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

A R T I C L E   I N F O
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

A B S T R A C T
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.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction
Consider the following augmented system of linear equations:

\[
\begin{pmatrix}
A & B \\
B^T & 0
\end{pmatrix}
\begin{pmatrix}
x \\
y
\end{pmatrix}
=
\begin{pmatrix}
b \\
q
\end{pmatrix},
\]

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
We obtain and Schur complement equivalent form 2. Optimal parameters of the GSSOR method

Let $A$ the GSSOR method can be written as

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

We consider the following splitting

$$
A = \begin{pmatrix}
A & B \\
-B^T & 0
\end{pmatrix} = D - A_L - A_U,
$$

where

$$
D = \begin{pmatrix}
A & 0 \\
0 & Q
\end{pmatrix}, \quad A_L = \begin{pmatrix}
0 & 0 \\
B^T & 0
\end{pmatrix}, \quad A_U = \begin{pmatrix}
0 & -B \\
0 & Q
\end{pmatrix},
$$

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 $B^T A^{-1} B$.

Let

$$
L = D^{-1} A_L, \quad U = D^{-1} A_U, \quad \Omega = \begin{pmatrix}
\omega I & 0 \\
0 & \tau I
\end{pmatrix}, \quad c = \begin{pmatrix}
b \\
-q
\end{pmatrix},
$$

where $\omega$ and $\tau$ 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 $n$th 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:

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

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

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

where

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

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

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

where

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

and

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

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

$$
\begin{aligned}
\begin{cases}
q^{(k+1)} = y^{(k)} + \frac{\tau (2 - \tau)}{1 - \tau} Q^{-1} B^T \left[(1 - \omega) x^{(k)} - \omega A^{-1} B y^{(k)} + \omega A^{-1} b\right] - \frac{\tau (2 - \tau)}{1 - \tau} Q^{-1} q, \\
x^{(k+1)} = (1 - \omega) q^{(k)} - \omega A^{-1} B y^{(k)} + (1 - \omega) q^{(k)} + \omega (2 - \omega) A^{-1} b.
\end{cases}
\end{aligned}
$$

Next, we discuss the optimal parameters of the GSSOR method. Denote the eigenset 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 $Q^{-1} B^T A^{-1} B$ by $\mu$, that is $\mu \in \sigma(Q^{-1} B^T A^{-1} B)$, and denote the largest and the smallest eigenvalues of $Q^{-1} B^T A^{-1} B$ by $\mu_{\text{max}}$ and $\mu_{\text{min}}$, 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 $Q^{-1} B^T A^{-1} B$ are all positive.
Lemma 2.2 ([26]). Suppose that $\mu$ is an eigenvalue of $Q^{-1}B^TA^{-1}B$, if $\lambda$ satisfies
\[\lambda^2 - \left(1 + (\omega - 1)^2 + \frac{\omega \tau (\omega - 2)(\tau - 2)}{\tau - 1}\right)\lambda + (\omega - 1)^2 = 0,\] (2.9)
then $\lambda$ is an eigenvalue of $\mathcal{H}_{(\omega, \tau)}$. Conversely, if $\lambda$ is an eigenvalue of $\mathcal{H}_{(\omega, \tau)}$ such that $\lambda \neq 1$ and $\lambda \neq (1 - \omega)^2$, and $\mu$ satisfies (2.9), then $\mu$ 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 $\omega$ satisfies $0 < \omega < 2$ and $\tau$ satisfies the following condition:
\[0 < \tau < \min\{\tau_1, 1\} \text{ or } 2 < \tau < \tau_1 + 2,\] (2.10)
where $\tau_1 = \min \left\{\frac{2 + 2(\omega - 1)^2}{\omega(2 - \omega)}, 1\right\}$.

