

Modified Iterative Methods for Consistent Linear Systems

Ananda D. Gunawardena

Department of Mathematics and Computer Science

University of Houston—Downtown

Houston, Texas 77002

and

S. K. Jain and Larry Snyder

Department of Mathematics

Ohio University

Athens, Ohio 45701

Submitted by Michael Neumann

ABSTRACT

In order to solve a linear system $Ax = b$, certain elementary row operations are performed on A before applying the Gauss-Seidel or Jacobi iterative methods. It is shown that when A is a nonsingular M -matrix or a singular tridiagonal M -matrix, the modified method yields considerable improvement in the rate of convergence for the iterative method. It is also shown that in some cases this method is superior to certain other modified iterative methods. The performance of this modified method on some matrices other than M -matrices is also investigated.

1. INTRODUCTION

Given a linear system $Ax = b$, it is often impractical to employ direct methods to obtain a solution when A is large and sparse. The use of iterative methods to solve large sparse systems is certainly not new, and in fact they date back to Gauss (1823). Iterative methods generate a sequence of approximate solutions $\{x^{(k)}\}$ to a linear system.

The modified iterative methods that we present in this paper were introduced by Mokari-Bolhassan and Trick [3]. At most $n - 1$ elementary row operations are applied to the system $Ax = b$ to obtain an equivalent system $\bar{A}x = \bar{b}$, where \bar{A} is a matrix with its first upper codiagonal zero. Then standard iterative methods are applied to the modified system $\bar{A}x = \bar{b}$.

Many other authors have studied various methods to accelerate the convergence of iterative methods. In particular, Milaszewicz [2] suggests that

LINEAR ALGEBRA AND ITS APPLICATIONS

©Elsevier Science Publishing Co., Inc., 1991
655 Avenue of the Americas, New York, NY 10010

0024-3795/91/\$3.50

if the original iteration matrix T is nonnegative and irreducible, then performing Gaussian elimination on a selected column of T to make it zero will improve the convergence of the iteration matrix. We compare our method with Milaszewicz's method by numerical examples with random value entries.

2. DEFINITIONS, NOTATION, AND KNOWN RESULTS

For any $n \times n$ matrix A , the *directed graph* $\Gamma(A)$ of A is defined to be the pair (V, E) where $V = \{1, 2, \dots, n\}$ is the set of *vertices* and $E = \{(i, j) : a_{ij} \neq 0, 1 \leq i, j \leq n\}$ is the set of *edges*. A *path* from i_1 to i_p is an ordered tuple of vertices (i_1, i_2, \dots, i_p) such that for each $k, (i_k, i_{k+1}) \in E$. A directed graph is said to be *strongly connected* if for each pair (i, j) of vertices, there is a path from i to j . The reflexive transitive closure of the graph $\Gamma(A)$ is a graph denoted by $\overline{\Gamma(A)}$. It is the smallest reflexive and transitive relation which includes the relation $\Gamma(A)$. The matrix A is said to be *irreducible* if $\Gamma(A)$ is strongly connected. A matrix A is called a *Z-matrix* if $a_{ij} \leq 0$ for all i, j such that $i \neq j$. A Z-matrix such that each column sum is equal to zero is called a *Q-matrix*. Any matrix A can be split in the form $A = D - L - U$ where D is a diagonal matrix, $-L$ is strictly lower triangular, and $-U$ is strictly upper triangular. We shall refer to this splitting as the *usual splitting* of A . A matrix A is said to be *nonnegative* if each entry of A is nonnegative, and is said to be a *positive matrix* if each entry is positive. We shall denote this by $A \geq 0$ and $A > 0$, respectively. The *spectral radius* of A is denoted by $\rho(A)$.

THEOREM 2.1 (Perron-Frobenius).

- (a) If A is a positive matrix, then $\rho(A)$ is a simple eigenvalue of A .
- (b) If A is nonnegative and irreducible, then $\rho(A)$ is a simple eigenvalue of A . Furthermore, any eigenvalue with the same modulus as $\rho(A)$ is also simple, and A has a positive eigenvector x corresponding to the eigenvalue $\rho(A)$. Any other positive eigenvector of A is a multiple of x .

