NUMERICAL LINEAR ALGEBRA WITH APPLICATIONS
Numer. Linear Algebra Appl. 2007; **14**:319–335
Published online 9 January 2007 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/nla.517

On successive-overrelaxation acceleration of the Hermitian and skew-Hermitian splitting iterations

Zhong-Zhi Bai[1,*,†], Gene H. Golub[2] and Michael K. Ng[3]

[1]*State Key Laboratory of Scientific/Engineering Computing, Institute of Computational Mathematics and Scientific/Engineering Computing, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, P.O. Box 2719, Beijing 100080, People's Republic of China*
[2]*Scientific Computing and Computational Mathematics Program, Department of Computer Science, Stanford University, Stanford, CA 94305-9025, U.S.A.*
[3]*Department of Mathematics, Hong Kong Baptist University, Kowloon Tong, Hong Kong*

SUMMARY

We further generalize the technique for constructing the Hermitian/skew-Hermitian splitting (HSS) iteration method for solving large sparse non-Hermitian positive definite system of linear equations to the normal/skew-Hermitian (NS) splitting obtaining a class of normal/skew-Hermitian splitting (NSS) iteration methods. Theoretical analyses show that the NSS method converges unconditionally to the exact solution of the system of linear equations. Moreover, we derive an upper bound of the contraction factor of the NSS iteration which is dependent solely on the spectrum of the normal splitting matrix, and is independent of the eigenvectors of the matrices involved. We present a successive-overrelaxation (SOR) acceleration scheme for the NSS iteration, which specifically results in an acceleration scheme for the HSS iteration. Convergence conditions for this SOR scheme are derived under the assumption that the eigenvalues of the corresponding block Jacobi iteration matrix lie in certain regions in the complex plane. A numerical example is used to show that the SOR technique can significantly accelerate the convergence rate of the NSS or the HSS iteration method. Copyright © 2006 John Wiley & Sons, Ltd.

Received 17 August 2005; Revised 17 June 2006; Accepted 20 June 2006

KEY WORDS: non-Hermitian matrix; normal matrix; Hermitian matrix; skew-Hermitian matrix; splitting iteration method; successive overrelaxation

*Correspondence to: Zhong-Zhi Bai, State Key Laboratory of Scientific/Engineering Computing, Institute of Computational Mathematics and Scientific/Engineering Computing, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, P.O. Box 2719, Beijing 100080, People's Republic of China.
†E-mail: bzz@lsec.cc.ac.cn

Contract/grant sponsor: The Special Funds for Major State Basic Research Projects; contract/grant number: G1999032803
Contract/grant sponsor: The National Basic Research Program; contract/grant number: 2005CB321702
Contract/grant sponsor: The China NNSF National Outstanding Young Scientist Foundation; contract/grant number: 10525102
Contract/grant sponsor: The National Natural Science Foundation of China; contract/grant number: 10471146
Contract/grant sponsor: DOE Grant; contract/grant number: DOE-FC02-01ER4177
Contract/grant sponsor: RGC Grants; contract/grant numbers: 7046/03P, 7035/04P, 7035/05P

1. INTRODUCTION

To solve the large sparse non-Hermitian and positive definite system of linear equations

$$Ax = b, \quad A \in \mathbb{C}^{n\times n} \quad \text{non-singular} \quad \text{and} \quad x, b \in \mathbb{C}^n \tag{1}$$

we observe that the coefficient matrix A naturally possesses the *Hermitian/skew-Hermitian* (HS) splitting [1–5]

$$A = H + S \tag{2}$$

where

$$H = \tfrac{1}{2}(A + A^*) \quad \text{and} \quad S = \tfrac{1}{2}(A - A^*)$$

with A^* being the conjugate transpose of A. Bai *et al.* [6] recently proposed the use of the *Hermitian/skew-Hermitian splitting* (HSS) iteration method. Theoretical analysis has shown that this HSS iteration converges unconditionally to the exact solution of the system of linear equations (1), with the bound on convergence speed about the same as that of the conjugate gradient method when applied to Hermitian matrices. Moreover, the upper bound of the contraction factor is dependent on the spectrum of the Hermitian part H, but is independent of the spectrum of the skew-Hermitian part S as well as the eigenvalues of the matrices H, S and A. Numerical experimentation has shown that the HSS iteration method is very efficient and robust for solving the non-Hermitian and positive definite linear systems.

However, even if we have determined the optimal value of the parameter α for the upper bound of the contraction factor of the HSS iteration from the lower and the upper eigenvalue bounds of the matrix H, a main difficulty is to find a more accurate optimal parameter α^* such that the spectral radius of the iteration matrix of the HSS method is minimized and, hence, the convergence speed of the HSS iteration is optimized.

In this paper, we consider another approach to accelerate the HSS iteration. To this end, we first generalize the HS splitting to the *normal/skew-Hermitian* (NS) splitting

$$A = N + S_0$$

where $N \in \mathbb{C}^{n\times n}$ is a normal matrix and $S_0 \in \mathbb{C}^{n\times n}$ a skew-Hermitian matrix, and establish a generalized version of the HSS iteration, called as the *normal/skew-Hermitian splitting* (NSS) iteration method. Then, by equivalently transforming the original system of linear equations (1) into a double-dimensional linear system through the fixed-point equations implied in the NSS iteration, we define a block *successive-overrelaxation* (SOR) iteration [7] for this new linear system. Because the coefficient matrix of this new linear system is of block 2×2 structure and, hence, possesses block Property A (see, for instance, [8]), we can show that this block SOR iteration accelerates the NSS and, in particular, the HSS iteration, in a similar fashion to the classical block SOR theory for matrices with block Property A.

Moreover, for the NSS method, theoretical analysis shows that it converges unconditionally to the exact solution of the system of linear equations (1). The upper bound of the contraction factor of the NSS iteration is dependent on the spectrum of the normal splitting matrix N, and is independent of the spectrum of the skew-Hermitian splitting matrix S_0 as well as the eigenvectors of N, S_0 and A.

DOI: 10.1002/nla

For the SOR acceleration scheme, convergence conditions are derived under the assumption that the eigenvalues μ of the corresponding block Jacobi iteration matrix lie in certain regions in the complex plane. The optimum value of ω is determined under various assumptions on the location of the eigenvalues μ of the block Jacobi matrix.

The organization of this paper is as follows. In Section 2, we study the convergence properties and analyse the convergence rate of the NSS iteration. In Section 3, we study the SOR acceleration of the NSS iteration. In Section 4, we consider the choice of the relaxation parameter ω of the SOR acceleration of the NSS iteration. A numerical example is presented in Section 5. And finally, in Section 6 we draw a brief conclusion and include some remarks.

2. THE HSS AND THE NSS ITERATIONS

Throughout the paper, the non-Hermitian matrix $A \in \mathbb{C}^{n\times n}$ (i.e. $A \neq A^*$) is positive definite if its Hermitian part $H = \frac{1}{2}(A + A^*)$ is Hermitian positive definite.

We first review the HSS iteration method presented in Bai *et al.* [6].

The HSS iteration method: Given an initial guess $x^{(0)} \in \mathbb{C}^n$. For $k = 0, 1, 2, \ldots$ until $\{x^{(k)}\}$ converges, compute

$$(\alpha I + H)x^{(k+1/2)} = (\alpha I - S)x^{(k)} + b$$

$$(\alpha I + S)x^{(k+1)} = (\alpha I - H)x^{(k+1/2)} + b$$

where α is a given positive constant.

