Journal of Computational and Applied Mathematics 266 (2014) 52–60

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

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

Optimal parameters of the generalized symmetric SOR method for augmented systems

Zhen Chao, Naimin Zhang *, Yunzeng Lu

School of Mathematics and Information Science, Wenzhou University, Wenzhou, 325035, People's Republic of China

ARTICLE INFO

Article history:
Received 26 August 2013
Received in revised form 13 December 2013

MSC:
65F08
65F10
65F35

Keywords:
Optimal parameters
Generalized symmetric SOR method
Augmented system
Convergence

ABSTRACT

For the augmented system of linear equations, Zhang and Lu (2008) recently studied the generalized symmetric SOR method (GSSOR) with two parameters. In this note, the optimal parameters of the GSSOR method are obtained, and numerical examples are given to illustrate the corresponding results.

1. Introduction

Consider the following augmented system of linear equations:

$$\begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} b \\ q \end{pmatrix}, \tag{1.1}$$

where $A \in \mathbb{R}^{n\times n}$ is a symmetric positive definite matrix, and $B \in \mathbb{R}^{n\times m}$ has full column rank, $b \in \mathbb{R}^n$ and $q \in \mathbb{R}^m$, B^T denotes the transpose of B. Under these assumptions, the above augmented system (1.1) has a unique solution. Systems of the form (1.1) are also termed as Karush–Kuhn–Tucker (KKT) systems, or saddle point problems, which arise in a number of scientific computing and engineering applications, such as computational fluid dynamics, mixed finite element approximation of elliptic partial differential equations, constrained optimization, and constrained least-squares problems, see, e.g., [1–6].

For the augmented system (1.1), there are many efficient iterative methods as well as their numerical properties which have been studied in the literature, such as the Uzawa methods [7–10], the Krylov subspace methods [11–13], the HSS methods and its variants [14–17], the SOR-like methods [18–20], the GSOR methods [21].

Darvishi et al. studied the SSOR iterative method [22] for solving the augmented systems, the modified SSOR (MSSOR) iterative method was discussed in [23,24], and the generalized MSSOR method was studied in [25]. Zhang and Lu discussed the generalized symmetric SOR (GSSOR) method [26] which has two parameters for the augmented systems.

Since the GSSOR method has two parameters, the choice of the optimal parameters which makes the fast convergence of GSSOR is very important for the efficiency of the method. In this paper, we study the optimal parameters of the GSSOR method for the augmented system (1.1).

* Corresponding author. Tel.: +86 577 86689262.
E-mail addresses: zhenchao1120@163.com (Z. Chao), nmzhang@wzu.edu.cn, nmzhang@aliyun.com (N. Zhang).

0377-0427/$ – see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.cam.2014.01.023

The rest of this paper is organized as follows. In Section 2, we give the optimal parameters of the generalized symmetric SOR method for augmented system (1.1). In Section 3, some numerical experiments are given to examine the feasibility and effectiveness of the GSSOR method with optimal parameters to solve augmented systems.

2. Optimal parameters of the GSSOR method

Firstly, we review the GSSOR method presented in [26]. The augmented system (1.1) can be written as the following equivalent form

$$\begin{pmatrix} A & B \\ -B^T & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} b \\ -q \end{pmatrix}. \tag{2.1}$$

We consider the following splitting

$$\mathbb{A} = \begin{pmatrix} A & B \\ -B^T & 0 \end{pmatrix} = D - \mathbb{A}_L - \mathbb{A}_U, \tag{2.2}$$

where

$$D = \begin{pmatrix} A & 0 \\ 0 & Q \end{pmatrix}, \qquad \mathbb{A}_L = \begin{pmatrix} 0 & 0 \\ B^T & 0 \end{pmatrix}, \qquad \mathbb{A}_U = \begin{pmatrix} 0 & -B \\ 0 & Q \end{pmatrix},$$

and $Q \in \mathbb{R}^{m\times m}$ is a symmetric positive definite matrix which generally is an approximate (preconditioning) matrix of the Schur complement $B^TA^{-1}B$.

Let

$$L = D^{-1}\mathbb{A}_L, \qquad U = D^{-1}\mathbb{A}_U, \qquad \Omega = \begin{pmatrix} \omega I & 0 \\ 0 & \tau I \end{pmatrix}, \qquad c = \begin{pmatrix} b \\ -q \end{pmatrix},$$

where ω and τ are two nonzero real numbers, and I is an identity matrix of proper order.

Let $z^{(n)} = ((x^{(n)})^T, (y^{(n)})^T)^T$ be the nth approximation for solving (2.1) by the GSSOR method using the splitting (2.2). We obtain $z^{(n+1/2)}$ by the GSSOR method [26] as follows:

$$z^{(n+1/2)} = \mathcal{L}_{(\omega,\tau)} z^{(n)} + (I - \Omega L)^{-1} D^{-1}\Omega c, \tag{2.3}$$

where

$$\mathcal{L}_{(\omega,\tau)} = (I - \Omega L)^{-1}[(I - \Omega) + \Omega U].$$

By backward generalized SOR, we compute $z^{(n+1)}$ from $z^{(n+1/2)}$ as

$$z^{(n+1)} = \mathcal{J}_{(\omega,\tau)} z^{(n+1/2)} + (I - \Omega U)^{-1} D^{-1}\Omega c, \tag{2.4}$$

where

$$\mathcal{J}_{(\omega,\tau)} = (I - \Omega U)^{-1}[(I - \Omega) + \Omega L].$$

After eliminating $z^{(n+1/2)}$ from (2.3) and (2.4), we obtain the GSSOR method as follows:

$$z^{(n+1)} = \mathcal{H}_{(\omega,\tau)} z^{(n)} + \mathcal{M}_{(\omega,\tau)} c, \tag{2.5}$$

where

$$\mathcal{H}_{(\omega,\tau)} = \mathcal{J}_{(\omega,\tau)}\mathcal{L}_{(\omega,\tau)} = (I - \Omega U)^{-1}[(I - \Omega) + \Omega L](I - \Omega L)^{-1}[(I - \Omega) + \Omega U], \tag{2.6}$$

and

$$\mathcal{M}_{(\omega,\tau)} = (I - \Omega U)^{-1}(2I - \Omega)(I - \Omega L)^{-1} D^{-1}\Omega. \tag{2.7}$$

