Ren et al. *Journal of Inequalities and Applications* (2016) 2016:279
DOI 10.1186/s13660-016-1196-8

Journal of Inequalities and Applications
a SpringerOpen Journal

RESEARCH

Open Access

A selected method for the optimal parameters of the AOR iteration

Luna Ren[1], Fujiao Ren[1*] and Ruiping Wen[2]

*Correspondence: ren123@126.com
[1]Department of Mathematics, Taiyuan Normal University, Taiyuan 030012, Shanxi, P.R. China
Full list of author information is available at the end of the article

Abstract

In this paper, we present an optimization technique to find the optimal parameters of the AOR iteration, which just needs to minimize the 2-norm of the residual vector and avoids solving the spectral radius of the iteration matrix of the SOR method. Meanwhile, numerical results are provided to indicate that the new method is more robust than the AOR method for larger intervals of the parameters ω and γ.

Keywords: linear systems; AOR iteration; optimal parameter; optimization

1 Introduction

In recent years, the iterative solvers of a large sparse linear system of equations

$$Ax = b \tag{1.1}$$

are considered in many scientific computing and engineering problems, where the coefficient matrix $A \in \mathcal{R}^{n\times n}$ is a nonsingular matrix, $b \in \mathcal{R}^n$ is a given right-hand vector, and $x \in \mathcal{R}^n$ is an unknown vector.

The accelerated overrelaxation (AOR) method, which have been proven to be a powerful tool for solving the linear system of equations (1.1), was introduced firstly by Hadjidimos [1]. In particular, he showed that when the two parameters are easily obtainable, the AOR method converges faster than the other methods of the same type. Thus, the matter about the determination of the optimal acceleration and overrelaxation parameters has to be further investigated. Lots of significant results about the optimal parameters of the AOR method were given by Hadjidimos [2]. Analytic formulas about the optimal parameters were put forward by Martins [3] in the cases where the coefficient matrix A is a weakly diagonally dominant matrix or an H-matrix. Besides, analytic formulas for optimal parameters were also provided by Hadjidimos [1, 4] in those cases where the coefficient matrix A is a consistently ordered 2-cyclic matrix, an irreducibly weakly diagonally dominant matrix, an L-matrix, or a real symmetric positive definite matrix. In order to compute the optimal parameters, the spectral radius of the corresponding SOR iteration matrix is required, which may greatly decrease the computing efficiently of the AOR iteration. In addition, the computation of this spectral radius is usually a difficult task. Furthermore, the choice of an analytic formula for a general nonsymmetric matrix is little known to us. Thus, applications of the AOR method to widespread real problems are seriously restricted.

The asymptotically optimal successive overrelaxation method of choosing the optimal factor in a dynamic fashion according to known information at the current iterate step was proposed by Bai and Chi [5]. Besides, a quasi-Chebyshev accelerated iteration method for solving a linear system was presented by Wen, Meng, and Wang [6], who obtained the optimal parameter by an optimization model. Similarly, a method of determining the optimal parameter of the SOR method was also introduced by Meng [7]. Based on the facts mentioned, an optimization technique relating to choosing the optimal parameters is put forward. Here, the optimal parameters ω and γ are computed by solving a lower-order nonlinear system that is determined by the residual vector and the coefficient matrix A. Furthermore, the optimal parameters are selected by the Newton iteration method instead of specific analytic formulas in [1, 3, 4]. In this study, applying this optimization technique to the AOR iteration, we present a modified method called the asymptotically optimal AOR (AOAOR) method, which is more stable and effective for large linear systems than the AOR method.

In Section 2, we first briefly review the AOR method and its properties. Then we put forward the AOAOR method in Section 3. In Section 4, we use numerical experiments to show the stability of the AOAOR method. We end the paper with conclusions in Section 5.

2 AOR method and its properties

Hadjidimos [1] proposed the following splitting method with two parameters of the coefficient matrix A:

$$A = \mathcal{M}_{\gamma,\omega} - \mathcal{N}_{\gamma,\omega}$$

with

$$\mathcal{M}_{\gamma,\omega} = \frac{1}{\omega}(A_D - \gamma A_L), \qquad \mathcal{N}_{\gamma,\omega} = \frac{1}{\omega}\bigl[(1-\omega)A_D + (\omega-\gamma)A_L + \omega A_U\bigr], \tag{2.1}$$

where $\gamma, \omega \neq 0$, A_D is the diagonal part of A, and $-A_L$ and $-A_U$ are strictly lower and strictly upper triangular parts of A, respectively. The iteration format of the AOR method for solving the linear systems (1.1) is

$$x^{p+1} = \mathcal{L}_{\gamma,\omega} x^p + g_{\gamma,\omega},$$

where

$$\mathcal{L}_{\gamma,\omega} = (A_D - \gamma A_L)^{-1}\bigl[(1-\omega)A_D + (\omega-\gamma)A_L + \omega A_U\bigr], \qquad g_{\gamma,\omega} = \omega(A_D - \gamma A_L)^{-1} b.$$

We observe the specific values of the parameters γ and ω in [1] when the AOR method can be reduced into:

- the Jacobi method if $\omega = 0$, $\gamma = 1$;
- the Simultaneous Overrelaxtion method if $\gamma = 0$;
- the Gauss-Seidel method if $\omega = 1$, $\gamma = 1$;
- the Successive Overrelaxation method if $\omega = \gamma$.

The various generalizations and/or modifications of the AOR method can be found in [1–4, 8]. More precisely, we survey the following conclusions.

Theorem 2.1 (Theorems 4.2, 4.4, 4.6, and 4.9 of [4]) *Let A be a nonsingular matrix with nonzero diagonal entries. Then $\{x^p\}_{p=0}^{\infty}$ generated by the AOR method converges to the unique solution x_* of the linear system* (1.1) *in the following cases*:

(a) *if A is an irreducibly weakly diagonally dominant matrix*; *for $0 < \gamma \leq 1$, a sufficient condition is $0 < \omega < 2\gamma/[1 + \rho(\mathcal{L}_{\gamma,\gamma})]$*;

(b) *if A is a real symmetric positive definite matrix*; *for $0 < \gamma < 2$, a sufficient condition is $0 < \omega < 2\gamma/[1 + \rho(\mathcal{L}_{\gamma,\gamma})]$*;

(c) *if A is an L-matrix*; *for $0 < \gamma \leq 1$, sufficient conditions are* (i) $0 < \omega < 2\gamma/[1 + \rho(\mathcal{L}_{\gamma,\gamma})]$, (ii) *$A$ is an M-matrix*;

(d) *if A is an M-matrix*; *for $0 \leq \gamma \leq 1$, a sufficient condition is $0 < \omega \leq \max\{1, \frac{2\gamma}{1+\rho(\mathcal{L}_{\gamma,\gamma})}\}$*.

Remark In order to compute the optimal parameters, the $\rho(\mathcal{L}_{\gamma,\omega})$ is quite expensive and may greatly decrease the computing efficiency of the AOR iteration method from Theorem 2.1. Hence, the formulas for the optimal parameters in Theorem 2.1 are only of theoretical meaning and far away from practical applications. To further derive a reasonably applicable rule for choosing the optimal parameters, the properties of the errors or residuals of the AOR method need to be still investigated.