THEOREM 2.2. Let A be a nonnegative matrix. Then:

- (a) If $\alpha x \leq Ax$ for some nonnegative vector $x, x \neq 0$, then $\alpha \leq \rho(A)$.
- (b) If $Ax \leq \beta x$ for some positive vector x , then $\rho(A) \leq \beta$. Moreover, if A is irreducible and if

$$0 \neq ax \leq Ax \leq \beta x$$

for some nonnegative vector x , then $\alpha < \rho(A) \leq \beta$ and x is a positive vector.

CONSISTENT LINEAR SYSTEMS

3

THEOREM 2.3. *Let A be irreducible. If S is the maximum row sum of A and s is the minimum row sum of A , then $s \leq \rho(A) \leq S$.*

THEOREM 2.4. *Let $A = I - L - U$ be a Z -matrix with the usual splitting. Let $T = (I - L)^{-1}U$ and $T_j = L + U$ be the Gauss-Seidel and Jacobi iteration matrices of A , respectively. Then exactly one of the following holds:*

- (a) $\rho(T) = \rho(T_j) = 0$,
- (b) $0 < \rho(T) < \rho(T_j) < 1$,
- (c) $1 = \rho(T) = \rho(T_j)$,
- (d) $1 < \rho(T_j) < \rho(T)$.

3. PRELIMINARY RESULTS

Although it is straightforward, we include the proof of the following for the sake of completeness.

LEMMA 3.1. *Let $A = I - L - U$ be a Q -matrix with the usual splitting. Suppose that T is the Gauss-Seidel iteration matrix defined by $T = (I - L)^{-1}U$ or the Jacobi iteration matrix defined by $T_j = L + U$. Then $\rho(T) = 1$.*

Proof. Let $e = (1, 1, \dots, 1)^T$. Then $e^T A = 0$, since A is a Q -matrix. This implies $e^T(I - L) = e^T U$ and therefore

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

This implies $\rho([U(I - L)^{-1}]^T) = 1$ by Theorem 2.2, and hence $\rho(U(I - L)^{-1}) = 1$. Since $\text{spectrum}(U(I - L)^{-1}) = \text{spectrum}((I - L)^{-1}U)$, we have $\rho((I - L)^{-1}U) = 1$. Thus $\rho(T) = 1$.

It is easy to see that $e^T(L + U) = e^T$, since A is a Q -matrix. Hence by Theorem 2.2, $\rho(L^T + U^T) = 1$ and therefore $\rho(L + U) = 1$. ■

REMARK 3.2. The usual splitting of A is $A = D - L - U$, where D , $-L$, and $-U$ are the diagonal, strictly lower, and strictly upper triangular parts of A . With the assumption that $a_{ii} \neq 0$ for all i , let us consider the new matrix $\hat{A} = D^{-1}A = I - D^{-1}L - D^{-1}U$. Clearly the system $\hat{A}x = \hat{b} = D^{-1}b$ is equivalent to the system $Ax = b$. Without loss of generality we may assume A has the splitting of the form $A = I - L - U$ when $a_{ii} \neq 0$.

4

A. D. GUNAWARDENA, S. K. JAIN, AND L. SNYDER

REMARK 3.3. Let $A = [a_{ij}]$ be a matrix with the splitting $A = I - L - U$. Then multiplication of A by $I + S$, where

$$S = \begin{pmatrix} 0 & -a_{12} & 0 & \cdots & 0 \\ 0 & 0 & -a_{23} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -a_{n-1n} \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix},$$

transforms the first upper codiagonal to zero. Let

$$\bar{A} = (I + S)A = I - L - SL - (U - S + SU).$$

It is easy to see that the strictly upper triangular part of \bar{A} (that is, $-U + S - SU$) has its first upper codiagonal zero.

Whenever $a_{i+1,i}a_{i,i+1} \neq 1$ for $i = 1, 2, \dots, n-1$, $(I - SL - L)^{-1}$ exists and hence it is possible to define the Gauss-Seidel iteration matrix for \bar{A} , namely

$$\bar{T} = (I - SL - L)^{-1}(U - S + SU).$$

If A is tridiagonal, then $\bar{T} = (I - SL - L)^{-1}U^2$.