Given initial vectors $x^{(0)} \in \mathbb{R}^n$ and $y^{(0)} \in \mathbb{R}^m$, and relaxation parameters ω and τ, for $k = 0, 1, 2, \ldots$, it is easy to see that the GSSOR method can be written as

$$\begin{cases} y^{(k+1)} = y^{(k)} + \dfrac{\tau(2-\tau)}{1-\tau} Q^{-1}B^T[(1-\omega)x^{(k)} - \omega A^{-1}By^{(k)} + \omega A^{-1}b] - \dfrac{\tau(2-\tau)}{1-\tau} Q^{-1}q, \\ x^{(k+1)} = (1-\omega)^2 x^{(k)} - \omega A^{-1}B[y^{(k+1)} + (1-\omega)y^{(k)}] + \omega(2-\omega)A^{-1}b. \end{cases} \tag{2.8}$$

Next, we discuss the optimal parameters of the GSSOR method. Denote the spectral set and the spectral radius of a square matrix H by $\sigma(H)$ and $\rho(H)$, respectively. In this paper, we also denote an eigenvalue of $Q^{-1}B^TA^{-1}B$ by μ, that is $\mu \in \sigma(Q^{-1}B^TA^{-1}B)$, and denote the largest and the smallest eigenvalues of $Q^{-1}B^TA^{-1}B$ by $\mu_{\max}$ and $\mu_{\min}$, respectively.

Lemma 2.1. *Let $A \in \mathbb{R}^{n\times n}$ and $Q \in \mathbb{R}^{m\times m}$ be symmetric positive definite, $B \in \mathbb{R}^{n\times m}$ be full column rank. Then, the eigenvalues of $Q^{-1}B^TA^{-1}B$ are all positive.*

Lemma 2.2 ([26]). *Suppose that μ is an eigenvalue of $Q^{-1}B^TA^{-1}B$, if λ satisfies*

$$\lambda^2 - \left(1 + (\omega-1)^2 + \frac{\omega\tau(\omega-2)(\tau-2)}{\tau-1}\mu\right)\lambda + (\omega-1)^2 = 0, \tag{2.9}$$

then λ is an eigenvalue of $\mathcal{H}_{(\omega,\tau)}$. Conversely, if λ is an eigenvalue of $\mathcal{H}_{(\omega,\tau)}$ such that $\lambda \neq 1$ and $\lambda \neq (1-\omega)^2$, and μ satisfies (2.9), then μ is a nonzero eigenvalue of $Q^{-1}B^TA^{-1}B$.

Lemma 2.3 ([26]). *Let $A \in \mathbb{R}^{n\times n}$ and $Q \in \mathbb{R}^{m\times m}$ be symmetric positive definite, $B \in \mathbb{R}^{n\times m}$ be full column rank. Then the GSSOR method is convergent if and only if ω satisfies $0 < \omega < 2$ and τ satisfies the following condition:*

$$0 < \tau < \min\{\tau_1, 1\} \quad \text{or} \quad 2 < \tau < \tau_1 + 2, \tag{2.10}$$

where $\tau_1 = \min\left\{\frac{2+2(\omega-1)^2}{\omega(2-\omega)\mu}\right\}$.

For any $\mu \in \sigma(Q^{-1}B^TA^{-1}B)$, the two roots of (2.9) or the two eigenvalues of the iteration matrix $\mathcal{H}_{(\omega,\tau)}$ are given by

$$\lambda_1(\omega,\tau,\mu) = \frac{1}{2}\left[f(\omega,\tau,\mu) + \sqrt{f(\omega,\tau,\mu)^2 - 4(\omega-1)^2}\right], \tag{2.11}$$

$$\lambda_2(\omega,\tau,\mu) = \frac{1}{2}\left[f(\omega,\tau,\mu) - \sqrt{f(\omega,\tau,\mu)^2 - 4(\omega-1)^2}\right], \tag{2.12}$$

where

$$f(\omega,\tau,\mu) = 1 + (\omega-1)^2 + \frac{\omega\tau(\omega-2)(\tau-2)}{\tau-1}\mu. \tag{2.13}$$

Let $\lambda(\omega,\tau,\mu)$ be the larger modulus of the two roots $\lambda_1(\omega,\tau,\mu)$ and $\lambda_2(\omega,\tau,\mu)$, that is

$$\lambda(\omega,\tau,\mu) = \max\{|\lambda_1(\omega,\tau,\mu)|, |\lambda_2(\omega,\tau,\mu)|\}. \tag{2.14}$$

Then, we have the following two cases.

Case I: If $\Delta \equiv f(\omega,\tau,\mu)^2 - 4(\omega-1)^2 \le 0$, then,

$$|\lambda_1(\omega,\tau,\mu)| = |\lambda_2(\omega,\tau,\mu)| = |\omega-1|,$$

that is

$$\lambda(\omega,\tau,\mu) = |\omega-1|. \tag{2.15}$$

Case II: If $\Delta > 0$, then both $\lambda_1(\omega,\tau,\mu)$ and $\lambda_2(\omega,\tau,\mu)$ are real, and it holds

$$\lambda(\omega,\tau,\mu) = \begin{cases} \lambda_1(\omega,\tau,\mu) & \text{if } f(\omega,\tau,\mu) > 0, \\ -\lambda_2(\omega,\tau,\mu) & \text{if } f(\omega,\tau,\mu) \le 0. \end{cases} \tag{2.16}$$

