International Journal of Engineering, Mathematical and Physical Sciences
ISSN: 2517-9934
Vol:4, No:8, 2010

A modified inexact Uzawa Algorithm for generalized saddle point problems

Shu-Xin Miao

Abstract—In this note, we discuss the convergence behavior of a modified inexact Uzawa algorithm for solving generalized saddle point problems, which is an extension of the result obtained in a recent paper [Z.H. Cao, Fast Uzawa algorithm for generalized saddle point problems, Appl. Numer. Math., 46 (2003) 157-171].

Keywords—Saddle point problem; Inexact Uzawa algorithm; Convergence behavior.

I. Introduction

In this note, we consider the generalized saddle point problems of the form

$$\begin{pmatrix} A & B^T \\ B & -C \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} p \\ q \end{pmatrix}, \tag{1}$$

where $A \in R^{n\times n}$ is symmetric positive definite, $B \in R^{m\times n}$ is of full row rank, and $C \in R^{m\times m}$ is symmetric positive semidefinite, $p \in R^n$ and $q \in R^m$ are given vectors, with $m \leq n$.

The generalized saddle point problems (1) arises in a wide variety of scientific and engineering applications, see [2] and references therein. Frequently, the matrices A and B are large and sparse. So iterative methods become more attractive than direct methods for solving the generalized saddle point problems (1). Many efficient iterative methods have been studied in the literature [1], [2], [8], [9], [12], [13], [15]. For example, Miao and Wang [12] studied a class of stationary iterative methods for (1) based on the work of Yun and Kim [14].

Uzawa-type algorithms are of interest because they are simple, efficient and have minimal computer memory requirements. Therefore, Uzawa-type algorithms are widely used in engineering community, especially, are used for solving saddle point problems [1], [3], [4], [5], [6], [7], [10], [11], [15].

Recently, Cao [5] consider the inexact Uzawa algorithm for solving generalized saddle point problems (1), which is an extension of the results obtained in [3]. In this note, a slight modification of the inexact Uzawa algorithm for solving generalized saddle point problems (1) (see [5]) is discussed, a bound of convergence rate is obtained.

Shu-Xin Miao is with the Department of Mathematics, Northwest Normal Univerisity, Lanzhou, 730070, P.R. China e-mail: shuxinmiao@gmail.com.
Manuscript received ****; revised ****.

II. Modified inexact Uzawa algorithm

Let Q_A and Q_B be symmetric, positive definite $n \times n$ and $m \times m$ matrix, respectively, satisfying

$$\begin{aligned} & (1-\delta)(Q_A u,\ u) \leq (Au,\ u) \\ < \quad & (Q_A u,\ u), \quad \forall\, u \in R^n, \\ & (1-\gamma)(Q_B v,\ v) \leq ((BA^{-1}B^T + C)v,\ v) \\ \leq \quad & (Q_B v,\ v), \quad \forall\, v \in R^m \end{aligned} \tag{2}$$

for some $0 < \delta < 1$ and $0 \leq \gamma < 1$. Here $(\cdot,\ \cdot)$ is the Euclidean inner product in R^n or R^m. Then the inexact Uzawa algorithm for solving (1) as follows:

Algorithm 1. (Inexact Uzawa Algorithm) *For* $x_0 \in R^n$ and $y_0 \in R^m$, *given, the iterative sequence* $\{(x_i,\ y_i)\}$ *is defined, for* $i = 1, 2, \cdots$, *by*

$$\begin{cases} x_{i+1} = x_i + Q_A^{-1}(p - (Ax_i + B^T y_i)), \\ y_{i+1} = y_i + Q_B^{-1}(Bx_{i+1} - Cy_i - q). \end{cases}$$

From (2), we can see that $Q_A - A$ and Q_B are symmetric and positive definite, therefore we can define an inner product in $R^n \times R^m$ by (cf. [3], [5])

$$\begin{aligned} & \left[\begin{pmatrix} u \\ v \end{pmatrix}, \begin{pmatrix} r \\ s \end{pmatrix}\right] \\ = \quad & \left(\begin{pmatrix} Q_A - A & \\ & Q_B \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix}, \begin{pmatrix} r \\ s \end{pmatrix}\right) \\ \equiv \quad & ((Q_A - A)u, r) + (Q_B v, s). \end{aligned} \tag{3}$$

