On generalized SSOR-like iteration method for saddle point problems

CHUNPING PAN
Department of Humanities and Social Sciences
Zhejiang Industry Polytechnic College
Shaoxing, Zhejiang 312000
P. R. CHINA
panchunping@126.com

Abstract: In this paper, we study the iterative algorithms for saddle point problems(SPP). We present a new symmetric successive over-relaxation method with three parameters, which is the extension of the SSOR iteration method. Under some suitable conditions, we give the convergence results. Numerical examples further confirm the correctness of the theory and the effectiveness of the method.

Key–Words: iterative method, saddle point problems, SOR-like, SSOR-like; symmetric and positive definite matrix

1 Introduction

We consider the iterative solutions of large sparse saddle point problems of the form

$$\begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} p \\ q \end{pmatrix} \tag{1}$$

where $A \in R^{m\times m}$ is a symmetric positive definite matrix, $B \in R^{m\times n}$ is a matrix of full column rank, $p \in R^m$ and $q \in R^n$ are two given vectors, here $m \geq n$. Denote by B^T the transpose of the matrix B. These assumptions guarantee the existence and uniqueness of the solution of the linear system.

This system arises as the first-order optimality conditions for the following equality-constrained quadratic programming problem:

$$\min J(x) = \frac{1}{2}x^T Ax - p^T x \tag{2}$$

$$\text{subject to} \quad Bx = q \tag{3}$$

In this case the variable y represents the vector of Lagrange multipliers. Any solution (x_*, y_*) of (1) is a saddle point for the Lagrangian

$$L(x,y) = \frac{1}{2}x^T Ax - p^T x + (Bx-q)^T y \tag{4}$$

hence the name "saddle point problem" given to (1). Recall that a saddle point is a point $(x_*, y_*) \in R^{n+m}$ that satisfies

$$L(x_*, y) \leq L(x_*, y_*) \leq L(x, y_*) \tag{5}$$

for any $x \in R^n$ and $y \in R^m$, or equivalently,

$$\min_x \max_y L(x,y) = L(x_*, y_*) = \max_y \min_x L(x,y) \tag{6}$$

Systems of the form (1) also arise in nonlinearly constrained optimization (sequential quadratic programming and interior point methods), in fluid dynamics ('Stokes' problem), incompressible elasticity, circuit analysis, structural analysis, and so forth[1].

Since the above problem is large and sparse, iterative methods for solving equation (1) are effective because of storage requirements and preservation of sparsity. The well-known SOR method, which is a simple iterative method that is popular in engineering applications, cannot be applied directly to system (1) because of the singularity of the block diagonal part of the coefficient matrix. Recently, more attention has been paid to a class of iterative methods namely splitting methods. The best known and the oldest methods is Uzawa and the preconditioned Uzawa algorithms[2]. For solving augmented systems (1) , Golub et al have presented several SOR-like algorithms and have considered the optimum choice for the iterative parameter by using some nonsingular preconditioning matrix Q instead of the null block in the coefficient matrix (1) [3]. Bai et al without extra cost per iteration step, developed SOR-like method and presented the GSOR method, parameterized Uzawa and the inexact parameterized Uzawa methods for solving singular and nonsingular saddle point problems[4, 5, 6, 12]. Additionally, these authors proved the convergence and semiconvergence of these methods under suitable conditions. Furthermore, Bai, Golub and Ng discussed the convergence and the preconditioning property of the Hermitian and skew-Hermitian splitting (HSS) iteration method when it is used to solve the saddle point problem[8]. Then, Bai et al. establish a class of preconditioned Hermitian/skew-Hermitian splitting (PHSS) iteration method for saddle point problems[9], and Pan et al. further proposed its two-parameter and four-parameter acceleration, called the generalized preconditioned Hermitian/skew-Hermitian splitting (GPHSS) iteration method, and studied the convergence of this iterative scheme[10, 16]. Both theory and experiments have shown that these methods are very robust and efficient for solving the saddle-point problems when they are used as either solvers or preconditioners (for the Krylov subspace iteration methods). Li et al.[17], have considered the Chebyshev acceleration of the SOR-like method by a proper choice of the auxiliary matrix Q. On the other hand, Martins et al.[18], presented a variant of the accelerated overrelaxation iterative method. Others[7, 13, 14, 15], have proposed SSOR-like models for this purpose.

Considering all above approaches, In this paper, We present a new symmetric successive overrelaxation method with three parameters. Under some suitable conditions, we give the convergence results. Numerical results show that the new methods are very effective. The rest of the paper is organized as follows. In Section 2, the outline of our SSOR-like method to solve (1) is provided. In Section 3, we obtain the convergence region for this method. In Section 4, some numerical computations are presented. Finally, conclusions are made for this paper.

2 The Three-parameter SSOR-like iteration methods

In this section, we review the SOR-like iteration method for solving the saddle-point problems presented by G. H. Golub et al[3].