From Eq. (2.9), we know

$$\lambda_1(\omega,\tau,\mu)\lambda_2(\omega,\tau,\mu) = (\omega-1)^2,$$

then, it is easy to see

$$\lambda(\omega,\tau,\mu) \ge |\omega-1| \quad \text{for } 0 < \omega < 2. \tag{2.17}$$

Hence, the spectral radius of the GSSOR iteration matrix can be defined by:

$$\rho(\mathcal{H}_{(\omega,\tau)}) = \max_{\mu\in\sigma(Q^{-1}B^TA^{-1}B)}\{\lambda(\omega,\tau,\mu)\}, \tag{2.18}$$

and the optimal parameters ω_{opt} and τ_{opt} satisfy

$$\rho(\mathcal{H}_{(\omega_{\text{opt}},\tau_{\text{opt}})}) = \min_{\omega \text{ and } \tau \text{ satisfy Lemma 2.3}}\{\rho(\mathcal{H}_{(\omega,\tau)})\}. \tag{2.19}$$

Let

$$\lambda_i(\omega,\tau) = \max_{\mu\in\sigma(Q^{-1}B^TA^{-1}B)}\{|\lambda_i(\omega,\tau,\mu)|\}, \quad i = 1, 2. \tag{2.20}$$

Then it holds

$$\rho(\mathcal{H}_{(\omega,\tau)}) = \max\{\lambda_1(\omega,\tau), \lambda_2(\omega,\tau)\}. \tag{2.21}$$

Thus, in order to investigate the characteristic form of the spectral radius $\rho(\mathcal{H}_{(\omega,\tau)})$, and determine the optimum parameters, we have to fully understand the properties of $\lambda_1(\omega,\tau)$ and $\lambda_2(\omega,\tau)$.

By Lemma 2.3, it holds $\omega(\omega-2)<0$, $\frac{\tau(\tau-2)}{\tau-1}>0$, which means

$$\frac{\omega\tau(\omega-2)(\tau-2)}{\tau-1}<0. \tag{2.22}$$

From Eqs. (2.11) and (2.12), we know $|\lambda_1(\omega,\tau,\mu)|\geq|\lambda_2(\omega,\tau,\mu)|$ while $\Delta>0$ and $f(\omega,\tau,\mu)>0$ (i.e., $f(\omega,\tau,\mu)>2|\omega-1|$), and $|\lambda_2(\omega,\tau,\mu)|\geq|\lambda_1(\omega,\tau,\mu)|$ while $\Delta>0$ and $f(\omega,\tau,\mu)\leq 0$ (i.e., $f(\omega,\tau,\mu)<-2|\omega-1|$). Then together with (2.11), (2.12), (2.13), (2.20), and (2.22) we have

$$\begin{cases}\lambda_1(\omega,\tau)=\dfrac{1}{2}\left[f(\omega,\tau,\mu_{\min})+\sqrt{f(\omega,\tau,\mu_{\min})^2-4(\omega-1)^2}\right],\\ \lambda_2(\omega,\tau)=\dfrac{1}{2}\left[-f(\omega,\tau,\mu_{\max})+\sqrt{f(\omega,\tau,\mu_{\max})^2-4(\omega-1)^2}\right].\end{cases} \tag{2.23}$$

Next, we analyze Eq. (2.23). By (2.22), there exist two variables μ_1,μ_2 satisfying the following equations:

$$1+(\omega-1)^2+\frac{\omega\tau(\omega-2)(\tau-2)}{\tau-1}\mu_1=2|\omega-1|, \tag{2.24}$$

and

$$1+(\omega-1)^2+\frac{\omega\tau(\omega-2)(\tau-2)}{\tau-1}\mu_2=-2|\omega-1|, \tag{2.25}$$

where $0\leq\mu_1\leq\mu_2$. Moreover, by (2.23), it holds $\mu_1,\mu_2\in[\mu_{\min},\mu_{\max}]$, actually, if $\mu_1<\mu_{\min}$ or $\mu_2>\mu_{\max}$, then it holds $-2|\omega-1|<f(\omega,\tau,\mu)<2|\omega-1|$, which means $\Delta<0$. So it is in contradiction with the condition $\Delta>0$ of (2.23).

From (2.24) and (2.25), by some computations, it holds

$$|\omega-1|=\frac{\sqrt{\mu_2}-\sqrt{\mu_1}}{\sqrt{\mu_1}+\sqrt{\mu_2}}, \tag{2.26}$$

$$\frac{\tau(\tau-2)}{\tau-1}=\frac{1}{\sqrt{\mu_1\mu_2}}. \tag{2.27}$$

So, $f(\omega,\tau,\mu)$ can be rewritten as

$$f(\omega,\tau,\mu)=\frac{2(\mu_1+\mu_2-2\mu)}{(\sqrt{\mu_1}+\sqrt{\mu_2})^2}. \tag{2.28}$$

Then, we have

$$\begin{cases}\lambda_1(\omega,\tau)=\dfrac{(\sqrt{\mu_2-\mu_{\min}}+\sqrt{\mu_1-\mu_{\min}})^2}{(\sqrt{\mu_1}+\sqrt{\mu_2})^2},\\ \lambda_2(\omega,\tau)=\dfrac{(\sqrt{\mu_{\max}-\mu_1}+\sqrt{\mu_{\max}-\mu_2})^2}{(\sqrt{\mu_1}+\sqrt{\mu_2})^2}.\end{cases} \tag{2.29}$$

For convenience, let $\lambda_1(\omega,\tau)=\lambda_1(\mu_1,\mu_2)$, and $\lambda_2(\omega,\tau)=\lambda_2(\mu_1,\mu_2)$. It is easy to see that the following results hold true.

