The alternate iterative Gauss-Seidel method for linear systems

Bingyuan Pu1,*, Xun Yuan2
1Chengdu Textile College, Department of Basic Science, Chengdu 611731, China
2Chengdu University of Information Technology, College of Applied Mathematics, Chengdu 611731, China

* Corresponding author: skypuby@163.com

Abstract. In this paper, we present an alternate iterative Gauss-Seidel method for linear systems. The spectral radius of the iteration matrix and the convergence of the proposed method are discussed. Finally, the numerical examples are provided to confirm our theoretical analysis and demonstrate the efficiency of the new method.

1. Introduction

Consider the following linear system

\[ Ax = b, \]

where \( A \) is an \( n \times n \) nonsingular square matrix, \( x \) and \( b \) are \( n \)-dimensional vectors.

A usual scheme for the solution of (1) is the splitting iterative method as

\[ Mx^{(k+1)} = Nx^{(k)} + b, k = 0, 1, \ldots, \]

where \( A = M - N \) and \( M \) is non-singular. Then (2) can also be written as

\[ x^{(k+1)} = Tx^{(k)} + c, k = 0, 1, \ldots, \]

where \( T = M^{-1}N, c = M^{-1}b \). Without loss of generality, assume that \( A = I - L - U \), where \( I \) is the identity matrix, -\( L \) and -\( U \) are the strictly lower and upper triangular parts of \( A \), respectively. Then the iteration matrix of the classical Gauss-Seidel method is

\[ T = (I - L)^{-1} U. \]

In [1], Benzi stated that preconditioning is the most critical ingredient in the development of efficient solvers for challenging problems in scientific computation and that the importance of preconditioning is destined to increase even further. But given that a general optimal purpose preconditioner is unlikely to exist, pursing the more efficient preconditioners is a great challenging subject in scientific computation. Till now, many researchers have paid great efforts to this field and done a lot of research over the past decades. The same is true in the field of solving linear systems as well as its application. See [1-13] for more details.

Transform the original system (1) into the preconditioned form:

\[ P A x = P b, k = 0, 1, \ldots, \]

where \( P \) is a preconditioner. Let \( PA = M_p - N_p \) with the \( M_p \) being nonsingular. The corresponding iterative scheme is given by

\[ x^{(k+1)} = T x^{(k)} + c, k = 0, 1, \ldots, \]

where \( T = M_p^{-1}N_p, c = M_p^{-1}Pb \).
In a simpler form, Milaszewicz [2] considered the preconditioner \( P_c = (I + C) \) to eliminate the elements of the first column below the diagonal of \( A \). Gunawardena et al. [3] considered the preconditioner \( P_s = (I + S) \) to eliminate the elements of the first upper codiagonal of \( A \). In [12], we combined these two preconditioning strategies and proposed a \((I + C + S)\) style preconditioned Gauss-Seidel method. Some convergence theorems and comparison theorems were derived therein. To put that another way, this paper introduces the alternate iterative Gauss-Seidel method to reach the better convergence performance.

The remainder of this paper is structured as follows. The alternate iterative Gauss-Seidel method is firstly introduced in Section 2. The spectral radius of the iteration matrix for this new iterative method and the theoretical analysis of convergence are also investigated in this section. In Section 3, we compare our methods with other classical Gauss-Seidel iterative methods in linear system, and some numerical examples are provided to confirm our theoretical analysis. Finally, conclusions and open questions are discussed in Section 4.

2. Alternate iterative Gauss-Seidel method

Suppose there are two different splittings of matrix \( A \): \( A = M_1 - N_1 = M_2 - N_2 \). Given the initial value of iteration \( x^{(0)} \), and then solving the systems of linear equations

\[
\begin{cases}
    x^{(k+1)} = M_1^{-1}N_1x^{(k)} + M_1^{-1}b \\
    x^{(k+1)} = M_2^{-1}N_2x^{(k+1)} + M_2^{-1}b, k = 0,1,\ldots.
\end{cases}
\]

(7)

Delete the expression \( x^{(k+1)} \) and get the following iteration

\[
x^{(k+1)} = T_c x^{(k)} + M_2^{-1}\left(N_2M_1^{-1} + I\right)b, k = 0,1,\ldots,
\]

(8)

where \( T_c = M_2^{-1}N_2M_1^{-1}N_1 \) is the iterative matrix of the alternate iterative method.