In matrix–vector form, the above HSS iteration method can be equivalently rewritten as

$$x^{(k+1)} = \mathscr{M}(\alpha)x^{(k)} + \mathscr{G}(\alpha)b, \quad k = 0, 1, 2, \ldots \tag{3}$$

where

$$\mathscr{M}(\alpha) = (\alpha I + S)^{-1}(\alpha I - H)(\alpha I + H)^{-1}(\alpha I - S)$$

and

$$\mathscr{G}(\alpha) = 2\alpha(\alpha I + S)^{-1}(\alpha I + H)^{-1}$$

Here, $\mathscr{M}(\alpha)$ is the iteration matrix of the HSS iteration. In fact, (3) also results from the splitting

$$A = B(\alpha) - C(\alpha)$$

of the coefficient matrix A, with

$$B(\alpha) = \frac{1}{2\alpha}(\alpha I + H)(\alpha I + S)$$

$$C(\alpha) = \frac{1}{2\alpha}(\alpha I - H)(\alpha I - S)$$

We note that

$$\mathscr{M}(\alpha) = B(\alpha)^{-1}C(\alpha) \quad \text{and} \quad \mathscr{G}(\alpha) = B(\alpha)^{-1}$$

DOI: 10.1002/nla

and $B(\alpha)$ can be served as a preconditioner, called the HSS preconditioner, to the system of linear equations (1).

The following theorem established in [6] describes the convergence property of the HSS iteration.

Theorem 2.1
Let $A \in \mathbb{C}^{n\times n}$ be a positive definite matrix, $H = \frac{1}{2}(A+A^*)$ and $S = \frac{1}{2}(A-A^*)$ be its Hermitian and skew-Hermitian parts, respectively, and α be a positive constant. Then the spectral radius $\rho(\mathscr{M}(\alpha))$ of the iteration matrix $\mathscr{M}(\alpha)$ of the HSS iteration is bounded by

$$\sigma(\alpha) = \max_{\lambda_j \in \lambda(H)} \frac{|\alpha - \lambda_j|}{|\alpha + \lambda_j|}$$

where $\lambda(H)$ is the spectral set of the matrix H. Therefore, it holds that

$$\rho(\mathscr{M}(\alpha)) \leqslant \sigma(\alpha) < 1 \quad \forall \alpha > 0$$

i.e. the HSS iteration converges to the exact solution $x^* \in \mathbb{C}^n$ of the system of linear equations (1).

Moreover, if $\gamma_{\min}$ and $\gamma_{\max}$ are the lower and the upper bounds of the eigenvalues of the matrix H, respectively, then

$$\alpha^* \equiv \arg\min_{\alpha} \left\{ \max_{\gamma_{\min} \leqslant \lambda \leqslant \gamma_{\max}} \left| \frac{\alpha - \lambda}{\alpha + \lambda} \right| \right\} = \sqrt{\gamma_{\min}\gamma_{\max}} \tag{4}$$

and

$$\sigma(\alpha^*) = \frac{\sqrt{\gamma_{\max}} - \sqrt{\gamma_{\min}}}{\sqrt{\gamma_{\max}} + \sqrt{\gamma_{\min}}} = \frac{\sqrt{\kappa(H)} - 1}{\sqrt{\kappa(H)} + 1}$$

where $\kappa(H) := \gamma_{\max}/\gamma_{\min}$ approximates the spectral condition number of H.

More generally, we split the coefficient matrix $A \in \mathbb{C}^{n\times n}$ into

$$A = N + S_0$$

where $N \in \mathbb{C}^{n\times n}$ is a normal matrix and $S_0 \in \mathbb{C}^{n\times n}$ a skew-Hermitian matrix. Briefly, we call this splitting as a *normal/skew-Hermitian* (NS) splitting of the matrix A.

We remark that, in actual applications, there are situations that a matrix may be more naturally split into its normal and skew-Hermitian parts rather than its Hermitian and skew-Hermitian parts, due to the concrete structure and property of the coefficient matrix of the linear system (1) inherited from the original physical or mathematical problem. The circulant-plus-diagonal matrix of the form $C + \imath D$ is exactly such an example, which often arises from the image processing and the numerical solution of the Helmholtz equation, etc., where $\imath$ is the imaginary unit.

Note that when the matrix A is positive definite, N must be a positive definite matrix as S_0 is skew-Hermitian and

$$N + N^* = A + A^*$$

Unlike the HS splitting, the NS splitting is not unique for a given matrix A. We note that any skew-Hermitian matrix $S_r \in \mathbb{C}^{n\times n}$ satisfying $HS_r = S_r H$, where $H = \frac{1}{2}(A + A^*)$, can result in a new NS splitting

$$A = (N + S_r) + (S_0 - S_r)$$

DOI: 10.1002/nla

As a matter of fact, the matrix $(S_0 - S_r)$ is obviously skew-Hermitian, and the matrix $(N + S_r)$ is normal.

Let us consider the following example. Suppose $H = Q^* \Lambda_h Q$ and $S = V^* \Lambda_s V$, where Q and V are unitary matrices, and Λ_h and Λ_s are diagonal matrices. If we choose the matrix $S_r = \imath c I$, where c is a real number, then it is easy to check that $H S_r = S_r H$ and $H + S_r$ is normal.

We remark that the set of normal matrices is a large class of matrices containing, for instance, the set of Hermitian, circulant and unitary matrices. We refer the readers to [9–12] for detailed discussion about normal matrices.

Similar to the HSS iteration method, we can define a so-called *normal/skew-Hermitian splitting* (NSS) iteration method as follows.

The NSS iteration method: Given an initial guess $x^{(0)} \in \mathbb{C}^n$. For $k = 0, 1, 2, \ldots$, compute

$$(\alpha I + N)x^{(k+1/2)} = (\alpha I - S_0)x^{(k)} + b$$
$$(\alpha I + S_0)x^{(k+1)} = (\alpha I - N)x^{(k+1/2)} + b$$

until $\{x^{(k)}\}$ converges, where α is a given positive constant.

Evidently, each iterate of the NSS iteration alternates between the normal matrix N and the skew-Hermitian matrix S_0, analogously to the classical alternating direction implicit (ADI) iteration for partial differential equations. In fact, we can reverse roles of the matrices N and S_0 in the above NSS iteration so that we may first solve the system of linear equations with coefficient matrix $\alpha I + S_0$ and then solve the system of linear equations with coefficient matrix $\alpha I + N$.

Note that both $\alpha I + N$ and $\alpha I + S_0$ are normal matrices. Therefore, the linear systems with the coefficient matrices $\alpha I + N$ and $\alpha I + S_0$ may be solved accurately and efficiently by some Krylov subspace iteration methods, e.g. GMRES, as now the GMRES method naturally reduces to an iterative process of three-term recurrence. See [13, 14] for other iteration methods about solving large sparse normal system of linear equations.