Let ε^p and r^p denote the error and residual vectors of the AOR method at the pth iterate step, respectively, that is, $\varepsilon^p = x^p - x_*$, $r^p = b - Ax^p$, and let

$$\mathcal{H}_{\gamma,\omega} = A\mathcal{M}_{\gamma,\omega}^{-1} - I, \tag{2.2}$$

where $\mathcal{M}_{\gamma,\omega}^{-1}$ has been defined by (2.1). Then we have the following results.

Theorem 2.2 *For the AOR method*,

(a) *if A is a symmetric positive definite matrix, then*:

$$\begin{aligned} \|\varepsilon^{p+1}\|_A^2 &= (r^p)^T \mathcal{H}_{\gamma,\omega}^T A^{-1} \mathcal{H}_{\gamma,\omega} r^p, \\ \frac{\partial}{\partial\gamma} \|\varepsilon^{p+1}\|_A^2 &= \frac{2}{\omega} (r^p)^T \mathcal{H}_{\gamma,\omega}^T \mathcal{M}_{\gamma,\omega}^{-1} A_L \mathcal{M}_{\gamma,\omega}^{-1} r^p, \\ \frac{\partial}{\partial\omega} \|\varepsilon^{p+1}\|_A^2 &= \frac{2}{\omega} (r^p)^T \mathcal{H}_{\gamma,\omega}^T \mathcal{M}_{\gamma,\omega}^{-1} r^p; \end{aligned}$$

(b) *if A is a general nonsymmetric matrix, then*:

$$\begin{aligned} \|r^{p+1}\|_2^2 &= (r^p)^T \mathcal{H}_{\gamma,\omega}^T \mathcal{H}_{\gamma,\omega} r^p, \\ \frac{\partial}{\partial\gamma} \|r^{p+1}\|_2^2 &= \frac{2}{\omega} (r^p)^T \mathcal{H}_{\gamma,\omega}^T A\mathcal{M}_{\gamma,\omega}^{-1} A_L \mathcal{M}_{\gamma,\omega}^{-1} r^p, \\ \frac{\partial}{\partial\omega} \|r^{p+1}\|_2^2 &= \frac{2}{\omega} (r^p)^T \mathcal{H}_{\gamma,\omega}^T A\mathcal{M}_{\gamma,\omega}^{-1} r^p. \end{aligned}$$

Proof By computing we have

$$\begin{aligned} \|\varepsilon^{p+1}\|_A^2 &= (\varepsilon^{p+1}, A\varepsilon^{p+1}) = (\varepsilon^p + \mathcal{M}_{\gamma,\omega}^{-1} r^p, A\varepsilon^p + A\mathcal{M}_{\gamma,\omega}^{-1} r^p) \\ &= (A^{-1}(-r^p + A\mathcal{M}_{\gamma,\omega}^{-1} r^p), -r^p + A\mathcal{M}_{\gamma,\omega}^{-1} r^p) \end{aligned}$$

$$\begin{aligned}
&= \left(A^{-1}\mathcal{H}_{\gamma,\omega}r^p, \mathcal{H}_{\gamma,\omega}r^p\right) \\
&= \left(r^p\right)^T \mathcal{H}_{\gamma,\omega}^T A^{-1}\mathcal{H}_{\gamma,\omega}r^p
\end{aligned}$$

and

$$\begin{aligned}
\left\|r^{p+1}\right\|_2^2 &= \left(r^p - A\mathcal{M}_{\gamma,\omega}^{-1}r^p, r^p - A\mathcal{M}_{\gamma,\omega}^{-1}r^p\right) \\
&= \left(-\mathcal{H}_{\gamma,\omega}r^p, -\mathcal{H}_{\gamma,\omega}r^p\right) \\
&= \left(r^p\right)^T \mathcal{H}_{\gamma,\omega}^T \mathcal{H}_{\gamma,\omega}r^p.
\end{aligned}$$

Because of the equalities

$$\begin{aligned}
&\frac{\partial(\mathcal{M}_{\gamma,\omega}^{-1})}{\partial\gamma} = -\mathcal{M}_{\gamma,\omega}^{-1}\frac{\partial(\mathcal{M}_{\gamma,\omega})}{\partial\gamma}\mathcal{M}_{\gamma,\omega}^{-1} \quad \text{and} \quad \frac{\partial(\mathcal{M}_{\gamma,\omega})}{\partial\gamma} = -\frac{1}{\omega}A_L, \\
&\frac{\partial(\mathcal{M}_{\gamma,\omega}^{-1})}{\partial\omega} = -\mathcal{M}_{\gamma,\omega}^{-1}\frac{\partial(\mathcal{M}_{\gamma,\omega})}{\partial\omega}\mathcal{M}_{\gamma,\omega}^{-1} \quad \text{and} \quad \frac{\partial(\mathcal{M}_{\gamma,\omega})}{\partial\omega} = -\frac{1}{\omega^2}(A_D - \gamma A_L),
\end{aligned}$$

we have

$$\begin{aligned}
&\frac{\partial(\mathcal{M}_{\gamma,\omega}^{-1})}{\partial\gamma} = \frac{1}{\omega}\mathcal{M}_{\gamma,\omega}^{-1}A_L\mathcal{M}_{\gamma,\omega}^{-1} \quad \text{and} \quad \frac{\partial(\mathcal{H}_{\gamma,\omega})}{\partial\gamma} = \frac{1}{\omega}A\mathcal{M}_{\gamma,\omega}^{-1}A_L\mathcal{M}_{\gamma,\omega}^{-1}, \\
&\frac{\partial(\mathcal{M}_{\gamma,\omega}^{-1})}{\partial\omega} = \frac{1}{\omega}\mathcal{M}_{\gamma,\omega}^{-1} \quad \text{and} \quad \frac{\partial(\mathcal{H}_{\gamma,\omega})}{\partial\omega} = \frac{1}{\omega}A\mathcal{M}_{\gamma,\omega}^{-1}.
\end{aligned}$$