Assume \( A = I - L - U \), and take \( P_c = (I + C) \) as the preconditioner, then we can get the following splitting of matrix \( A_c = (I + C)A \):

\[
A_c = (I + C)(I - L - U) = (I - L - CL - (U - C + CU),
\]

and the corresponding Gauss-Seidel iterative matrix \( T_c = (I - L - CL)^{-1}(U - C + CU) \).

Similarly, while taking \( P_s = (I + S) \) as the preconditioner, we can get the following splitting of matrix \( A_s = (I + S)A \):

\[
A_s = (I - U - SU) - (L - S + SL),
\]

and the corresponding Gauss-Seidel iterative matrix \( T_s = (I - U - SU)^{-1}(L - S + SL) \).

Thus, an alternate iterative method is developed and the corresponding Gauss-Seidel iterative matrix is \( T_d = T_cT_s \).

In the next part of this effort, we will discuss the convergence property of the proposed method. Here are some essential definitions and preliminaries. We call a vector and the corresponding Gauss-Seidel iterative matrix \( x \in \mathbb{R}^n \) positive (nonnegative) and write \( x > 0(x \geq 0) \) if its all entries \( x_i \) are positive (nonnegative). Similarly, a matrix \( A \in \mathbb{R}^{n \times n} \) is called positive (nonnegative) and write \( A > 0(A \geq 0) \) if all its entries \( a_{ij} \) are positive (nonnegative) for all \( i, j \). Let \( B = (b_{ij}) \in \mathbb{R}^{n \times n} \), then we write \( A > B \) if \( a_{ij} > b_{ij} \) for all \( i, j \).

A matrix \( A \) is a \( Z \)-matrix if \( a_{ij} \leq 0 \), for all \( i, j = 1,\ldots,n \), such that \( i \neq j \). A matrix \( A \) is called an \( M \)-matrix if \( A = sI - B, B \geq 0 \) and \( s > \rho(B) \), where \( \rho(B) \) denotes the spectral radius of \( B \).
A = M - N is said to be a splitting of A if M is nonsingular. A splitting A = M - N is said to be convergent if the iteration matrix M⁻¹N is convergent \( i.e., \rho(M^{-1}N) < 1 \), regular if \( M^{-1} \geq 0 \) and \( N \geq 0 \), weak regular if \( M^{-1} \geq 0 \) and \( M^{-1} \geq 0 \) and M-splitting if M is a nonsingular M-matrix and \( N \geq 0 \).

**Lemma 2.1** [13] Let A be a Z-matrix, then the following statements are equivalent:
1. A is a nonsingular M-matrix.
2. There exists a vector \( x > 0 \) such that \( Ax > 0 \).

**Lemma 2.2** [14] Let A be a nonnegative matrix. If \( Ax \leq \alpha \cdot x \) for some positive vector \( x \) and positive real number \( \alpha \), then \( \rho(A) \leq \alpha \). Moreover, \( \rho(A) < \alpha \) if \( Ax < \alpha x \).

**Lemma 2.3** [15] If the square matrix A and \( I - T \) are nonsingular, then there exists a only pair of matrix B and C such that \( B = A(I - T)^{-1} \) is non-singular, where \( C = B - A \) and \( T = B^{-1}C \).

**Lemma 2.4** [15] Let a square matrix \( A \) be nonsingular and \( 1 \leq \alpha \leq 0 \). If the splitting \( A = M - N = P - Q \) is weakly regular, then the only splitting \( A = B - C \) is also weakly regular, which is derived from \( T = P^{-1}QM^{-1}N \) by Lemma 2.3.

**Lemma 2.5** [16] Let \( A = M_1 - N_1 = M_2 - N_2 \) be the different weakly regular splittings and \( M_1^{-1} \geq M_2^{-1} \). Then it holds that
\[
0 \leq \rho(M_1^{-1}N_1) \leq \rho(M_2^{-1}N_2) < 1. \tag{9}
\]

Next, we will give corresponding convergence and comparison results of the presented method.