The corresponding norm is denoted by

$$[|q|] = [q, q]^{1/2}, \quad \forall\, q \in R^n \times R^m. \tag{4}$$

For the inexact Uzawa algorithm 1, Cao [5] provide the following convergence theorem:

Theorem 2. Assume that (2) hold. Let x, y be the solution pair for (1), x_i, y_i be defined by the inexact Uzawa algorithm 1, and set

$$e_i = \begin{pmatrix} e_i^x \\ e_i^y \end{pmatrix} = \begin{pmatrix} x - x_i \\ y - y_i \end{pmatrix}.$$

Then for $i = 1, 2, \cdots$

$$[|e_i|] \leq \rho^i [|e_0|],$$

where

$$\rho = \frac{\gamma(1-\delta) + \sqrt{\gamma^2(1-\delta^2) + 4\delta}}{2}.$$

In this note, we discuss the following slight modification of the inexact Uzawa algorithm 1 for solving generalized saddle point problem (1).

Algorithm 3. (Modified Inexact Uzawa Algorithm) For $x_0 \in R^n$ and $y_0 \in R^m$, *given, the iterative sequence* $\{(x_i,\ y_i)\}$ *is defined, for* $i = 1, 2, \cdots$, *by*

$$\begin{cases} x_{i+1} = x_i + Q_A^{-1}(p - (Ax_i + B^T y_i)), \\ y_{i+1} = y_i + \omega Q_B^{-1}(Bx_{i+1} - Cy_i - q), \end{cases}$$

where $\omega \in (0, 1]$ *is a real parameter.*

Remark 4. Algorithm 3 is an extension of Algorithm 1. It is also an extension of the inexact Uzawa algorithm considered in [6].

III. Convergence analysis

In what follows, we consider the convergence of the modified inexact Uzawa algorithm 2. Similar to (3), we can define inner product as

$$\left\lfloor \begin{pmatrix} u \\ v \end{pmatrix}, \begin{pmatrix} r \\ s \end{pmatrix} \right\rfloor = ((Q_A - A)u, r) + \omega^{-1}(Q_B v, s).$$

Therefore the corresponding norm for e_i can be defined as

$$|||e_i||| = [((Q_A - A)e_i^x, e_i^x) + \omega^{-1}(Q_B e_i^y, e_i^y)]^{1/2}. \quad (5)$$

We have the following convergence result for Algorithm 2.

Theorem 5. Assume that (2) hold. Then for $i = 1, 2, \cdots$

$$|||e_i||| \le \rho_\omega^i |||e_0|||,$$

where $\rho_\omega = \max\{r_1(\omega), r_2(\omega)\}$ and

$$\begin{aligned} r_1(\omega) &= \frac{1}{2}[(1-\delta)(1-\omega(1-\gamma)) \\ &\quad + \sqrt{(1-\delta)^2(1-\omega(1-\gamma))^2 + 4\delta}\ ], \\ r_2(\omega) &= \sqrt{\delta}. \end{aligned}$$

Proof. Denote $S_a = BA^{-1}B^T + C$ and $Q_B(\omega) = \omega^{-1}Q_B$, it is easy to see that the iterative error equation can be expressed as (cf. [5])

$$\begin{pmatrix} -e_{i+1}^x \\ e_{i+1}^y \end{pmatrix} = M_1 \begin{pmatrix} e_i^x \\ e_i^y \end{pmatrix}, \quad (6)$$

where

$$M_1 = \begin{pmatrix} -(I - Q_A^{-1}A) & Q_A^{-1}B^T \\ Q_B(\omega)^{-1}B(I - Q_A A) & I - Q_B(\omega)^{-1}S_a \end{pmatrix}.$$

From [5] and (6), we know that

$$|||e_{i+1}||| \le \sigma(M_1)|||e_i|||,$$

where $\sigma(M_1)$ is the spectrum of matrix M_1. Since M_1 is symmetric with respect to the $\lfloor\cdot,\cdot\rfloor$ inner product, its eigenvalues are real. We shall bound the positive and negative eigenvalue of M_1 in what follows.