Thus,

$$\begin{aligned}
\frac{\partial}{\partial\omega}\left\|\varepsilon^{p+1}\right\|_A^2 &= \left(r^p\right)^T\left(\frac{\partial\mathcal{H}_{\gamma,\omega}}{\partial\omega}\right)^T A^{-1}\mathcal{H}_{\gamma,\omega}r^p + \left(r^p\right)^T\mathcal{H}_{\gamma,\omega}^T A^{-1}\frac{\partial\mathcal{H}_{\gamma,\omega}}{\partial\omega}r^p \\
&= \frac{1}{\omega}\left(\left(r^p\right)^T\mathcal{M}_{\gamma,\omega}^{-T}\mathcal{H}_{\gamma,\omega}r^p + \left(r^p\right)^T\mathcal{H}_{\gamma,\omega}^T\mathcal{M}_{\gamma,\omega}^{-1}r^p\right) \\
&= \frac{2}{\omega}\left(r^p\right)^T\mathcal{H}_{\gamma,\omega}^T\mathcal{M}_{\gamma,\omega}^{-1}r^p, \\
\frac{\partial}{\partial\gamma}\left\|\varepsilon^{p+1}\right\|_A^2 &= \left(r^p\right)^T\left(\frac{\partial\mathcal{H}_{\gamma,\omega}}{\partial\gamma}\right)^T A^{-1}\mathcal{H}_{\gamma,\omega}r^p + \left(r^p\right)^T\mathcal{H}_{\gamma,\omega}^T A^{-1}\frac{\partial\mathcal{H}_{\gamma,\omega}}{\partial\gamma}r^p \\
&= \frac{1}{\omega}\left(\left(r^p\right)^T\mathcal{M}_{\gamma,\omega}^{-T}A_L^T\mathcal{M}_{\gamma,\omega}^{-T}\mathcal{H}_{\gamma,\omega}r^p + \left(r^p\right)^T\mathcal{H}_{\gamma,\omega}^T\mathcal{M}_{\gamma,\omega}^{-1}A_L\mathcal{M}_{\gamma,\omega}^{-1}r^p\right) \\
&= \frac{2}{\omega}\left(r^p\right)^T\mathcal{H}_{\gamma,\omega}^T\mathcal{M}_{\gamma,\omega}^{-1}A_L\mathcal{M}_{\gamma,\omega}^{-1}r^p, \\
\frac{\partial}{\partial\omega}\left\|r^{p+1}\right\|_2^2 &= \left(r^p\right)^T\left(\frac{\partial\mathcal{H}_{\gamma,\omega}}{\partial\omega}\right)^T\mathcal{H}_{\gamma,\omega}r^p + \left(r^p\right)^T\mathcal{H}_{\gamma,\omega}^T\frac{\partial\mathcal{H}_{\gamma,\omega}}{\partial\omega}r^p \\
&= \frac{1}{\omega}\left(\left(r^p\right)^T\mathcal{M}_{\gamma,\omega}^{-T}A^T\mathcal{H}_{\gamma,\omega}r^p + \left(r^p\right)^T\mathcal{H}_{\gamma,\omega}^T A\mathcal{M}_{\gamma,\omega}^{-1}r^p\right) \\
&= \frac{2}{\omega}\left(r^p\right)^T\mathcal{H}_{\gamma,\omega}^T A\mathcal{M}_{\gamma,\omega}^{-1}r^p,
\end{aligned}$$

$$
\begin{aligned}
\frac{\partial}{\partial\gamma}\left\|r^{p+1}\right\|_2^2 &= \left(r^p\right)^T\left(\frac{\partial\mathcal{H}_{\gamma,\omega}}{\partial\gamma}\right)^T\mathcal{H}_{\gamma,\omega}r^p + \left(r^p\right)^T\mathcal{H}_{\gamma,\omega}^T\frac{\partial\mathcal{H}_{\gamma,\omega}}{\partial\gamma}r^p \\
&= \frac{1}{\omega}\left(\left(r^p\right)^T\mathcal{M}_{\gamma,\omega}^{-T}A_L^T\mathcal{M}_{\gamma,\omega}^{-T}A^T\mathcal{H}_{\gamma,\omega}r^p + \left(r^p\right)^T\mathcal{H}_{\gamma,\omega}^TA\mathcal{M}_{\gamma,\omega}^{-1}A_L\mathcal{M}_{\gamma,\omega}^{-1}r^p\right) \\
&= \frac{2}{\omega}\left(r^p\right)^T\mathcal{H}_{\gamma,\omega}^TA\mathcal{M}_{\gamma,\omega}^{-1}A_L\mathcal{M}_{\gamma,\omega}^{-1}r^p.
\end{aligned}
$$

The theorem is proved. □

The asymptotically optimal AOR (AOAOR) method for both cases of symmetric positive definite and general nonsymmetric linear systems can be established by Theorem 2.2.

3 The asymptotically optimal AOR method

In this section, we establish the asymptotically optimal AOR method by using the idea of [5–7].

Since $(A_D - \gamma A_L)^{-1} = (I - \gamma L)^{-1}A_D^{-1}$ with $L = A_D^{-1}A_L$ is a strictly lower triangular matrix, $L^n = O$ (the zero matrix), and $\rho(\gamma L) < 1$, $(I - \gamma L)^{-1}$ can be written in the form of Taylor expansion ([9]). Then

$$
(I-\gamma L)^{-1} = \sum_{k=0}^{n-1}(\gamma L)^k. \tag{3.1}
$$

Therefore, $\mathcal{M}_{\gamma,\omega}^{-1}$ can be expressed as

$$
\mathcal{M}_{\gamma,\omega}^{-1} = \omega(A_D-\gamma A_L)^{-1} = \omega(I-\gamma L)^{-1}A_D^{-1} = \omega\sum_{k=0}^{n-1}(\gamma L)^kA_D^{-1}. \tag{3.2}
$$

Evidently, $\mathcal{M}_{\gamma,\omega}^{-1}$ can be approximated by a lower-order truncation of the matrix series on the right-hand side of (3.2). In general,

$$
\mathcal{M}_{\gamma,\omega}^{-1} \approx \omega\left(I+\alpha\gamma L+\beta^2\gamma^2L^2\right)A_D^{-1} \equiv \mu(\gamma,\omega,\alpha,\beta), \tag{3.3}
$$

where α and β are two real parameters.

In terms of (2.2) and (3.3), we have

$$
\mathcal{H}_{\gamma,\omega} = A\mathcal{M}_{\gamma,\omega}^{-1} - I = \omega A\left(I+\alpha\gamma L+\beta^2\gamma^2L^2\right)A_D^{-1} - I \equiv \nu(\gamma,\omega,\alpha,\beta). \tag{3.4}
$$

Now, applying (3.2) and (3.3) to Theorem 2.2, we get the following results:

(a) when A is a symmetric positive definite matrix,

$$
\begin{aligned}
\frac{\partial}{\partial\gamma}\left\|\varepsilon^{p+1}\right\|_A^2 &\approx \frac{2}{\omega}\left(r^p\right)^T\nu(\gamma,\omega,\alpha,\beta)^T\mu(\gamma,\omega,\alpha,\beta)A_L\mu(\gamma,\omega,\alpha,\beta)r^p \\
&= 2\big(-\omega\left(r^p\right)^TLA_D^{-1}r^p - 2\omega\gamma\alpha\left(r^p\right)^TL^2A_D^{-1}r^p + \omega^2\left(r^p\right)^TA_D^{-1}ALA_D^{-1}r^p \\
&\quad + \omega^2\gamma\left(2\alpha\left(r^p\right)^TA_D^{-1}AL^2A_D^{-1}r^p + \alpha\left(r^p\right)^TA_D^{-1}L^TALA_D^{-1}r^p\right) \\
&\quad + \omega^2\gamma^2\left(2\alpha^2+\beta^2\right)\left(r^p\right)^TA_D^{-1}\left(L^T\right)^2ALA_D^{-1}r^p
\end{aligned}
$$