We shall call \bar{T} the *modified Gauss-Seidel iteration matrix*.

REMARK 3.4. The first column of the standard Gauss-Seidel iteration matrix $T = (I - L)^{-1}U$ is zero, whereas the first two columns of the modified Gauss-Seidel matrix

$$\bar{T} = (I - L - SL)^{-1}(U - S - SU)$$

are zero. Thus we may partition T and \bar{T} so that

$$T = \begin{pmatrix} 0 & T_0 \\ 0 & T_1 \end{pmatrix} \quad \text{and} \quad \bar{T} = \begin{pmatrix} 0 & \bar{T}_0 \\ 0 & \bar{T}_1 \end{pmatrix},$$

where T_1 and \bar{T}_1 are $(n-1) \times (n-1)$ and $(n-2) \times (n-2)$ matrices respectively.

CONSISTENT LINEAR SYSTEMS

5

In what follows, when $A = (a_{ij})$ is a Z-matrix we shall assume that $a_{ii} = 1$ for all i .

The following lemma may be obtained as a special case of Lemma 3.4 in [5].

LEMMA 3.5. *Let A be a Z-matrix such that $0 < a_{ii+1}a_{i+1i} < 1$ for $i = 1, 2, \dots, n-1$. Then T_1 and \bar{T}_1 are irreducible matrices (where T_1 and \bar{T}_1 are defined as in Remark 3.4).*

Proof. We will show that T_1 is irreducible by showing that $(2, 3, \dots, n, 2)$ is a cycle in $\Gamma(T)$; consequently it is also a cycle in $\Gamma(T_1)$, so $\Gamma(T_1)$ is strongly connected. Since L is a nonnegative nilpotent matrix, $\rho(L) = 0$. Thus

$$(I - L)^{-1} = I + L + L^2 + L^3 + \dots + L^{n-1}.$$

Therefore $T = (I + L + L^2 + L^3 + \dots + L^{n-1})U$ and

$$\Gamma(T) = \overline{\Gamma(L)} \Gamma(U).$$

The condition $0 < a_{ii+1}a_{i+1i}$ implies that $(1, 2, \dots, n)$ is a path in $\Gamma(U)$ and $(n, n-1, \dots, 1)$ is a path in $\Gamma(L)$. Also $\Gamma(U) \subseteq \Gamma(T)$, since

$$\Gamma(T) = \Gamma(U) \cup \Gamma(LU) \cup \Gamma(L^2U) \cup \dots \cup \Gamma(L^{n-1}U).$$

In particular, $(2, 3, \dots, n)$ is a path in $\Gamma(T_1)$, so we only need to show that $(n, 2)$ is an edge in $\Gamma(T)$. For this note that $(n, 1)$ is an edge in $\overline{\Gamma(L)}$ and $(1, 2)$ is an edge in $\Gamma(U)$. Consequently $(n, 2)$ is an edge in $\Gamma(T)$.

Similarly it can be shown that \bar{T}_1 is irreducible. ■

LEMMA 3.6. *Let A be a Z-matrix such that $0 < a_{ii+1}a_{i+1i} < 1$ for $i = 1, 2, \dots, n-1$. Then $\rho(T) = 1$ implies $\rho(\bar{T}) = 1$. (All matrices considered here are defined as in Remark 3.4.)*

Proof. Clearly $\rho(T) = 1$ implies $\rho(T_1) = 1$. Lemma 3.5, T_1 is irreducible. By Theorem 2.1 there exists a positive vector ω' such that $T_1\omega' = \omega'$. Now define

$$\omega = \begin{pmatrix} T_0\omega' \\ \omega' \end{pmatrix}.$$

6

A. D. GUNAWARDENA, S. K. JAIN, AND L. SNYDER

Then clearly $T\omega = \omega$. Since $T_0 \neq 0$, $T_0\omega'$ is a positive scalar and hence ω is a positive vector. Consider

$$\tilde{T}\omega = (I - L - SL)^{-1}(U - S + SU)\omega.$$

By factoring $I - L$ from the right hand side we get

$$\begin{aligned}\tilde{T}\omega &= [I - (I - L)^{-1}SL]^{-1}(I - L)^{-1}(U - S + SU)\omega \\ &= [I - (I - L)^{-1}SL]^{-1}[(I - L)^{-1}U\omega - (I - L)^{-1}S\omega + (I - L)^{-1}SU\omega].\end{aligned}$$

