NUMERICAL LINEAR ALGEBRA WITH APPLICATIONS
Numer. Linear Algebra Appl. 2016; **23**:693–705
Published online 13 April 2016 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/nla.2050

Towards an optimal condition number of certain augmented Lagrangian-type saddle-point matrices

R. Estrin and C. Greif*,†

The University of British Columbia, Department of Computer Science, 2366 Main Mall, Vancouver, BC V6T 1Z4, Canada

SUMMARY

We present an analysis for minimizing the condition number of nonsingular parameter-dependent 2×2 block-structured saddle-point matrices with a maximally rank-deficient (1,1) block. The matrices arise from an augmented Lagrangian approach. Using quasidirect sums, we show that a decomposition akin to simultaneous diagonalization leads to an optimization based on the extremal nonzero eigenvalues and singular values of the associated block matrices. Bounds on the condition number of the parameter-dependent matrix are obtained, and we demonstrate their tightness on some numerical examples. Copyright © 2016 John Wiley & Sons, Ltd.

Received 25 September 2015; Revised 28 February 2016; Accepted 12 March 2016

KEY WORDS: saddle-point matrices; condition number; singular values; eigenvalues

1. INTRODUCTION

Consider the saddle-point system

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

where $A \in \mathbb{R}^{n \times n}$ is assumed to be symmetric-positive semidefinite with rank $n - m$, $B \in \mathbb{R}^{m \times n}$ with $m < n$, and $u, f \in \mathbb{R}^n$, $p, g \in \mathbb{R}^m$. Let us denote the coefficient matrix of (1) by

$$K = \begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix}. \tag{2}$$

We will assume throughout that B has full row rank and that K is nonsingular. The requirement $\operatorname{rank}(A) = n - m$ is rather significant, as it limits us to consider a very specific class of problems. We say that A is *maximally rank-deficient* because it has the minimal rank that can still allow a nonsingular K. We have recently shown in [1] that there are several applications that lead to saddle-point matrices of this type, and these matrices have unique properties. For example, their inverse has a special nonzero structure, and specialized preconditioners can be designed for solving the corresponding linear systems.

Suppose $W \in \mathbb{R}^{m \times m}$ is a nonsingular (typically symmetric positive-definite) weight matrix. Then (1) can be reformulated as follows:

$$\begin{pmatrix} A + B^T W^{-1} B & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f + B^T W^{-1} g \\ g \end{pmatrix}. \tag{3}$$

*Correspondence to: C. Greif, The University of British Columbia, Department of Computer Science, 2366 Main Mall, Vancouver, BC V6T 1Z4, Canada

†E-mail: greif@cs.ubc.ca

We will denote the matrix of (3) as follows:

$$K(W) = \begin{pmatrix} A + B^T W^{-1} B & B^T \\ B & 0 \end{pmatrix}.$$

Notice that $K(0)$ is identical to the matrix K defined in (2) and associated with (1).

While (3) is mathematically equivalent to (1), from a numerical point of view, $K(W)$ and K may be very different in terms of conditioning, spectral structure, and other aspects [2]. Under the assumptions we make in this paper, A is singular, whereas $A + B^T W^{-1} B$ is nonsingular, because a necessary condition for the nonsingularity of K is that the null spaces of A and B do not intersect except at the zero vector [3, Section 3]. This, in turn, potentially enriches the family of solution methods that may be used for solving (3) in comparison with solution methods for solving (1); specifically, the Schur complement $B(A + B^T W^{-1} B)^{-1} B^T$ is defined (which has a simple structure as shown in [1]), whereas the analogous Schur complement of K associated with A is undefined owing to the singularity of A; for solution methods and eigenvalue estimates based on Schur complements see, for example, [3–7] and the references therein.

A desirable goal is to find a numerically good choice for W. In certain applications, such a choice may often be based on the underlying application (e.g., [8]), where a scalar Laplacian is used. In the absence of specific characteristics of the underlying matrices, a possible consideration may be to seek to improve the conditioning of the linear system using a simple choice of W. To that end, we will assume that W is a scaled identity matrix,

$$W^{-1} = \gamma I_m,$$

and study when we can expect the condition number of the leading block and the saddle-point matrix to improve (desirably simultaneously) as a function of γ. A reduction in the condition number may lead to more accurate numerical solutions. In the case of iterative solvers, it may also result in faster convergence, although factors other than the condition number (such as clustering of eigenvalues) are just as important.

With a slight abuse of notation, let us denote the associated matrix $K(W) = K(\gamma I)$ as

$$K(\gamma) = \begin{pmatrix} A + \gamma B^T B & B^T \\ B & 0 \end{pmatrix}$$

and the leading block as

$$A(\gamma) = A + \gamma B^T B.$$

Notice that $A(0) \equiv A$.

The approach based on (3) or its simplified form with $K(\gamma)$ has been extensively explored in the literature. It is related to the technique of augmented Lagrangian in constrained optimization [9–11], and has been studied in [2] and other places. Related methods have been successfully applied in the solution of saddle-point systems arising from numerical solution of partial differential equations with constraints, notably in fluid flow [12–14] and time-harmonic Maxwell equations [8] (see [2, 3] for additional references). The need to deal with linear systems involving $A(\gamma)$ may arise either as a subproblem within the augmented saddle-point problem or independently. For example, in the numerical solution of partial differential equations arising in electromagnetics, the discrete operator A may represent a curl-curl operator, and it is possible to remove the singularity associated with this operator by adopting the strategy discussed here [8, 15]. In the constrained optimization front, the need to solve systems associated with $A(\gamma)$ has arisen in various articles as part of the revised interest in the alternating direction method of multipliers [16].

However, to the best of our knowledge, the specific setting where A has rank $n - m$ has not been studied.

In Section 2 we derive a decompositional relation that allows us to tie the choice of γ to the extremal nonzero eigenvalues of A and the extremal singular values of B. We show that our choice optimizes the condition number of $A(\gamma)$. In Section 3, we optimize the condition number of $K(\gamma)$. Our observations are accompanied by numerical experiments in Section 4 that validate our analysis. Finally, in Section 5, we draw some conclusions.