Let
\[
A_\alpha = (I + C_\alpha)A
\]
the splitting and iterative matrix of \( A_\alpha \) be
\[
A_\alpha = (I - L - C_\alpha L)(U - C_\alpha + C_\alpha U),
\]
\[
T_\alpha = (I - L - C_\alpha L)^{-1}(U - C_\alpha + C_\alpha U).
\]

**Theorem 2.1** Let A be a nonsingular M-matrix and \( \alpha_i \in (0,1], i = 1,2,n-1 \), then when \( 0 < C_\alpha \leq C_\alpha^2 \), it holds that
\[
\rho(T_{\alpha_i}^2) \leq \rho(T_{\alpha_i}^1) < 1. \tag{10}
\]

Besides, this theorem also holds in the similar preconditioning case of \( S_\alpha \).

**Proof.** Assume \( A_c = (I - L - CL)(U - C + CU) \) and let \( E_c = I - L - CL \) and \( -F_c = U - C + CU \) be the lower triangular matrix and strictly upper triangular matrix of \( A_c \), respectively. Then \( A_c = E_c - F_c \) and the corresponding Gauss-Seidel iterative matrix \( T_c = E_c^{-1}F_c \).

Because A is an M-matrix, so we can get \( E_c^{-1} \geq 0 \). Since \( I + C \geq 0, M_c^{-1} = E_c^{-1}(I + C) \geq 0 \) and \( N_c = (I + C)F_c^{-1} \geq 0 \), and thereby \( A = M_c - N_c \) is regular. In addition, from the assumption \( 0 < C_\alpha \leq C_\alpha^2 \), it holds that \( E_c^1 = I - L - C_\alpha L \geq I - L - C_\alpha^2 L = E_c^2 \). Thus, \( (M_c^1)^{-1} \leq (M_c^2)^{-1} \). From \( T_c = (M_c^1)^{-1}N_c \), we can get \( \rho(T_{\alpha_i}^2) \leq \rho(T_{\alpha_i}^1) < 1 \). The situation of \( S_\alpha \) can be deduced in the same way.
Theorem 2.2 Let $A$ be a nonsingular $M$-matrix and $(\overline{L}_k, \overline{U}_k, E_k)$ is multisplitting matrix set, where $E_k$ defined as Theorem 2.1 is positive diagonal matrix, $\overline{L}_k \geq 0$ is the strictly lower triangular matrix of $A$, $\overline{U}_k$ is the strictly upper triangular matrix of $A$ and $A = I - \overline{U}_k - \overline{L}_k, k = 1, 2, \cdots, m$. Let $(E_a)_k = I - \overline{L}_k - C_a \overline{L}_k$, $(F_a)_k = \overline{U}_k - C_a + C_a \overline{U}_k$. Then for $\alpha_i \in (0, 1], i = 1, 2, \cdots, n - 1$, the alternate multisplitting iterative matrix $\overline{T}_a = \sum_{k=1}^{m} E_k (E_a)_k^{-1} (F_a)_k$ satisfies $\rho(\overline{T}_a) < 1$.

Proof. Similar to Theorem 2.1, it’s easy to prove that $(E_a)_k, (E_a)_k^{-1}$ and $(F_a)_k$ are all nonnegative. Then $\overline{T}_a \geq 0$, and

$$\overline{T}_a = \sum_{k=1}^{m} E_k (E_a)_k^{-1} (F_a)_k$$

$$= I - \sum_{k=1}^{m} E_k (E_a)_k^{-1} (I + C_a) A.$$

Since $A$ is a nonsingular $M$-matrix, it follows easily from Lemma 2.1 that there exists a positive vector $x \in \mathbb{R}^n$, such that $Ax > 0$.

Then, $\overline{T}_a x = x - \sum_{k=1}^{m} E_k (E_a)_k^{-1} (I + C_a) A x > 0$, and

$$\overline{T}_a x < x.$$

It follows from Lemma 2.2 that $\rho(\overline{T}_a) < 1$.