To establish the convergence properties of iterative method for the saddle-point problems, It need to begin by writing the saddle-point problem (1) in nonsymmetric form:

$$\mathcal{A}Z = b \tag{7}$$

where

$$\mathcal{A} = \begin{pmatrix} A & B \\ -B^T & 0 \end{pmatrix}, Z = \begin{pmatrix} x \\ y \end{pmatrix}, b = \begin{pmatrix} p \\ -q \end{pmatrix} \tag{8}$$

For the coefficient matrix of the augmented linear system (7), we make the following splitting

$$\mathcal{A} = D - L - U, \tag{9}$$

where

$$D = \begin{pmatrix} A & 0 \\ 0 & Q \end{pmatrix},\ L = \begin{pmatrix} 0 & 0 \\ B^T & 0 \end{pmatrix},$$
$$U = \begin{pmatrix} 0 & -B \\ 0 & Q \end{pmatrix} \tag{10}$$

and $Q \in R^{n\times n}$ is a prescribed nonsingular and symmetric matrix. Let ω be a nonzero real. Golub et al considered the following generalized SOR iteration scheme for the augmented linear system (7):

$$\begin{pmatrix} x^{(k+1)} \\ y^{(k+1)} \end{pmatrix} = (D - \omega L)^{-1}[(1-\omega)D + \omega U] \begin{pmatrix} x^{(k)} \\ y^{(k)} \end{pmatrix} + \omega(D - \omega L)^{-1} \begin{pmatrix} p \\ -q \end{pmatrix}$$

More precisely, we have the following algorithmic description of the SOR-like iteration method.

Algorithm 1. *(The SOR-like iteration method)*

Let $Q \in R^{n\times n}$ be a nonsingular and symmetric matrix. Given initial vectors

$$x^{(0)} \in R^m, \ ; \ y^{(0)} \in R^n$$

and a relaxation factors $\omega > 0$. For $k = 0, 1, 2, \ldots$ until the iteration sequence $\{(x^{(k)^T}, y^{(k)^T})^T\}$ is convergent, compute

$$\begin{cases} x^{(k+1)} & = (1-\omega)x^{(k)} + \omega A^{-1}(p - By^{(k)}), \\ y^{(k+1)} & = y^{(k)} + \omega Q^{-1}(B^T x^{(k+1)} - q). \end{cases} \tag{11}$$

here Q is an approximate (preconditioning) matrix of the Schur complement matrix $B^T A^{-1} B$.

Let ω and τ be two nonzero reals, $I_m \in R^{m\times m}$ and $I_n \in R^{n\times n}$ be the m-by-m and n-by-n identity matrices, respectively, and

$$\Omega = \begin{pmatrix} \omega I_m & 0 \\ 0 & \tau I_n \end{pmatrix}$$

Bai et al [4] considered the following generalized SOR iteration scheme for the augmented linear system (7):

$$\begin{pmatrix} x^{(k+1)} \\ y^{(k+1)} \end{pmatrix} = (D - \Omega L)^{-1}[(I-\Omega)D + \Omega U] \begin{pmatrix} x^{(k)} \\ y^{(k)} \end{pmatrix} + (D - \Omega L)^{-1}\Omega \begin{pmatrix} p \\ -q \end{pmatrix}.$$

More precisely, we have the following algorithmic description of the GSOR iteration method.

Algorithm 2. *(The GSOR iteration method)*

Let $Q \in R^{n\times n}$ be a nonsingular and symmetric matrix. Given initial vectors

$$x^{(0)} \in R^m, \ \ y^{(0)} \in R^n$$

and two relaxation factors $\omega, \tau \neq 0$. For $k = 0, 1, 2, \dots$ until the iteration sequence $\{(x^{(k)^T}, y^{(k)^T})^T\}$ is convergent, compute

$$\begin{cases} x^{(k+1)} & = (1-\omega)x^{(k)} + \omega A^{-1}(p - By^{(k)}), \\ y^{(k+1)} & = y^{(k)} + \tau Q^{-1}(B^T x^{(k+1)} - q). \end{cases} \tag{12}$$

here Q is an approximate (preconditioning) matrix of the Schur complement matrix $B^T A^{-1} B$.

Most recently, a growing interest has been noticed in solving augmented systems with symmetric SOR method (SSOR). Some have considered the SSOR iteration scheme for the saddle point problems (7)(see [7]):

$$\begin{cases} \begin{pmatrix} x^{(k+\frac{1}{2})} \\ y^{(k+\frac{1}{2})} \end{pmatrix} = (D-\omega L)^{-1}[(1-\omega)D+\omega U] \begin{pmatrix} x^{(k)} \\ y^{(k)} \end{pmatrix} \\ \qquad + \omega(D-\omega L)^{-1} \begin{pmatrix} p \\ -q \end{pmatrix} \\ \begin{pmatrix} x^{(k+1)} \\ y^{(k+1)} \end{pmatrix} = (D-\omega U)^{-1}[(1-\omega)D+\omega L] \begin{pmatrix} x^{(k+\frac{1}{2})} \\ y^{(k+\frac{1}{2})} \end{pmatrix} \\ \qquad + \omega(D-\omega U)^{-1} \begin{pmatrix} p \\ -q \end{pmatrix} \end{cases}$$

Whereas others [14, 15] applied the nonsingular symmetric matrix Q with the following splitting:

$$\begin{pmatrix} A & B \\ -B^T & 0 \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & Q \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ B^T & \frac{Q}{2} \end{pmatrix} - \begin{pmatrix} 0 & -B \\ 0 & \frac{Q}{2} \end{pmatrix} \tag{13}$$

The authors of [13], considered the following splitting:

$$\begin{pmatrix} A & B \\ -B^T & 0 \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & Q \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ B^T & \alpha Q \end{pmatrix} - \begin{pmatrix} 0 & -B \\ 0 & \beta Q \end{pmatrix} \tag{14}$$

where Q being nonsingular symmetric, $\alpha, \beta \in R, \alpha + \beta = 1$, and applied the symmetric SOR method to solve the the saddle point problems (7).