DOI: 10.1002/nla

Notation. Throughout the paper, we use for matrices the norm notation $\|.\|$ to mean the induced two-norm, $\|.\|_2$.

2. OPTIMIZING THE CONDITION NUMBER OF $A(\gamma)$

A small condition number of $A(\gamma)$ and $K(\gamma)$ may be beneficial in the derivation of numerically stable solution methods. We start our quest of optimizing the condition number by constructing a decomposition that is related to a simultaneous diagonalization of the matrices involved. The following result uses the *quasidirect sum* of matrices [17].

Proposition 2.1
Let $M, N \in \mathbb{R}^{n\times n}$ and let $\text{rank}(M) = r$, $\text{rank}(N) = n-r$, such that $M+N$ is nonsingular. Then there exist nonsingular matrices $P, Q \in \mathbb{R}^{n\times n}$ and nonsingular $S \in \mathbb{R}^{r\times r}$, $T \in \mathbb{R}^{(n-r)\times(n-r)}$ such that

$$M = P\begin{pmatrix} S & 0 \\ 0 & 0\end{pmatrix} Q^T\,; \quad N = P\begin{pmatrix} 0 & 0 \\ 0 & T\end{pmatrix} Q^T.$$

Proof
The proof is straightforward, using SVD, and the decomposition may not be unique. We simply consider the reduced (economy size) SVD of M and N:

$$M = USV^T\,; \quad N = WTZ^T,$$

where $U, V \in \mathbb{R}^{n\times r}$, $W, Z \in \mathbb{R}^{n\times(n-r)}$, all with orthonormal columns, and $S \in \mathbb{R}^{r\times r}$, $T \in \mathbb{R}^{(n-r)\times(n-r)}$ are diagonal. We can then construct the desired decomposition with $P = [U\ W]$, $Q = [V\ Z]$. Because $M+N$ is nonsingular, it follows that P, Q are nonsingular. □

Proposition 2.1 does not require symmetry, and thus, it applies to settings that go beyond the assumptions we make in this paper. The proposition leads to a couple of interesting observations that are unique to matrices with the rank structure we are interested in:

Corollary 2.1
For matrices $M, N \in \mathbb{R}^{n\times n}$, with $\text{rank}(M) = r$, $\text{rank}(N) = n-r$ and $M+N$ nonsingular, we have that

$$M(M+N)^{-1}N = 0, \tag{4}$$

and $(M+N)^{-1}M$ has eigenvalues $\lambda = 0, 1$ with multiplicities $n-r$ and r, respectively.

Proof
We decompose M, N according to Proposition 2.1 and observe that

$$\begin{aligned} M\,(M+N)^{-1}\,N &= P\begin{pmatrix} S & 0 \\ 0 & 0\end{pmatrix} Q^T\left(P\begin{pmatrix} S & 0 \\ 0 & T\end{pmatrix} Q^T\right)^{-1} P\begin{pmatrix} 0 & 0 \\ 0 & T\end{pmatrix} Q^T \\ &= P\begin{pmatrix} S & 0 \\ 0 & 0\end{pmatrix}\begin{pmatrix} S & 0 \\ 0 & T\end{pmatrix}^{-1}\begin{pmatrix} 0 & 0 \\ 0 & T\end{pmatrix} Q^T \\ &= 0. \end{aligned}$$

The multiplicity of the eigenvalues follows because $\text{rank}((M+N)^{-1}M) = r$, and so it is possible to show that this matrix has a projection property:

$$\begin{aligned} \left((M+N)^{-1}M\right)^2 &= (M+N)^{-1}(M+N-N)(M+N)^{-1}M \\ &= (M+N)^{-1}M - (M+N)^{-1}N(M+N)^{-1}M \\ &= (M+N)^{-1}M, \end{aligned}$$

DOI: 10.1002/nla

where we used (4) to transition from the second to the third equation. Because $\text{null}(M) = n - r$, we obtain the desired eigenvalue multiplicities. □

The two-norms of the matrices P, Q, which are concatenations of rectangular matrices with orthonormal columns, are bounded by a constant. Indeed, let $P = [U_1\ U_2\ \dots\ U_k]$, where each $U_i \in \mathbb{R}^{n\times r_i}$, $\text{rank}(U_i) = r_i$, $U_i^T U_i = I$ and $\sum_{i=1}^k r_i = n$. By the triangular inequality for matrix norms

$$\begin{aligned} \|PP^T\| &= \left\| \sum_{i=1}^k U_i U_i^T \right\| \\ &\leqslant \sum_{i=1}^k \|U_i U_i^T\| \\ &= k, \end{aligned} \tag{5}$$

from which it follows that $\|P\| \leqslant \sqrt{k}$. However, the norm of the inverse, namely $\|P^{-1}\|$, cannot be similarly bounded. We note here that we are particularly interested in the case $k = 2$.

Let us turn our attention to the conditioning of $A(\gamma)$. Let

$$\lambda_1 > \lambda_2 > \cdots > \lambda_{n-m} > 0 \qquad \text{and} \qquad \sigma_1 > \sigma_2 > \cdots > \sigma_m > 0$$

be the eigenvalues of A and singular values of B, respectively. Define

$$\alpha = \min\left\{ \frac{\|A\|}{\|B\|^2}, \frac{\|B^\dagger\|^2}{\|A^\dagger\|} \right\} = \min\left\{ \frac{\lambda_1}{\sigma_1^2}, \frac{\lambda_{n-m}}{\sigma_m^2} \right\}; \tag{6}$$

$$\beta = \max\left\{ \frac{\|A\|}{\|B\|^2}, \frac{\|B^\dagger\|^2}{\|A^\dagger\|} \right\} = \max\left\{ \frac{\lambda_1}{\sigma_1^2}, \frac{\lambda_{n-m}}{\sigma_m^2} \right\}, \tag{7}$$

where the superscript † denotes pseudo-inverse. It was experimentally observed in [2] that the optimal γ, in terms of minimizing the condition number of $A(\gamma)$, typically lies in a neighborhood of $\frac{\|A\|}{\|B\|^2}$. We now show that if $\alpha < \gamma < \beta$, then the condition number of $A(\gamma)$ is reduced to near-optimality.