In matrix–vector form, the NSS iteration method can be equivalently rewritten as

$$x^{(k+1)} = \mathscr{M}_0(\alpha)x^{(k)} + \mathscr{G}_0(\alpha)b, \quad k = 0, 1, 2, \ldots \tag{5}$$

where

$$\mathscr{M}_0(\alpha) = (\alpha I + S_0)^{-1}(\alpha I - N)(\alpha I + N)^{-1}(\alpha I - S_0)$$

and

$$\mathscr{G}_0(\alpha) = 2\alpha(\alpha I + S_0)^{-1}(\alpha I + N)^{-1}$$

Here, $\mathscr{M}_0(\alpha)$ is the iteration matrix of the NSS iteration. In fact, (5) may also result from the splitting

$$A = B_0(\alpha) - C_0(\alpha)$$

of the coefficient matrix A, with

$$B_0(\alpha) = \frac{1}{2\alpha}(\alpha I + N)(\alpha I + S_0)$$

$$C_0(\alpha) = \frac{1}{2\alpha}(\alpha I - N)(\alpha I - S_0)$$

DOI: 10.1002/nla

Obviously,

$$\mathscr{M}_0(\alpha) = B_0(\alpha)^{-1} C_0(\alpha) \quad \text{and} \quad B_0(\alpha) = \mathscr{G}_0(\alpha)^{-1}$$

We should point out that $B_0(\alpha)$ can be served as a preconditioner, called the NSS preconditioner, to the system of linear equations (1).

Following the proof of Theorem 2.1 (see [6]), we can demonstrate the convergence property of the above NSS iteration.

Theorem 2.2
Let $A \in \mathbb{C}^{n\times n}$ be a positive definite matrix, $N \in \mathbb{C}^{n\times n}$ be a normal matrix and $S_0 \in \mathbb{C}^{n\times n}$ be a skew-Hermitian matrix such that $A = N + S_0$, and α be a positive constant. Then the spectral radius $\rho(\mathscr{M}_0(\alpha))$ of the iteration matrix $\mathscr{M}_0(\alpha)$ of the NSS iteration is bounded by

$$\sigma_0(\alpha) = \max_{\lambda_j \in \lambda(N)} \frac{|\alpha - \lambda_j|}{|\alpha + \lambda_j|} = \max_{\gamma_j + \imath\eta_j \in \lambda(N)} \sqrt{\frac{(\alpha - \gamma_j)^2 + \eta_j^2}{(\alpha + \gamma_j)^2 + \eta_j^2}}$$

Therefore, it holds that

$$\rho(\mathscr{M}_0(\alpha)) \leqslant \sigma_0(\alpha) < 1 \quad \forall \alpha > 0$$

i.e. the NSS iteration converges to the exact solution $x^* \in \mathbb{C}^n$ of the system of linear equations (1). Moreover, if $\gamma_{\min}$ and $\gamma_{\max}$, $\eta_{\min}$ and $\eta_{\max}$ are the lower and the upper bounds of the real, the absolute values of the imaginary parts of the eigenvalues of the matrix N, respectively, and $\Omega = [\gamma_{\min}, \gamma_{\max}] \times [\eta_{\min}, \eta_{\max}]$, then

$$\alpha^* \equiv \arg\min_{\alpha} \left\{ \max_{(\gamma,\eta)\in\Omega} \sqrt{\frac{(\alpha - \gamma)^2 + \eta^2}{(\alpha + \gamma)^2 + \eta^2}} \right\}$$

$$= \begin{cases} \sqrt{\gamma_{\min}\gamma_{\max} - \eta_{\max}^2} & \text{for } \eta_{\max} < \sqrt{\gamma_{\min}\gamma_{\max}} \\ \sqrt{\gamma_{\min}^2 + \eta_{\max}^2} & \text{for } \eta_{\max} \geqslant \sqrt{\gamma_{\min}\gamma_{\max}} \end{cases}$$

and

$$\sigma_0(\alpha^*) = \begin{cases} \left(\dfrac{\gamma_{\min} + \gamma_{\max} - 2\sqrt{\gamma_{\min}\gamma_{\max} - \eta_{\max}^2}}{\gamma_{\min} + \gamma_{\max} + 2\sqrt{\gamma_{\min}\gamma_{\max} - \eta_{\max}^2}} \right)^{1/2} & \text{for } \eta_{\max} < \sqrt{\gamma_{\min}\gamma_{\max}} \\ \left(\dfrac{\sqrt{\gamma_{\min}^2 + \eta_{\max}^2} - \gamma_{\min}}{\sqrt{\gamma_{\min}^2 + \eta_{\max}^2} + \gamma_{\min}} \right)^{1/2} & \text{for } \eta_{\max} \geqslant \sqrt{\gamma_{\min}\gamma_{\max}} \end{cases}$$

Proof
By the similarity invariance of the matrix spectrum, we have

$$\rho(\mathscr{M}_0(\alpha)) = \rho((\alpha I - N)(\alpha I + N)^{-1}(\alpha I - S_0)(\alpha I + S_0)^{-1})$$

DOI: 10.1002/nla

$$\leqslant \|(\alpha I - N)(\alpha I + N)^{-1}(\alpha I - S_0)(\alpha I + S_0)^{-1}\|_2$$

$$\leqslant \|(\alpha I - N)(\alpha I + N)^{-1}\|_2 \|(\alpha I - S_0)(\alpha I + S_0)^{-1}\|_2$$

Letting $Q_0(\alpha) = (\alpha I - S_0)(\alpha I + S_0)^{-1}$ and noting that $S_0^* = -S_0$, we see that

$$\begin{aligned} Q_0(\alpha)^* Q_0(\alpha) &= (\alpha I - S_0)^{-1}(\alpha I + S_0)(\alpha I - S_0)(\alpha I + S_0)^{-1} \\ &= (\alpha I - S_0)^{-1}(\alpha I - S_0)(\alpha I + S_0)(\alpha I + S_0)^{-1} = I \end{aligned}$$

That is to say, $Q_0(\alpha)$ is a unitary matrix ($Q_0(\alpha)$ is also called the Cayley transform of S_0). Therefore, $\|Q_0(\alpha)\|_2 = 1$. It then follows that

$$\rho(\mathscr{M}_0(\alpha)) \leqslant \|(\alpha I - N)(\alpha I + N)^{-1}\|_2$$

Because N is a normal matrix, there exist a unitary matrix $U \in \mathbb{C}^{n\times n}$ and a complex diagonal matrix $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) \in \mathbb{C}^{n\times n}$ such that $N = U^*\Lambda U$. Hence, we can further get

$$\rho(\mathscr{M}_0(\alpha)) \leqslant \max_{\lambda_j \in \lambda(N)} \left|\frac{\alpha - \lambda_j}{\alpha + \lambda_j}\right| = \max_{\lambda_j = \gamma_j + \imath\eta_j \in \lambda(N)} \left|\frac{\alpha - (\gamma_j + \imath\eta_j)}{\alpha + (\gamma_j + \imath\eta_j)}\right|$$

$$= \max_{\gamma_j + \imath\eta_j \in \lambda(N)} \sqrt{\frac{(\alpha - \gamma_j)^2 + \eta_j^2}{(\alpha + \gamma_j)^2 + \eta_j^2}}$$

Since the real parts γ_j of λ_j, $j = 1, 2, \ldots, n$, are positive and α is a positive constant, it is easy to see that $\rho(\mathscr{M}_0(\alpha)) \leqslant \sigma_0(\alpha) < 1$.

We now compute the optimal parameter α^* which minimizes the upper bound $\sigma_0(\alpha)$ of the convergence factor $\rho(\mathscr{M}_0(\alpha))$ of the NSS iteration.

