andy Wathen and the referees for their helpful input during the process of the work.

References

[1] A. Altman and J. Gondzio, *Regularized symmetric indefinite systems in interior point methods for linear and quadratic optimization*, Optim. Methods Softw., 11/12 (1999), pp. 275{302. Interior point methods.

[2] O. Axelsson and V. A. Barker, *Finite element solution of boundary value problems. Theory and computation*, vol. 35 of Classics in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2001. Reprint of the 1984 original.

[3] M. Benzi, G. H. Golub, and J. Liesen, *Numerical solution of saddle point problems*, Acta Numerica, 14 (2005), pp. 1{137.

[4] L. Bergamaschi, J. Gondzio, M. Venturin, and G. Zilli, *Inexact constraint preconditioners for linear systems arising in interior point methods*, Tech. Report MS-005-002, University of Edinburgh, 2005.

[5] G. Biros and O. Ghattas, *A Lagrange-Newton-Krylov-Schur method for PDE-constrained optimization*, SIAG/Optimization Views-and-News, 11 (2000), pp. 12{18.

[6] H. S. Dollar, *Iterative Linear Algebra for Constrained Optimization*, Doctor of Philosophy, Oxford University, 2005.

[7] H. S. Dollar, N. I. M. Gould, W. H. A. Schilders, and A. J. Wathen, *Implicit-factorization 66(as)]TJ/F49.96Tf131Tf-300(in)3(O.)-5iorative0-523(Line1ion)-v9*

for

[11] N. I. M. Gould, *Iterative methods for ill-conditioned linear systems from optimization*, in Nonlinear Optimization and Related Topics, G. DiPillo and F. Giannessi, eds., Dordrecht, The Netherlands, 1999, Kluwer Academic Publishers, pp. 123{142.

[12] N. I. M. Gould, M. E. Hribar, and J. Nocedal, *On the solution of equality constrained quadratic programming problems arising in optimization*

[24] H. A. van der Vorst, *Iterative Krylov Methods for Large Linear Systems*, Cambridge Monographs on Applied and Computational Mathematics, Cambridge University Press, Cambridge, United Kingdom, 1 ed., 2003.

[25] R. J. Vanderbei, *Symmetric quaside¯nite matrices*, SIAM J. Optim., 5 (1995), pp. 100{113.

[26] S. J. Wright, *Primal-dual interior-point methods*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997.