Here we make the following splitting

$$\mathcal{A} \equiv \begin{pmatrix} A & B \\ -B^T & 0 \end{pmatrix} = D - L' - U' \tag{15}$$

where

$$D = \begin{pmatrix} A & 0 \\ 0 & Q \end{pmatrix}, L' = \begin{pmatrix} 0 & 0 \\ B^T & \alpha Q \end{pmatrix}$$

and

$$U' = \begin{pmatrix} 0 & -B \\ 0 & (1-\alpha)Q \end{pmatrix}$$

We consider the the following SSOR iteration scheme for the saddle point problems (7):

$$\begin{cases} \begin{pmatrix} x^{(k+\frac{1}{2})} \\ y^{(k+\frac{1}{2})} \end{pmatrix} = (D-\Omega L')^{-1}[(I-\Omega)D+\Omega U'] \begin{pmatrix} x^{(k)} \\ y^{(k)} \end{pmatrix} \\ \qquad + (D-\Omega L')^{-1}\Omega \begin{pmatrix} p \\ -q \end{pmatrix} \\ \begin{pmatrix} x^{(k+1)} \\ y^{(k+1)} \end{pmatrix} = (D-\Omega U')^{-1}[(I-\Omega)D+\Omega L'] \begin{pmatrix} x^{(k+\frac{1}{2})} \\ y^{(k+\frac{1}{2})} \end{pmatrix} \\ \qquad + (D-\Omega U')^{-1}\Omega \begin{pmatrix} p \\ -q \end{pmatrix} \end{cases}$$

More precisely, we have the following algorithmic description of this three-parameter symmetric SOR method method(3-SSOR-like).

Algorithm 3. *(The 3-SSOR-like iteration method)*

Let $Q \in R^{n\times n}$ be a nonsingular and symmetric matrix. Given initial vectors

$$x^{(0)} \in R^m, \ \ y^{(0)} \in R^n$$

and three relaxation factors $\omega, \tau > 0$, $0 \leq \alpha \leq 1$. For $k = 0, 1, 2, \dots$ until the iteration sequence $\{(x^{(k)^T}, y^{(k)^T})^T\}$ is convergent, compute

$$\begin{cases} x^{(k+\frac{1}{2})} = (1-\omega)x^{(k)} + \omega A^{-1}(p - By^{(k)}) \\ y^{(k+\frac{1}{2})} = y^{(k)} + \frac{\tau}{1-\alpha\tau} Q^{-1}(B^T x^{(k+\frac{1}{2})} - q) \\ y^{(k+1)} = y^{(k+\frac{1}{2})} + \frac{\tau}{1-\tau+\alpha\tau} Q^{-1}(B^T x^{(k+\frac{1}{2})} - q) \\ x^{(k+1)} = (1-\omega)x^{(k+\frac{1}{2})} + \omega A^{-1}(p - By^{(k+1)}) \end{cases} \tag{16}$$

here Q is an approximate (preconditioning) matrix of the Schur complement matrix $B^T A^{-1} B$.

Obviously, when $\alpha = 0$, Algorithm 3 reduces to `GSSOR-like` method in [7]; when $\alpha = \frac{1}{2}$, Algorithm 3 reduces to `GMSSOR` method in[15]; when

$\alpha = 0, \omega = \tau$, it becomes the SSOR method studied in Darvishi and Hessari [13]; when $\alpha = \frac{1}{2}, \omega = \tau$, it is the same as the method studied by Wu et al in[14].

By selecting different matrix Q, we can get some useful 3-SSOR-like iterative algorithm. Such as $Q = \theta I, (\theta \neq 0), Q = B^T A^{-1} B, Q = B^T B$, in addition to the above special selection method, as long as you keep Q symmetric positive definite, can also have other method.

Evidently, the 3-SSOR-like iteration method can be equivalently rewritten as

$$\begin{pmatrix} x^{(k+1)} \\ y^{(k+1)} \end{pmatrix} = M_{(\alpha,\omega,\tau)} \begin{pmatrix} x^{(k)} \\ y^{(k)} \end{pmatrix} + H^{-1}_{(\alpha,\omega,\tau)} \begin{pmatrix} p \\ -q \end{pmatrix} \quad (17)$$

where I is the identity matrix,

$$M_{(\alpha,\omega,\tau)} = G_{(\alpha,\omega,\tau)} F_{(\alpha,\omega,\tau)} \quad (18)$$

$$= \begin{pmatrix} \mathcal{M}_{11} & \mathcal{M}_{12} \\ \mathcal{M}_{21} & \mathcal{M}_{22} \end{pmatrix} \quad (19)$$

$$= P^{-1}_{(\alpha,\omega,\tau)} N^{-1}_{(\alpha,\omega,\tau)} S_{(\alpha,\omega,\tau)} \quad (20)$$

and

$$H_{(\alpha,\omega,\tau)} = \Omega^{-1}(D - \Omega L')D^{-1}(2I - \Omega)^{-1}(D - \Omega U') = \begin{pmatrix} \frac{1}{\omega(2-\omega)} A & \frac{1}{2-\omega} B \\ -\frac{1}{2-\omega} B^T & -\frac{\omega}{2-\omega} B^T A^{-1} B + \frac{(1-\alpha\tau)(1-\tau+\alpha\tau)}{\tau(2-\tau)} Q \end{pmatrix}$$

where

$$G_{(\alpha,\omega,\tau)} = (D - \Omega U')^{-1}[(I - \Omega)D + \Omega L'] = \begin{pmatrix} (1-\omega)I_m - \frac{\omega\tau}{1-\tau+\alpha\tau} A^{-1} B Q^{-1} B^T & -\omega A^{-1} B \\ \frac{\tau}{1-\tau+\alpha\tau} Q^{-1} B^T & I_n \end{pmatrix},$$