$$\begin{cases}\lambda_1(\mu_1,\mu_2)=\lambda_2(\mu_1,\mu_2) & \text{if } \mu_1+\mu_2=\mu_{\max}+\mu_{\min},\\ \lambda_1(\mu_1,\mu_2)>\lambda_2(\mu_1,\mu_2) & \text{if } \mu_1+\mu_2>\mu_{\max}+\mu_{\min},\\ \lambda_1(\mu_1,\mu_2)<\lambda_2(\mu_1,\mu_2) & \text{if } \mu_1+\mu_2<\mu_{\max}+\mu_{\min}.\end{cases} \tag{2.30}$$

Then by (2.21), we have

$$\rho(\mathcal{H}_{(\omega,\tau)})=\begin{cases}\lambda_1(\mu_1,\mu_2) & \text{if } \mu_1+\mu_2\geq\mu_{\max}+\mu_{\min},\\ \lambda_2(\mu_1,\mu_2) & \text{if } \mu_1+\mu_2<\mu_{\max}+\mu_{\min}.\end{cases} \tag{2.31}$$

Remark 2.4. If $\mu_{\max}=\mu_{\min}$, since $\mu_1,\mu_2\in[\mu_{\min},\mu_{\max}]$, it holds $\mu_1=\mu_2=\mu_{\max}=\mu_{\min}$, which yields $\lambda_1(\mu_1,\mu_2)=\lambda_2(\mu_1,\mu_2)=0$, and $\rho(\mathcal{H}_{(\omega,\tau)})=0$. That is, the GSSOR iteration matrix has a zero spectral radius. So, without loss of generality, in the following sections, we assume that $\mu_{\max}$ and $\mu_{\min}$ are not equal, that is, $0<\mu_{\min}<\mu_{\max}$.

Theorem 2.5. *Suppose the conditions of Lemma 2.3 are satisfied. Then the optimal parameters of the GSSOR method are given by*

$$\omega_{\text{opt}}=1\pm\frac{\sqrt{\mu_{\max}}-\sqrt{\mu_{\min}}}{\sqrt{\mu_{\max}}+\sqrt{\mu_{\min}}},\qquad \tau_{\text{opt}}=1+\frac{1\pm\sqrt{1+4\mu_{\max}\mu_{\min}}}{2\sqrt{\mu_{\max}\mu_{\min}}}, \tag{2.32}$$

and the corresponding optimal convergence factor of the GSSOR method is

$$\rho(\mathcal{H}_{(\omega_{\text{opt}},\tau_{\text{opt}})})=\frac{\sqrt{\mu_{\max}}-\sqrt{\mu_{\min}}}{\sqrt{\mu_{\max}}+\sqrt{\mu_{\min}}}. \tag{2.33}$$

Proof. Since

$$\begin{cases} \dfrac{\partial\lambda_1(\mu_1,\mu_2)}{\partial\mu_1} = \dfrac{\sqrt{\mu_2-\mu_{\min}}+\sqrt{\mu_1-\mu_{\min}}}{\sqrt{\mu_1}+\sqrt{\mu_2}}\cdot\dfrac{\sqrt{\mu_1\mu_2}+\mu_{\min}-\sqrt{(\mu_1-\mu_{\min})(\mu_2-\mu_{\min})}}{\sqrt{\mu_1(\mu_1-\mu_{\min})}(\sqrt{\mu_1}+\sqrt{\mu_2})^2}, \\ \dfrac{\partial\lambda_2(\mu_1,\mu_2)}{\partial\mu_1} = -\dfrac{\sqrt{\mu_{\max}-\mu_1}+\sqrt{\mu_{\max}-\mu_2}}{\sqrt{\mu_1}+\sqrt{\mu_2}}\cdot\dfrac{\sqrt{\mu_1\mu_2}+\mu_{\max}-\sqrt{(\mu_{\max}-\mu_1)(\mu_{\max}-\mu_2)}}{\sqrt{\mu_1(\mu_{\max}-\mu_1)}(\sqrt{\mu_1}+\sqrt{\mu_2})^2}, \end{cases}$$

it holds

$$\frac{\partial\lambda_1(\mu_1,\mu_2)}{\partial\mu_1} > 0 \quad \text{and} \quad \frac{\partial\lambda_2(\mu_1,\mu_2)}{\partial\mu_1} < 0. \tag{2.34}$$

Analogously it also holds

$$\frac{\partial\lambda_1(\mu_1,\mu_2)}{\partial\mu_2} > 0 \quad \text{and} \quad \frac{\partial\lambda_2(\mu_1,\mu_2)}{\partial\mu_2} < 0. \tag{2.35}$$

Now, we prove $\rho(\mathcal{H}_{(\omega,\tau)})$ has no minimum at $\mu_1+\mu_2 \neq \mu_{\max}+\mu_{\min}$, see the following two cases:

Case I: Assume $\rho(\mathcal{H}_{(\omega,\tau)})$ has minimum at $\mu_1+\mu_2 > \mu_{\max}+\mu_{\min}$. By (2.31), it holds $\rho(\mathcal{H}_{(\omega,\tau)}) = \lambda_1(\mu_1,\mu_2)$. Let $\hat{\mu} = \mu_{\max}+\mu_{\min}-\mu_2$. Then $\mu_1 > \hat{\mu}$. By the above monotone property of the function $\lambda_1(\mu_1,\mu_2)$, we have $\lambda_1(\mu_1,\mu_2) > \lambda_1(\hat{\mu},\mu_2)$, which contradicts the assumption.

Case II: Assume $\rho(\mathcal{H}_{(\omega,\tau)})$ has minimum at $\mu_1+\mu_2 < \mu_{\max}+\mu_{\min}$. By (2.31), it holds $\rho(\mathcal{H}_{(\omega,\tau)}) = \lambda_2(\mu_1,\mu_2)$. Let $\check{\mu} = \mu_{\max}+\mu_{\min}-\mu_2$. Then $\mu_1 < \check{\mu}$. By the above monotone property of the function $\lambda_2(\mu_1,\mu_2)$, we have $\lambda_2(\mu_1,\mu_2) > \lambda_2(\check{\mu},\mu_2)$, which contradicts the assumption.