This theorem is proved. □

Theorem 2.3 Let $A = M_1 - N_1 = M_2 - N_2$ be weakly regular splitting of matrix $A$. Then the corresponding alternate iterative matrix $T = M_1^{-1} N_1 M_2^{-1} N_2$ satisfies

$$\rho(T) \leq \min \left( \rho \left( M_1^{-1} N_1 \right), \rho \left( M_2^{-1} N_2 \right) \right) < 1.$$  \hspace{1cm} (11)

Proof. From the assumption that $A = M_1 - N_1 = M_2 - N_2$ is a weakly regular splitting and by Lemma 2.4, we see the splitting $A = B - C$ is also a weakly regular one. Besides, $B^{-1} = M_1^{-1} \left( N_1 M_2^{-1} + I \right) = M_1^{-1} + M_1^{-1} N_1 M_2 \geq M_1^{-1}$ and

$$B^{-1} = M_2^{-1} + M_2^{-1} N_2 M_1 \geq M_2^{-1}.$$  

Then by Lemma 2.5, we have

$$\rho(T) = \rho \left( B^{-1} C \right) \leq \rho \left( M_2^{-1} N_2 \right) < 1$$

and

$$\rho(T) \leq \rho \left( M_2^{-1} N_2 \right) < 1.$$

Thus, we have completed the proof of the theorem.

3. Numerical examples

In this section, we will report the numerical experiments to examine the efficiency of our new modified Gauss-Seidel iterative method in linear system. The following examples are conducted to compare the different preconditioned Gauss-Seidel methods. All the numerical results are obtained with a MATLAB R2012a implementation on Windows 10 with 2.5GHz i5 processor and 4GB memory.
The initial approximation $x^{(0)}$ is taken as a zero vector and all the iterations are terminated when \[ \|x^{(k+1)} - x^{(k)}\|/\|x^{(k)}\| < 10^{-6}. \] And for convenience, we use “$n$” to denote the matrix size, “$\rho$” to the spectral radius, “$ite$” to the number of iteration, and “$CPU$” to the CPU time in seconds.

**Example 1.** This example only focuses on the spectral radii of the iterative matrices in different Gauss-Seidel methods to compare the rate of convergence. The experimental $M$-matrices are generated randomly. Numerical results are presented in Table 1.

| $n$ | $\rho(T)$ | $\rho(T_c)$ | $\rho(T_s)$ | $\rho(T_{cs})$ | $\rho(T_r)$ |
|-----|------------|-------------|-------------|----------------|-------------|
| 50  | 0.7193     | 0.5465      | 0.6611      | 0.4732         | 0.5341      |
| 100 | 0.7369     | 0.5862      | 0.5909      | 0.4844         | 0.5472      |
| 150 | 0.7176     | 0.6421      | 0.6079      | 0.5348         | 0.5765      |
| 200 | 0.7183     | 0.5985      | 0.6256      | 0.5492         | 0.5893      |

It can be seen from Table 1 that our proposed alternate iterative method and (I+C+S)-style preconditioned method are superior to the other three methods. For instance, when $n=100$, comparing with the results by $T$, $T_c$ and $T_s$, the spectral radius of $T_{cs}$ has been reduced by 34.27%, 17.37% and 18.02%, respectively, and that by $T_r$ has been reduced by 25.74%, 6.65% and 7.40%, respectively.

**Example 2.** Consider the linear system: $Ax = b$, where

$$ A = \begin{pmatrix} 1 & q & r & s & q & \cdots \\ s & 1 & q & r & s & \cdots \\ r & s & \cdots & \cdots & s \\ q & \cdots & 1 & q & r \\ s & \cdots & r & s & 1 \\ \cdots & s & q & r & s & 1 \end{pmatrix} $$

Here, $q = -p/n, r = -p/(n+1), s = -p/(n+2)$ and $b$ is chosen so that $x = (1,1,\cdots,1)^T$ is the solution of the system. The numerical results are reported in Table 2.