$$F_{(\alpha,\omega,\tau)} = (D - \Omega L')^{-1}[(I - \Omega)D + \Omega U'] = \begin{pmatrix} (1-\omega)I_m & -\omega A^{-1} B \\ \frac{(1-\omega)\tau}{1-\alpha\tau} Q^{-1} B^T & I_n - \frac{\omega\tau}{1-\alpha\tau} Q^{-1} B^T A^{-1} B \end{pmatrix};$$

$$\mathcal{M}_{11} = (1-\omega)^2 I_m - \frac{\omega\tau(2-\tau)(1-\omega)}{(1-\tau+\alpha\tau)(1-\alpha\tau)} A^{-1} B Q^{-1} B^T;$$

$$\mathcal{M}_{12} = (\omega^2 - 2\omega) A^{-1} B + \frac{\omega^2 \tau(2-\tau) A^{-1} B Q^{-1} B^T A^{-1} B}{(1-\tau+\alpha\tau)(1-\alpha\tau)};$$

$$\mathcal{M}_{21} = \frac{(1-\omega)\tau(2-\tau)}{(1-\tau+\alpha\tau)(1-\alpha\tau)} Q^{-1} B^T;$$

$$\mathcal{M}_{22} = I_n - \frac{\omega\tau(2-\tau)}{(1-\tau+\alpha\tau)(1-\alpha\tau)} Q^{-1} B^T A^{-1} B$$

and

$$P_{(\alpha,\omega,\tau)} = \begin{pmatrix} A & \omega B \\ 0 & (1-\tau+\alpha\tau)Q \end{pmatrix}$$

$$N_{(\alpha,\omega,\tau)} = \begin{pmatrix} \frac{1}{1-\omega} A^{-1} & 0 \\ -\frac{(1-\alpha\tau)\tau}{(1-\omega)(1-\tau+\alpha\tau)} Q^{-1} B^T A^{-1} & \frac{1-\alpha\tau}{1-\tau+\alpha\tau} Q^{-1} \end{pmatrix};$$

$$S_{(\alpha,\omega,\tau)} = \begin{pmatrix} (1-\omega)I_m & -\omega A^{-1} B \\ (1-\omega)\tau Q^{-1} B^T & (1-\alpha\tau)I_n - \omega\tau Q^{-1} B^T A^{-1} B \end{pmatrix}$$

Here, $M_{(\alpha,\omega,\tau)}$ is the iteration matrix of the 3-SSOR-like iteration. In fact, (17) may also result from the splitting

$$\mathcal{A} = H_{(\alpha,\omega,\tau)} - \mathcal{N}_{(\alpha,\omega,\tau)} \quad (21)$$

of the coefficient matrix $\mathcal{A}$, with

$$\mathcal{N}_{(\alpha,\omega,\tau)} = H_{(\alpha,\omega,\tau)} - \mathcal{A} = \begin{pmatrix} \frac{(1-\omega)^2}{\omega(2-\omega)} A & \frac{\omega-1}{2-\omega} B \\ -\frac{1-\omega}{2-\omega} B^T & -\frac{\omega}{2-\omega} B^T A^{-1} B + \frac{(1-\alpha\tau)(1-\tau+\alpha\tau)}{\tau(2-\tau)} Q \end{pmatrix}$$

Easily, we see that

$$M_{(\alpha,\omega,\tau)} = H^{-1}_{(\alpha,\omega,\tau)} \mathcal{N}_{(\alpha,\omega,\tau)} \quad (22)$$

is the iteration matrix of the 3-SSOR-like method. See also (19). Therefore, the 3-SSOR-like method is convergent if an only if the spectral radius of the matrix $M_{(\alpha,\omega,\tau)}$, define in (19) or (22), is less than one, i.e., $\rho(M_{(\alpha,\omega,\tau)}) < 1$. See[11].

3 Convergence analysis

In this section, we will analyze convergence region for parameters α, ω and τ, in the 3-SSOR-like method to solve the saddle point problems (7).

Note that

$$D - \Omega L' = \begin{pmatrix} A & 0 \\ -\tau B^T & (1-\alpha\tau)Q \end{pmatrix}$$

$$D - \Omega U' = \begin{pmatrix} A & \omega B \\ 0 & (1-\tau+\alpha\tau)Q. \end{pmatrix}$$

Since $\mathcal{A}$ is symmetric and positive definite and Q is nonsingular, so

$$\det(D - \Omega L') = (1-\alpha\tau)^n \det(A)\det(Q) \neq 0;$$
$$\det(D - \Omega U') = (1-\tau+\alpha\tau)^n \det(A)\det(Q) \neq 0$$

if and only if $1 - \alpha\tau \neq 0$ and $1 - \tau + \alpha\tau \neq 0$, i.e.,

$$\tau \neq \left\{ \frac{1}{\alpha}, \frac{1}{1-\alpha} \right\}. \quad (23)$$

Lemma 4. *[11] Consider the quadratic equation $x^2 - bx + c = 0$, when b and c are real numbers, Both roots of the equation are less than one in modulus if and only if $|c| < 1$ and $|b| < 1 + c$.*

Lemma 5. *Suppose that μ is an eigenvalue of $J = Q^{-1}B^TA^{-1}B$. If λ satisfies*

$$\begin{aligned}&(1-\tau+\alpha\tau)(1-\alpha\tau)(\lambda-1)[(1-\omega)^2-\lambda]\\ &=\lambda\tau\omega(2-\tau)(2-\omega)\mu\end{aligned} \tag{24}$$

then λ is an eigenvalue of $M_{(\alpha,\omega,\tau)}$. Conversely, if $\lambda \neq 1$ is an eigenvalue of $M_{(\alpha,\omega,\tau)}$, and $\lambda \neq (1-\omega)^2$, and μ satisfies (24), then μ is a nonzero eigenvalue of J.