It is easy to see from the above two cases that $\rho(\mathcal{H}_{(\omega,\tau)})$ may have minimum only at $\mu_1+\mu_2 = \mu_{\max}+\mu_{\min}$.

When $\mu_1+\mu_2 = \mu_{\max}+\mu_{\min}$, from (2.30) and (2.31), it holds

$$\begin{aligned} \rho(\mathcal{H}_{(\omega,\tau)}) &= \lambda_1(\mu_1,\mu_2) = \lambda_2(\mu_1,\mu_2) \\ &= \frac{(\sqrt{\mu_{\max}-\mu_1}+\sqrt{\mu_{\max}-\mu_2})^2}{(\sqrt{\mu_1}+\sqrt{\mu_2})^2} \\ &= \frac{(\sqrt{\mu_{\max}-\mu_1}+\sqrt{\mu_1-\mu_{\min}})^2}{(\sqrt{\mu_1}+\sqrt{\mu_{\max}+\mu_{\min}-\mu_1})^2} \\ &= \frac{\mu_{\max}-\mu_{\min}+2\sqrt{-\mu_1^2+(\mu_{\max}+\mu_{\min})\mu_1-\mu_{\max}\mu_{\min}}}{\mu_{\max}+\mu_{\min}+2\sqrt{-\mu_1^2+(\mu_{\max}+\mu_{\min})\mu_1}}. \end{aligned}$$

Let $t = -\mu_1^2+(\mu_{\max}+\mu_{\min})\mu_1 \geq \mu_{\max}\mu_{\min}$, and define a function

$$f(t) = \rho(\mathcal{H}_{(\omega,\tau)}) = \frac{\mu_{\max}-\mu_{\min}+2\sqrt{t-\mu_{\max}\mu_{\min}}}{\mu_{\max}+\mu_{\min}+2\sqrt{t}}. \tag{2.36}$$

When $t = \mu_{\max}\mu_{\min}$, i.e., $-\mu_1^2+(\mu_{\max}+\mu_{\min})\mu_1 = \mu_{\max}\mu_{\min}$, it holds

$$(\mu_{\max}-\mu_1)(\mu_1-\mu_{\min}) = 0,$$

so $\mu_1 = \mu_{\min}$, and $\mu_2 = \mu_{\max}$.

When $t > \mu_{\max}\mu_{\min}$, it holds

$$\frac{f(t)}{dt} = \frac{(\mu_{\max}+\mu_{\min})\sqrt{t}-(\mu_{\max}-\mu_{\min})\sqrt{t-\mu_{\max}\mu_{\min}}+2\mu_{\max}\mu_{\min}}{\sqrt{t(t-\mu_{\max}\mu_{\min})}(\mu_{\max}+\mu_{\min}+2\sqrt{t})^2} > 0,$$

which means $f(t)$ is increasing with respect to t.

From the above analysis, we see that $f(t)$ has minimum at $t = \mu_{\max}\mu_{\min}$, which means $\rho(\mathcal{H}_{(\omega,\tau)})$ has minimum

$$\frac{\sqrt{\mu_{\max}}-\sqrt{\mu_{\min}}}{\sqrt{\mu_{\max}}+\sqrt{\mu_{\min}}}$$

at $\mu_1 = \mu_{\min}$ and $\mu_2 = \mu_{\max}$. And from (2.26) and (2.27), we obtain the corresponding optimal parameters

$$\omega_{\text{opt}} = 1 \pm \frac{\sqrt{\mu_{\max}}-\sqrt{\mu_{\min}}}{\sqrt{\mu_{\max}}+\sqrt{\mu_{\min}}}, \qquad \tau_{\text{opt}} = 1 + \frac{1 \pm \sqrt{1+4\mu_{\max}\mu_{\min}}}{2\sqrt{\mu_{\max}\mu_{\min}}},$$

which completes the proof. $\square$

Remark 2.6. The splitting of the GSSOR method is different from that of the GSOR method [21], and the iteration matrices of them are different. However, accidentally, we find that the two methods have the same optimal convergence factors, that is, the minimal spectral radii of their iteration matrices are equal.

3. Numerical experiments

In this section, we use two examples to compare the optimal parameters and the corresponding optimal spectral radii of the SOR-like method, the GSOR method, the MSSOR method and the GSSOR method. We denote the number of iteration steps by "IT", denote elapsed CPU time in seconds by "CPU", and denote the norm of absolute residual vectors by "RES". Here, the "RES" is defined as

$$\mathrm{RES} := \sqrt{\|b - Ax^{(k)} - By^{(k)}\|_2^2 + \|q - B^T x^{(k)}\|_2^2},$$

with $({x^{(k)}}^T, {y^{(k)}}^T)^T$ the final approximate solution.

All the computations are implemented in MATLAB on a PC computer with Intel (R) Core (TM) i3 CPU 2.27 GHz, and 2.00 GB memory.

In actual computation, we choose the right hand side vector $(b^T, q^T)^T \in \mathbb{R}^{n+m}$ such that the exact solution of linear system (1.1) is $({x^{(*)}}^T, {y^{(*)}}^T)^T = (1, 1, \dots, 1)^T \in \mathbb{R}^{n+m}$, and all runs are started from the initial vector $({x^{(0)}}^T, {y^{(0)}}^T)^T = (0, 0, \dots, 0)^T \in \mathbb{R}^{n+m}$, and terminated if the current iteration satisfies $\mathrm{ERR} \le 10^{-9}$, where

$$\mathrm{ERR} := \frac{\sqrt{\|x^{(k)} - x^{(*)}\|_2^2 + \|y^{(k)} - y^{(*)}\|_2^2}}{\sqrt{\|x^{(0)} - x^{(*)}\|_2^2 + \|y^{(0)} - y^{(*)}\|_2^2}}.$$

From Theorem 2.5, there are two pairs parameters $(\omega_{\mathrm{opt}}, \tau_{\mathrm{opt}})$, which have the same corresponding optimal convergence factor and numerical results. So, in our numerical experiments, we choose one pair of them

$$\omega_{\mathrm{opt}} = 1 - \frac{\sqrt{\mu_{\max}} - \sqrt{\mu_{\min}}}{\sqrt{\mu_{\max}} + \sqrt{\mu_{\min}}}, \qquad \tau_{\mathrm{opt}} = 1 + \frac{1 - \sqrt{1 + 4\mu_{\max}\mu_{\min}}}{2\sqrt{\mu_{\max}\mu_{\min}}}$$

for the GSSOR method. The corresponding numerical results are listed in the following tables.