Now $(I - L)^{-1}U\omega = \omega$ implies $U\omega = (I - L)\omega$. Therefore we get

$$\begin{aligned}\tilde{T}\omega &= [I - (I - L)^{-1}SL]^{-1}[\omega - (I - L)^{-1}S\omega + (I - L)^{-1}S(I - L)\omega] \\ &= [I - (I - L)^{-1}SL]^{-1}[\omega - (I - L)^{-1}SL\omega] \\ &= [I - (I - L)^{-1}SL]^{-1}[I - (I - L)^{-1}SL]\omega \\ &= \omega.\end{aligned}$$

Thus by Theorem 2.2, $\rho(\tilde{T}) = 1$. ■

LEMMA 3.7. *Let A be a Z -matrix such that $0 < a_{ii+1}a_{i+1i} < 1$. Then $\rho(T) = \rho(\tilde{T}) = \lambda$ if and only if $\lambda = 1$.*

Proof. To show necessity, suppose that T and \tilde{T} are partitioned as follows:

$$T = \begin{pmatrix} 0 & T_0 \\ 0 & T_1 \end{pmatrix} \quad \text{and} \quad \tilde{T} = \begin{pmatrix} 0 & \tilde{T}_0 \\ 0 & \tilde{T}_1 \end{pmatrix}.$$

First note that $\rho(T) = \rho(T_1) = \lambda$. Since T_1 is irreducible, there exists a positive vector ω' such that $T_1\omega' = \lambda\omega'$.

CONSISTENT LINEAR SYSTEMS

7

Let

$$\omega = \begin{pmatrix} \lambda^{-1}T_0\omega' \\ \omega' \end{pmatrix}.$$

Then ω is a positive vector and $T\omega = \lambda\omega$. Similarly we can obtain a positive vector v such that $v^T = \lambda v^T$, since \bar{T}_1 is irreducible and $\rho(\bar{T}_1) = \lambda$.

Consider

$$v^T \bar{T} \omega = v^T (I - L - SL)^{-1} (U - S + SU) \omega$$

and

$$v^T \bar{T} \omega = \lambda v^T \omega.$$

This implies

$$v^T (I - L - SL)^{-1} (U - S + SU) \omega - \lambda v^T \omega = 0,$$

that is,

$$v^T (I - L - SL)^{-1} [(U - S + SU) \omega - \lambda (I - L - SL) \omega] = 0.$$

Since

$$U\omega = \lambda(I - L)\omega, \quad (1)$$

we have

$$v^T (I - L - SL)^{-1} [\lambda(I - L)\omega - S\omega + SU\omega - \lambda(I - L)\omega + \lambda SL\omega] = 0,$$

$$v^T (I - L - SL)^{-1} (-S\omega + SU\omega + \lambda SL\omega) = 0,$$

$$(I - L - SL)^{-1} [-S\omega + \lambda S(I - L)\omega + \lambda SL\omega] = 0 \quad (2)$$

by (1)

$$v^T (I - L - SL)^{-1} (-S\omega + \lambda S\omega) = 0,$$

$$v^T (I - L - SL)^{-1} (\lambda - 1)S\omega = 0.$$

Since $(I - L - SL)^{-1}$ is a nonnegative lower triangular matrix, we can write

$$(I - L - SL)^{-1} = D + L',$$

where D is a positive diagonal matrix and L' is a nonnegative strictly lower triangular matrix. Then

$$\begin{aligned} v^T(I - L - SL)^{-1}S\omega &= v^T(D + L')S\omega \\ &= v^TDS\omega + v^TL'S\omega. \end{aligned}$$

Since $DS\omega \neq 0$ and v^T is a positive vector,

$$v^T(I - L - SL)^{-1}S\omega \neq 0.$$

Therefore, by Equation (2), $\lambda = 1$.