Proof. Evidently, the eigenvalue μ of J are real and nonzero if Q is a nonsingular symmetric positive definite matrix. Let λ ($\lambda \neq (1-\omega)^2$ and $\lambda \neq 1$) be a nonzero eigenvalue of $M_{(\alpha,\omega,\tau)}$ with eigenvector $\begin{pmatrix} x \\ y \end{pmatrix} \in R^{m+n}$. Then we can obtain

$$\begin{pmatrix} \mathcal{M}_{11} & \mathcal{M}_{12} \\ \mathcal{M}_{21} & \mathcal{M}_{22} \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \lambda\begin{pmatrix} x \\ y \end{pmatrix} \tag{25}$$

or

$$P^{-1}_{(\alpha,\omega,\tau)}N^{-1}_{(\alpha,\omega,\tau)}S_{(\alpha,\omega,\tau)}\begin{pmatrix} x \\ y \end{pmatrix} = \lambda\begin{pmatrix} x \\ y \end{pmatrix} \tag{26}$$

so

$$S_{(\alpha,\omega,\tau)}\begin{pmatrix} x \\ y \end{pmatrix} = \lambda N_{(\alpha,\omega,\tau)}P_{(\alpha,\omega,\tau)}\begin{pmatrix} x \\ y \end{pmatrix} \tag{27}$$

that is,

$$\begin{aligned}&\begin{pmatrix} (1-\omega)I_m & -\omega A^{-1}B \\ (1-\omega)\tau Q^{-1}B^T & (1-\alpha\tau)I_n - \omega\tau J \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}\\ &= \lambda\begin{pmatrix} \frac{1}{1-\omega}A^{-1} & 0 \\ \frac{-(1-\alpha\tau)\tau}{(1-\omega)(1-\tau+\alpha\tau)}Q^{-1}B^TA^{-1} & \frac{1-\alpha\tau}{1-\tau+\alpha\tau}Q^{-1} \end{pmatrix}\\ &\times\begin{pmatrix} A & \omega B \\ 0 & (1-\tau+\alpha\tau)Q \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}\end{aligned}$$

so

$$\begin{aligned}&\begin{pmatrix} (1-\omega)I_m & -\omega A^{-1}B \\ (1-\omega)\tau Q^{-1}B^T & (1-\alpha\tau)I_n - \omega\tau J \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \\ &\lambda\begin{pmatrix} \frac{1}{1-\omega}I_m & \frac{\omega}{1-\omega}A^{-1}B \\ \frac{-(1-\alpha\tau)\tau Q^{-1}B^T}{(1-\omega)(1-\tau+\alpha\tau)} & (1-\alpha\tau)I_n - \frac{\omega\tau(1-\alpha\tau)J}{(1-\omega)(1-\tau+\alpha\tau)} \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}.\end{aligned}$$

From above we have the following system of two equations

$$\begin{cases} (1-\omega-\frac{\lambda}{1-\omega})x = (\omega+\frac{\lambda\omega}{1-\omega})A^{-1}By; \\ [(1-\omega)\tau + \frac{\lambda\tau(1-\alpha\tau)}{(1-\omega)(1-\tau+\alpha\tau)}]Q^{-1}B^Tx = \\ (\lambda-1)(1-\alpha\tau)y + (\omega\tau - \frac{\lambda\omega\tau(1-\alpha\tau)}{(1-\omega)(1-\tau+\alpha\tau)})Jy. \end{cases} \tag{28}$$

From the first equation, we can obtain

$$x = \frac{\omega(\lambda+1-\omega)}{(\omega-1)^2-\lambda}A^{-1}By, \tag{29}$$

which means $y \neq 0$. Taking the place of x in the second equation yields

$$\begin{aligned}&[(1-\omega)\tau + \frac{\lambda\tau(1-\alpha\tau)}{(1-\omega)(1-\tau+\alpha\tau)}]\frac{\omega(\lambda+1-\omega)}{(\omega-1)^2-\lambda}y = \\ &(\lambda-1)(1-\alpha\tau)y + (\omega\tau - \frac{\lambda\omega\tau(1-\alpha\tau)}{(1-\omega)(1-\tau+\alpha\tau)})Jy.\end{aligned}$$

By simple manipulations, it is easy to get that

$$\begin{aligned}&(1-\tau+\alpha\tau)(\lambda-1)(1-\alpha\tau)[(1-\omega)^2-\lambda]y\\ &=\lambda\tau\omega(2-\tau)(2-\omega)Jy.\end{aligned}$$

Since μ is an eigenvalue of $J = Q^{-1}B^TA^{-1}B$, then we have

$$\begin{aligned}&(1-\tau+\alpha\tau)(1-\alpha\tau)(\lambda-1)[(1-\omega)^2-\lambda]\\ &=\lambda\tau\omega(2-\tau)(2-\omega)\mu\end{aligned}$$

We can prove the second assertion by reversing the process. □

We let $\mu_k (k = 1, 2, \cdots, n)$ be the eigenvalues of the matrix J, and denote by $\mu_{min} = \min_{1\le k\le n}\mu_k$ and $\mu_{max} = \max_{1\le k\le n}\mu_k$. Moreover, without loss of generality, we assume that

$$0 < \mu_{min} = \mu_1 \le \mu_2 \le \cdots \le \mu_{n-1} \le \mu_n = \mu_{max}$$

The following theorem presents a sufficient condition for guaranteeing the convergence of the 3-SSOR-like method.