Example 3.1 (*[17]*)**.** Consider the Stokes equations: find μ and ω such that

$$\begin{cases} -\mu\triangle\mathbf{u} + \nabla\mathbf{w} = \tilde{f}, & \text{in } \Omega, \\ \nabla\cdot\mathbf{u} = \tilde{g}, & \text{in } \Omega, \\ \mathbf{u} = 0, & \text{on } \partial\Omega, \\ \displaystyle\int_\Omega \mathbf{w}(x)dx = 0, & \end{cases} \tag{3.1}$$

where $\Omega = (0, 1) \times (0, 1) \subset \mathbb{R}^2$, $\partial\Omega$ is the boundary of Ω, $\mathbf{u}$ is a vector-valued function representing the velocity, $\triangle$ is the componentwise Laplace operator, and $\mathbf{w}$ is a scalar function representing the pressure. By discretizing (3.1) with the upwind scheme, it obtains the linear equation (1.1) with the matrix blocks of the following form:

$$A = \begin{pmatrix} I\otimes T + T\otimes I & 0 \\ 0 & I\otimes T + T\otimes I \end{pmatrix} \in \mathbb{R}^{2l^2\times 2l^2}, \qquad B = \begin{pmatrix} I\otimes F \\ F\otimes I \end{pmatrix} \in \mathbb{R}^{2l^2\times l^2},$$

where

$$T = \frac{\mu}{h^2}\cdot \mathrm{tridiag}(-1, 2, -1) \in \mathbb{R}^{l\times l}, \qquad F = \frac{1}{h}\cdot \mathrm{tridiag}(-1, 1, 0) \in \mathbb{R}^{l\times l},$$

with $\otimes$ denotes the Kronecker product symbol and $h = \frac{1}{l+1}$ the discretization meshsize, and $S = \mathrm{tridiag}(a, b, c)$ is a tridiagonal matrix with $S_{i-1,i} = a$, $S_{i,i} = b$, $S_{i,i+1} = c$ for appropriate i. Let $n = 2l^2$ and $m = l^2$ in this example.

Example 3.2 (*[27]*)**.** Consider the Oseen equations:

$$\begin{cases} -\nu\triangle\mathbf{u} + (\mathbf{w}\cdot\nabla)\mathbf{u} + \nabla p = \tilde{f}, & \text{in } \Omega, \\ -\nabla\cdot\mathbf{u} = 0, \quad \text{in } \Omega, \\ \nabla\cdot\mathbf{w} = 0. \end{cases} \tag{3.2}$$

The test problem is a leaky two-dimensional lid-driven cavity problem in the square domain: $\Omega = (0 < x < 1 : 0 < y < 1)$, where $\mathbf{u} = (u, v)^T$ denotes the velocity field, and $\mathbf{w} = (a, b)^T$ denotes the wind. The boundary conditions are $u = v = 0$

Table 1
Choice of the matrix Q.

Case no.	Matrix Q	Description
I	$B^T\hat{A}^{-1}B$	$\hat{A} = \text{tridiag}(A)$
II	$B^T\hat{A}^{-1}B$	$\hat{A} = \text{diag}(A)$

Table 2
Optimal parameter(s) versus spectral radius for Example 3.1.

n			512	1152	2048	4608
m			256	576	1024	2304
Case I	SOR-like	ω_{opt}	0.3657	0.2620	0.2037	0.1409
		ρ_{opt}	0.7964	0.8591	0.8923	0.9269
	MSSOR	ω_{opt}	0.1848	0.1316	0.1022	0.0705
		ρ_{opt}	0.8152	0.8684	0.8978	0.9295
	GSOR	ω_{opt}	0.4429	0.3307	0.2635	0.1872
		τ_{opt}	0.2854	0.1985	0.1619	0.1033
		ρ_{opt}	0.7464	0.8181	0.8582	0.9016
	GSSOR	ω_{opt}	0.2536	0.1819	0.1418	0.0984
		τ_{opt}	0.1326	0.0943	0.0731	0.0503
		ρ_{opt}	0.7464	0.8181	0.8582	0.9016
Case II	SOR-like	ω_{opt}	0.2720	0.1915	0.1476	0.1013
		ρ_{opt}	0.8533	0.8992	0.9232	0.9480
	MSSOR	ω_{opt}	0.1367	0.0960	0.074	0.0507
		ρ_{opt}	0.8633	0.9040	0.926	0.9493
	GSOR	ω_{opt}	0.3419	0.2489	0.1956	0.1368
		τ_{opt}	0.2066	0.1423	0.1084	0.0735
		ρ_{opt}	0.8112	0.8667	0.8969	0.9291
	GSSOR	ω_{opt}	0.1888	0.1333	0.1031	0.0709
		τ_{opt}	0.0980	0.0686	0.0528	0.0361
		ρ_{opt}	0.8112	0.8667	0.8969	0.9291

Table 3
IT, CPU and RES for Example 3.1.