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
\[\lambda_1(\omega, \tau, \mu) = \frac{1}{2} \left[f(\omega, \tau, \mu) + \sqrt{f(\omega, \tau, \mu)^2 - 4(\omega - 1)^2}\right],\] \[\lambda_2(\omega, \tau, \mu) = \frac{1}{2} \left[f(\omega, \tau, \mu) - \sqrt{f(\omega, \tau, \mu)^2 - 4(\omega - 1)^2}\right],\] (2.11, 2.12)
where
\[f(\omega, \tau, \mu) = 1 + (\omega - 1)^2 + \frac{\omega \tau (\omega - 2)(\tau - 2)}{\tau - 1}\mu.\] (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
\[\lambda(\omega, \tau, \mu) = \max\{\lambda_1(\omega, \tau, \mu), \lambda_2(\omega, \tau, \mu)\}.\] (2.14)

Then, we have the following two cases.
Case I: If $\Delta \equiv f(\omega, \tau, \mu)^2 - 4(\omega - 1)^2 \leq 0$, then,
\[|\lambda_1(\omega, \tau, \mu)| = |\lambda_2(\omega, \tau, \mu)| = |\omega - 1|;\] (2.15)
that is
\[\lambda(\omega, \tau, \mu) = |\omega - 1|.\] (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
\[\lambda(\omega, \tau, \mu) = \begin{cases} \lambda_1(\omega, \tau, \mu) & \text{if } f(\omega, \tau, \mu) > 0, \\ -\lambda_2(\omega, \tau, \mu) & \text{if } f(\omega, \tau, \mu) \leq 0. \end{cases}\] (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
\[\lambda(\omega, \tau, \mu) \geq |\omega - 1| \quad \text{for } 0 < \omega < 2.\] (2.17)

Hence, the spectral radius of the GSSOR iteration matrix can be defined by:
\[\rho(\mathcal{H}_{(\omega, \tau)}) = \max_{\mu \in \sigma(Q^{-1}B^TA^{-1}B)} \{\lambda(\omega, \tau, \mu)\},\] (2.18)
and the optimal parameters $\omega_{\text{opt}}$ and $\tau_{\text{opt}}$ satisfy
\[\rho(\mathcal{H}_{(\omega_{\text{opt}}, \tau_{\text{opt}})}) = \min_{\omega \text{ and } \tau \text{ satisfy Lemma 2.3}} \{\rho(\mathcal{H}_{(\omega, \tau)})\}.\] (2.19)

Let
\[\lambda_i(\omega, \tau) = \max_{\mu \in \sigma(Q^{-1}B^TA^{-1}B)} \{\lambda_i(\omega, \tau, \mu)\}, \quad i = 1, 2.\] (2.20)
Then it holds
\[\rho(\mathcal{H}_{(\omega, \tau)}) = \max\{\lambda_1(\omega, \tau), \lambda_2(\omega, \tau)\}.\] (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 $\omega(\omega - 2) < 0$, \( \frac{r(t-2)}{r-1} > 0 \), which means
\[
\frac{\omega \tau (\omega - 2)(\tau - 2)}{\tau - 1} < 0.
\] (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
\[
\begin{aligned}
\lambda_1(\omega, \tau) &= \frac{1}{2} \left[ f(\omega, \tau, \mu_{\min}) + \sqrt{f(\omega, \tau, \mu_{\min})^2 - 4(\omega - 1)^2} \right] , \\
\lambda_2(\omega, \tau) &= \frac{1}{2} \left[ -f(\omega, \tau, \mu_{\max}) + \sqrt{f(\omega, \tau, \mu_{\max})^2 - 4(\omega - 1)^2} \right] .
\end{aligned}
\] (2.23)

Next, we analyze Eq. (2.23). By (2.22), there exist two variables $\mu_1$, $\mu_2$ satisfying the following equations:
\[
1 + (\omega - 1)^2 + \frac{\omega \tau (\omega - 2)(\tau - 2)}{\tau - 1} \mu_1 = 2|\omega - 1| ,
\] (2.24)
and
\[
1 + (\omega - 1)^2 + \frac{\omega \tau (\omega - 2)(\tau - 2)}{\tau - 1} \mu_2 = -2|\omega - 1| .
\] (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
\[
|\omega - 1| = \frac{\sqrt{\mu_2 - \mu_1}}{\sqrt{\mu_1} + \sqrt{\mu_2}} .
\] (2.26)
\[
\frac{\tau (\tau - 2)}{\tau - 1} = \frac{1}{\sqrt{\mu_1 \mu_2}} .
\] (2.27)

So, $f(\omega, \tau, \mu)$ can be rewritten as
\[
f(\omega, \tau, \mu) = \frac{2(\mu_1 + \mu_2 - 2\mu)}{(\sqrt{\mu_1} + \sqrt{\mu_2})^2} .
\] (2.28)