Noticing that

$$\frac{(\alpha - \gamma)^2 + \eta^2}{(\alpha + \gamma)^2 + \eta^2}$$

is an increasing function with respect to the variable η, we have

$$\begin{aligned} &\max_{(\gamma,\eta)\in\Omega} \left\{\frac{(\alpha - \gamma)^2 + \eta^2}{(\alpha + \gamma)^2 + \eta^2}\right\} \\ &= \max_{\gamma_{\min} \leqslant \gamma \leqslant \gamma_{\max}} \left\{\frac{(\alpha - \gamma)^2 + \eta_{\max}^2}{(\alpha + \gamma)^2 + \eta_{\max}^2}\right\} \\ &= \begin{cases} \max\left\{\dfrac{(\alpha - \gamma_{\min})^2 + \eta_{\max}^2}{(\alpha + \gamma_{\min})^2 + \eta_{\max}^2}, \dfrac{(\alpha - \gamma_{\max})^2 + \eta_{\max}^2}{(\alpha + \gamma_{\max})^2 + \eta_{\max}^2}\right\} & \text{for } \eta_{\max} < \sqrt{\gamma_{\min}\gamma_{\max}} \\ \dfrac{(\alpha - \gamma_{\min})^2 + \eta_{\max}^2}{(\alpha + \gamma_{\min})^2 + \eta_{\max}^2} & \text{for } \eta_{\max} \geqslant \sqrt{\gamma_{\min}\gamma_{\max}} \end{cases} \end{aligned}$$

DOI: 10.1002/nla

If α^* is the minimum point of $\sigma_0(\alpha)$, then it must satisfy

$$\frac{(\alpha-\gamma_{\min})^2+\eta_{\max}^2}{(\alpha+\gamma_{\min})^2+\eta_{\max}^2}=\frac{(\alpha-\gamma_{\max})^2+\eta_{\max}^2}{(\alpha+\gamma_{\max})^2+\eta_{\max}^2}$$

when $\eta_{\max}<\sqrt{\gamma_{\min}\gamma_{\max}}$, and

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}\left(\frac{(\alpha-\gamma_{\min})^2+\eta_{\max}^2}{(\alpha+\gamma_{\min})^2+\eta_{\max}^2}\right)=0$$

when $\eta_{\max}\geqslant\sqrt{\gamma_{\min}\gamma_{\max}}$. Therefore,

$$\alpha^*=\begin{cases}\sqrt{\gamma_{\min}\gamma_{\max}-\eta_{\max}^2} & \text{for } \eta_{\max}<\sqrt{\gamma_{\min}\gamma_{\max}}\\ \sqrt{\gamma_{\min}^2+\eta_{\max}^2} & \text{for } \eta_{\max}\geqslant\sqrt{\gamma_{\min}\gamma_{\max}}\end{cases}$$

and straightforward derivations yield

$$\begin{aligned}\sigma_0(\alpha^*)&=\sqrt{\frac{(\alpha^*-\gamma_{\min})^2+\eta_{\max}^2}{(\alpha^*+\gamma_{\min})^2+\eta_{\max}^2}}\\
&=\begin{cases}\left(\dfrac{\left(\sqrt{\gamma_{\min}\gamma_{\max}-\eta_{\max}^2}-\gamma_{\min}\right)^2+\eta_{\max}^2}{\left(\sqrt{\gamma_{\min}\gamma_{\max}-\eta_{\max}^2}+\gamma_{\min}\right)^2+\eta_{\max}^2}\right)^{1/2} & \text{for } \eta_{\max}<\sqrt{\gamma_{\min}\gamma_{\max}}\\ \left(\dfrac{\left(\sqrt{\gamma_{\min}^2+\eta_{\max}^2}-\gamma_{\min}\right)^2+\eta_{\max}^2}{\left(\sqrt{\gamma_{\min}^2+\eta_{\max}^2}+\gamma_{\min}\right)^2+\eta_{\max}^2}\right)^{1/2} & \text{for } \eta_{\max}\geqslant\sqrt{\gamma_{\min}\gamma_{\max}}\end{cases}\\
&=\begin{cases}\left(\dfrac{\gamma_{\min}+\gamma_{\max}-2\sqrt{\gamma_{\min}\gamma_{\max}-\eta_{\max}^2}}{\gamma_{\min}+\gamma_{\max}+2\sqrt{\gamma_{\min}\gamma_{\max}-\eta_{\max}^2}}\right)^{1/2} & \text{for } \eta_{\max}<\sqrt{\gamma_{\min}\gamma_{\max}}\\ \left(\dfrac{\gamma_{\min}^2+\eta_{\max}^2-\gamma_{\min}\sqrt{\gamma_{\min}^2+\eta_{\max}^2}}{\gamma_{\min}^2+\eta_{\max}^2+\gamma_{\min}\sqrt{\gamma_{\min}^2+\eta_{\max}^2}}\right)^{1/2} & \text{for } \eta_{\max}\geqslant\sqrt{\gamma_{\min}\gamma_{\max}}\end{cases}\\
&=\begin{cases}\left(\dfrac{\gamma_{\min}+\gamma_{\max}-2\sqrt{\gamma_{\min}\gamma_{\max}-\eta_{\max}^2}}{\gamma_{\min}+\gamma_{\max}+2\sqrt{\gamma_{\min}\gamma_{\max}-\eta_{\max}^2}}\right)^{1/2} & \text{for } \eta_{\max}<\sqrt{\gamma_{\min}\gamma_{\max}}\\ \left(\dfrac{\sqrt{\gamma_{\min}^2+\eta_{\max}^2}-\gamma_{\min}}{\sqrt{\gamma_{\min}^2+\eta_{\max}^2}+\gamma_{\min}}\right)^{1/2} & \text{for } \eta_{\max}\geqslant\sqrt{\gamma_{\min}\gamma_{\max}}\end{cases}\end{aligned}$$

□

DOI: 10.1002/nla

Theorem 2.2 shows that the convergence speed of the NSS iteration is bounded by $\sigma_0(\alpha)$, which depends on the spectrum of the normal matrix N, but does not depend on the spectrum of the skew-Hermitian matrix S_0, and neither on the eigenvectors of the matrices N, S_0 and A. In particular, if N is Hermitian positive definite, then Theorem 2.2 recovers Theorem 2.1 (see also [6, Theorem 2.2 and Corollary 2.3]) about the convergence of the HSS iteration method.

Moreover, we can see from Theorem 2.2 that the NSS iteration method can achieve fast convergence speed when the area of $\Omega = [\gamma_{\min}, \gamma_{\max}] \times [\eta_{\min}, \eta_{\max}]$ is small, in particular, when $\gamma_{\max} \approx \gamma_{\min}$ and $\eta_{\max} \approx 0$. This shows that first preconditioning the linear system (1) with a good preconditioner P and then employing the NSS iteration to the preconditioned linear system $P^{-1}Ax = P^{-1}b$ can lead to a fast convergent iteration scheme for solving the system of linear equations (1).