We decompose A and $B^T B$ as in Proposition 2.1. Let

$$A = USU^T, \quad B = WTZ^T, \tag{8}$$

be the economy-size SVD of A and B, respectively. Note that for A, this is not quite the spectral decomposition, because zero eigenvalues are not included; the matrix S is smaller in dimensions than A. It follows that the columns of Z form the eigenvectors of $B^T B$. Define

$$P = [U\ Z]. \tag{9}$$

Then we have

$$A(\gamma) = P\Sigma P^T,$$

where

$$\Sigma \equiv \begin{pmatrix} S & 0 \\ 0 & \gamma T^2 \end{pmatrix}. \tag{10}$$

Because Σ is the only matrix that depends on γ, this simplifies our analysis for $\|A(\gamma)\|$. We can write

$$A^{-1}(\gamma) = P^{-T} \begin{pmatrix} S^{-1} & 0 \\ 0 & \gamma^{-1} T^{-2} \end{pmatrix} P^{-1},$$

and can then study the effect of γ on $\left\| A^{-1}(\gamma) \right\|$.

Although P is not orthogonal, we can still use $\kappa(\Sigma)$ to bound $\kappa(A(\gamma))$ to within a constant from above and below, as we now show.

Theorem 2.1
Given A, B, P, Σ be as defined earlier, we have

$$\kappa^{-2}(P)\,\kappa(\Sigma) \leqslant \kappa(A(\gamma)) \leqslant \kappa^{2}(P)\,\kappa(\Sigma),$$

so the trend of the growth of $\kappa(A(\gamma))$ is determined by $\kappa(\Sigma)$.

Proof
This result follows immediately from the fact that

$$\left\| P \Sigma P^{T} \right\| \geqslant \left\| P^{-1} \right\|^{-2} \|\Sigma\|,$$

and by the triangle inequality on norms for matrix multiplication. □

The aforementioned bound is tight in the sense that it holds as an equality when P is orthogonal. If P is not 'near-orthogonal', that is, A and $B^T B$ span subspaces of $\mathbb{R}^{n\times n}$ that greatly overlap, then this bound may be weaker, because the condition number of P might be large.

Noting that γ affects the conditioning of Σ, not of P, we now study the effect of γ on Σ for various cases. As we shall see, different regions of values of γ relative to α and β give us a different characterization of the growth of $\|A(\gamma)\|$ and $\|A^{-1}(\gamma)\|$.

We partition $\mathbb{R}^+$ into three regions, for γ in intervals $(0,\alpha)$, (β,∞), and $[\alpha,\beta]$, and proceed to analyze each of these cases.

Case 1. $\underline{0 < \gamma < \alpha}$. In this case, we have that $\lambda_1 > \gamma\sigma_1^2$ and $\lambda_{n-m}^{-1} < \gamma^{-1}\sigma_m^{-2}$, and thus

$$\begin{aligned} \|\Sigma\| &= \|A\|; \\ \left\|\Sigma^{-1}\right\| &= \gamma^{-1} \left\| \left(B^T B\right)^{\dagger} \right\|. \end{aligned}$$

We can see that for small γ, $\|A(\gamma)\|$ will stay relatively constant, while the norm of the inverse would grow asymptotically like γ^{-1}, as γ goes to infinity. Then $\kappa(A(\gamma)) \in \Theta(\gamma^{-1})$, which shrinks as γ increases.

Case 2. $\underline{\gamma > \beta}$. In this case, we have that $\lambda_1 < \gamma\sigma_1^2$ and $\lambda_{n-m}^{-1} > \gamma^{-1}\sigma_m^{-2}$, and thus

$$\begin{aligned} \|\Sigma\| &= \gamma \left\| B^T B \right\|; \\ \left\|\Sigma^{-1}\right\| &= \left\| A^{\dagger} \right\|. \end{aligned}$$

We can see that for large γ, $\left\| A(\gamma)^{-1} \right\|$ will stay relatively constant, while $\|A(\gamma)\|$ would grow like γ. Then $\kappa(A(\gamma)) \in \Theta(\gamma)$, which grows as γ increases.

Case 3. $\underline{\alpha < \gamma < \beta}$. Note that we may have either $\frac{\lambda_1}{\sigma_1^2} > \frac{\lambda_{n-m}}{\sigma_m^2}$ or the reverse, depending on the problem. In the following, we cover both scenarios.

First, suppose that $\frac{\lambda_{n-m}}{\sigma_m^2} < \gamma < \frac{\lambda_1}{\sigma_1^2}$. In this case, we have that $\lambda_1 > \gamma\sigma_1^2$ and $\lambda_{n-m}^{-1} > \gamma^{-1}\sigma_m^{-2}$, and thus

$$\begin{aligned} \|\Sigma\| &= \|A\|; \\ \left\|\Sigma^{-1}\right\| &= \left\| A^{\dagger} \right\|. \end{aligned}$$

In this case, we see that $\kappa(A(\gamma)) \simeq \kappa(\Sigma) = \kappa(S)$ is constant.

DOI: 10.1002/nla

Next, we consider $\frac{\lambda_1}{\sigma_1^2} < \gamma < \frac{\lambda_{n-m}}{\sigma_m^2}$. In this case, we have that $\lambda_1 < \gamma\sigma_1^2$ and $\lambda_{n-m}^{-1} < \gamma^{-1}\sigma_m^{-2}$, and thus

$$\|\Sigma\| = \gamma \left\|B^T B\right\|;$$
$$\left\|\Sigma^{-1}\right\| = \gamma^{-1} \left\|\left(B^T B\right)^{\dagger}\right\|.$$

Again, in this case, we see that $\kappa\left(A\left(\gamma\right)\right) \simeq \kappa\left(\Sigma\right) = \kappa\left(T^2\right)$.