Theorem 6. *Consider the system of linear equations (7). Let $A \in R^{m\times m}$ and $Q \in R^{n\times n}$ be symmetric positive definite matrix, and $B \in R^{m\times n}$ be of full column rank. Denote the smallest and the largest eigenvalues of the matrix $J = Q^{-1}B^TA^{-1}B$ by μ_{min} and μ_{max}, respectively. Then the 3-SSOR-like method is convergent, if ω satisfies $0 < \omega < 2$, and α satisfies $0 \le \alpha \le 1$, and τ satisfies the following condition:*

$$0 < \frac{\tau(2-\tau)}{(1-\tau+\alpha\tau)(1-\alpha\tau)} < \frac{(2-\omega)^2+\omega^2}{(2-\omega)\omega\mu_{max}} \tag{30}$$

Proof. After some manipulations on lemma 5, we have

$$\lambda^2 + b\lambda + c = 0 \tag{31}$$

where

$$b = -\omega^2 + 2\omega - 2 + \frac{\tau\omega(2-\tau)(2-\omega)\mu}{(1-\tau+\alpha\tau)(1-\alpha\tau)}$$

and

$$c = (1-\omega)^2.$$

By lemma 4, $|\lambda| < 1$ if and only if

$$|(1-\omega)^2| < 1, \tag{32}$$

and

$$\left| -\omega^2 + 2\omega - 2 + \frac{\tau\omega(2-\tau)(2-\omega)\mu}{(1-\tau+\alpha\tau)(1-\alpha\tau)} \right| < (1-\omega)^2 + 1. \tag{33}$$

From (32) we have

$$0 < \omega < 2 \tag{34}$$

and the relation (35) becomes to the following inequalities:

$$-1 - (\omega-1)^2 < -\omega^2 + 2\omega - 2 + \frac{\tau\omega(2-\tau)(2-\omega)\mu}{(1-\tau+\alpha\tau)(1-\alpha\tau)} < (1-\omega)^2 + 1. \tag{35}$$

It follows that

$$\frac{\tau\omega(2-\tau)(2-\omega)\mu}{(1-\tau+\alpha\tau)(1-\alpha\tau)} > 0 \tag{36}$$

and

$$2 + 2(\omega-1)^2 - \frac{\tau\omega(2-\tau)(2-\omega)\mu}{(1-\tau+\alpha\tau)(1-\alpha\tau)} > 0. \tag{37}$$

obviously, we have

$$0 < \frac{\tau(2-\tau)}{(1-\tau+\alpha\tau)(1-\alpha\tau)} < \frac{(2-\omega)^2+\omega^2}{(2-\omega)\omega\mu}.$$

Then, we have

$$0 < \frac{\tau(2-\tau)}{(1-\tau+\alpha\tau)(1-\alpha\tau)} < \frac{(2-\omega)^2+\omega^2}{(2-\omega)\omega\mu_{max}}.$$

This completes the proof. □

Note that

$$(2-\omega)^2 + \omega^2 \geq 2\omega(2-\omega),$$

so

$$\frac{2}{\mu_{max}} \leq \frac{(2-\omega)^2+\omega^2}{(2-\omega)\omega\mu_{max}}$$

The following corollary presents a new condition for guaranteeing the convergence of the 3-SSOR-like method.

Corollary 7. *Consider the system of linear equations (7). Let $A \in R^{m\times m}$ and $Q \in R^{n\times n}$ be symmetric positive definite matrix, and $B \in R^{m\times n}$ be of full column rank. Denote the smallest and the largest eigenvalues of the matrix $J = Q^{-1}B^TA^{-1}B$ by μ_{min} and μ_{max},respectively. Then the 3-SSOR-like method is convergent, if ω satisfies $0 < \omega < 2$, and α satisfies $0 \leq \alpha \leq 1$, and τ satisfies the following condition:*

$$0 < \frac{\tau(2-\tau)}{(1-\tau+\alpha\tau)(1-\alpha\tau)} < \frac{2}{\mu_{max}} \tag{38}$$

Theorem 8. *[15] Consider the system of linear equations (7). Let $A \in R^{m\times m}$ and $Q \in R^{n\times n}$ be symmetric positive definite matrix, and $B \in R^{m\times n}$ be of full column rank. Denote the smallest and the largest eigenvalues of the matrix $J = Q^{-1}B^TA^{-1}B$ by μ_{min} and μ_{max},respectively. Then the 3-SSOR-like method is convergent, if ω satisfies $0 < \omega < 2$, and α satisfies $\alpha = \frac{1}{2}$, and τ satisfies the following condition:*

$$0 < \tau < \frac{2 + 2(\omega-1)^2}{2\omega(2-\omega)\mu_{max} + 1 + (\omega-1)^2} \tag{39}$$

Proof. Since $\alpha = \frac{1}{2}$, after some manipulations on lemma 5,we have

$$\lambda^2 + b\lambda + c = 0 \tag{40}$$

where

$$b = -\omega^2 + 2\omega - 2 + \frac{4\tau\omega(2-\omega)\mu}{2-\tau}$$

and

$$c = (1-\omega)^2$$

By lemma 4, $|\lambda| < 1$ if and only if

$$|(1-\omega)^2| < 1, \tag{41}$$

and

$$\left| -\omega^2 + 2\omega - 2 + \frac{4\tau\omega(2-\omega)\mu}{2-\tau} \right| < (1-\omega)^2 + 1. \tag{42}$$