Sufficiency follows from Lemma 3.6. ■

LEMMA 3.8. *Let $A = I - L - U$ be a Z-matrix such that $0 < a_{i+1,i}a_{i,i+1} < 1$, where $-L$ and $-U$ are the strictly lower and strictly upper triangular parts of A , respectively. Then the standard Jacobi iteration matrix $T_j = L + U$ and the modified Jacobi iteration matrix $\tilde{T}_j = (I - SL)^{-1}(L + U - S + SU)$ are both irreducible.*

Proof. It follows from the condition $0 < a_{i+1,i}a_{i,i+1}$ that $(1, 2, \dots, n - 1, n, n - 1, \dots, 2, 1)$ is a path in $\Gamma(L + U)$. Hence $\Gamma(T_j)$ is strongly connected, and so T_j is irreducible.

Next we show that \tilde{T}_j is an irreducible matrix. Note that

$$\begin{aligned} \Gamma(\tilde{T}_j) &\supseteq \Gamma(L + U - S + SU) \\ &= \Gamma(L) \cup \Gamma(U - S) \cup \Gamma(SU), \end{aligned}$$

where the first line follows from the fact that $(I - SL)^{-1}$ is a nonnegative matrix with a positive diagonal. Recall that $\Gamma(L)$ contains the path $(n, n - 1, \dots, 1)$, and note that $\Gamma(SU)$ contains the edges $(1, 3), (2, 4), \dots, (n - 2, n)$. Hence $\Gamma(\tilde{T}_j)$ is strongly connected and \tilde{T}_j is irreducible. ■

LEMMA 3.9. *Let $A = I - L - U$ be a Z-matrix such that $0 < a_{i+1,i}a_{i,i+1} < 1$ and $\rho(T) = 1$, where $T = (I - L)^{-1}U$. Define $T_\epsilon = [(1 + \epsilon)I - L]^{-1}U$ for $\epsilon \geq 0$. Then $\rho(T_\epsilon) < 1/(1 + \epsilon)$.*

CONSISTENT LINEAR SYSTEMS

9

Proof. $\rho(T) = 1$ implies that there exists a positive vector ω such that $T\omega = \omega$. That is,

$$U\omega = (I - L)\omega. \quad (3)$$

Consider

$$\begin{aligned} T_\epsilon \omega - \frac{1}{1+\epsilon} \omega &= [(1+\epsilon)I - L]^{-1} U\omega - \frac{1}{1+\epsilon} \omega \\ &= [(1+\epsilon)I - L]^{-1} \left(U\omega - [(1+\epsilon)I - L] \frac{1}{1+\epsilon} \omega \right) \\ &= [(1+\epsilon)I - L]^{-1} \left((I - L)\omega - \omega + \frac{L\omega}{1+\epsilon} \right), \quad \text{by (3),} \\ &= [(1+\epsilon)I - L]^{-1} \frac{-\epsilon}{1+\epsilon} L\omega. \end{aligned} \quad (4)$$

Note that $[(1+\epsilon)I - L]^{-1} L\omega \geq 0$, but $[(1+\epsilon)I - L]^{-1} L\omega \neq 0$. Therefore

$$T_\epsilon \omega - \frac{1}{1+\epsilon} \omega \leq 0,$$

by (4). Hence

$$T_\epsilon \omega \leq \frac{1}{1+\epsilon} \omega. \quad (5)$$

As in T , the first column of T_ϵ is zero, and so we may partition

$$T_\epsilon = \begin{pmatrix} 0 & T^* \\ 0 & T'_\epsilon \end{pmatrix},$$

and following the methods employed in the proof of Lemma 3.5, it is easy to show that T'_ϵ is a nonnegative irreducible matrix. Now write

$$\omega = \begin{pmatrix} \omega_0 \\ \bar{\omega} \end{pmatrix}.$$

Then

$$T'_\epsilon \bar{\omega} \leq \frac{1}{1+\epsilon} \bar{\omega} \quad \text{by (5).}$$

Since T'_ϵ is irreducible, by Theorem 2.2 we have $\rho(T'_\epsilon) < 1/(1 + \epsilon)$. This together with $\rho(T'_\epsilon) = \rho(T'_\epsilon)$ gives us the desired result. ■