$$
\begin{aligned}
&\quad + 2\omega^2\gamma^3\alpha\beta^2\left(r^p\right)^T A_D^{-1}\left(L^T\right)^2 AL^2 A_D^{-1} r^p\big) \\
&= 2\big(-\eta_3\omega - 2\alpha\eta_4\omega\gamma + \eta_5\omega^2 + (2\alpha\eta_6 + \alpha\eta_7)\omega^2\gamma \\
&\quad + \left(2\alpha^2 + \beta^2\right)\eta_8\omega^2\gamma^2 + 2\alpha\beta^2\eta_9\omega^2\gamma^3\big),
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial}{\partial\omega}\left\|\varepsilon^{p+1}\right\|_A^2 &\approx \frac{2}{\omega}\left(r^p\right)^T \nu(\gamma,\omega,\alpha,\beta)^T \mu(\gamma,\omega,\alpha,\beta) r^p \\
&= 2\big(-\left(r^p\right)^T A_D^{-1} r^p + \omega\left(r^p\right)^T A_D^{-1} A A_D^{-1} r^p - \gamma\alpha\left(r^p\right)^T L A_D^{-1} r^p \\
&\quad - \gamma^2\beta^2\left(r^p\right)^T L^2 A_D^{-1} r^p + 2\omega\gamma\alpha\left(r^p\right)^T A_D^{-1} A L A_D^{-1} r^p \\
&\quad + \omega\gamma^2\left(2\beta^2\left(r^p\right)^T A_D^{-1} A L^2 A_D^{-1} r^p + \alpha^2\left(r^p\right)^T A_D^{-1} L^T A L A_D^{-1} r^p\right) \\
&\quad + 2\omega\gamma^3\alpha\beta^2\left(r^p\right)^T A_D^{-1}\left(L^T\right)^2 A L A_D^{-1} r^p \\
&\quad + \omega\gamma^4\beta^4\left(r^p\right)^T A_D^{-1}\left(L^T\right)^2 A L^2 A_D^{-1} r^p\big) \\
&= 2\big(-\eta_1 + \eta_2\omega - \alpha\eta_3\gamma - \beta^2\eta_4\gamma^2 + 2\alpha\eta_5\omega\gamma + \left(2\beta^2\eta_6 + \alpha^2\eta_7\right)\omega\gamma^2 \\
&\quad + 2\alpha\beta^2\eta_8\omega\gamma^3 + \beta^4\eta_9\omega\gamma^4\big),
\end{aligned}
$$

where

$$
\begin{cases}
\eta_1 = (r^p)^T A_D^{-1} r^p, \\
\eta_2 = (r^p)^T A_D^{-1} A A_D^{-1} r^p, \\
\eta_3 = (r^p)^T L A_D^{-1} r^p, \\
\eta_4 = (r^p)^T L^2 A_D^{-1} r^p, \\
\eta_5 = (r^p)^T A_D^{-1} A L A_D^{-1} r^p, \\
\eta_6 = (r^p)^T A_D^{-1} A L^2 A_D^{-1} r^p, \\
\eta_7 = (r^p)^T A_D^{-1} L^T A L A_D^{-1} r^p, \\
\eta_8 = (r^p)^T A_D^{-1} (L^T)^2 A L A_D^{-1} r^p, \\
\eta_9 = (r^p)^T A_D^{-1} (L^T)^2 A L^2 A_D^{-1} r^p;
\end{cases}
$$

(b) when A is a general nonsymmetric matrix,

$$
\begin{aligned}
\frac{\partial}{\partial\gamma}\left\|r^{p+1}\right\|_2^2 &\approx \frac{2}{\omega}\left(r^p\right)^T \nu(\gamma,\omega,\alpha,\beta)^T A\mu(\gamma,\omega,\alpha,\beta) A_L \mu(\gamma,\omega,\alpha,\beta) r^p \\
&= 2\big(-\omega\left(r^p\right)^T A L A_D^{-1} r^p - 2\omega\gamma\alpha\left(r^p\right)^T A L^2 A_D^{-1} r^p \\
&\quad + \omega^2\left(r^p\right)^T A_D^{-1} A^T A L A_D^{-1} r^p \\
&\quad + \omega^2\gamma\left(2\alpha\left(r^p\right)^T A_D^{-1} A^T A L^2 A_D^{-1} r^p + \alpha\left(r^p\right)^T A_D^{-1} L^T A^T A L A_D^{-1} r^p\right) \\
&\quad + \omega^2\gamma^2\left(2\alpha^2 + \beta^2\right)\left(r^p\right)^T A_D^{-1}\left(L^T\right)^2 A^T A L A_D^{-1} r^p \\
&\quad + 2\omega^2\gamma^3\alpha\beta^2\left(r^p\right)^T A_D^{-1}\left(L^T\right)^2 A^T A L^2 A_D^{-1} r^p\big) \\
&= 2\big(-\xi_3\omega - 2\alpha\xi_4\omega\gamma + \xi_5\omega^2 + (2\alpha\xi_6 + \alpha\xi_7)\omega^2\gamma \\
&\quad + \left(2\alpha^2 + \beta^2\right)\xi_8\omega^2\gamma^2 + 2\alpha\beta^2\xi_9\omega^2\gamma^3\big),
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial}{\partial\omega}\left\|r^{p+1}\right\|_2^2 &\approx \frac{2}{\omega}\left(r^p\right)^T \nu(\gamma,\omega,\alpha,\beta)^T A\mu(\gamma,\omega,\alpha,\beta) r^p \\
&= 2\big(-\left(r^p\right)^T A A_D^{-1} r^p + \omega\left(r^p\right)^T A_D^{-1} A^T A A_D^{-1} r^p - \gamma\alpha\left(r^p\right)^T A L A_D^{-1} r^p
\end{aligned}
$$

$$
\begin{aligned}
&- \gamma^2\beta^2\left(r^p\right)^T AL^2A_D^{-1}r^p + 2\omega\gamma\alpha\left(r^p\right)^T A_D^{-1}A^TALA_D^{-1}r^p \\
&+ \omega\gamma^2\left(2\beta^2\left(r^p\right)^T A_D^{-1}A^TAL^2A_D^{-1}r^p + \alpha^2\left(r^p\right)^T A_D^{-1}L^TA^TALA_D^{-1}r^p\right) \\
&+ 2\omega\gamma^3\alpha\beta^2\left(r^p\right)^T A_D^{-1}\left(L^T\right)^2A^TALA_D^{-1}r^p \\
&+ \omega\gamma^4\beta^4\left(r^p\right)^T A_D^{-1}\left(L^T\right)^2A^TAL^2A_D^{-1}r^p\big) \\
= {}& 2\big(-\xi_1 + \xi_2\omega - \alpha\xi_3\gamma - \beta^2\xi_4\gamma^2 + \alpha\xi_5\omega\gamma + \left(2\beta^2\xi_6 + \alpha^2\xi_7\right)\omega\gamma^2 \\
&+ 2\alpha\beta^2\xi_8\omega\gamma^3 + \beta^4\xi_9\omega\gamma^4\big),
\end{aligned}
$$

where

$$
\begin{cases}
\xi_1 = (r^p)^T AA_D^{-1}r^p, \\
\xi_2 = (r^p)^T A_D^{-1}A^TAA_D^{-1}r^p, \\
\xi_3 = (r^p)^T ALA_D^{-1}r^p, \\
\xi_4 = (r^p)^T AL^2A_D^{-1}r^p, \\
\xi_5 = (r^p)^T A_D^{-1}A^TALA_D^{-1}r^p, \\
\xi_6 = (r^p)^T A_D^{-1}(L^T)^2A^TAA_D^{-1}r^p, \\
\xi_7 = (r^p)^T A_D^{-1}L^TA^TALA_D^{-1}r^p, \\
\xi_8 = (r^p)^T A_D^{-1}(L^T)^2A^TALA_D^{-1}r^p, \\
\xi_9 = (r^p)^T A_D^{-1}(L^T)^2A^TAL^2A_D^{-1}r^p.
\end{cases}
$$

Our discussion can be summarized as the following theorems.