Then, we have
\[
\begin{aligned}
\lambda_1(\omega, \tau) &= \frac{\sqrt{\mu_2 - \mu_1} + \sqrt{\mu_1 - \mu_{\min}}}{(\sqrt{\mu_1} + \sqrt{\mu_2})^2} , \\
\lambda_2(\omega, \tau) &= \frac{\sqrt{\mu_{\max} - \mu_1} + \sqrt{\mu_{\max} - \mu_2}}{(\sqrt{\mu_1} + \sqrt{\mu_2})^2} .
\end{aligned}
\] (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.
\[
\begin{aligned}
\lambda_1(\mu_1, \mu_2) &= \lambda_2(\mu_1, \mu_2) & \text{if } \mu_1 + \mu_2 &= \mu_{\max} + \mu_{\min}, \\
\lambda_1(\mu_1, \mu_2) &> \lambda_2(\mu_1, \mu_2) & \text{if } \mu_1 + \mu_2 > \mu_{\max} + \mu_{\min}, \\
\lambda_1(\mu_1, \mu_2) &< \lambda_2(\mu_1, \mu_2) & \text{if } \mu_1 + \mu_2 < \mu_{\max} + \mu_{\min} .
\end{aligned}
\] (2.30)

Then by (2.21), we have
\[
\rho(H_{(\omega, \tau)}) = \begin{cases} 
\lambda_1(\mu_1, \mu_2) & \text{if } \mu_1 + \mu_2 \geq \mu_{\max} + \mu_{\min}, \\
\lambda_2(\mu_1, \mu_2) & \text{if } \mu_1 + \mu_2 < \mu_{\max} + \mu_{\min} .
\end{cases}
\] (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 $\rho(H_{(\omega, \tau)}) = 0$. 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
\[
\omega_{\text{opt}} = 1 \pm \frac{\sqrt{\mu_{\max} - \mu_{\min}} - \sqrt{\mu_{\min}}}{\sqrt{\mu_{\max} + \mu_{\min}}} , \quad \tau_{\text{opt}} = 1 + \frac{1 \pm \sqrt{1 + 4\mu_{\max}\mu_{\min}}}{2\sqrt{\mu_{\max}\mu_{\min}}} .
\] (2.32)
and the corresponding optimal convergence factor of the GSSOR method is
\[
\rho(H_{(\omega_{\text{opt}}, \tau_{\text{opt}})}) = \frac{\sqrt{\mu_{\max} - \mu_{\min}}}{\sqrt{\mu_{\max} + \mu_{\min}}} .
\] (2.33)
Proof. Since
\begin{align*}
\frac{\partial \lambda_1(\mu_1, \mu_2)}{\partial \mu_1} &= \sqrt{\mu_2 - \mu_{\min}} + \sqrt{\mu_1 - \mu_{\min}}, \quad \frac{\sqrt{\mu_1 \mu_2} - \mu_{\min} - \sqrt{(\mu_1 - \mu_{\min})(\mu_2 - \mu_{\min})}}{\sqrt{\mu_1 + \mu_2}}, \\
\frac{\partial \lambda_2(\mu_1, \mu_2)}{\partial \mu_1} &= -\frac{\sqrt{\mu_{\max} - \mu_1 + \sqrt{\mu_{\max} - \mu_2}}}{\sqrt{\mu_1 + \mu_2}}, \quad \frac{\sqrt{\mu_1 \mu_2} + \mu_{\max} - \sqrt{(\mu_{\max} - \mu_1)(\mu_{\max} - \mu_2)}}{\sqrt{\mu_1 + \mu_2}},
\end{align*}

it holds
\begin{equation}
\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}
\end{equation}

Analogously it also holds
\begin{equation}
\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}
\end{equation}

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

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

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

It is easy to see from the above two cases that $\rho(\mathcal{H}_{(t, \tau)})$ 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{equation*}
\rho(\mathcal{H}_{(t, \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 + \mu_2}} = \frac{(\sqrt{\mu_{\max} - \mu_1} + \sqrt{\mu_{\min} - \mu_1})^2}{\sqrt{\mu_1 + \mu_2}},
\end{equation*}
\begin{equation*}
= \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{equation*}

Let $t = -\mu_1^2 + (\mu_{\max} + \mu_{\min})\mu_1 \geq \mu_{\max}\mu_{\min}$, and define a function
\begin{equation}
f(t) = \rho(\mathcal{H}_{(t, \tau)}) = \frac{\mu_{\max} - \mu_{\min} + 2\sqrt{-\mu_{\max}\mu_{\min}}}{\mu_{\max} + \mu_{\min} + 2\sqrt{t}}. \tag{2.36}
\end{equation}