Now, if we introduce a vector norm $|||x||| = \|(\alpha I + S_0)x\|_2$ $(\forall x \in \mathbb{C}^n)$ and represent the induced matrix norm by $|||X||| = \|(\alpha I + S_0)X(\alpha I + S_0)^{-1}\|_2$ $(\forall X \in \mathbb{C}^{n\times n})$, then we have

$$|||\mathscr{M}_0(\alpha)||| = \|(\alpha I - N)(\alpha I + N)^{-1}(\alpha I - S_0)(\alpha I + S_0)^{-1}\|_2 \leqslant \sigma_0(\alpha)$$

and it follows that

$$|||x^{(k+1)} - x^*||| \leqslant \sigma_0(\alpha)|||x^{(k)} - x^*|||, \quad k = 0, 1, 2, \ldots$$

where $\{x^{(k)}\}$ is the iteration sequence generated by the NSS method. Therefore, $\sigma_0(\alpha)$ is also an upper bound of the contraction factor of the NSS iteration in the sense of the $|||\cdot|||$-norm. Furthermore, if the lower and the upper bounds $\gamma_{\min}$, $\eta_{\min}$ and $\gamma_{\max}$, $\eta_{\max}$ are known, then the optimal parameter α^* for $\sigma_0(\alpha)$ (or the upper bound of $\rho(\mathscr{M}_0(\alpha))$ or $|||\mathscr{M}_0(\alpha)|||$) can be obtained. By employing this optimal parameter α^*, we particularly have

$$|||\mathscr{M}_0(\alpha^*)||| \leqslant \sigma_0(\alpha^*)$$

and

$$|||x^{(k+1)} - x^*||| \leqslant \sigma_0(\alpha^*)|||x^{(k)} - x^*|||, \quad k = 0, 1, 2, \ldots$$

in light of Theorem 2.2.

We remark that with respect to the variable $\eta_{\max}$ the function $\sigma_0(\alpha^*)$ has a discontinuity at $\eta_{\max} = \sqrt{\gamma_{\min}\gamma_{\max}}$. In addition, when $\eta_{\max} < \sqrt{\gamma_{\min}\gamma_{\max}}$, it holds that

$$\left(\frac{\gamma_{\min} + \gamma_{\max} - 2\sqrt{\gamma_{\min}\gamma_{\max} - \eta_{\max}^2}}{\gamma_{\min} + \gamma_{\max} + 2\sqrt{\gamma_{\min}\gamma_{\max} - \eta_{\max}^2}}\right)^{1/2} \geqslant \frac{\sqrt{\gamma_{\max}} - \sqrt{\gamma_{\min}}}{\sqrt{\gamma_{\max}} + \sqrt{\gamma_{\min}}}$$

and when $\eta_{\max} \geqslant \sqrt{\gamma_{\min}\gamma_{\max}}$, it holds that

$$\left(\frac{\sqrt{\gamma_{\min}^2 + \eta_{\max}^2} - \gamma_{\min}}{\sqrt{\gamma_{\min}^2 + \eta_{\max}^2} + \gamma_{\min}}\right)^{1/2} \geqslant \left(\frac{\sqrt{\gamma_{\max}} - \sqrt{\gamma_{\min}}}{\sqrt{\gamma_{\max}} + \sqrt{\gamma_{\min}}}\right)^{1/2} \geqslant \frac{\sqrt{\gamma_{\max}} - \sqrt{\gamma_{\min}}}{\sqrt{\gamma_{\max}} + \sqrt{\gamma_{\min}}}$$

By Theorems 2.1 and 2.2, the above inequality implies that among all NSS iterations for solving the non-Hermitian positive definite linear system (1), the optimal upper bound of the contraction factor of the HSS iteration is the smallest. However, since the upper bound may be much overestimated, and the actual convergence factor may be far from the corresponding upper bound of the contraction

DOI: 10.1002/nla

factor, we could not assert that the HSS iteration is the fastest one among all NSS iterations for solving the non-Hermitian positive definite linear system (1). In addition, a structured splitting of the coefficient matrix A of the linear system (1) utilizing its intrinsic property can often lead to the sub-systems of linear equations with coefficient matrices $\alpha I + N$ and $\alpha I + S_0$, which admit fast linear solvers. This may considerably save the computational cost of the NSS iteration. As a result, the NSS iteration method may be more efficient than the HSS iteration method in actual computations, provided both of them have comparable convergence rates for the referred linear system.

From the proof of Theorem 2.2, we see that when the coefficient matrix A is normal, we have $NS_0 = S_0N$ and, therefore, $\rho(\mathscr{M}_0(\alpha)) = |||\mathscr{M}_0(\alpha)||| = \sigma_0(\alpha)$. The optimal parameter α^* then minimizes all of these three quantities.

3. THE SOR ACCELERATION

Because the HSS iteration is a special case of the NSS iteration, in this section we will establish the SOR acceleration scheme and the associated theory mainly for the NSS iteration. The corresponding results for the HSS iteration are then straightforward.

From the definition of the NSS iteration we can obtain the fixed-point equations

$$\begin{aligned}(\alpha I + N)x &= (\alpha I - S_0)y + b\\ (\alpha I + S_0)y &= (\alpha I - N)x + b\end{aligned} \tag{6}$$

These two fixed-point equations have the following equivalence relationships with the original system of linear equations (1).

Theorem 3.1
If $x^* \in \mathbb{C}^n$ is the exact solution of (1), then it is the fixed points of both equations in (6). Conversely, if $x^* \in \mathbb{C}^n$ is a fixed point of either of the two equations in (6), then it is the exact solution of (1).

Proof
The result directly follows from the equalities

$$\begin{aligned}Ax^* &\equiv (N + S_0)x^*\\ &= (\alpha I + N)x^* - (\alpha I - S_0)x^*\\ &= (\alpha I + S_0)x^* - (\alpha I - N)x^* \equiv b\end{aligned}$$

□

Theorem 2.2 immediately implies the following fact.

Theorem 3.2
Let $A \in \mathbb{C}^{n\times n}$ be a positive definite matrix, and $N \in \mathbb{C}^{n\times n}$ be a normal matrix and $S_0 \in \mathbb{C}^{n\times n}$ be a skew-Hermitian matrix such that $A = N + S_0$. Then the matrix

$$A_0(\alpha) = \begin{bmatrix} \alpha I + N & -(\alpha I - S_0) \\ -(\alpha I - N) & \alpha I + S_0 \end{bmatrix} \in \mathbb{C}^{2n\times 2n}$$

is non-singular for any constant $\alpha > 0$.

DOI: 10.1002/nla

Proof
Because $N \in \mathbb{C}^{n\times n}$ is normal and $S_0 \in \mathbb{C}^{n\times n}$ is skew-Hermitian, the matrices $\alpha I + N$ and $\alpha I + S_0$ are positive definite and, hence, non-singular, for all $\alpha>0$. Moreover, it holds that

$$A_0(\alpha) = \begin{bmatrix} I & 0 \\ -(\alpha I - N)(\alpha I + N)^{-1} & I \end{bmatrix} \begin{bmatrix} \alpha I + N & -(\alpha I - S_0) \\ 0 & \mathscr{S}_0(\alpha) \end{bmatrix}$$

where

$$\mathscr{S}_0(\alpha) = (\alpha I + S_0) - (\alpha I - N)(\alpha I + N)^{-1}(\alpha I - S_0)$$

is the Schur complement of the matrix $A_0(\alpha)$.