Theorem 3.1 *Let A be a symmetric positive definite matrix. Then reasonable approximations γ^p, ω^p satisfying* $\arg\min_{2>\omega\geq\gamma>0}\|\varepsilon^{p+1}\|_A$ *are given by solving the nonlinear system*

$$
\begin{cases}
-\hat{\eta}_3\omega - 2\alpha\hat{\eta}_4\omega\gamma + \hat{\eta}_5\omega^2 + (2\alpha\hat{\eta}_6 + \alpha\hat{\eta}_7)\omega^2\gamma + (2\alpha^2 + \beta^2)\hat{\eta}_8\omega^2\gamma^2 \\
\quad + 2\alpha\beta^2\hat{\eta}_9\omega^2\gamma^3 = 0, \\
-\hat{\eta}_1 + \hat{\eta}_2\omega - \alpha\hat{\eta}_3\gamma - \beta^2\hat{\eta}_4\gamma^2 + 2\alpha\hat{\eta}_5\omega\gamma + (2\beta^2\hat{\eta}_6 + \alpha^2\hat{\eta}_7)\omega\gamma^2 \\
\quad + 2\alpha\beta^2\hat{\eta}_8\omega\gamma^3 + \beta^4\hat{\eta}_9\omega\gamma^4 = 0,
\end{cases}
\tag{3.5}
$$

where

$$
\begin{aligned}
&\rho_1 = A_D^{-1}r^p, \qquad \rho_2 = L\rho_1, \qquad \rho_3 = A\rho_1, \\
&\rho_4 = A\rho_2, \qquad \rho_5 = L\rho_2, \qquad \rho_6 = A\rho_5.
\end{aligned}
\tag{3.6}
$$

Hence,

$$
\begin{cases}
\hat{\eta}_1 = (r^p)^T\rho_1, \qquad \hat{\eta}_2 = \rho_1^T\rho_3, \qquad \hat{\eta}_3 = (r^p)^T\rho_2, \\
\hat{\eta}_4 = (r^p)^T\rho_5, \qquad \hat{\eta}_5 = \rho_1^T\rho_4, \qquad \hat{\eta}_6 = \rho_1^T\rho_6, \\
\hat{\eta}_7 = \rho_4^T\rho_4, \qquad \hat{\eta}_8 = \rho_5^T\rho_4, \qquad \hat{\eta}_9 = \rho_5^T\rho_6.
\end{cases}
\tag{3.7}
$$

Theorem 3.2 *Let A be a general nonsymmetric matrix. Then reasonable approximations γ^p, ω^p satisfying* $\arg\min_{2>\omega\geq\gamma>0}\|r^{p+1}\|_2$ *are given by solving the nonlinear system*

$$\begin{cases} -\hat{\xi}_3\omega - 2\alpha\hat{\xi}_4\omega\gamma + \hat{\xi}_5\omega^2 + (2\alpha\hat{\xi}_6 + \alpha\hat{\xi}_7)\omega^2\gamma + (2\alpha^2 + \beta^2)\hat{\xi}_8\omega^2\gamma^2 \\ \quad + 2\alpha\beta^2\hat{\xi}_9\omega^2\gamma^3 = 0, \\ -\hat{\xi}_1 + \hat{\xi}_2\omega - \alpha\hat{\xi}_3\gamma - \beta^2\hat{\xi}_4\gamma^2 + 2\alpha\hat{\xi}_5\omega\gamma + (2\beta^2\hat{\xi}_6 + \alpha^2\hat{\xi}_7)\omega\gamma^2 \\ \quad + 2\alpha\beta^2\hat{\xi}_8\omega\gamma^3 + \beta^4\hat{\xi}_9\omega\gamma^4 = 0, \end{cases} \tag{3.8}$$

where

$$\begin{aligned} &\delta_1 = A_D^{-1}r^p, \qquad \delta_2 = A\delta_1, \qquad \delta_3 = L\delta_1, \\ &\delta_4 = A\delta_3, \qquad \delta_5 = L\delta_3, \qquad \delta_6 = A\delta_5. \end{aligned} \tag{3.9}$$

Hence,

$$\begin{cases} \hat{\xi}_1 = (r^p)^T\delta_2, \qquad \hat{\xi}_2 = \delta_2^T\delta_2, \qquad \hat{\xi}_3 = (r^p)^T\delta_4, \\ \hat{\xi}_4 = (r^p)^T\delta_6, \qquad \hat{\xi}_5 = \delta_2^T\delta_4, \qquad \hat{\xi}_6 = \delta_2^T\delta_6, \\ \hat{\xi}_7 = \delta_4^T\delta_4, \qquad \hat{\xi}_8 = \delta_6^T\delta_4, \qquad \hat{\xi}_9 = \delta_6^T\delta_6. \end{cases} \tag{3.10}$$

By Theorems 3.1-3.2 the asymptotically optimal AOR (AOAOR) method can be constructed for the cases where A is a symmetric positive definite matrix or a general nonsymmetric matrix.

Method 3.1 (AOAOR method for a symmetric positive definite matrix)

S0. Given an initial vector $x^0 \in \mathcal{R}^n$, a precision ε_1, and two parameters α, β, for $p = 0, 1, 2, \ldots$:
S1. Compute $r^p = b - Ax^p$.
S2. Compute (3.6).
S3. Compute (3.7).
S4. Solve (3.5) by the Newton method (see [10]) for obtaining γ^p, ω^p.
S5. Determine γ^p, ω^p that make the Hessian matrix positive definite.
S6. Solve $(A_D - \gamma A_L)y^p = r^p$ to get y^p.
S7. Compute $x^{p+1} = x^p + \omega y^p$.

Method 3.2 (AOAOR method for a nonsymmetric matrix)

S0. Given an initial vector $x^0 \in \mathcal{R}^n$, a precision ε_2, and two parameters α, β, for $p = 0, 1, 2, \ldots$:
S1. Compute $r^p = b - Ax^p$.
S2. Compute (3.9).
S3. Compute (3.10).
S4. Solve (3.8) by the Newton method (see [10]) for obtaining γ^p, ω^p.
S5. Determine γ^p, ω^p that make the Hessian matrix positive definite.
S6. Solve $(A_D - \gamma A_L)y^p = r^p$ to get y^p.
S7. Compute $x^{p+1} = x^p + \omega y^p$.