When $t = \mu_{\max}\mu_{\min}$, i.e., $-\mu_1^2 + (\mu_{\max} + \mu_{\min})\mu_1 = \mu_{\max}\mu_{\min}$, it holds
\begin{equation*}
(\mu_{\max} - \mu_1)(\mu_1 - \mu_{\min}) = 0,
\end{equation*}
so $\mu_1 = \mu_{\min}$ and $\mu_2 = \mu_{\max}$.

When $t > \mu_{\max}\mu_{\min}$, it holds
\begin{equation*}
f(t) = \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,
\end{equation*}
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 $\rho(\mathcal{H}_{(t, \tau)})$ has minimum
\begin{equation*}
\frac{\sqrt{\mu_{\max} - \mu_{\min}}}{\sqrt{\mu_{\max} + \sqrt{\mu_{\min}}}}
\end{equation*}
at $\mu_1 = \mu_{\min}$ and $\mu_2 = \mu_{\max}$. And from (2.26) and (2.27), we obtain the corresponding optimal parameters
\begin{equation*}
\alpha_{opt} = 1 \pm \frac{\sqrt{\mu_{\max} - \mu_{\min}}}{\sqrt{\mu_{\max} + \sqrt{\mu_{\min}}}}, \quad \tau_{opt} = 1 + \frac{1 \pm \sqrt{1 + 4\mu_{\max}\mu_{\min}}}{2\sqrt{\mu_{\max}\mu_{\min}}},
\end{equation*}
which completes the proof. \( \Box \)
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

\[ RES := \sqrt{\|b - Ax^{(k)} - By^{(k)}\|_2^2 + \|q - Bx^{(k)}\|_2^2}, \]

with \((x^{(k)}, y^{(k)})^T\) 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 \((x^{(s)})^T, y^{(s)})^T\) = \((1, 1, \ldots, 1)^T \in \mathbb{R}^{n+m}\), and all runs are started from the initial vector \((x^{(0)})^T, y^{(0)})^T\) = \((0, 0, \ldots, 0)^T \in \mathbb{R}^{n+m}\), and terminated if the current iteration satisfies \(ERR \leq 10^{-9}\), where

\[ ERR := \frac{\sqrt{\|x^{(k)} - x^{(s)}\|_2^2 + \|y^{(k)} - y^{(s)}\|_2^2}}{\sqrt{\|x^{(0)} - x^{(s)}\|_2^2 + \|y^{(0)} - y^{(s)}\|_2^2}}. \]

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

\[ \omega_{opt} = 1 - \frac{1 + \sqrt{\mu_{max}} - \sqrt{\mu_{min}}}{\mu_{max} + \mu_{min}}, \quad \tau_{opt} = 1 + \frac{1 - \sqrt{1 + 4\mu_{max}\mu_{min}}}{2\sqrt{\mu_{max}\mu_{min}}} \]

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

Example 3.1 ([17]). Consider the Stokes equations: find \(\mu\) and \(\omega\) such that

\[
\begin{cases}
-\mu \Delta \mathbf{u} + \nabla \mathbf{w} = \tilde{f}, & \text{in } \Omega, \\
\nabla \cdot \mathbf{u} = \tilde{g}, & \text{in } \Omega, \\
\mathbf{u} = 0, & \text{on } \partial \Omega, \\
\int_{\Omega} \mathbf{w}(x) dx = 0,
\end{cases}
\]

where \(\Omega = (0, 1) \times (0, 1) \subset \mathbb{R}^2\), \(\partial \Omega\) is the boundary of \(\Omega\), \(\mathbf{u}\) is a vector-valued function representing the velocity, \(\Delta\) 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:

\[
A = \begin{pmatrix} T + I & 0 \\
0 & T + I \end{pmatrix}, \quad B = \begin{pmatrix} F & I \\
F & I \end{pmatrix} \in \mathbb{R}^{2l^2 \times 2l^2},
\]

where

\[
T = \frac{\mu}{h^2} \cdot \text{tridiag}(-1, 2, -1) \in \mathbb{R}^{l^2 \times l}, \quad F = \frac{1}{h} \cdot \text{tridiag}(-1, 1, 0) \in \mathbb{R}^{l^2 \times l},
\]

with \(\otimes\) denotes the Kronecker product symbol and \(h = \frac{1}{l+1}\) the discretization meshsize, and \(S = \text{tridiag}(a, b, c)\) 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:

\[
\begin{cases}
-\nu \Delta \mathbf{u} + (\mathbf{w} \cdot \nabla) \mathbf{u} + \nabla p = \tilde{f}, & \text{in } \Omega, \\
-\nabla \cdot \mathbf{u} = 0, & \text{in } \Omega, \\
\nabla \cdot \mathbf{w} = 0.
\end{cases}
\]