From (41) we have

$$0 < \omega < 2, \tag{43}$$

and the relation (42) becomes the following inequalities:

$$-1 - (\omega-1)^2 < -\omega^2 + 2\omega - 2 + \frac{4\tau\omega(2-\omega)\mu}{2-\tau} < (1-\omega)^2 + 1. \tag{44}$$

It follows that

$$\frac{4\tau\omega(2-\omega)\mu}{2-\tau} > 0 \tag{45}$$

and

$$2+2(\omega-1)^2 - \frac{4\tau\omega(2-\omega)\mu}{2-\tau} > 0. \tag{46}$$

Then, we have

$$0 < \tau < \frac{2+2(\omega-1)^2}{2\omega(2-\omega)\mu_{max}+1+(\omega-1)^2}.$$

This completes the proof. □

Theorem 9. *Consider the system of linear equations (7). Let $A \in R^{m\times m}$ and $Q \in R^{n\times n}$ be symmetric positive definite matrix, and $B \in R^{m\times n}$ be of full column rank. Denote the smallest and the largest eigenvalues of the matrix $J = Q^{-1}B^TA^{-1}B$ by μ_{min} and μ_{max}, respectively. Then the 3-SSOR-like method is convergent, if $\omega = \tau$ and α satisfies $\alpha = 0$, and ω satisfies the following condition:*

$$0 < \omega < \frac{2}{\sqrt{4\mu_{max}+1}+1} < 1 \tag{47}$$

Proof. Since $\alpha = 0$, after some manipulations on lemma 5, then, we have

$$\lambda^2 + b\lambda + c = 0 \tag{48}$$

where

$$b = -\omega^2 + 2\omega - 2 - (\frac{\omega^3 - 2\omega^2}{1-\omega} - 2\omega^2 + \omega^3)\mu$$

and

$$c = (1-\omega)^2.$$

By lemma 4, $|\lambda| < 1$ if and only if

$$|(1-\omega)^2| < 1 \tag{49}$$

and

$$\left|-\omega^2 + 2\omega - 2 - \left(\frac{\omega^3 - 2\omega^2}{1-\omega} - 2\omega^2 + \omega^3\right)\mu\right| < (1-\omega)^2 + 1 \tag{50}$$

From (49) we have

$$0 < \omega < 2 \tag{51}$$

and the relation (50) changes to the following inequalities:

$$\begin{aligned} &-1-(\omega-1)^2 \\ &< -\omega^2 + 2\omega - 2 - \left(\frac{\omega^3 - 2\omega^2}{1-\omega} - 2\omega^2 + \omega^3\right)\mu \\ &< (1-\omega)^2 + 1. \end{aligned}$$

It leads to

$$\frac{\omega^2(2-\omega)^2\mu}{1-\omega} > 0 \tag{52}$$

and

$$\frac{\omega^2(2-\omega)^2\mu}{1-\omega} < (2-\omega)^2 + \omega^2. \tag{53}$$

If

$$2\frac{\omega^2(2-\omega)^2\mu}{1-\omega} < (2-\omega)^2,$$

then we have

$$\frac{\omega^2(2-\omega)^2\mu}{1-\omega} < (2-\omega)^2 < (2-\omega)^2 + \omega^2,$$

which follows that

$$\omega^2\mu + \omega - 1 < 0.$$

Therefore, we have

$$0 < \omega < \frac{2}{\sqrt{4\mu_{max}+1}+1} < 1.$$

This completes the proof. □

4 Numerical examples

In this section, we use a numerical example to further examine the effectiveness and show the advantages of the 3-SSOR-like method over the SOR-like method, GSOR method, SSOR-like and GSSOR-like method.

This example is a system of purely algebraic equations discussed in [12]. The matrices A and B are defined as follows:

$$A = (a_{i,j})_{m\times m} = \begin{cases} i+1, & i = j, \\ 1, & |i-j| = 1, \\ 0, & \text{otherwise.} \end{cases}$$

$$B = (b_{i,j})_{m\times n} = \begin{cases} j, & i = j+m-n, \\ 0, & \text{otherwise.} \end{cases}$$

We report the corresponding the number of iterations (denoted by IT), the spectral radius (denoted by ρ), the time needed for convergence (denoted by CPU) and the norm of absolute error vectors (denoted by RES) by choosing $Q = B^TB$ for all the SOR-like, 3-SSOR-like, GSOR, SSOR-like and GSSOR-like methods. The stopping criterion are used in the computations,

$$\frac{||r_k||_2}{||r_0||_2} < 10^{-6}$$

where

$$r_k = \begin{pmatrix} p \\ -q \end{pmatrix} - \begin{pmatrix} A & B \\ -B^T & 0 \end{pmatrix} \begin{pmatrix} x^{(k)} \\ y^{(k)} \end{pmatrix}$$

and $\{(x^{(k)^T}, y^{(k)^T})^T\}$ is the k-th iteration for each of the methods.

The optimum parameter for the SOR-like and G-SOR method were determined according to had and given the results. We chose the parameters for the SSOR-like, GSSOR-like and 3-SSOR-like method by trial and error. All the computations were performed on an Intel E2180 2.0GHZ CPU, 2.0G Memory, Windows XP system using Matlab 7.0.