It remains to show that $\kappa(\Sigma)$ is minimized when $\gamma \in [\alpha, \beta]$. Because $\kappa(\Sigma)$ is continuous in γ, and because $\kappa(\Sigma)$ is decreasing in γ for $\gamma < \alpha$ and increasing in γ for $\gamma > \beta$, necessarily it must be minimized within $[\alpha, \beta]$. With $\kappa(\Sigma)$ approximating $\kappa(A+\gamma B^T B)$, we would expect the true condition number to be minimized within that interval.

Because the condition number of P does not depend on γ, the aforementioned observations are valid for a sufficiently large γ. However, when P is very ill-conditioned, it may indeed take a large γ to identify the asymptotic behavior that we have characterized earlier.

In summary, we see that our analysis of $\kappa\left(A\left(\gamma\right)\right)$ boils down to dealing merely with the algebraic relationships among extremal nonzero eigenvalues and singular values, simplifying previous attempts to analyze this problem. Theorem 2.1 can determine the regions of γ that are non-optimal, although the bounds require P to be orthogonal, in order for them to be tight. As we will see in Section 4, our bounds are remarkably accurate when $\kappa(P)$ is modest. When P is ill-conditioned, we cannot expect anymore the bounds to be very tight, but the trend is still fully captured. The ranges of α, β provided in (6), (7) provide useful bounds on practical choices for γ.

3. OPTIMIZING THE CONDITION NUMBER OF $K(\gamma)$

We now perform a similar analysis to minimize the condition number of $K(\gamma)$. Recalling the eigendecomposition of A and the SVD of B from (8) and P defined in (9), we can decompose $K(\gamma)$ as follows:

$$K(\gamma) = P'R\left(P'\right)^T \tag{11}$$

where

$$P' = \begin{pmatrix} P & 0 \\ 0 & W \end{pmatrix} \quad \text{and} \quad R = \begin{pmatrix} \Sigma & (T')^T \\ T & 0 \end{pmatrix},$$

with Σ from (10) and $T' = [0\ T] \in \mathbb{R}^{m\times n}$. Similarly to our analysis of $A(\gamma)$ in Section 2, γ does not affect the conditioning of P', and thus, we are concerned with minimizing the condition number of the middle matrix in (11), R. That said, an ill-conditioned P' may require a larger γ to enter the asymptotic behavior that we are set out to characterize. We now study R by seeking a result analogous to Theorem 2.1.

Theorem 3.1
Let P, $K(\gamma)$, and R be defined in (9) and (11). Then

$$\kappa^{-2}\left(P\right)\kappa\left(R\right) \leqslant \kappa\left(K(\gamma)\right) \leqslant \kappa^{2}\left(P\right)\kappa\left(R\right),$$

so the trend of the growth of $\kappa\left(K(\gamma)\right)$ is determined by $\kappa\left(R\right)$.

The proof of this theorem is exactly the same as that of Theorem 2.1, with the additional observation that because P' is block-diagonal and W is orthogonal, then $\kappa\left(P'\right) = \kappa\left(P\right)$. As before, this bound can be simplified using (5).

To analyze R, we first apply a symmetric permutation as was carried out in [2, Lemma 2.6]. We define a permutation vector in MATLAB notation as

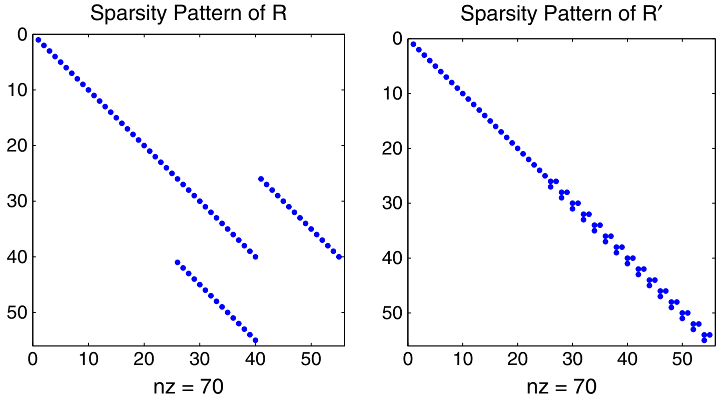

Figure 1. Sparsity patterns of R and R': the original matrix is on the left, and the permuted one is on the right.

$$\bar{p} = [1 : n-m, n-m+1, n+1, n-m+2, n+2, n-m+3, n+3, \dots, n, n+m],$$

and apply the symmetric permutation to R. Let the permuted matrix be

$$R' \equiv R(\bar{p}, \bar{p}) = \begin{pmatrix} S & 0 \\ 0 & T^d \end{pmatrix},$$

where $S = \operatorname{diag}(\lambda_i)$ is the matrix of eigenvalues for A and T^d is a block-diagonal matrix with m 2×2 blocks such that

$$T^d = \operatorname{diag}\begin{pmatrix} \gamma\sigma_i^2 & \sigma_i \\ \sigma_i & 0 \end{pmatrix}. \tag{12}$$

The sparsity patterns of R and R' can be found in Figure 1.

Having permuted R into a matrix with a more favorable sparsity pattern, we can begin a similar analysis of $\|R\| = \|R'\|$ and $\left\|R^{-1}\right\| = \left\|(R')^{-1}\right\|$.

Note that the eigenvalues of T^d in (12) are

$$\mu_i(\gamma)^{\pm} = \frac{1}{2}\left(\gamma\sigma_i^2 \pm \sqrt{\gamma^2\sigma_i^4 + 4\sigma_i^2}\right), \tag{13}$$

while the eigenvalues of the inverse are

$$\left(\mu_i^{-1}(\gamma)\right)^{\pm} = \frac{1}{2}\left(-\gamma \pm \sqrt{\gamma^2 + \frac{4}{\sigma_i^2}}\right). \tag{14}$$

Remark 3.1
From here henceforth, unless otherwise noted, define $\mu_m^{-1}(\gamma) = \left|\mu_m^{-1}(\gamma)^-\right|$ and $\mu_1(\gamma) = \mu_1(\gamma)^+$.