We first provide a bound for the positive eigenvalues of M_1. Let

$$N = \begin{pmatrix} -\delta I & \delta^{1/2}L \\ \delta^{1/2}L^* & I - L^*L - Q_B(\omega)^{-1}C \end{pmatrix},$$

where $L = (I - Q_A^{-1}A)^{-1/2}Q_A^{-1}B^T$ and $L^* = Q_B(\omega)^{-1}B(I - Q_A^{-1}A)^{1/2}$. Then the largest eigenvalue of M_1 is bounded by the largest eigenvalue of N (see [5]). Let λ be a positive eigenvalue of N with corresponding eigenvector $\{\psi_1,\ \psi_2\}$, i.e.,

$$\begin{aligned} &-\delta\psi_1 + \delta^{1/2}L\psi_2 = \lambda\psi_1, \\ &\delta^{1/2}L^*\psi_1 + (I - L^*L - Q_B(\omega)^{-1}C)\psi_2 = \lambda\psi_2. \end{aligned} \quad (7)$$

Eliminating ψ_1 gives

$$-\lambda L^*L\psi_2 = (\lambda+\delta)Q_B(\omega)^{-1}C\psi_2 + (\lambda+\delta)(\lambda-1)\psi_2.$$

From the first equation of (7), we can see that $\psi_2 \ne 0$, and hence

$$\begin{aligned} &-\lambda(Q_B(\omega)L^*L\psi_2, \psi_2) \quad (8) \\ = \ &(\lambda+\delta)(C\psi_2, \psi_2) + (\lambda+\delta)(\lambda-1)(Q_B(\omega)\psi_2, \psi_2). \end{aligned}$$

By the first equation of (2) and the definition of L and L^* it follows that

$$\begin{aligned} (Q_B(\omega)L^*L\psi_2, \psi_2) &= (Q_A^{-1}B^T\psi_2, B^T\psi_2) \\ &\ge (1-\delta)(BA^{-1}B^T\psi_2, \psi_2). \end{aligned}$$

Now (8) imply

$$\begin{aligned} 0 &= \lambda(Q_B(\omega)L^*L\psi_2, \psi_2) + (\lambda+\delta)(C\psi_2, \psi_2) \\ &\quad + (\lambda+\delta)(\lambda-1)(Q_B(\omega)\psi_2, \psi_2) \\ &\ge \lambda(1-\delta)(BA^{-1}B^T\psi_2, \psi_2) + (\lambda+\delta)(C\psi_2, \psi_2) \\ &\quad + (\lambda+\delta)(\lambda-1)(Q_B(\omega)\psi_2, \psi_2) \\ &= \lambda(1-\delta)(S_a\psi_2, \psi_2) + \delta(1+\lambda)(C\psi_2, \psi_2) \\ &\quad + (\lambda+\delta)(\lambda-1)(Q_B(\omega)\psi_2, \psi_2) \\ &\ge \lambda(1-\delta)(1-\gamma)(Q_B\psi_2, \psi_2) \\ &\quad + (\lambda+\delta)(\lambda-1)(Q_B(\omega)\psi_2, \psi_2) \\ &= [\lambda(1-\delta)(1-\gamma) + \omega(\lambda+\delta)(\lambda-1)](Q_B\psi_2, \psi_2). \end{aligned}$$

Since Q_B is symmetric positive definite and $\psi_2 \ne 0$, we get

$$\lambda(1-\delta)(1-\gamma) \le -\omega(\lambda+\delta)(\lambda-1),$$

From which we have

$$\lambda \le r_1(\omega),$$

where

$$\begin{aligned} r_1(\omega) &= \frac{1}{2}[(1-\delta)(1-\omega(1-\gamma)) \\ &\quad + \sqrt{(1-\delta)^2(1-\omega(1-\gamma))^2 + 4\delta}\ ]. \end{aligned}$$