4. MAIN RESULTS IN Z-MATRICES

THEOREM 4.1. *Suppose $A = I - L - U$ is a Z-matrix such that $0 < a_{i,i+1}a_{i+1,i} < 1$, where $-L$ and $-U$ are strictly lower and strictly upper triangular parts of A respectively. Let $T = (I - L)^{-1}U$ and $\bar{T} = (I - L - SL)^{-1}(U - S + SU)$ be the standard and modified Gauss-Seidel iteration matrices, respectively. Then*

- (a) $\rho(\bar{T}) < \rho(T)$ if $\rho(T) < 1$,
- (b) $\rho(\bar{T}) = \rho(T)$ if $\rho(T) = 1$,
- (c) $\rho(\bar{T}) > \rho(T)$ if $\rho(T) > 1$.

Proof. Part (b) follows from Lemma 3.6. Now to prove (a) and (c), first we note that there exists a positive vector ω such that

$$T\omega = \lambda\omega, \tag{6}$$

where $\lambda = \rho(T)$. Now consider

$$\begin{aligned} \bar{T}\omega &= (I - L - SL)^{-1}(U - S + SU)\omega \\ &= (I - L - SL)^{-1}(I + S)U\omega - (I - L - SL)^{-1}S\omega \\ &= (I - L - SL)^{-1}(I + S)\lambda(I - L)\omega - (I - L - SL)^{-1}S\omega \quad \text{by (6)}. \end{aligned}$$

Therefore

$$\begin{aligned} \bar{T}\omega - T\omega &= (I - L - SL)^{-1} \\ &\quad \times [\lambda(I + S)(I - L)\omega - S\omega - (I - L - SL)(I - L)^{-1}U\omega] \\ &= (I - L - SL)^{-1}[\lambda\omega - \lambda L\omega + \lambda S\omega - S\omega - U\omega] \\ &= (I - L - SL)^{-1}[\lambda(I - L)\omega + (\lambda - 1)S\omega - U\omega] \\ &= (I - L - SL)^{-1}(\lambda - 1)S\omega. \end{aligned}$$

CONSISTENT LINEAR SYSTEMS

11

Write $(I - L - SL)^{-1} = D + L'$ for some positive diagonal matrix D and a nonnegative strictly lower triangular matrix L' . Then $(I - L - SL)^{-1}S\omega = (D + L')S\omega \geq 0$, since $DS\omega \geq 0$. Also, since $DS\omega \neq 0$, $(I - L - SL)^{-1}S\omega$ is a nonzero, nonnegative vector.

If $\lambda < 1$, then $\tilde{T}\omega - T\omega \leq 0$. Therefore

$$\tilde{T}\omega \leq \lambda\omega.$$

By using the partitioned form of

$$\tilde{T} = \begin{pmatrix} 0 & \tilde{T}_0 \\ 0 & \tilde{T}_1 \end{pmatrix}$$

introduced in Remark 3.4, we get $\rho(\tilde{T}) < \lambda$, by Theorem 2.2.

If $\lambda > 1$, then $\tilde{T}\omega - T\omega \geq 0$ but not equal to 0. Hence $\tilde{T}\omega \geq \lambda\omega$, and this implies $\tilde{T}_1\bar{\omega} \geq \lambda\bar{\omega}$, where

$$\omega = \begin{pmatrix} \alpha \\ \bar{\omega} \end{pmatrix}, \quad \bar{\omega} > 0.$$

Therefore $\rho(\tilde{T}_1) > \lambda$ by Theorem 2.2. Hence $\rho(\tilde{T}) > \lambda$. ■

REMARK 4.2. We recall that when the iteration matrix T is convergent, $\rho(T) < 1$. Theorem 4.1 shows that the modified iteration matrix has a faster convergence rate when the standard iteration matrix is convergent, and the modified iteration matrix diverges even faster when the standard iteration matrix is divergent.

At this point one might ask: How much faster is the convergence of the modified iteration matrix than of the standard iteration matrix when they are both convergent?

When A is a Z -matrix which is not tridiagonal, the answer to the above question seems to depend on the magnitude of $\rho(T)$. We have tested many examples with random entries and noted the following:

(a) When $\rho(T) < 1$ and close to 1, $\rho(T) - \rho(\tilde{T})$ seems to be relatively small, and hence the improvement seems to be rather slight.