It should be noted that for fixed γ, $\mu_i(\gamma)^+$ decreases monotonically and $\mu_i^{-1}(\gamma)^-$ increases in magnitude (that is, becomes more negative) as i increases, $0 \leqslant i \leqslant m$. For a fixed σ_i, $\mu_i(\gamma)^+$ and $\mu_i^{-1}(\gamma)^-$ grow monotonically with γ. Thus, we have that $\|R\| = \max\left\{\lambda_1, \mu_1^+(\gamma)\right\}$ and $\left\|R^{-1}\right\| = \min\left\{\lambda_{n-m}^{-1}, \mu_m^{-1}(\gamma)^-\right\}$, because R' is block-diagonal and each block is symmetric. Define

$$\begin{aligned} \psi &= \min\left\{\frac{\|A\|}{\|B\|^2} - \|A\|^{-1}, \left\|A^{\dagger}\right\| - \frac{\left\|B^{\dagger}\right\|^2}{\left\|A^{\dagger}\right\|}\right\} \\ &= \min\left\{\frac{\lambda_1}{\sigma_1^2} - \frac{1}{\lambda_1}, \frac{1}{\lambda_{n-m}} - \frac{\lambda_{n-m}}{\sigma_m^2}\right\}; \end{aligned} \tag{15}$$

$$\begin{aligned}\omega &= \max\left\{\frac{\|A\|}{\|B\|^2} - \|A\|^{-1}, \left\|A^{\dagger}\right\| - \frac{\left\|B^{\dagger}\right\|^2}{\left\|A^{\dagger}\right\|}\right\} \\ &= \max\left\{\frac{\lambda_1}{\sigma_1^2} - \frac{1}{\lambda_1}, \frac{1}{\lambda_{n-m}} - \frac{\lambda_{n-m}}{\sigma_m^2}\right\}.\end{aligned} \tag{16}$$

Again, we partition $\mathbb{R}$ into three intervals, $(0, \psi)$, $[\psi, \omega]$, and (ω, ∞), and examine $\kappa(R)$ for each interval. Note that it is possible for $\frac{\lambda_1}{\sigma_1^2} - \frac{1}{\lambda_1}, \frac{1}{\lambda_{n-m}} - \frac{\lambda_{n-m}}{\sigma_m^2}$ to be negative, and we restrict $\gamma > 0$.

Case 1. $\underline{0 < \gamma < \psi}$. For $\gamma < \psi$, it can be verified that $\lambda_1 > \mu_1(\gamma)$ and that $\lambda_{n-m}^{-1} > \mu_m^{-1}(\gamma)$. Thus, we have

$$\|R\| = \|A\|;$$
$$\left\|R^{-1}\right\| = \left\|A^{\dagger}\right\|.$$

Thus, for small γ, $\kappa(R) = \kappa(A)$ and thus remains constant. As will be shown in the following cases, it is for small γ (in fact, for $\gamma = 0$), that the condition number of R is minimized. Although using small γ would be tempting, it was seen in Section 2 that $A + \gamma B^T B$ is poorly conditioned. Note that assuming $\psi > 0$ imposes a constraint on the eigenvalues of A and the singular values of B.

Case 2. $\underline{\gamma > \omega}$. In this case, it can be verified that $\lambda_1 < \mu_1(\gamma)$ and that $\lambda_{n-m}^{-1} < \mu_m^{-1}(\gamma)$. Then

$$\|R\| = \|\mu_1(\gamma)\|;$$
$$\left\|R^{-1}\right\| = \left\|\mu_m^{-1}(\gamma)\right\|.$$

It can be seen that $\kappa(R) \in \Theta(\gamma^2)$ for large γ, agreeing with the results from [2].

Case 3. $\underline{\psi < \gamma < \omega}$. We again split this case based on whether $\frac{\lambda_1}{\sigma_1^2} - \frac{1}{\lambda_1} < \frac{1}{\lambda_{n-m}} - \frac{\lambda_{n-m}}{\sigma_m^2}$ or the reverse.

Firstly, let $\frac{\lambda_1}{\sigma_1^2} - \frac{1}{\lambda_1} < \gamma < \frac{1}{\lambda_{n-m}} - \frac{\lambda_{n-m}}{\sigma_m^2}$. Then $\lambda_1 < \mu_1(\gamma)$, while $\lambda_{n-m}^{-1} > \mu_m^{-1}(\gamma)$, leading to

$$\|R\| = \|\mu_1(\gamma)\|;$$
$$\left\|R^{-1}\right\| = \left\|A^{\dagger}\right\|.$$

In the second case, $\frac{\lambda_1}{\sigma_1^2} - \frac{1}{\lambda_1} > \gamma > \frac{1}{\lambda_{n-m}} - \frac{\lambda_{n-m}}{\sigma_m^2}$. Then $\lambda_1 > \mu_1(\gamma)$, while $\lambda_{n-m}^{-1} < \mu_m^{-1}(\gamma)$, leading to

$$\|R\| = \|A\|;$$
$$\left\|R^{-1}\right\| = \left\|\mu_m^{-1}(\gamma)\right\|.$$

Thus in both subcases, we find that $\kappa(R) \in \Theta(\gamma)$.

After analyzing the condition numbers of both Σ and R, which asymptotically correlate (for γ sufficiently large) with the condition numbers of $A + \gamma B^T B$ and $K(\gamma)$, respectively, we see that we are interested in $\gamma \in (0, \psi] \cap [\alpha, \beta]$. It is entirely possible that the intersection is empty, and so one will want to choose γ that falls near the endpoints of one of the desired intervals.

Let us establish conditions under which $(0, \psi] \cap [\alpha, \beta] \neq \emptyset$. The following proposition captures the conditions under which we may obtain a non-empty intersection.