By noticing

$$\mathscr{S}_0(\alpha) = (\alpha I + S_0)(I - \mathscr{M}_0(\alpha))$$

and

$$\rho(\mathscr{M}_0(\alpha))<1 \quad \forall \alpha>0$$

from Theorem 2.2, we know that $\mathscr{S}_0(\alpha)$ is a non-singular matrix for any constant $\alpha>0$. Therefore, $A_0(\alpha)$ is a non-singular matrix for any constant $\alpha>0$. □

By making use of Theorem 3.2, the two fixed-point equations in (6) now readily form a block 2×2 linear system, which is equivalent to the original linear system (1). This fact is precisely stated as follows.

Theorem 3.3
Let $A \in \mathbb{C}^{n\times n}$ be a positive definite matrix, $N \in \mathbb{C}^{n\times n}$ be a normal matrix and $S_0 \in \mathbb{C}^{n\times n}$ be a skew-Hermitian matrix such that $A = N + S_0$. If $x^* \in \mathbb{C}^n$ is the exact solution of (1), then $z^* = \begin{bmatrix} x^* \\ x^* \end{bmatrix} \in \mathbb{C}^{2n}$ is the exact solution of the linear system

$$A_0(\alpha)z \equiv \begin{bmatrix} \alpha I + N & -(\alpha I - S_0) \\ -(\alpha I - N) & \alpha I + S_0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} b \\ b \end{bmatrix} \equiv c \tag{7}$$

Conversely, if $z^* = \begin{bmatrix} x^* \\ y^* \end{bmatrix} \in \mathbb{C}^{2n}$ is the exact solution of the linear system (7), then it must hold $x^* = y^*$ and x^* is the exact solution of (1).

Proof
The result follows from direct verification and Theorem 3.2. □

The block Jacobi iteration for the fixed-point equations of (6), or for the block 2×2 linear system (7), is

$$x^{(k+1)} = (\alpha I + N)^{-1}[(\alpha I - S_0)y^{(k)} + b]$$

$$y^{(k+1)} = (\alpha I + S_0)^{-1}[(\alpha I - N)x^{(k)} + b]$$

DOI: 10.1002/nla

or equivalently,

$$z^{(k+1)} = \mathscr{J}(\alpha)z^{(k)} + c(\alpha) \tag{8}$$

where $z^{(k)} = \begin{bmatrix} x^{(k)} \\ y^{(k)} \end{bmatrix}$ and

$$\mathscr{J}(\alpha) = \begin{bmatrix} 0 & (\alpha I + N)^{-1}(\alpha I - S_0) \\ (\alpha I + S_0)^{-1}(\alpha I - N) & 0 \end{bmatrix} \tag{9}$$

And the block SOR iteration for the fixed-point equations of (6), or for the block 2×2 linear system (7), is

$$\begin{aligned} x^{(k+1)} &= (1-\omega)x^{(k)} + \omega(\alpha I + N)^{-1}[(\alpha I - S_0)y^{(k)} + b] \\ y^{(k+1)} &= (1-\omega)y^{(k)} + \omega(\alpha I + S_0)^{-1}[(\alpha I - N)x^{(k+1)} + b] \end{aligned}$$

with ω the relaxation parameter, or equivalently,

$$z^{(k+1)} = \mathscr{L}_\omega(\alpha)z^{(k)} + c_\omega(\alpha)$$

where

$$\mathscr{L}_\omega(\alpha) = \begin{bmatrix} (1-\omega)I & \omega(\alpha I + N)^{-1}(\alpha I - S_0) \\ \omega(1-\omega)(\alpha I + S_0)^{-1}(\alpha I - N) & (1-\omega)I + \omega^2 \mathscr{M}_0(\alpha) \end{bmatrix} \tag{10}$$

and $\mathscr{M}_0(\alpha)$ is the NSS matrix.

The choice of $\omega = 1$ in the SOR method (10) results in the block Gauss–Seidel iteration for solving the block 2×2 linear system (7). Correspondingly, $\mathscr{L}_\omega(\alpha)$ reduces to the block Gauss–Seidel matrix $\mathscr{L}_1(\alpha)$.

The relationships among the eigenvalues of the block Jacobi matrix $\mathscr{J}(\alpha)$, the block Gauss–Seidel matrix $\mathscr{L}_1(\alpha)$ and the NSS matrix $\mathscr{M}_0(\alpha)$ are straightforward, and are summarized in the following theorem.

Theorem 3.4
The following statements hold true:

(a) If μ is an eigenvalue of the block Jacobi matrix $\mathscr{J}(\alpha)$ of (9), then μ^2 is an eigenvalue of the block Gauss–Seidel matrix $\mathscr{L}_1(\alpha)$. If θ is a non-zero eigenvalue of $\mathscr{L}_1(\alpha)$ and $\mu^2 = \theta$, then μ is an eigenvalue of $\mathscr{J}(\alpha)$. Thus, the block Jacobi iteration method converges if and only if the block Gauss–Seidel iteration method converges, and if both converge, then $\rho(\mathscr{L}_1(\alpha)) = (\rho(\mathscr{J}(\alpha)))^2 < 1$.

(b) The sets of non-zero eigenvalues of the block Gauss–Seidel matrix $\mathscr{L}_1(\alpha)$ and the NSS matrix $\mathscr{M}_0(\alpha)$ are exactly the same. Thus, the block Gauss–Seidel iteration method converges if and only if the NSS iteration method converges, and if both converge, then $\rho(\mathscr{L}_1(\alpha)) = \rho(\mathscr{M}_0(\alpha)) < 1$.

As an immediate consequence of Theorems 2.2 and 3.4, we have the following corollary.

DOI: 10.1002/nla

Corollary 3.1
Let $A \in \mathbb{C}^{n\times n}$ be a positive definite matrix, $N \in \mathbb{C}^{n\times n}$ be a normal matrix and $S_0 \in \mathbb{C}^{n\times n}$ be a skew-Hermitian matrix such that $A = N + S_0$. Then the block Jacobi, the block Gauss–Seidel and the NSS matrices are convergent, and it holds that

$$\rho(\mathscr{L}_1(\alpha)) = \rho(\mathscr{M}_0(\alpha)) = (\rho(\mathscr{J}(\alpha)))^2 < 1$$

This corollary clearly shows that we can theoretically conclude that the asymptotic convergence rates of the block Gauss–Seidel iteration and the NSS iteration are the same, and are twice that of the block Jacobi iteration. However, numerically, the error vectors of these iterations may be explicitly determined by a specific initial error vector.

Because the coefficient matrix $A_0(\alpha)$ of the 2×2 linear system (7) possesses block Property A, there exists a functional relationship between the eigenvalues of the block Jacobi matrix $\mathscr{J}(\alpha)$ and the block SOR matrix $\mathscr{L}_\omega(\alpha)$.

Theorem 3.5
If $\omega \neq 0$, if λ is a non-zero eigenvalue of the matrix $\mathscr{L}_\omega(\alpha)$ of (10) and if μ satisfies

$$(\lambda + \omega - 1)^2 = \lambda\omega^2\mu^2 \tag{11}$$

then μ is an eigenvalue of the block Jacobi matrix $\mathscr{J}(\alpha)$ of (9). Conversely, if μ is an eigenvalue of $\mathscr{J}(\alpha)$ and λ satisfies (11), then λ is an eigenvalue of $\mathscr{L}_\omega(\alpha)$.