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 \(u = v = 0\)
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 differences scheme [28] to discretize (3.2). Then we obtain the matrix representation of the Oseen equations (3.2),

\[
\begin{pmatrix}
A & B \\
B^T & 0
\end{pmatrix}
\begin{pmatrix}
u \\
p
\end{pmatrix} =
\begin{pmatrix}
f \\
0
\end{pmatrix}
\]  

(3.3)
Table 4
Optimal parameter(s) versus spectral radius for Example 3.2.

<table>
<thead>
<tr>
<th>n</th>
<th>m</th>
<th>(\omega_{\text{opt}})</th>
<th>(\rho_{\text{opt}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>480</td>
<td>255</td>
<td>0.3662</td>
<td>0.7961</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.2564</td>
<td>0.8623</td>
</tr>
<tr>
<td>1104</td>
<td>575</td>
<td>0.1970</td>
<td>0.8961</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.1344</td>
<td>0.9304</td>
</tr>
<tr>
<td>1984</td>
<td>1023</td>
<td>0.1850</td>
<td>0.0983</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.0673</td>
<td>0.0663</td>
</tr>
<tr>
<td>4512</td>
<td>2303</td>
<td>0.8150</td>
<td>0.9012</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.9327</td>
<td>0.9327</td>
</tr>
</tbody>
</table>

Case I SOR-like

<table>
<thead>
<tr>
<th>n</th>
<th>m</th>
<th>(\omega_{\text{opt}})</th>
<th>(\rho_{\text{opt}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>480</td>
<td>255</td>
<td>0.2726</td>
<td>0.8529</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.1875</td>
<td>0.9014</td>
</tr>
<tr>
<td>1104</td>
<td>575</td>
<td>0.1428</td>
<td>0.9259</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.0966</td>
<td>0.9505</td>
</tr>
<tr>
<td>1984</td>
<td>1023</td>
<td>0.1370</td>
<td>0.8630</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.0715</td>
<td>0.9060</td>
</tr>
<tr>
<td>4512</td>
<td>2303</td>
<td>0.1288</td>
<td>0.8961</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.0483</td>
<td>0.9304</td>
</tr>
</tbody>
</table>

Case II SOR-like

Table 5
IT, CPU and RES for Example 3.2.

<table>
<thead>
<tr>
<th>n</th>
<th>m</th>
<th>IT</th>
<th>CPU</th>
<th>RES</th>
</tr>
</thead>
<tbody>
<tr>
<td>480</td>
<td>255</td>
<td>126</td>
<td>0.427</td>
<td>3.699e−9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3.168</td>
<td>13.529</td>
</tr>
<tr>
<td>1104</td>
<td>575</td>
<td>200</td>
<td>8.181</td>
<td>3.665e−9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5.563</td>
<td>23.018</td>
</tr>
<tr>
<td>1984</td>
<td>1023</td>
<td>276</td>
<td>142</td>
<td>3.590e−9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>217</td>
<td>13.529</td>
</tr>
<tr>
<td>4512</td>
<td>2303</td>
<td>431</td>
<td>217</td>
<td>3.593e−9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>217</td>
<td>23.018</td>
</tr>
</tbody>
</table>

Case I MSSOR

with the matrix blocks of the following form:

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

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

And \( A \) is nonsymmetric and positive real, \( \text{rank}(B) = l^2 - 1 \). For convenience, let \( n = 2(l-1) \) and \( m = l^2 - 1 \). 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:

\[
\begin{pmatrix}
\tilde{A} & \tilde{B} \\
-B^T & 0
\end{pmatrix}
\]

where \( \tilde{A} = \frac{1}{2} (A + A^T) \), \( \tilde{B} \) is obtained by dropping the first column of \( B \). (3.4)
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