Proposition 3.1
Let A, B, be as given in previous sections. Then if

1. $\alpha = \frac{\lambda_1}{\sigma_1^2}$, $\psi = \frac{1}{\lambda_{n-m}} - \frac{\lambda_{n-m}}{\sigma_m^2}$ and $\frac{\lambda_1}{\sigma_1^2} + \frac{\lambda_{n-m}}{\sigma_m^2} < \frac{1}{\lambda_{n-m}}$
2. $\alpha = \frac{\lambda_{n-m}}{\sigma_m^2}$, $\psi = \frac{\lambda_1}{\sigma_1^2} - \frac{1}{\lambda_1}$ and $\frac{\lambda_1}{\sigma_1^2} - \frac{\lambda_{n-m}}{\sigma_m^2} > \frac{1}{\lambda_1}$, $\lambda_1 > \sigma_1$
3. $\alpha = \frac{\lambda_{n-m}}{\sigma_m^2}$, $\psi = \frac{1}{\lambda_{n-m}} - \frac{\lambda_{n-m}}{\sigma_m^2}$ and $\sqrt{2}\lambda_{n-m} < \sigma_m$

then $(0, \psi] \cap [\alpha, \beta] \neq \emptyset$.

Thus we have conditions where we can potentially find optimal condition numbers for both $A(\gamma)$ and $K(\gamma)$ *simultaneously*. If A or B are scaled individually, it is possible to shift α, β, ψ, ω to more favorable positions, although the effectiveness of this is problem-dependent.

We note that while our bounds depend only on two pairs of extremal (nonzero) eigenvalues and singular values, two of these four quantities may be tough to compute. Indeed, while the largest eigenvalue of A, λ_1, and the largest singular value σ_1 of B are expected to be relatively easy to compute at least when they are well separated from their second largest counterparts (in which case, we can effectively use, for example, a power method-type method), it is expected that λ_{n-m} and σ_m would be significantly harder to compute.

4. NUMERICAL EXPERIMENTS

We gauge the accuracy of our estimation of the condition number for varying γ and the desired value of the optimal γ by presenting two numerical examples. The first one deals with a well-conditioned matrix, and our bounds are shown to be remarkably tight. The second example is one of an ill-conditioned matrix, where we see that while the bounds are not as tight, they still capture the trend.

Example 4.1
The matrices A and B arise from the finite element method being applied to solving the time-harmonic Maxwell's equation [8]; A is 6080×6080 and represents a discrete curl-curl operator, and B represents a discrete divergence operator and is of dimensions 1985×6080. In the results presented in the following, we have scaled A and B by modest multiplicative factors to better illustrate the merits of our analysis. We have $\kappa(P) = 1.67$, so we see that A and $B^T B$ have nearly orthogonal column spaces, allowing P to be well conditioned. For our augmented matrix,

$$\alpha = 3.88,$$
$$\beta = 1734.$$

We then plot $\kappa\left(A + \gamma B^T B\right)$ against our estimation of the condition number as derived in Section 2. The lower bound is taken to be $\kappa^{-2}(P)\,\kappa(\Sigma)$, while the upper bound is taken to be $\kappa^2(P)\,\kappa(\Sigma)$.

We can see in Figure 2 the estimation of $\kappa\left(A + \gamma B^T B\right)$ matches the actual condition number well within an order of magnitude for both the upper and lower bounds. In theory, we cannot expect the variation of $\kappa\left(A + \gamma B^T B\right)$ to be constant over the interval $\alpha < \gamma < \beta$, because our bound is affected by the spectrum of P. At the same time, the condition number of Σ is definitely constant over this interval, because we have $\lambda_{n-m} < \gamma\sigma_i^2 < \lambda_1$. Pleasingly, we are seeing in Figure 2 that the variation in the condition number is minor.

We report a similar situation for our estimation of $\kappa\left(K(\gamma)\right)$. We have

$$\psi = 62.5,$$
$$\omega = 1734.$$

As can be seen in Figure 3, the eigenvalues of A and singular values of B allow for three distinct regions where $\kappa(R)$ is flat, growing linearly, and growing quadratically. Therefore, we would be

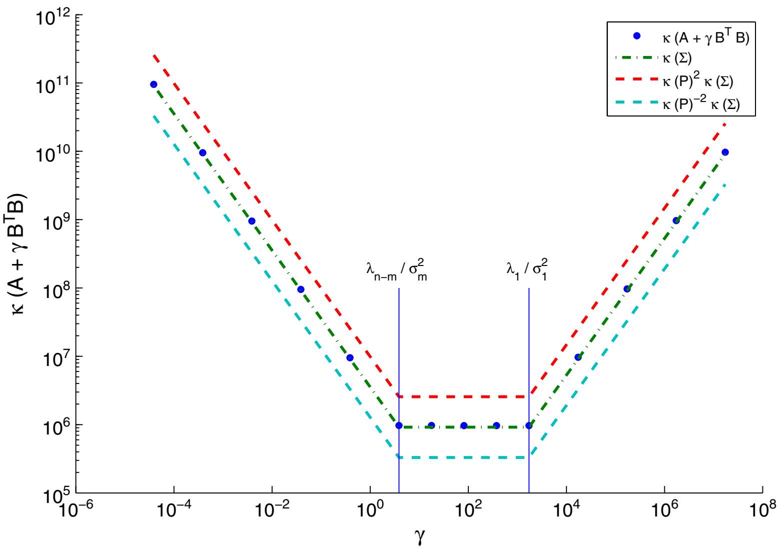

Figure 2. Condition number of augmented leading block as a function of γ, for a problem arising from the discretized time-harmonic Maxwell equations.