Methods 3.1-3.2 both shared the same cost. Method 3.1 requires nine vectors x, y, r, ρ_i $(i = 1, 2, \ldots, 6)$ stored, and each iteration requires seven matrix-vector products (Ax, Lx),

Table 1 Operation forms and flops at each step of the iteration

Oper. form	Number of iteration steps						Total flops
	Step 1	Step 2	Step 3	Step 6	Step 7	Total	
Ax	1	3	0	0	0	4	$4[(2m-1)n]$
Lx	0	2	0	1	0	3	$3[(2m_l-1)n-m_l(m_l-1)]$
(x,y)	0	0	9	0	0	9	$9[(2n-1)]$
ζy	0	0	0	2	1	3	$3[n]$
$x+y$	1	0	0	1	1	3	$3[n]$
$\zeta\cdot\kappa$	0	0	0	0	0	0	0

nine inner products (S3 of Method 3.1), three operations of the form ζx, three operations of the form $x+y$, and no operations of the form $\zeta\kappa$, where ζ and κ are scalars. Thus, if the number of nonzero entries on each row of the matrix A is m and that of the matrix L is m_l, then the details are listed in Table 1.

4 Numerical experiments

The test example is the following two-dimensional partial differential equation with Dirichlet boundary condition ([5]):

$$\begin{cases} -\frac{\partial^2 u}{\partial t_1^2} - \frac{\partial^2 u}{\partial t_2^2} + \xi \frac{\partial u}{\partial t_1} + \zeta \frac{\partial u}{\partial t_1} + 4\sigma u = f(t_1,t_2), & (t_1,t_2) \in \Omega, \\ u(t_1,t_2) = 0, \quad (t_1,t_2) \in \partial\Omega, \end{cases} \tag{4.1}$$

where ξ, ζ, and σ are all constants, Ω is the unit square $(0,1)\times(0,1)$ in $\mathcal{R}^2$, $\partial\Omega$ is the boundary of the domain Ω, and $f(t_1,t_2):\Omega\to\mathcal{R}^1$ is a given function.

If $u_{i,j}$ and $f_{i,j}$ denote the approximate solution of (4.1) and an approximation of the function $f(t_1,t_2)$ at the grid point (ih,jh), respectively, then a discretized approximation of (4.1) is the following linear system of equations:

$$\mu_1 u_{i+1,j} + \eta_1 u_{i-1,j} + \mu_2 u_{i,j+1} + \eta_2 u_{i,j-1} + \mu_0 u_{i,j} = h^2 f_{i,j}, \quad i,j = 1,2,\ldots,N, \tag{4.2}$$

where $(N+1)h = 1$, and

$$\begin{cases} \mu_0 = 4(1+\sigma h^2), \quad \mu_1 = -(1-\frac{1}{2}\xi h), \quad \mu_2 = -(1-\frac{1}{2}\zeta h), \\ \eta_1 = -(1+\frac{1}{2}\xi h), \quad \eta_2 = -(1+\frac{1}{2}\zeta h). \end{cases}$$

Let

$$x^T = (u_{1,1},\ldots,u_{1,N},u_{2,1},\ldots,u_{2,N},\ldots,u_{N,1},\ldots,u_{N,N}).$$

The linear system (4.2) can be written as the linear system (1.1) with

$$A = \begin{pmatrix} T & \mu_2 I & & & \\ \eta_2 I & T & \mu_2 I & & \\ & \ddots & \ddots & \ddots & \\ & & \eta_2 I & T & \mu_2 I \\ & & & \eta_2 I & T \end{pmatrix} \in \mathcal{R}^{n\times n}$$

with

$$T = \begin{pmatrix} \mu_0 & \mu_1 & & & \\ \eta_1 & \mu_0 & \mu_1 & & \\ & \ddots & \ddots & \ddots & \\ & & \eta_1 & \mu_0 & \mu_1 \\ & & & \eta_1 & \mu_0 \end{pmatrix} \in \mathcal{R}^{N\times N}$$

and

$$b^T = h^2(f_{1,1},\dots,f_{1,N},f_{2,1},\dots,f_{2,N},\dots,f_{N,1},\dots,f_{N,N}),$$

where $n = N \times N$.

In practical computations, the right-hand side b is generated by $b = Ae$, in which $e = (1,1,\dots,1)^T \in \mathcal{R}^n$, the initial vector $x^0 \in \mathcal{R}^n$ is taken to be zero, and all runs are terminated if the current iteration satisfies either the value of residual

$$\mathrm{RES} \equiv \|r^p\|_2 \le \varepsilon \|r^0\|_2 \tag{4.3}$$

or the number of iteration steps exceeds 20,000. The iteration index p, namely, the number of iteration steps satisfying (4.3), is particularly denoted as "IT", and the running time in seconds is denoted as "CPU". In addition, all experiments are performed in MATLAB R2012a on PC with 3.4 GHz processor, 4 GB memory, and 32-bit operating system, "†" indicates that the number of iterations is greater than 20,000, that is, the iteration fails, "-" indicates that the computing is out of memory.

With different selections of ξ, ζ, σ, different forms of the coefficient matrix are given. Particularly, the coefficient matrix A is a symmetric matrix when $\xi = \zeta = \sigma = 0.0$ or $\xi = \zeta = 0.0$, $\sigma = 2.5$, and it is a nonsymmetric matrix for $\xi = 30.0$, $\zeta = 0.0$, $\sigma = 10.0$ or $\xi = 0.0$, $\zeta = 30.0$, $\sigma = 10.0$. In Tables 2-5, we list the numbers of iteration steps, the 2-norms of the residual vectors, and the computing times for different choices of ξ, ζ, σ.

From Tables 2-5 we can see that although the IT and CPU of the AOR method are less than those of the AOAOR method when $h^{-1} < 128$, the AOR method is out of memory, whereas the AOAOR method can continue to run for $h^{-1} \ge 128$. Moreover, in the iteration process, $(A_D - \gamma A_L)^{-1}$ of dimension N^2 requires to be solved for the AOR method. The AOAOR method, however, only needs to solve the inverse of N-dimensional T. Thus, the AOAOR method is better than the AOR method with respect to the increase of the size of the coefficient matrix.

The rates of convergence of the AOR and AOAOR methods are determined by the choice of the initial guesses ω and γ. If $\omega \in (0,2)$ and $\gamma \le \omega$, then the AOR method is convergent. Therefore, we choose $\omega \in (0,2)$ and $\gamma \le \omega$, where ω changes with step size of 0.1, and γ changes with step size of 0.01 in computation. It is well known that $\omega > 1$ is a wise choice since it is the overrelaxation iteration. Hence, Tables 6-9 only list the ITs of the AOAOR and AOR methods for $\omega \in (1,2)$. We further choose two points $\omega = 1.9$ and $\omega = 1$ as examples to approach two extremal points and $\omega = 1.4$ close to an intermediate point to explain that the IT of the AOR method changes greatly and that of the AOAOR method is almost unchanged for the larger intervals of the parameters ω and γ. In Tables 6-9, we have provided the gap of the iteration numbers with respect to different ξ, ζ, σ, and ε.