(b) When $\rho(T) < 1$ and close to 0.5, the difference $\rho(T) - \rho(\tilde{T})$ seems to be relatively large and the modified method should be preferred over the standard method.

We present some matrices in Example 4.3 to illustrate observations (a) and (b) above.

EXAMPLE 4.3.

(a) Let

$$A = \begin{pmatrix} 1 & -0.2 & -0.1 & -0.4 & -0.2 \\ -0.2 & 1 & -0.3 & -0.1 & -0.6 \\ -0.3 & -0.2 & 1 & -0.1 & -0.6 \\ -0.1 & -0.1 & -0.1 & 1 & -0.01 \\ -0.2 & -0.3 & -0.4 & -0.3 & 1 \end{pmatrix} = I - L - U,$$

where $-L$ and $-U$ are the strictly lower and strictly upper triangular parts of A respectively. If

$$T = (I - L)^{-1}U \quad \text{and} \quad \tilde{T} = (I - L - SL)^{-1}(U - S + SU),$$

where S is the first upper codiagonal of $-U$, then $\rho(T) = 0.9611$ and $\rho(\tilde{T}) = 0.9505$.

(b) Let

$$A = \begin{pmatrix} 1 & -0.0089 & -0.1305 & -0.0679 & -0.0252 \\ -0.2891 & 1 & -0.4724 & -0.2938 & -0.3628 \\ -0.1424 & -0.3383 & 1 & -0.0972 & -0.0290 \\ -0.3454 & -0.3384 & -0.4843 & 1 & -0.2982 \\ -0.0363 & -0.1415 & -0.3680 & -0.1266 & 1 \end{pmatrix}.$$

If T and \tilde{T} are defined as above, then

$$\rho(T) = 0.6897 \quad \text{and} \quad \rho(\tilde{T}) = 0.5610.$$

REMARK 4.4. In the examples we tested, it seems that a reduction of the spectral radius by 0.1 results in an average saving of about six iterations for a convergence criterion of 0.1 percent accuracy.

In the following theorem, we compare the modified and standard *Jacobi iteration matrices* for a Z -matrix A .

THEOREM 4.5. Let $A = I - L - U$ be a Z -matrix, where $-L$ and $-U$ are the strictly lower and strictly upper triangular parts of A . Let $T_j = L + U$

and $\bar{T}_j = (I - SL)^{-1}(L + U - S + SU)$ be the standard and modified Jacobi iteration matrices, respectively. Further assume that T_j and \bar{T}_j are irreducible matrices. Then

- (a) $\rho(\bar{T}_j) < \rho(T_j)$ if $\rho(T_j) < 1$,
- (b) $\rho(\bar{T}_j) > \rho(T_j)$ if $\rho(T_j) > 1$,
- (c) $\rho(\bar{T}_j) = \rho(T_j)$ if $\rho(T_j) = 1$.

Proof. Since T_j is irreducible, by Theorem 2.1 there exists a positive vector ω such that $T_j\omega = \lambda\omega$, where $\lambda = \rho(T_j)$. This implies

$$(L + U)\omega = \lambda\omega. \quad (7)$$

Now consider

$$\begin{aligned} \bar{T}_j\omega - T_j\omega &= (I - SL)^{-1}[L + U - S + SU - (I - SL)(L + U)]\omega \\ &= (I - SL)^{-1}[-S + SU + SL^2 + SLU]\omega \\ &= (I - SL)^{-1}S[-I + U + L^2 + LU]\omega \\ &= (I - SL)^{-1}S[-I + U + \lambda L]\omega \quad \text{by (7)} \\ &= (I - SL)^{-1}S(-\omega + \lambda\omega - L\omega + \lambda L\omega) \\ &= (I - SL)^{-1}S(I + L)(\lambda - 1)\omega. \end{aligned} \quad (8)$$

If $\rho(T_j) < 1$, then $\bar{T}_j\omega \leq \rho(T_j)\omega$ by (8), and so, by using Theorem 2.2, we obtain $\rho(\bar{T}_j) < \rho(T_j)$.