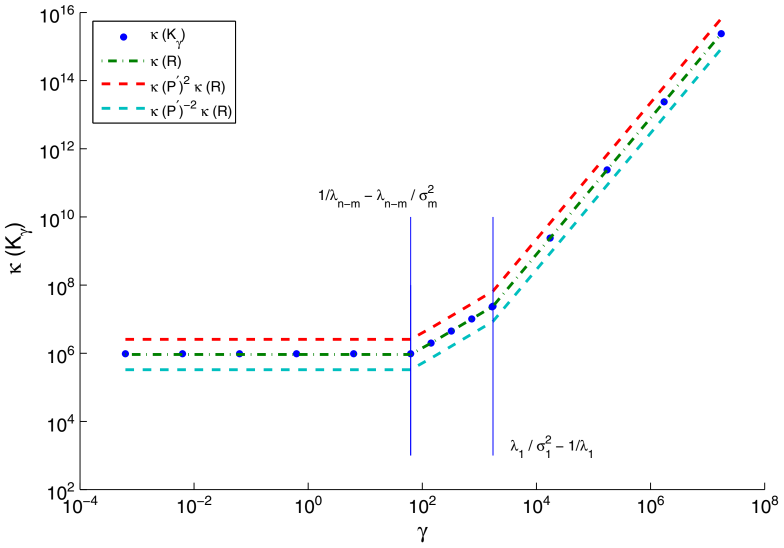

Figure 3. Condition number of augmented saddle-point matrix as a function of γ, for a problem arising from the discretized time-harmonic Maxwell equations.

interested in keeping $\gamma < \psi$. We note here that ψ might be zero or negative in certain instances. The practical meaning of such cases is that γ should be kept as small as possible for the conditioning of the saddle-point matrix to be minimized. The choice $\gamma = \alpha$ may work well in this case.

In both Figures 2 and 3, it can be seen that while Theorems 2.1 and 3.1 give true bounds on the condition numbers, $\kappa(\Sigma)$ and $\kappa(R)$ provide excellent estimates. Because $[\alpha, \beta] \cap [0, \psi] \neq \emptyset$, one would ideally choose γ in the range $[\alpha, \psi]$, which would result in minimizing the condition number of both $A + \gamma B^T B$ and the saddle-point matrix itself.

Example 4.2

Next, we consider a case where P is (relatively) ill conditioned and investigate its effect on our estimates. We note here that we cannot take P to have a condition number that is overly large,

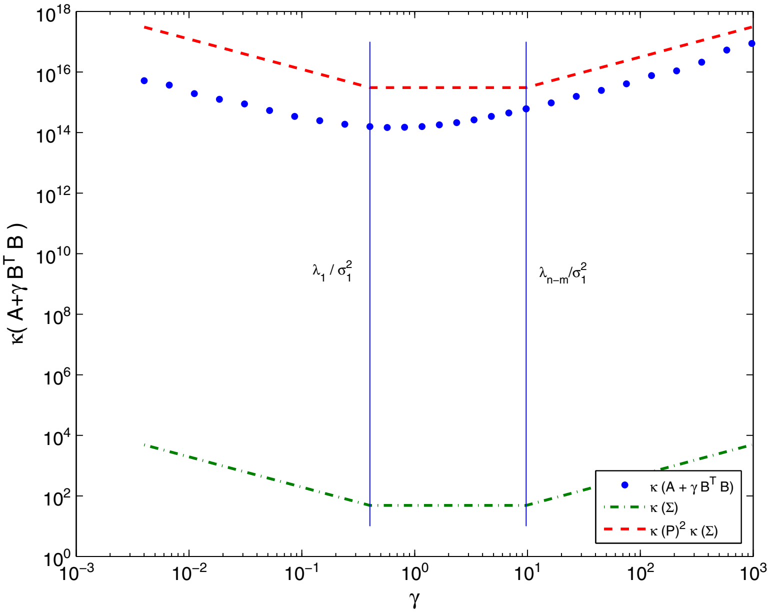

Figure 4. Condition number of ill-conditioned $A + \gamma B^T B$ for random A and B over varying γ.

because this quantity appears in squared form in the bounds for the condition numbers of $A(\gamma)$ and $K(\gamma)$ in Theorems 2.1 and 3.1, respectively. Therefore, to avoid considering numerically singular $A(\gamma)$ and $K(\gamma)$, we need to settle for P whose condition number is smaller than the square root of the roundoff unit.

We take a random 1000×1000 matrix for A and a random 300×1000 matrix for B, such that $\kappa(P) \simeq 8 \cdot 10^6$. We have that $\kappa\left(A + \gamma B^T B\right) \geqslant 10^{14}$. For these A and B, we have

$$\alpha = 0.4,$$

$$\beta = 9.7,$$

and we plot $\kappa(A + \gamma B^T B)$ in Figure 4. (Note that α and β are independent of the condition number of P.)

We can make some observations about the quality of the estimation for ill-conditioned P in Figure 4 as compared with a well-conditioned P in Figure 2. As expected, the large condition number of P results in the true condition number better estimated by the upper bound of $\kappa(\Sigma)\kappa(P)^2$ than just $\kappa(\Sigma)$. The asymptotics for small and large γ still grow like γ^{-1} and γ, respectively; however, the transition regions for the condition numbers are not as well defined as they are in Figure 2. With $\alpha \leqslant \gamma \leqslant \beta$, we expect $\kappa\left(A + \gamma B^T B\right)$ to be fairly constant, but instead, optimality occurs at the slight dip where $\gamma \approx \frac{\alpha+\beta}{2}$.

Next, we plot $\kappa(K(\gamma))$ in Figure 5. For this case, we have

$$\psi = 0,$$

$$\omega = 4.9.$$

These values are small and do not depend on γ or on $\kappa(P)$, and because $\psi = 0$, no constraint is imposed in relation to the eigenvalues of A and the singular values of B.

Similarly to the situation for $A(\gamma)$, we see that while $\kappa(R)$ no longer accurately captures $\kappa(K(\gamma))$, our bounds very well trap the condition number. Interestingly, the slope of the computed condition number no longer follows that of the bounds.