n			512	1152	2048	4608
m			256	576	1024	2304
Case I	SOR-like	IT	130	200	272	420
		CPU	0.401	2.786	11.660	135.704
		RES	7.308e−8	1.162e−7	1.528e−7	2.145e−7
	MSSOR	IT	147	218	290	438
		CPU	0.725	5.235	21.423	254.695
		RES	5.955e−8	9.209e−8	1.316e−7	1.995e−7
	GSOR	IT	99	149	199	301
		CPU	0.434	2.533	10.003	81.406
		RES	8.181e−8	1.043e−7	1.452e−7	2.282e−7
	GSSOR	IT	100	150	200	303
		CPU	0.607	3.908	15.765	179.662
		RES	6.194e−8	8.601e−8	1.254e−7	1.817e−7
Case II	SOR-like	IT	191	293	398	611
		CPU	0.555	4.026	16.574	201.230
		RES	7.462e−8	1.133e−7	1.418e−7	2.153e−7
	MSSOR	IT	208	311	416	630
		CPU	1.034	7.334	30.374	357.707
		RES	6.732e−8	9.931e−8	1.312e−7	1.971e−7
	GSOR	IT	142	213	286	434
		CPU	0.519	3.319	13.454	136.772
		RES	7.711e−8	1.177e−7	1.468e−7	2.189e−7
	GSSOR	IT	143	214	287	435
		CPU	0.825	5.627	22.303	235.253
		RES	6.318e−8	1.027e−7	1.322e−7	2.036e−7

on the three fixed walls ($x = 0, y = 0, x = 1$), and $u = 1, v = 0$ on the moving wall ($y = 1$). We take constant "wind " $a = 1, b = 2$, and use the "marker and cell" (MAC) finite difference scheme [28] to discretize (3.2). Then we obtain the matrix representation of the Oseen equations (3.2),

$$\begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix} \tag{3.3}$$

Table 4
Optimal parameter(s) versus spectral radius for Example 3.2.

n			480	1104	1984	4512
m			255	575	1023	2303
Case I	SOR-like	ω_{opt}	0.3662	0.2564	0.1970	0.1344
		ρ_{opt}	0.7961	0.8623	0.8961	0.9304
	MSSOR	ω_{opt}	0.1850	0.1288	0.0983	0.0673
		ρ_{opt}	0.8150	0.8712	0.9012	0.9327
	GSOR	ω_{opt}	0.4437	0.3246	0.2555	0.1791
		τ_{opt}	0.2856	0.1939	0.1466	0.0984
		ρ_{opt}	0.7459	0.8218	0.8628	0.9060
	GSSOR	ω_{opt}	0.2541	0.1782	0.1372	0.0940
		τ_{opt}	0.1326	0.0923	0.0706	0.0480
		ρ_{opt}	0.7459	0.8218	0.8628	0.9060
Case II	SOR-like	ω_{opt}	0.2726	0.1875	0.1428	0.0966
		ρ_{opt}	0.8529	0.9014	0.9259	0.9505
	MSSOR	ω_{opt}	0.1370	0.0940	0.0715	0.0483
		ρ_{opt}	0.8630	0.9060	0.9285	0.9517
	GSOR	ω_{opt}	0.3427	0.2442	0.1895	0.1308
		τ_{opt}	0.2070	0.1392	0.1047	0.0700
		ρ_{opt}	0.8107	0.8694	0.9003	0.9323
	GSSOR	ω_{opt}	0.1893	0.1306	0.0997	0.0677
		τ_{opt}	0.0982	0.0672	0.0510	0.0344
		ρ_{opt}	0.8107	0.8694	0.9003	0.9323

Table 5
IT, CPU and RES for Example 3.2.

n			480	1104	1984	4512
m			255	575	1023	2303
Case I	SOR-like	IT	126	200	276	431
		CPU	0.427	3.168	13.529	104.581
		RES	3.699e−9	3.665e−9	3.590e−9	3.693e−9
	MSSOR	IT	142	217	293	449
		CPU	0.818	5.563	23.018	207.584
		RES	3.352e−9	3.167e−9	3.316e−9	3.359e−9
	GSOR	IT	96	148	201	309
		CPU	0.427	2.756	10.903	83.996
		RES	3.925e−9	3.708e−9	3.610e−9	3.609e−9
	GSSOR	IT	97	149	202	310
		CPU	0.465	3.554	15.078	133.348
		RES	2.949e−9	3.063e−9	3.128e−9	3.279e−9
Case II	SOR-like	IT	185	292	401	625
		CPU	0.924	5.935	24.312	144.444
		RES	3.865e−9	3.920e−9	4.115e−9	4.083e−9
	MSSOR	IT	202	309	418	643
		CPU	1.366	10.721	30.997	274.851
		RES	3.261e−9	3.619e−9	3.960e−9	3.870e−9
	GSOR	IT	137	212	288	446
		CPU	0.612	3.699	14.619	121.355
		RES	4.436e−9	3.990e−9	3.986e−9	3.368e−9
	GSSOR	IT	138	213	289	447
		CPU	0.669	4.991	22.117	203.743
		RES	3.617e−9	3.484e−9	3.599e−9	3.140e−9

with the matrix blocks of the following form:

$$A = \begin{pmatrix} F_1 & 0 \\ 0 & F_2 \end{pmatrix} \in \mathbb{R}^{2l(l-1)\times 2l(l-1)}, \qquad B^T = (B_1, B_2) \in \mathbb{R}^{l^2\times 2l(l-1)},$$

$$F_i = \nu A_i + N_i \in \mathbb{R}^{l(l-1)\times l(l-1)}, \quad (i = 1, 2).$$

And A is nonsymmetric and positive real, $\operatorname{rank}(B) = l^2 - 1$. For convenience, let $n = 2l(l-1)$ and $m = l^2 - 1$ in this example. To ensure the $(1, 1)$-block matrix is symmetric positive definite, $(1, 2)$-block has full column rank, finally we take the test coefficient matrix as follows:

$$\begin{pmatrix} \bar{A} & \bar{B} \\ -\bar{B}^T & 0 \end{pmatrix} \tag{3.4}$$

where $\bar{A} = \frac{1}{2}(A + A^T)$, $\bar{B}$ is obtained by dropping the first column of B.