Proof
See Theorem 4.5 in [15]. □

The convergence of the block SOR iteration straightforwardly follows from Corollary 3.1 and Theorem 3.5.

Theorem 3.6
Let $A \in \mathbb{C}^{n\times n}$ be a positive definite matrix, $N \in \mathbb{C}^{n\times n}$ be a normal matrix and $S_0 \in \mathbb{C}^{n\times n}$ be a skew-Hermitian matrix such that $A = N + S_0$.

(a) When all of the eigenvalues of the block Jacobi matrix $\mathscr{J}(\alpha)$ are real, the block SOR matrix $\mathscr{L}_\omega(\alpha)$ is convergent if and only if $0<\omega<2$.
(b) When some of the eigenvalues of the block Jacobi matrix $\mathscr{J}(\alpha)$ is complex, the block SOR matrix $\mathscr{L}_\omega(\alpha)$ is convergent if for some positive number $\tau \in (0, 1)$ and each eigenvalue $\mu = \delta + \imath\beta$ of $\mathscr{J}(\alpha)$, the point (δ, β) lies in the interior of the ellipse

$$\mathscr{E}(1, \tau) := \left\{ (\delta, \beta) : \delta^2 + \frac{\beta^2}{\tau^2} = 1 \right\}$$

and ω satisfies

$$0<\omega<\frac{2}{1+\tau} \tag{12}$$

Conversely, if the block SOR matrix converges, then all eigenvalues of $\mathscr{J}(\alpha)$ lie inside $\mathscr{E}(1, \tau)$ for some $\tau \in (0, 1)$. Moreover, if some μ lies on $\mathscr{E}(1, \tau)$ and if the block SOR matrix converges, then (12) holds.

DOI: 10.1002/nla

Proof
From Corollary 3.1 we know that $\rho(\mathscr{J}(\alpha))<1$. Therefore, Theorems 2.2 and 4.1 in Chapter 6 of [8] immediately result in (a) and (b), respectively. □

The functional relationship (11) is the basis for the precise determination of the value of ω which minimizes $\rho(\mathscr{L}_\omega(\alpha))$.

4. THEORETICAL DETERMINATION OF AN OPTIMUM ω

We now consider the choice of the relaxation parameter ω which minimizes $\rho(\mathscr{L}_\omega(\alpha))$ under various assumptions on the eigenvalues μ of the block Jacobi matrix $\mathscr{J}(\alpha)$. We only summarize some results in this section. However, detailed discussions can be found in Theorem 4.7 in Varga [15], and Section 6.4, Chapter 6 in Young [8]; see also [16].

In general, if a matrix is two-cyclic consistently ordered and the spectrum $\lambda(J)$ of the associated (point or block) Jacobi iteration matrix lies in the infinite unit strip $\mathscr{S}_\infty := \{z \in \mathbb{C} \mid |\mathrm{Re}(z)|<1\}$ of the complex plane then a *real* optimal value of the SOR parameter ω_{opt} and the spectral radius $\rho(\mathscr{L}_{\omega_{\mathrm{opt}}})$ of the associated SOR iteration matrix $\mathscr{L}_{\omega_{\mathrm{opt}}}$ can be obtained *via* the formulas:

$$\omega_{\mathrm{opt}} = \frac{2}{1+\sqrt{1+b^2-a^2}}, \quad \rho(\mathscr{L}_{\omega_{\mathrm{opt}}}) = \left(\frac{a+b}{1+\sqrt{1+b^2-a^2}}\right)^2 \tag{13}$$

Here, $\mathrm{Re}(\cdot)$ and $\mathrm{Im}(\cdot)$ denote the real and the imaginary parts of the corresponding complex, respectively; and a and b are the real and the imaginary semiaxes, respectively, of the best ellipse $\mathscr{E}_{\mathrm{opt}}$ that encloses the spectrum, is symmetric with respect to the real and the imaginary axes and lies in $\mathscr{S}_\infty$ (formulas (4.14) and (4.15) in [8]). In the degenerate cases where $\lambda(J)$ is real, with $\lambda(J) \subset [-a, a]$, $a<1$, $(b=0)$, purely imaginary, with $\lambda(J) \subset [-\imath b, \imath b]$, $b>0$, $(a=0)$, or lies in the union of a real and an imaginary segment, $\lambda(J) \subset [-a, a] \cup [-\imath b, \imath b] \subset \mathscr{S}_\infty$, the application of the formula in (13) is straightforward. In case $\lambda(J)$ lies in a rectangle $\mathscr{R} \subset \mathscr{S}_\infty$ and is symmetric with respect to the real and the imaginary axes, the semiaxes of the best $\mathscr{E}_{\mathrm{opt}}$ can be found by means of the co-ordinates (α, β) of the vertex of $\mathscr{R}$ in the first quadrant (formulas (4.24)–(4.26) in [8]). If $\lambda(J)$ lies in a convex polygon $\mathscr{H} \subset \mathscr{S}_\infty$, which is symmetric with respect to both axes, the semiaxes of $\mathscr{E}_{\mathrm{opt}}$ can be found *via* the co-ordinates of the vertices of $\mathscr{H}$ in the first quadrant by applying the Young-Eidson's algorithm [17]. It is noted that the cases where two or three of the vertices of $\mathscr{H}$ are in the first quadrant are treated analytically in Young's textbook [8] (*Note:* For simply convergent SORs it suffices to determine an ellipse $\mathscr{E}$ symmetric with respect to the real and the imaginary axes and such that $\lambda(J) \subset \mathscr{E} \subset \mathscr{S}_\infty$. Then, if a and b are its real and imaginary semiaxes the formula (13) will give again the optimal SOR parameter ω_{opt} and the corresponding optimal spectral radius $\rho(\mathscr{L}_{\omega_{\mathrm{opt}}})$ of a convergent SOR.).

5. NUMERICAL EXAMPLE

In this section, a numerical example is given to illustrate the usefulness of the SOR acceleration of the HSS iteration.

DOI: 10.1002/nla

Table I. The spectral radii of different iteration matrices.

$n=64$	$qh=1$	$qh=10$	$qh=100$	$qh=1000$
$\rho(\mathscr{J}(\alpha_t))$	0.7353	0.7725	0.9128	0.9713
$\rho(\mathscr{M}(\alpha_t))$	0.5406	0.5967	0.8332	0.9429
$\rho(\mathscr{L}_{\omega_s}(\alpha_s))$	0.5086	0.4583	0.4370	0.3990
$\rho(\mathscr{L}_{\omega_t}^{(\mathrm{SOR})})$	0.6629	>1	>1	>1
$n=128$				
$\rho(\mathscr{J}(\alpha_t))$	0.7855	0.7715	0.9196	0.9740
$\rho(\mathscr{M}(\alpha_t))$	0.5973	0.5952	0.8439	0.9486
$\rho(\mathscr{L}_{\omega_s}(\alpha_s))$	0.5655	0.4628	0.4581	0.4398
$\rho(\mathscr{L}_{\omega_t}^{(\mathrm{SOR})})$	0.6667	>1	>1	>1
$n=256$				
$\rho(\mathscr{J}(\alpha_t))$	0.8483	0.7759	0.9354	0.9796
$\rho(\mathscr{M}(\alpha_t))$	0.7196	0.5969	0.8715	0.9585
$\rho(\mathscr{L}_{\omega_s}(\alpha_s))$	0.6802	0.4627	0.4516	0.4349
$\rho(\mathscr{L}_{\omega_t}^{(\mathrm{SOR})})$	0.7246	>1	>1	>1

We consider the $n \times n$ coefficient matrix

$$A = \mathrm{tridiag}(-1 + qh/2, 2, -1 - qh/2)$$

arising from the centred finite difference of the differential equation

$$-u'' + qu' = f$$

with the homogeneous boundary condition using the centred difference scheme.