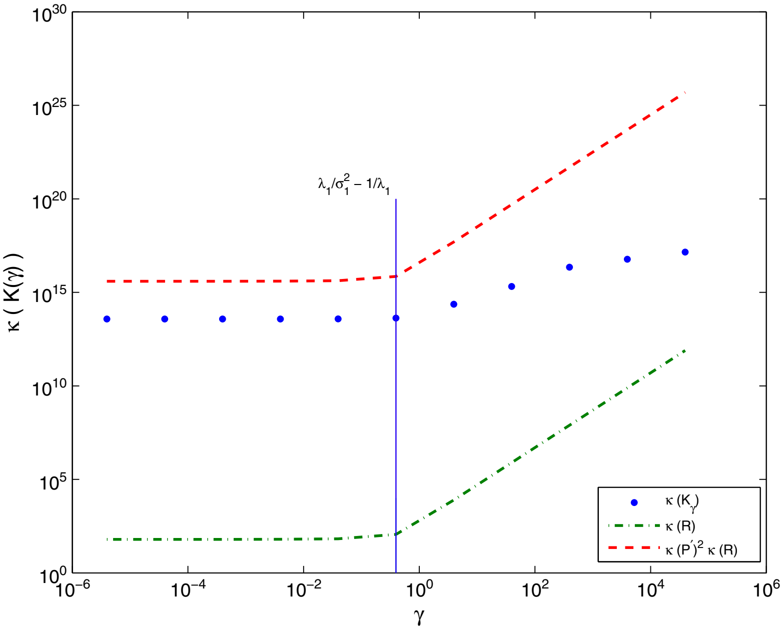

Figure 5. Condition number of ill-conditioned $\kappa(K(\gamma))$ for random A and B over varying γ.

5. CONCLUDING REMARKS

We have developed conditions for minimizing the condition numbers of $A(\gamma)$ and $K(\gamma)$, which depend on the extremal nonzero eigenvalues of A and singular values of B. We have also established conditions for a joint domain where both condition numbers are minimized.

Our approach applies to the case of maximal nullity of the leading block. While this is a restriction, it does represent a number of interesting applications [1], and in our experiments, the region of minimized condition numbers is predicted in a tight and precise fashion.

Our analysis straightforwardly extends to the nonsymmetric case, under the same rank assumptions on the matrices A and B. If A were to be nonsymmetric, then its eigendecomposition in (8) would be replaced by the SVD $A = USV^T$, and we would have $A(\gamma) = P\Sigma Q^T$, where $P = [U\ Z]$ and $Q = [V\ Z]$. Theorems 2.1 and 3.1 would then be adjusted accordingly in a seamless fashion.

Another potentially interesting issue to explore may be a situation where the rank of the leading block is slightly larger than $n - m$. Under that scenario, our analysis (and our reliance on quasi-direct sums) can no longer be carried out, but using eigenvalue interlacing arguments may still provide a way to find an effective approximation to the optimal γ. We leave this as an item for future investigation.

ACKNOWLEDGEMENTS

This manuscript has greatly benefited from a set of sharp and helpful comments made by two anonymous referees. The work of the second author was supported in part by NSERC Discovery grant 261539.

REFERENCES

1. Estrin R, Greif C. On nonsingular saddle-point systems with a maximally rank deficient leading block. *SIAM Journal on Matrix Analysis and Applications* 2015; **36**(2):367–384.
2. Golub GH, Greif C. On solving block-structured indefinite linear systems. *SIAM Journal on Scientific Computing* 2003; **24**(6):2076–2092.
3. Benzi M, Golub GH, Liesen J. Numerical solution of saddle point problems. *Acta Numerica* 2005; **14**:1–137.

4. Axelsson O, Neytcheva M. Eigenvalue estimates for preconditioned saddle point matrices. *Numerical Linear Algebra with Applications* 2006; **13**(4):339–360.
5. Elman HC, Silvester DJ, Wathen AJ. *Finite Elements and Fast Iterative Solvers: With Applications in Incompressible Fluid Dynamics* (Second), Numerical Mathematics and Scientific Computation. Oxford University Press: Oxford, 2014.
6. Olshanskii MA, Vassilevski YV. Pressure Schur complement preconditioners for the discrete Oseen problem. *SIAM Journal on Scientific Computing* 2007; **29**(6):2686–2704.
7. Saad Y, Suchomel B. ARMS: an algebraic recursive multilevel solver for general sparse linear systems. *Numerical Linear Algebra with Applications* 2002; **9**(5):359–378.
8. Greif C, Schötzau D. Preconditioners for the discretized time-harmonic Maxwell equations in mixed form. *Numerical Linear Algebra with Applications* 2007; **14**(4):281–297.
9. Fletcher R. *Practical Methods of Optimization* (Second), A Wiley-Interscience Publication. John Wiley & Sons, Ltd.: Chichester, 1987.
10. Hestenes MR. *Conjugate direction methods in optimization*, Applications of Mathematics, vol. 12. Springer-Verlag: New York-Berlin, 1980.
11. Powell MJD. A method for nonlinear constraints in minimization problems. In *Optimization (Symposium, University of Keele, Keele, 1968)*. Academic Press: London, 1969; 283–298.
12. Benzi M, Olshanskii MA. An augmented Lagrangian-based approach to the Oseen problem. *SIAM Journal on Scientific Computing* 2006; **28**(6):2095–2113.
13. Fortin M, Glowinski R. *Augmented Lagrangian Methods*, Studies in Mathematics and its Applications, vol. 15. North-Holland Publishing Co.: Amsterdam, 1983. Applications to the numerical solution of boundary value problems, Translated from the French by B. Hunt and D. C. Spicer.
14. Powell CE, Silvester D. Optimal preconditioning for Raviart–Thomas mixed formulation of second-order elliptic problems. *SIAM Journal on Matrix Analysis and Applications* 2003; **25**(3):718–738.
15. Hiptmair R. Finite elements in computational electromagnetism. *Acta Numerica* 2002; **11**:237–339.
16. Boyd S, Farikh N, Chu E, Peleato B, Eckstein J. Distributed optimization and statistical learning via the alternating direction methods of multipliers. *Foundations and Trends in Machine Learning* 2010; **3**(1):1–122.
17. Fiedler M. Remarks on the Schur complement. *Linear Algebra and its Applications* 1981; **39**:189–195.

DOI: 10.1002/nla