Applied Mathematics and Computation 217 (2011) 5596–5602

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

An efficient numerical method for preconditioned saddle point problems

Dongping Li [a], Jingyu Zhao [b], Guofeng Zhang [c,*]

[a] School of Mathematics, Changchun Normal University, Changchun 130032, PR China
[b] ZTE Corporation, Shenzhen 518120, PR China
[c] School of Mathematics and Statistics, Lanzhou University, Lanzhou 730000, PR China

ARTICLE INFO

Keywords:
Saddle point problems
Matrix splitting
Iteration method

ABSTRACT

In this paper, we consider the solution of linear systems of saddle point type by a preconditioned numerical method. We first transform the original linear system into two sub-systems with small size by a preconditioning strategy, then employ the conjugate gradient (CG) method to solve the linear system with a *SPD* coefficient matrix, and a splitting iteration method to solve the other sub-system, respectively. Numerical experiments show that the new method can achieve faster convergence than several effective preconditioners published in the recent literature in terms of total runtime and iteration steps.

1. Introduction

We consider the solution of large sparse saddle point problems with the structure

$$\mathcal{A}\mathbf{x} \equiv \begin{bmatrix} A & B \\ B^T & O \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix} \equiv \mathbf{b}, \tag{1.1}$$

where $A \in R^{m\times m}$ is symmetric positive definite, $B \in R^{m\times n}$ has full column rank, $m > n$, $f \in R^m$, and $g \in R^n$. These assumptions guarantee the existence and uniqueness of the solution of the linear system (1.1); see [1]. Systems of the form (1.1) arise in a variety of science and engineering applications, including constrained optimization, mixed finite element formulations of partial differential equations, circuit analysis, and so on; see [2,3]. Due to indefiniteness of the coefficient matrix and often poor spectral properties, such linear systems represent a significant challenge for solver developers. In recent years there has been a surge of interest in saddle point problems, and numerous solution techniques have been proposed for this type of systems; see [4–9] and the references therein. The aim of this paper is to present a new selection of solution methods for the saddle point problem (1.1), with an emphasis on iteration methods for large and sparse problems. This approach proposed avoids to compute the inverse of matrices by some modification.

The paper is organized as follows. In Section 2, the new iteration algorithm is presented, and some of its convergence properties are studied. In Section 3, the results of numerical experiments with our algorithm are reported. Finally, we offer some concluding remarks in Section 4.

Throughout the paper, we use $\lambda_{\min}$ and $\lambda_{\max}$ to denote the smallest and largest eigenvalues of the matrix A, and $\mu_{\min}$ the smallest singular value of the matrix B.

* Corresponding author.
E-mail address: gf_zhang@lzu.edu.cn (G.-F. Zhang).

0096-3003/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.amc.2010.12.037

2. Iteration algorithm and analysis of convergence

Let

$$\mathcal{P}_\alpha = \begin{bmatrix} I_m & -B(B^TB)^{-1} \\ O & I_n \end{bmatrix} \begin{bmatrix} I_m & O \\ B^T & -\alpha B^TB \end{bmatrix},$$

where α is a positive constant.

Preconditioning for the system (1.1) from the left with $\mathcal{P}_\alpha$ yields a system of the form

$$\mathcal{P}_\alpha \mathcal{A} \begin{bmatrix} y \\ z \end{bmatrix} = \mathcal{P}_\alpha \begin{bmatrix} f \\ g \end{bmatrix}. \tag{2.1}$$

An explicit calculation reveals that the system (2.1) can be rewritten as

$$\begin{bmatrix} A_3 & O \\ B_3 & B_2 \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} b_2 \\ q_2 \end{bmatrix}, \tag{2.2}$$

where $A_3 = A + \alpha BB^T - B(B^TB)^{-1}B^TA, B_2 = B^TB, B_3 = B^TA - \alpha B^TBB^T, b_2 = f - B(B^TB)^{-1}B^Tf + \alpha Bg, q_2 = B^Tf - \alpha B^TBg$. Obviously, A_3 is nonsingular.

We can get the solution of the system (2.2) by solving the coupled linear systems of the form

$$\begin{cases} A_3 y = b_2, \\ B_3 y + B_2 z = q_2, \end{cases} \tag{2.3}$$

which can be solved by first computing y from

$$A_3 y = b_2 \tag{2.4}$$

and then computing z from

$$B_2 z = q_2 - B_3 y. \tag{2.5}$$

The system (2.5) is *SPD*, and any solver for *SPD* systems can be applied. This could be a Cholesky factorization, or a preconditioned conjugate gradient (PCG) method, or some specialized solvers [9].

For (2.4), generally, the coefficient matrix A_3 is large and dense, so direct computations are very costly and impractical in actual implementations. To overcome this disadvantage, we can solve the system iteratively by splitting technology.

For matrix A_3, we make the following splitting

$$A_3 = M_\alpha - N,$$

where $M_\alpha = A + \alpha BB^T$ is a Hermitian positive matrix, and $N = B(B^TB)^{-1}B^TA$.

Then we consider the following parameterized iteration scheme for solving the system (2.4):

$$M_\alpha x_{k+1} = Nx_k + b_2. \tag{2.6}$$

Let $\mathcal{G}_\alpha = M_\alpha^{-1}N, c = M_\alpha^{-1}b_2$. The iteration sequence $\{x_k\}$ generated by (2.6) converges to the solution $x = A_3^{-1}b_2$ for arbitrary initial guesses x_0 and right-hand sides b_2 if and only if $\rho(\mathcal{G}_\alpha) < 1$, where $\rho(\mathcal{G}_\alpha)$ denotes the spectral radius of $\mathcal{G}_\alpha$.

Theorem 2.1. *Let $A \in R^{m\times m}$ be SPD, and $B \in R^{m\times n}$ be of full column rank. Then, the iteration (2.6) is convergent when $\alpha > \alpha^*$, where $\alpha^* = \frac{\lambda_{max}^2}{\mu_{min}\lambda_{min}}$. Furthermore, we have* $\lim_{\alpha\to\infty}\rho(\mathcal{G}_\alpha) = 0$.

Proof. Let $\lambda \in \mathbb{C}$ be an eigenvalue of $\mathcal{G}_\alpha$, x be a corresponding eigenvector with $\|x\|_2 = 1$, and δ be a positive constant. Evidently, $\lambda \neq 0$.

Note that $\mathcal{G}_\alpha x = \lambda x$, it follows that $\lambda M_\alpha x = Nx$, i.e.

$$\lambda(A + \alpha BB^T)x = B(B^TB)^{-1}B^TAx.$$

Then

$$\lambda x^*(A + \alpha BB^T)x = x^*B(B^TB)^{-1}B^TAx.$$

Therefore

$$|\lambda| = \frac{|x^*B(B^TB)^{-1}B^TAx|}{|x^*(A + \alpha BB^T)x|} = \frac{|x^*B(B^TB)^{-1}B^TAx|}{x^*Ax + \alpha x^*BB^Tx}. \tag{2.7}$$

If $B^Tx = \mathbf{0}$, then $\lambda = 0$, a contradiction. Hence, it must hold $B^Tx \neq \mathbf{0}$.

By (2.7), we can easily conclude that $|\lambda| \to 0$ as $\alpha \to \infty$, which implies $\rho(\mathcal{G}_\alpha) \to 0$.