It’s easy to see from Table 2 that our proposed modified Gauss-Seidel methods are more effective than the classic one and the method preconditioned by $(I+C)$ or $(I+S)$, in terms of spectral radius, iteration counts and CPU time, regardless of the matrix size. For example, when $n=20$, the iteration counts for $T_{cs}$ and $T_r$ have been reduced by 13.64% and 12.82%, respectively, comparing with the classic Gauss-Seidel method. It can also be seen that the same phenomenon goes for spectral radius and CPU time. But at the same time, we also have another sign that there is no distinctive relation between the $(I+C)$-style method and $(I+S)$-style method.

| $n$ | $\rho(T)$ | $\rho(T_c)$ | $\rho(T_s)$ | $\rho(T_{cs})$ | $\rho(T_r)$ |
|-----|------------|-------------|-------------|----------------|-------------|
| 10  | 0.6864     | 0.6582      | 0.6305      | 0.6013         | 0.6275      |
| 20  | 0.8239     | 0.8156      | 0.8070      | 0.7985         | 0.7985      |

| $n$ | CPU          |
|-----|--------------|
| 10  | 1.6996e-004  |
| 20  | 3.5716e-004  |
The results of the two examples above show that our proposed methods can improve the speed of the convergence process greatly compared with the classical Gauss-Seidel methods. But the alternate iterative Gauss-Seidel method has not obvious advantages over the \((I + C + S)\)-style one and this is likely related to the specific of examples. Next, we will choose two examples with very strong practical background to further illustrate the advantages of our proposed methods and focus on the comparison between our two modified Gauss-Seidel methods.

**Example 3.** In this example, we consider the linear system \(T_x = b\), which is discretized from Weiner-Hopf equation.

Where \(T = \left[ \frac{1}{\alpha_j i_j} \right]_{i,j=1}^{n} \), Numerical results are listed in Table 3, and the distribution of eigenvalues of iterative matrix is shown in Figure 1.

It can be clearly seen from Table 3 that our proposed two methods perform better than the other three methods. Furthermore, the alternate iterative method is the optimal one in all the listed methods for the Weiner-Hopf problem. When \(n = 200\), \(\rho(T)\) is only 63.37\%, 63.14\%, 32.79\%, 31.25\% of \(\rho(T), \rho(T_C), \rho(T_S), \rho(T_{CS})\), respectively. This is similar to the situation in iteration and CPU time.

| \(\rho\) | \(n=40\) | \(n=80\) | \(n=160\) |
|---|---|---|---|
| \(\rho\) | 0.9063 | 0.9040 | 0.9017 |
| ite | 165 | 161 | 157 |
| CPU | 7.5620e-004 | 7.6728e-004 | 6.2367e-004 |

Table 3. The comparissons of numerical results for Example 3.
Figure 1. The distribution of eigenvalues of iterative matrix for Example 3.

Figure 1 shows that the eigenvalues of the iterative matrix of our proposed methods are more focused than the other ones. Especially in this respect the alternate iterative method is more prominent, the eigenvalues become increasingly concentrated towards the coordinate origin. This in turn gives proof the advantages of our alternate iterative method compared with the other classical Gauss-Seidel methods.

Example 4. Consider the following integral equation [5]:

$$\sigma x(s) + \int_0^1 k(s,t)x(t) \, dt = 1, \ 0 \leq s \leq 1.$$ 

Discretize it and get the Toeplitz linear system $T_x = b$. Where $\sigma = 0.005$. 

![Figure 1(a)](image1.png)

![Figure 1(b)](image2.png)

![Figure 1(c)](image3.png)

![Figure 1(d)](image4.png)

![Figure 1(e)](image5.png)
\[ T_{n,i} = \sigma \delta_{i,j} + \frac{1}{k} \frac{1}{k(s,t)} \delta_{i,j} \delta_{i,j} = \frac{1}{1+|00(i+j)|} \]

\( \delta_{i,j} \) is Kronecker symbol, \( \delta_{i,j} = 1 \) when \( i = j \) and otherwise \( \delta_{i,j} = 0 \).

In this example, we only focus on the spectral radii of our proposed two methods to compare the convergence speed. Numerical results are reported in Table 4.