From the below numerical results, we can see that the iteration number and the time in the GSOR method, SSOR-like method, GSSOR-like method and 3-SSOR-like method are less than that in the SOR-like method. From the IT and CPU two rows, we know that we can decrease the number of iterations and the time needed for convergence by choosing three suitable parameters. However, the relaxed parameters of the 3-SSOR-like method are not optimal and only lie in the convergence region of the method. Furthermore, the determination of optimum values of the parameters needs further studies.

Table 1: Iteration number for the `SOR-like` method

m	n	m+n	ω_{opt}	IT	CPU(s)
50	40	90	1.8201	292	0.422
200	150	350	1.9533	1032	32.75
400	300	700	1.9759	2066	546.281

Table 2: Iteration number for the `GSOR` method

m	n	m+n	ω_{opt}	τ_{opt}	IT	CPU(s)
50	40	90	0.8668	24.0711	18	0.031
200	150	350	0.6461	51.2419	30	0.922
400	300	700	0.8901	201	17	4.219

Table 3: `SSOR-like` method's iteration number

m	n	m+n	ω	IT	CPU(s)
50	40	90	0.945	25	0.0712
200	150	350	0.988	23	1.218
400	300	700	0.993	23	7.358

Table 4: `GSSOR-like` method's iteration number

m	n	m+n	ω	τ	IT	CPU(s)
50	40	90	1.246	20.8	18	0.0289
200	150	350	0.6461	95	15	0.473
400	300	700	0.65	190	16	4.032

Table 5: `3-SSOR-like(GMSSOR)` method

m	50	200	400
n	40	150	300
m+n	90	350	700
α	0.5	0.5	0.5
ω	1.4	1.2	1.4
τ	1.7	1.9	1.95
IT	16	17	18
CPU(s)	0.0122	0.1396	0.7562
RES	6.6444E-7	6.7564E-7	6.0273E-7

Table 6: `3-SSOR-like` method

m	50	200	400
n	40	150	300
m+n	90	350	700
α	0.0006	0.005	0.004
ω	1.4	1.2	1.4
τ	25	58	100
IT	16	16	17
CPU(s)	0.0202	0.1308	0.7482
RES	4.1509E-7	6.8419E-7	7.3920E-7

References:

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

[2] J. H. Bramble, J. E. Pasciak and A. T. Vassilev, Analysis of the inexact Uzawa algorithm for saddle point problem, *SIAM J. Numer. Anal.*, vol.34, 1997, No.3, pp. 1072-1092.

[3] G. H. Golub, X. Wu, J.-Y. Yuan, SOR-like methods for augmented systems, *BIT*, vol.41, 2001, pp. 71-85.

[4] Z. Z. Bai, B. N. Parlett., Z.-Q. Wang, On generalized successive overrelaxation methods for augmented linear systems, *Numer. Math.*, vol.102, 2005, pp. 1-38.

[5] J. C. Li, X. Kong, Optimum parameters of GSOR-1ike methods for the augmented systems, *Applied Mathematics and Computation*, vol.204, 2008, No.1, pp. 150-161.

[6] C. P. Pan, H. Y. Wang, The polynomial Acceleration of GSOR Iterative Algorithms for Large Sparse Saddle point problems, *Chinese Journal of Engineering Mathematics*, vol.28, 2011, No.3, pp. 307-314.

[7] C. P. Pan, H. Y.Wang, W. L. Zhao, On generalized SSOR iterative method for Saddle point

problems, *Journal of Mathematics*, vol.31, 2011, No.3, pp. 569-574.

[8] Z. Z. Bai, G. H. Golub and M. K. Ng, Hermitian and skew-Hermitian splitting methods for non-Hermitian positive definite linear systems, *SIAM journal of Matrix Analysis and Applications*, vol.24, 2003, pp. 603-626.

[9] Z. Z. Bai, G. H. Golub and J. Y. Pan, Preconditioned Hermitian and skew-Hermitian splitting methods for non-Hermitian positive definite linear systems, *Numer. Math.*, vol.98, 2004, pp. 1-32.

[10] C. P. Pan, H. Y. Wang, On generalized Preconditioned Hermitian and skew-Hermitian splitting methods for saddle point problems, *Journal On Numerical Methods and Computer Applications*, vol.32, 2011, No.3, pp. 174-182.

[11] D. M. Young, *Iterative Solutions of Large Linear Systems*, Academic Press, New York, 1971.

[12] Q. Hu and J. Zou, An iterative method with variable relaxation parameters for saddle-point problems, *SIAM Journal of Matrix Analysis and Applications*, vol.23, 2001, pp. 317-338.

[13] M. T. Darvishi and P. Hessari, Symmetric SOR method for augmented systems, *Appl. Math. Comput.*, vol.183, 2006, pp. 409-415.

[14] S. L. Wu, T. Z. Huang, X. L. Zhao, A modified SSOR iterative method for augmented systems, *J. Comput. Appl. Math.*, vol.228, 2009, pp. 424-433.

[15] L. T. Zhang, T. Z. Huang, S. H. Cheng, Y. P. Wang, Convergence of a generalized MSSOR method for augmented systems, *J. Comput. Appl. Math.*, vol.236, 2012, pp. 1841-1850.

[16] C. P. Pan, On generalized preconditioned Hermitian and skew-Hermitian splitting methods for saddle point problems, *WSEAS Transactions on Mathematics*, vol.11,2012, pp.1147-1156.

[17] Z. Li, C. Li and D. J. Evans, Chebyshev acceleration for SOR-like method,*Int. J. Comput. Math.*, vol.82, 2005, pp.583-593.

[18] M. M. Martins, W. Yousif and J. L. Santos, A variant of the AOR method for augmented systems, *Math. Comput*, vol.81, 2012, pp.399-417.