In Example 3.1, we choose $\mu = 1$, the matrix Q is an approximation to the matrix $B^T A^{-1} B$, according to the two cases listed in Table 1.

In Tables 2 and 4, we list the optimal parameters and the corresponding optimal convergence factors of the different iterative methods. When the optimal parameters are employed, it is clear that all methods have reasonably small convergence factors, and the asymptotic convergence factor of the GSSOR method is the same as the GSOR method [17], which is much smaller than that of SOR-like method [18] and MSSOR method [24]. In particular, for these methods, we find that when n and m increase, the optimal parameters decrease, and all the corresponding optimal convergence factors of these methods increase gradually.

In Tables 3 and 5, we list numerical results with respect to IT, CPU and RES for the testing methods for Examples 3.1 and 3.2, with respect to varying m and n. From these tables, we see that the GSSOR method almost has the same efficiency as that of the GSOR method considerably in iteration steps and residual errors. Moreover, the GSSOR method always outperforms the SOR-like method and MSSOR method considerably in iteration steps.

Acknowledgments

The second author is supported by Zhejiang Provincial Natural Science Foundation of China under Grant No. Y1110451 and National Natural Science Foundation of China under Grant No. 61002039.

References

[1] M. Benzi, G.H. Golub, J. Liesen, Numerical solution of saddle point problems, Acta Numer. 14 (2005) 1–137.
[2] A. Björck, Numerical Methods for Least Squares Problems, SIAM, Philadelphia, PA, 1996.
[3] F. Brezzi, M. Fortin, Mixed and Hybrid Finite Element Methods, Springer, New York, 1991.
[4] H.C. Elman, Preconditioners for saddle point problems arising in computational fluid dynamics, Appl. Numer. Math. 43 (2002) 75–89.
[5] V. Girault, P.A. Raviart, Finite Element Approximation of the Navier Stokes Equations, Springer, Berlin, New York, 1979.
[6] J. Nocedal, S. Wright, Numerical Optimization, Springer, New York, 1999.
[7] Z.-Z. Bai, Z.-Q. Wang, On parameterized inexact Uzawa methods for generalized saddle point problems, Linear Algebra Appl. 428 (2008) 2900–2932.
[8] 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.
[9] F. Chen, Y.-L. Jiang, A generalization of the inexact parameterized Uzawa methods for saddle point problems, Appl. Math. Comput. 206 (2008) 765–771.
[10] H.C. Elman, G.H. Golub, Inexact and preconditioned Uzawa algorithms for saddle point problems, SIAM J. Numer. Anal. 31 (1994) 1645–1661.
[11] Z.-Z. Bai, Structured preconditioners for nonsingular matrices of block two-by-two structures, Math. Comp. 75 (2006) 791–815.
[12] Z.-Z. Bai, G.-Q. Li, Restrictively preconditioned conjugate gradient methods for systems of linear equations, IMA J. Numer. Anal. 23 (2003) 561–580.
[13] Z.-Z. Bai, J.-Y. Pan, M.K. Ng, New preconditioners for saddle point problems, Appl. Math. Comput. 172 (2006) 762–771.
[14] Z.-Z. Bai, G.H. Golub, Accelerated Hermitian and skew-Hermitian splitting iteration methods for saddle point problems, IMA J. Numer. Anal. 27 (2007) 1–23.
[15] Z.-Z. Bai, G.H. Golub, C.-K. Li, Optimal parameter in Hermitian and skew-Hermitian splitting method for certain two-by-two block matrices, SIAM J. Sci. Comput. 28 (2006) 583–603.
[16] Z.-Z. Bai, G.H. Golub, M.K. Ng, Hermitian and skew-Hermitian splitting methods for non-Hermitian positive defitine linear systems, SIAM J. Matrix Anal. Appl. 24 (2003) 603–626.
[17] Z.-Z. Bai, G.H. Golub, J.-Y. Pan, Preconditioned Hermitian and skew-Hermitian splitting methods for non-Hermitian positive semidefinite linear systems, Numer. Math. 98 (2004) 1–32.
[18] G.H. Golub, X. Wu, J.-Y. Yuan, SOR-like methods for augmented systems, BIT 41 (2001) 71–85.
[19] C.-J. Li, Z. Li, X.-H. Shao, Y.-Y. Nie, D.J. Evans, Optimum parameter for the SOR-like method for augmented systems, Int. J. Comput. Math. 81 (2004) 749–763.
[20] X.-H. Shao, Z. Li, C.-J. Li, Modified SOR-like method for the augment system, Int. J. Comput. Math. 84 (2007) 1653–1662.
[21] Z.-Z. Bai, B.N. Parlett, Z.-Q. Wang, On generalized successive overrelaxation methods for augmented linear systems, Numer. Math. 102 (2005) 1–38.
[22] M.T. Darvishi, P. Hessari, Symmetric SOR method for augmented systems, Appl. Math. Comput. 173 (2006) 404–420.
[23] M.T. Darvishi, P. Hessari, A modified symmetric successive overrelaxation method for augmented systems, Comput. Math. Appl. 61 (2011) 3128–3135.
[24] S.-L. Wu, T.-Z. Huang, X.-L. Zhao, A modified SSOR iterative method for augmented systems, J. Comput. Appl. Math. 228 (2009) 424–433.
[25] 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. 236 (2012) 1841–1850.
[26] G.-F. Zhang, Q.-H. Lu, On generalized symmetric SOR method for augmented systems, J. Comput. Appl. Math. 219 (2008) 51–58.
[27] N.-M. Zhang, T.-T. Lu, Y.-M. Wei, Semi-convergence analysis of Uzawa methods for singular saddle point problems, J. Comput. Appl. Math. 255 (2014) 334–345.
[28] H.C. Elman, Preconditioning for the steady-state Navier–Stokes equations with low viscosity, SIAM J. Sci. Comput. 20 (4) (1999) 1299–1316.