Table 2 Numerical results when $\xi = \zeta = \sigma = 0.0$ ($\varepsilon = h^2/5$)

Methods		h^{-1}					
		32	**64**	**128**	**256**	**288**	**300**
AOR	IT	41	89	-	-	-	-
	RES	1.9e–3	7.38e–4				
	CPU	0.20	4.18				
AOAOR	IT	216	895	3,816	16,200	†	†
	RES	2.1e–3	7.68e–4	2.76e–4	9.71e–5		
	CPU	10.84	58.96	275.80	2,300.30		

Table 3 Numerical results when $\xi = \zeta = 0$, $\sigma = 2.5$ ($\varepsilon = h^2/5$)

Methods		h^{-1}					
		32	**64**	**128**	**256**	**288**	**300**
AOR	IT	36	78	-	-	-	-
	RES	1.86e–3	7.0e–4				
	CPU	0.18	4.08				
AOAOR	IT	154	641	2,688	11,613	†	†
	RES	2.1e–3	7.8e–4	2.7e–4	9.75e–5		
	CPU	7.74	40.31	189.08	1,560.23		

Table 4 Numerical results when $\xi = 30.0$, $\zeta = 0.0$, $\sigma = 10.0$ ($\varepsilon = h^2$)

Methods		h^{-1}					
		32	**64**	**128**	**256**	**288**	**300**
AOR	IT	23	51	-	-	-	-
	RES	9.99e–3	3.42e–3				
	CPU	0.18	3.64				
AOAOR	IT	61	259	1,078	4,526	5,772	6,279
	RES	1.18e–2	3.9e–3	1.38e–3	4.8e–4	4.1e–3	3.86e–4
	CPU	2.54	13.21	52.06	288.58	379.42	434.97

Table 5 Numerical results when $\xi = 0.0$, $\zeta = 30.0$, $\sigma = 10.0$ ($\varepsilon = h^2$)

Methods		h^{-1}					
		32	**64**	**128**	**256**	**288**	**300**
AOR	IT	10	20	-	-	-	-
	RES	9.36e–3	2.6e–3				
	CPU	0.28	5.70				
AOAOR	IT	30	126	537	2,261	2,884	3,137
	RES	1.09e–2	3.8e–3	1.37e–3	4.8e–4	4.09e–4	3.86e–4
	CPU	3.80	31.29	132.69	400.02	560.39	1,092.20

From Tables 6-9 we note that:

(a) When ω is a specific value, as for $\omega = 1.9$ and $h^{-1} = 32$, the IT of the AOR method is thrown into [199, 1,824], much more than the IT of the AOAOR method in Table 8. Thus, the AOAOR method is superior to the AOR method. In Table 7, the IT increased from 95 to 529 for the AOR method, but it is 154 for the AOAOR method. Although a lot of ITs of the AOR method are less than 154, the AOR method is more efficient only by choosing an appropriate γ. Consequently, for the same ω, the AOAOR method does not depend on γ;

Table 6 The gap of IT when $\xi = \zeta = \sigma = 0.0$ ($\varepsilon = h^2/5$)

Methods	ω	γ	h^{-1}			
			32	64	128	256
AOR	1.9	[1.9 : –0.01 : 0]	[95, 782]	[120, 2,822]		
	1.4	[1.4 : –0.01 : 0]	[121, 1,699]	[513, 8,265]		
	1	[1 : –0.01 : 0]	[282, 573]	[1,197, 2,438]		
AOAOR	1.9	[1.9 : –0.01 : 0]	216	[895, 916]	[3,904, 3,907]	16,200
	1.4	[1.4 : –0.01 : 0]	216	[896, 903]	3,907	16,200
	1	[1 : –0.01 : 0]	216	[896, 906]	[3,901, 3,907]	16,200

Table 7 The gap of IT when $\xi = \zeta = 0.0$, $\sigma = 2.5$ ($\varepsilon = h^2/5$)

Methods	ω	γ	h^{-1}			
			32	64	128	256
AOR	1.9	[1.9 : –0.01 : 0]	[95, 529]	[120, 4,585]		
	1.4	[1.4 : –0.01 : 0]	[87, 7,872]	[365, 1,305]		
	1	[1 : –0.01 : 0]	[202, 405]	[852, 1,715]		
AOAOR	1.9	[1.9 : –0.01 : 0]	154	641	[2,689, 2,709]	11,613
	1.4	[1.4 : –0.01 : 0]	154	641	[2,693, 2,701]	11,613
	1	[1 : –0.01 : 0]	154	641	[2,688, 2,703]	11,613

Table 8 The gap of IT when $\xi = 30.0$, $\zeta = 0.0$, $\sigma = 10.0$ ($\varepsilon = h^2$)

Methods	ω	γ	h^{-1}				
			32	64	128	288	300
AOR	1.9	[1.9 : –0.01 : 0]	[199, 1,824]	[215, 526]			
	1.4	[1.4 : –0.01 : 0]	[23, 2,765]	[84, 1,137]			
	1	[1 : –0.01 : 0]	[50, 96]	[196, 387]			
AOAOR	1.9	[1.9 : –0.01 : 0]	61	259	[1,078, 1,081]	5,772	6,279
	1.4	[1.4 : –0.01 : 0]	61	259	[1,079, 1,081]	5,772	6,279
	1	[1 : –0.01 : 0]	61	259	[1,079, 1,081]	5,772	6,279

Table 9 The gap of IT when $\xi = 0.0$, $\zeta = 30.0$, $\sigma = 10.0$ ($\varepsilon = h^2$)

Methods	ω	γ	h^{-1}				
			32	64	128	288	300
AOR	1.9	[1.9 : –0.01 : 0]	[68, 1,860]	[82, 534]			
	1.4	[1.4 : –0.01 : 0]	[10, 1,039]	[48, 3,191]			
	1	[1 : –0.01 : 0]	[31, 92]	[160, 388]			
AOAOR	1.9	[1.9 : –0.01 : 0]	30	126	537	2,884	[3,137, 3,139]
	1.4	[1.4 : –0.01 : 0]	30	126	537	2,884	3,137
	1	[1 : –0.01 : 0]	30	126	537	2,884	3,137

(b) When ω is variable and γ is a suitable fixed value, the minimal IT of the AOR method is 199 for $\omega = 1.9$ and $h^{-1} = 32$, and it is 50 for $\omega = 1$ and $h^{-1} = 32$ in Table 8. Therefore, the result for $\omega = 1$ is better. But whether $\omega = 1$ or $\omega = 1.9$, the minimal IT ($\omega = 1.4$, IT = 23) is not reached. Thus, the IT of the AOR method can reach the minimum point only when ω is appropriate, whereas the ITs of the AOAOR method for $\omega = 1.9$, $\omega = 1.4$, and $\omega = 1$ are fixed at 61. Hence, the AOAOR method is independent of the choice of ω.