Next we estimate the negative eigenvalue of M_1, let $\lambda < 0$ be an eigenvalue of M_1 with corresponding eigenvector $\{\phi_1, \phi_2\}$, i.e.,

$$\begin{cases} -(I - Q_A^{-1}A)\phi_1 + Q_A^{-1}B^T\phi_2 = \lambda\phi_1, \\ Q_B(\omega)^{-1}B(I - Q_A^{-1}A)\phi_1 \\ +(I - Q_B(\omega)^{-1}BQ_A^{-1}B^T + C)\phi_2 = \lambda\phi_2. \end{cases} \quad (9)$$

From (9), we can see that $\phi_1 \ne 0$ (cf. [5]). Thus, any eigenvector of M_1 corresponding to a negative eigenvalue must have a nonzero component ϕ_1.

From (9) we have

$$((1-\lambda)I - Q_B(\omega)^{-1}C)\phi_2 = \lambda Q_B(\omega)^{-1}B\phi_1.$$

By (2) and noting $\lambda < 0$, it follows that $(1-\lambda)I - Q_B(\omega)^{-1}C$ is invertible. Thus, we get

$$\phi_2 = \lambda((1-\lambda)I - Q_B(\omega)^{-1}C)^{-1}Q_B(\omega)^{-1}B\phi_1. \qquad (10)$$

Substituting (10) into the first equation in (9) and taking an inner product with $Q_A\phi_1$ gives

$$\begin{aligned} & (1+\lambda)(Q_A\phi_1, \phi_1) && (11) \\ = \; & (A\phi_1, \phi_1) \\ & +\lambda(((1-\lambda)I - Q_B(\omega)^{-1}C)^{-1}Q_B(\omega)^{-1}B\phi_1, B\phi_1). \end{aligned}$$

For $\phi_1 \in R^n$, we have

$$\begin{aligned} & (((1-\lambda)I - Q_B(\omega)^{-1}C)^{-1}Q_B(\omega)^{-1}B\phi_1, B\phi_1) \\ = \; & (((1-\lambda)Q_B(\omega) - C)^{-1}B\phi_1, B\phi_1) \\ = \; & \sup_{v\in R^m} \frac{(((1-\lambda)Q_B(\omega) - C)^{-1}B\phi_1, v)^2}{(((1-\lambda)Q_B(\omega) - C)^{-1}v, v)} \\ = \; & \sup_{v\in R^m} \frac{(\phi_1, B^T v)^2}{(((1-\lambda)Q_B(\omega) - C)v, v)} \\ \le \; & \frac{1}{1-\lambda} \sup_{v\in R^m} \frac{(\phi_1, B^T v)^2}{((Q_B(\omega) - C)v, v)} \\ \le \; & \frac{1}{1-\lambda} \sup_{v\in R^m} \frac{(A\phi_1, \phi_1)(BA^{-1}B^T v, v)}{((Q_B(\omega) - C)v, v)}. && (12) \end{aligned}$$

As $\omega \in (0, 1]$, the following inequality hold

$$\begin{aligned} & \sup_{v\in R^m} \frac{(A\phi_1, \phi_1)(BA^{-1}B^T v, v)}{((Q_B(\omega) - C)v, v)} \\ \le \; & \omega \sup_{v\in R^m} \frac{(A\phi_1, \phi_1)(BA^{-1}B^T v, v)}{((Q_B - C)v, v)}. \end{aligned}$$

Then from (12) and (2), we have

$$\begin{aligned} & (((1-\lambda)I - Q_B(\omega)^{-1}C)^{-1}Q_B(\omega)^{-1}B\phi_1, B\phi_1) \\ \le \; & \frac{\omega}{1-\lambda} \sup_{v\in R^m} \frac{(A\phi_1, \phi_1)(BA^{-1}B^T v, v)}{((Q_B - C)v, v)} \\ \le \; & \frac{\omega}{1-\lambda}(A\phi_1, \phi_1). \end{aligned}$$

Now (11) becomes

$$(1+\lambda)(Q_A\phi_1, \phi_1) \ge (A\phi_1, \phi_1) + \frac{\lambda\omega}{1-\lambda}(A\phi_1, \phi_1).$$

Note that $\omega \le 1$ and $\lambda < 0$, we therefore have

$$[\lambda^2 - \delta](Q_A\phi_1, \phi_1) \le 0.$$

Q_A is symmetric positive definite, therefore we obtain the bound for the negative eigenvalue of M_1 as

$$-\sqrt{\delta} \le \lambda < 0.$$

We complete the proof.