**Table 4. Spectral radii of the iterative matrices for Example 4.**

| n     | 20   | 50   | 100  | 200  | 500  | 1000 |
|-------|------|------|------|------|------|------|
| \( \rho(T_{CS}) \) | 0.7463 | 0.8298 | 0.8911 | 0.9253 | 0.9469 | 0.9542 |
| \( \rho(T_{R}) \)  | 0.5682 | 0.6932 | 0.7958 | 0.8568 | 0.8968 | 0.9106 |

It’s easy to see from Table 4 that the alternate iterative Gauss-Seidel method is superior to the \( (I+C+S) \)-style Gauss-Seidel method. For instance, when \( n=100 \), \( \rho(T_{R}) \) is only 89.30% of \( \rho(T_{CS}) \).

**4. Conclusions**

In this paper, we have proposed an alternate iterative Gauss-Seidel method for solving linear systems. The convergence of the proposed method and some comparison theories are discussed in details. By conducting numerical experiments, we have presented convincing numerical results to demonstrate the superior performance of our proposed method. However, there are some open questions for further discussions. Which is better between \( (I+C) \) and \( (I+S) \), or \( (I+C+S) \) and \( (I+R+S) \)? There is still short of further information and that leaves a lot of room for fact-checking. In addition, can we consider the overlays of different preconditioning? In this light, the development of re-preconditioning is a subject for future study.

**5. References**

[1] M. Benzi, Preconditioning Techniques for Large Linear Systems: A Survey, J. Comput. Phys., 182(2002)418-477.
[2] J.P. Milaszewicz, Improving Jacobi and Gauss-Seidel iterations, Linear Algebra Appl. 93(1987)161-170.
[3] A.D. Gunawardena, S.K. Jain, L. Snyder, Modified methods for consistent linear systems, Linear Algebra Appl. 154-156(1991)123-143.
[4] X.Z. Wang, T.Z. Huang and Y.D. Fu, Comparison results on preconditioned SOR-type iterative method for Z-matrices Linear Systems, J. Comput. Appl. Math. 206(2007)726-732.
[5] X.Q. Jin, F.R. Lin, Block preconditioners with circulant blocks for general linear systems, Comput. Math. Appl., 58(2009)1309-1319.
[6] A.Y. Özban, On Accelerated Iterative Methods for the Solution of Systems of Linear Equations, Int. J. Computer Math., 79(2002)765-773.
[7] R.P. Wen, X.H. Yan, C.L. Wang, On the convergence of a new splitting iterative method for non-Hermitian positive definite linear systems, Appl. Math. Comput.248(2014)118-130.
[8] R. Li, Y. Saad, GPU-accelerated preconditioned iterative linear solvers, J. Supercomput., 63(2013)443-466.
[9] C.S. Liu, A two-side equilibration method to reduce the condition number of an ill-posed linear system, Comput. Model. Eng. Sci., 91(2013)17-42.
[10] A.N. Langville, C.D. Meyer, Google’s PageRank and Beyon: The Science of Search Engine Rankings, Princeton University Press, Princeton, NJ, 2006.
[11] B.Y. Pu, T.Z. Huang, C. Wen, A preconditioned and extrapolation-accelerated GMRES method for PageRank, Appl.Math. Lett., 37(2014)95-100.
[12] B.Y. Pu, C. Wen, A new preconditioned Gauss-Seidel method for linear systems, Int. J. Appl. Math. Stat., 53(2015)65-72.
[13] D.M. Young, Iterative Solution of Large Linear Systems, Academic Press, Elsevier, 2014.
[14] D.B. Szyld, M.T. Jones, Two-stage and multisplitting methods for the parallel solution of linear systems, SIAM J. Matrix Anal. Appl., 13(1992)671-679.
[15] M Benzi, D.B. Szyld, Existence and uniqueness of splittings for stationary iterative methods with applications to alternating methods, Numerische Mathematik, 76(1997)309-321.
[16] R.S. Varga, Matrix Iterative Analysis, Springer Science & Business, Berlin, 2009.

Acknowledgments
This research was supported by Key Fun Project of Sichuan Provincial Department of Education(17za0003).