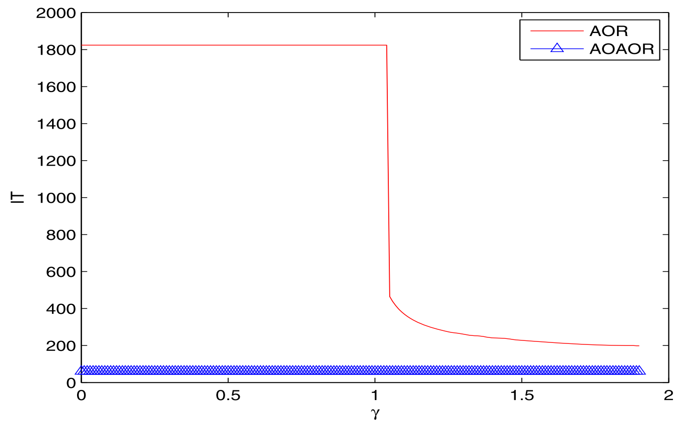

Figure 1 Curves of the IT for $\omega = 1.9$, $h^{-1} = 32$ in Table 8.

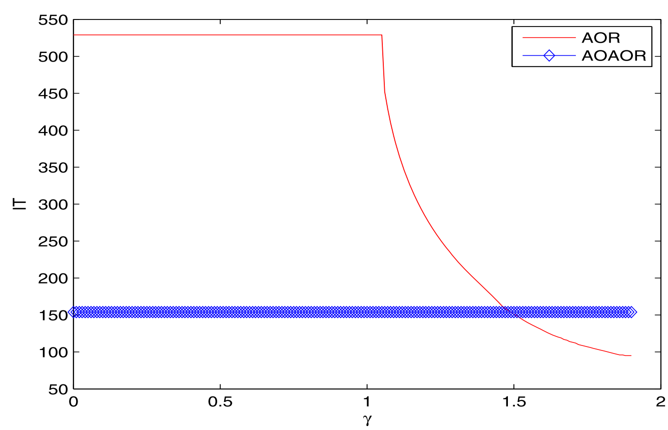

Figure 2 Curves of the IT for $\omega = 1.9$, $h^{-1} = 32$ in Table 7.

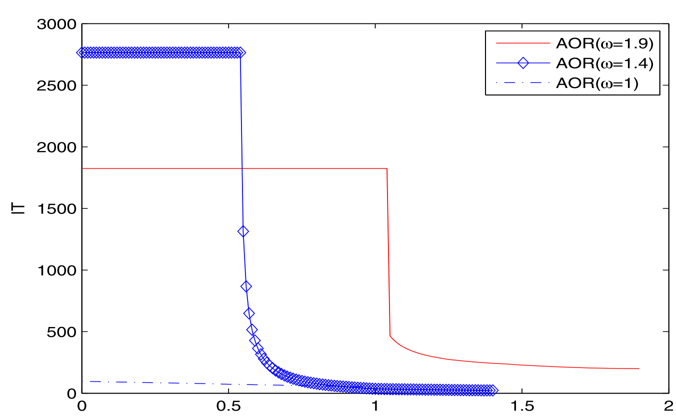

Figure 3 Curves of the IT for $\omega = 1.9$, $\omega = 1.4$, and $\omega = 1$ with $h^{-1} = 32$ in Table 8.

Consequently, we obtain the obvious conclusion that the IT of the AOR method is sensitive to the guesses of ω and γ. Furthermore, Figures 1-3 according to Tables 6-9 are given to distinctly illustrate the sensitivity of the AOR and AOAOR methods.

Figure 1 clearly depicts the variation with respect to ω. We notice that the graph of the AOR method is vertically descending, whereas that of the AOAOR method is completely flat, which shows that the IT of the AOAOR method is independent of the choice of γ. So does Figure 2.

When ω is not an exact value, the numbers of IT of the AOR method are dramatically different for invariable γ, that is, it depends on the choice of ω. However, the IT of the AOAOR method remains at 61. Therefore, the conclusion that the IT of the AOAOR method is independent of the choice of ω is illustrated in Figure 3.

Thus, from Figures 1-3 we see that both AOR and AOAOR methods are sensitive to the parameters ω and γ. For the larger intervals of the parameters ω and γ, the IT of the AOR method varies according to the curve wave; however, that of the AOAOR method is almost at the same level. This means that the AOR method is more sensitive to the initial guesses of the optimal parameters ω and γ, whereas the AOAOR method is more stable.

5 Conclusions

The asymptotically optimal accelerated overrelaxation (AOAOR) method for linear systems (1.1) has been presented. Especially, the optimal parameters are discussed. The numerical experiments have shown that the AOAOR method is efficient when the dimension of the coefficient matrix is over 2^{12}. Furthermore, the AOAOR method is more stable with respect the initial guesses of relaxation factors ω and γ.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
All authors contributed equally to the writing of this paper. All authors read and approved the final manuscript.

Author details
[1]Department of Mathematics, Taiyuan Normal University, Taiyuan 030012, Shanxi, P.R. China. [2]Higher Education Key Laboratory of Engineering and Scientific Computing, Taiyuan Normal University, Taiyuan 030012, Shanxi, P.R. China.

Acknowledgements
The authors are very much indebted to the anonymous referees for their helpful comments and suggestions, which greatly improved the original manuscript of this paper. This work is supported by grants from the NSF of Shanxi Province (201601D011004).

Received: 30 June 2016 Accepted: 4 October 2016 Published online: 10 November 2016

References

1. Hadiiidmos, A: Accelerated overrelaxation method. Math. Comput. **32**, 149-157 (1978)
2. Hadiiidmos, A: A survey of the iterative methods for the solution of linear systems by extrapolation, relaxation and other techniques. J. Comput. Appl. Math. **20**, 37-51 (1987)
3. Martins, MM: Generalized diagonal dominance in connection with the accelerated overrelaxation method. BIT Numer. Math. **22**, 73-78 (1982)
4. Hadiiidmos, A, Yeyios, A: The principle of extrapolation in connection with the accelerated overrelaxation method. Linear Algebra Appl. **30**, 115-128 (1980)
5. Bai, ZZ Chi, XB: Asymptotically optimal successive overrelaxation method for systems of linear equations. J. Comput. Math. **21**(5), 503-612 (2003)
6. Wen, RP, Meng, GY, Wang, CL: Quasi-Chebyshev accelerated iteration methods based on optimization for linear systems. Comput. Math. Appl. **66**, 934-942 (2013)
7. Meng, GY: A practical asymptotical optimal SOR method. Appl. Math. Comput. **242**, 707-715 (2014)
8. Hadiiidmos, A, Yeyios, A: On some extensions of the accelerated overrelaxation (AOR) theory. J. Math. Sci. **5**, 49-60 (1982)
9. Higham, NJ: Functions of Matrices Theory and Computation. SIAM, Philadelphia (2008)
10. Gill, PE, Golub, GH, Murray, W, Saunders, MA: Methods for modifying matrix factorizations. Math. Comput. **28**(126), 505-535 (1974)