We compute the spectral radii of different iteration matrices $\mathscr{L}_{\omega}^{(\mathrm{SOR})}$, $\mathscr{J}(\alpha)$, $\mathscr{M}(\alpha)$ and $\mathscr{L}_{\omega}(\alpha)$ for different α and qh. The optimal relaxation parameter ω_t is used in the block SOR iteration matrix. Here, $\mathscr{L}_{\omega}^{(\mathrm{SOR})}$ is the SOR iteration matrix for the original matrix A. The optimal SOR parameter ω_t is found computationally. In Table I, we list the spectral radii of different iteration matrices for different sizes n of the coefficient matrix A. In the table, the optimal parameter α_t:

$$\alpha_t \equiv \arg \min_{\alpha} \{\rho(\mathscr{M}(\alpha))\}$$

and the optimal parameters α_s and ω_s:

$$(\alpha_s, \omega_s) \equiv \arg \min_{(\alpha,\omega)} \{\rho(\mathscr{L}_{\omega}(\alpha))\}$$

are used.

We observe from the table that the spectral radii of the HSS iteration matrices are always less than those of the block Jacobi iteration matrices, and the spectral radii of the HSS iteration matrices are always greater than those of block SOR iteration matrices. Especially, when qh is large (the skew-Hermitian part is dominant), the spectral radii of the block SOR iteration matrices are significantly less than those of the HSS iteration matrices. In the table, we also report the spectral radii of the SOR iteration matrices for the original linear system. We find that when qh is large, the classical SOR iteration is divergent, but the HSS is convergent.

DOI: 10.1002/nla

6. CONCLUDING REMARKS

We have generalized the Hermitian/skew-Hermitian splitting iteration method to the normal/skew-Hermitian splitting iteration method, and studied the corresponding convergence as well as SOR acceleration theories. In fact, we can further relax the skew-Hermitian constraint on the splitting matrix S_0 and only assume that it is positive semidefinite. This results in a new splitting iteration method for solving the non-Hermitian positive definite system of linear equations (1). All theory established in this paper is equally applicable to this new splitting iteration method.

Our numerical results have demonstrated the effectiveness of the SOR acceleration scheme for the HSS/NSS iteration. Since the storage requirements for SOR are very small, the proposed method is competitive with the GMRES method.

Finally, we remark that the two-half steps at each NSS or SOR iterate require exact solutions with the $n \times n$ matrices $\alpha I + N$ and $\alpha I + S_0$. However, this is very costly and impractical in actual implementations. To further improve computing efficiency of the NSS or SOR iteration, we can employ inner and outer iterations to implement the algorithm. Convergence theories for the correspondingly resulted inexact NSS or SOR iterations can be established following an analogous analysis to that in [6] with slight and technical modifications. According to the practical choice of the relaxation parameter ω in the SOR acceleration scheme, the technique proposed in [18] may be adopted in actual applications.

ACKNOWLEDGEMENTS

The first author sincerely thanks professor Gene H. Golub for his generous and kind invitation for a visit of SCCM, Department of Computer Science, Stanford University, during January to April in 2002. The authors are very much indebted to the referees for their constructive and useful comments and suggestions which greatly improved the original manuscript of this paper.

Z.-Z. Bai's research is supported by The Special Funds for Major State Basic Research Projects, The National Basic Research Program, The China NNSF National Outstanding Young Scientist Foundation, and The National Natural Science Foundation of China; (G1999032803, 2005CB321702, 10525102, 10471146). G. H. Golub's research is supported, in part, by DOE-FC02-01ER4177 (DOE-FC02-01ER4177). M. K. Ng's research is supported by RGC Grants (7046/03P, 7035/04P and 7035/05P).

REFERENCES

1. Golub G, Vanderstraeten D. On the preconditioning of matrices with a dominant skew-symmetric component. *Numerical Algorithms* 2000; **25**:223–239.
2. Golub G, Wathen A. An iteration for indefinite systems and its application to the Navier–Stokes equations. *SIAM Journal on Scientific Computing* 1998; **19**:530–539.
3. Wang C-L, Bai Z-Z. Sufficient conditions for the convergent splittings of non-Hermitian positive definite matrices. *Linear Algebra and its Applications* 2001; **330**:215–218.
4. Axelsson O, Bai Z-Z, Qiu S-X. A class of nested iteration schemes for linear systems with a coefficient matrix with a dominant positive definite symmetric part. *Numerical Algorithms* 2004; **35**:351–372.
5. Bai Z-Z, Qiu S-X. Splitting-MINRES methods for linear systems with the coefficient matrix with a dominant indefinite symmetric part. *Mathematica Numerica Sinica* 2002; **24**:113–128.
6. Bai Z-Z, Golub G, Ng M. Hermitian and skew-Hermitian splitting methods for non-Hermitian positive definite linear systems. *SIAM Journal on Matrix Analysis and Applications* 2003; **24**:603–626.
7. Golub G, Van Loan C. *Matrix Computations* (3rd edn). The Johns Hopkins University Press: Baltimore, London, 1996.
8. Young D. *Iterative Solution of Large Linear Systems*. Academic Press: New York, London, 1971.
9. Elsner L, Ikramov Kh. Normal matrices: an update. *Linear Algebra and its Applications* 1998; **285**:291–303.

DOI: 10.1002/nla

10. Grone R, Johnson C, Sa E, Wolkowicz H. Normal matrices. *Linear Algebra and its Applications* 1987; **87**:213–225.
11. Huhtanen M, Nevanlinna O. Minimal decompositions and iterative methods. *Numerische Mathematik* 2000; **86**:257–281.
12. Ruhe A. Closest normal matrix finally found. *British Information Technology* 1987; **27**:585–598.
13. Huhtanen M. A Hermitian Lanczos method for normal matrices. *SIAM Journal on Matrix Analysis and Applications* 2002; **23**:1092–1108.
14. Huhtanen M. Aspects of nonnormality for iterative methods. *Linear Algebra and its Applications* 2005; **394**: 119–144.
15. Varga R. *Matrix Iterative Analysis* (2nd edn). Springer: New York, London, 2000.
16. Bai Z-Z, Golub G, Ng M. On successive-overrelaxation acceleration of the Hermitian and skew-Hermitian splitting iterations. *Technical Report No. SCCM-02-06*, Scientific Computing and Computational Mathematics Program, Department of Computer Science, Stanford University, 2002.
17. Young D, Eidson H. On the determination of the optimum relaxation factor for the SOR method when the eigenvalues of the Jacobi method are complex. *Report CNA-1*, Center for Numerical Analysis, University of Texas, Austin, Texas, 1970.
18. Bai Z-Z, Chi X-B. Asymptotically optimal successive overrelaxation methods for systems of linear equations. *Journal of Computational Mathematics* 2003; **21**:603–612.

DOI: 10.1002/nla