Remark 6. In particular, if $\omega = 1$, then Algorithm 3 becomes Algorithm 1. Therefore, we can obtain the convergence result of Algorithm 1 (Theorem 2) directly from Theorem 5.

Remark 7. We remark that $r_2(\omega) < 1$ as $0 < \delta < 1$. It is elementary to see that $r_1(\omega) < 1 - \frac{1}{2}\omega(1-\delta)(1-\gamma)$. Therefore $\rho_\omega < 1$, that is to say that the modified inexact Uzawa method (Algorithm 3) converges if (2) hold.

ACKNOWLEDGEMENTS

This work was supported by NWNU-KJCXGC-3-47.

REFERENCES

[1] Z.-Z. Bai, Z.-Q. Wang, On parameterized inexact Uzawa methods for generalized saddle point problems, *Linear Algebra Appl.*, 428 (2008) 2900-2932.
[2] M. Benzi, G.H. Golub, J. Liesen, Numerical solution of saddle point problems, *Acta Numer.*, 14 (2005) 1-137.
[3] J.H. Bramble, J.E. Pasciak, A.T. Vassilev, Analysis of the inexact Uzawa algorithm for saddle point problems, *SIAM J. Numer. Anal.*, 34 (1997) 1072-1092.
[4] Y. Cao, Y. Lin, Y. Wei, Nolinear Uzawa methods for solving nonsymmetric saddle point problems, *J. Appl. Math. Comput.*, 21 (2006) 1-21.
[5] Z.-H. Cao, Fast Uzawa algorithm for generalized saddle point problems, *Appl. Numer. Math.*, 46 (2003) 157-171.
[6] X.-L. Cheng, The inexact Uzawa algorithm for saddle point problem, *Appl. Math. Letters*, 13 (2000) 1-3.
[7] H.C. Elman, G.H. Golub, Inexact and preconditioned Uzawa algorithms for saddle point problems, *SIAM J. Numer. Anal.*, 31(1994) 1645-1661.
[8] B. Fischer, R. Ramage, D.J. Silvester, A.J. Wathen, Minimum residual methods for augmented systems, *BIT Numer. Math.*, 38 (1998) 527-543.
[9] G.H. Golub, X. Wu, J.-Y. Yuan, SOR-like methods for augmented systems, *BIT Numer. Math.*, 41 (2001) 71-85.
[10] Y. Lin, Y. Wei, A convergence analysis of the nonlinear Uzawa algorithm for saddle point problems, *Appl. Math. Letters*, 20 (2007) 1094-1098.
[11] Y. Lin, Y. Wei, Fast corrected Uzawa methods for solving symmetric saddle point problems, *Calcolo*, 43 (2006) 65-82.
[12] S.-X. Miao, K. Wang, On generalized stationary iterative method for solving the saddle point problems, *J. Appl. Math. Comput.*, DOI: 10.1007/s12190-009-0369-8.
[13] X. Wu, B.P.B. Silva, J.-Y. Yuan, Conjugate gradient method for rank deficient saddle point problems, *Numer. Algorithms*, 35 (2004) 139-154.
[14] J.H. Yun, S.W. Kim, Generalized stationary iterative method for solving linear systems, *Korean J. Comput. Appl. Math.*, 5 (1998) 341-349.
[15] B. Zheng, Z.-Z. Bai, X. Yang, On Semi-Convergence of Parameterized Uzawa Methods for Singular Saddle Point Problems, *Linear Algebra Appl.*, 431 (2009) 808-817.