Similarly we can get $\rho(\bar{T}_j) > \rho(T_j)$ if $\rho(T_j) > 1$, and (c) also follows from Theorem 2.2. ■

COROLLARY 4.6. Let $A = I - L - U$, T_j, \bar{T}_j be as defined in Theorem 4.5. Replace " T_j, \bar{T}_j are irreducible matrices" by the condition " $0 < a_{i+1,i}a_{i,i+1} < 1$." Then the conclusion of Theorem 4.5 holds.

The proof follows from Lemma 3.8.

5. FURTHER RESULTS ON Z-MATRICES

In this section we consider Z-matrices A such that $\rho(T) = 1$, where T is the Gauss-Seidel iteration matrix or Jacobi iteration matrix. In particular, we focus our attention on Q -matrices.

We have seen in Lemma 3.1 that when A is a Q -matrix, both the Gauss-Seidel and Jacobi iteration matrices have spectral radius equal to 1. In this case the iteration matrix T is semiconvergent if and only if $\rho(T) = 1$ is the only eigenvalue of T with modulus 1 and the Jordan blocks associated with the eigenvalue 1 are all 1×1 matrices [1]. In Lemma 3.6 and Theorem 4.5, we have that $\rho(T) = 1$ implies $\rho(\tilde{T}) = 1$, where T and \tilde{T} are either the standard and modified Gauss-Seidel iteration matrices or the standard and modified Jacobi iteration matrices respectively. In such cases the rate of convergence of the iteration matrix is determined by the second largest modulus of the eigenvalues. We call the second largest modulus the *subdominant eigenvalue* of the iteration matrix. Let

$$\gamma(T) = \max\{|\lambda| : \lambda \in \text{spectrum}(T), \lambda \neq 1\}.$$

We provide the following example to show that when $\rho(T) = 1$, the modified iterative method may not always be faster than the standard iterative method.

EXAMPLE 5.1. Let

$$A = \begin{pmatrix} 1 & -0.2 & 0 & -0.1 \\ -0.5 & 1 & -0.999 & -0.6 \\ 0 & -0.8 & 1 & -0.3 \\ -0.5 & 0 & -0.001 & 1 \end{pmatrix}.$$

We note the following facts:

- (a) A is a Q -matrix, and hence A is singular.
- (b) $\text{spectrum}(T) = \{0, 1, -0.025 + 0.1713i, -0.025 - 0.1713i\}$, where $T = (I - L)^{-1}U$.
- (c) $\text{spectrum}(\tilde{T}) = \{0, 0, 1, -0.3980\}$ where $\tilde{T} = (I - L - SL)^{-1}(U - S + SU)$.
- (d) $\gamma(T) = 0.1731$ and $\gamma(\tilde{T}) = 0.3980$.
- (e) The matrix A satisfies the hypothesis of Theorem 4.1, but \tilde{T} fails to give a better convergence rate than T .

REMARK 5.2. As we have seen in Example 5.1, in some cases the modified method applied to singular Z-matrices may fail to give a faster

convergence rate than the standard method. However, when A is a tridiagonal Q -matrix, the modified Gauss-Seidel method seems to increase the convergence rate of the iteration matrix dramatically, as evidenced by the many examples we tested with random entries. We have been able to establish only partial results confirming this.

This task is made more difficult by the limited availability of research material on the subdominant eigenvalue of a nonnegative matrix.

LEMMA 5.3. *Let $A = I - L - U$ be an irreducible tridiagonal Q -matrix of order $n > 2$ with the usual splitting. Let $T = (I - L)^{-1}U$, and define $T_\epsilon = [(1 + \epsilon)I - L]^{-1}U$ for some $\epsilon \geq 0$. Then there exists $\epsilon_0 > 0$ such that $\rho(T_{\epsilon_0}) = \gamma(T)$. Furthermore $\gamma(T) \geq \rho(T_\epsilon)$ for all $\epsilon \geq \epsilon_0$.*

Proof. Since A is a Q -matrix, $\rho(T) = 1$. Therefore by Lemma 3.9, $\rho(T_\epsilon) < 1/(1 + \epsilon)$ and so $\lim \rho(T_\epsilon) = 0$ as $\epsilon \rightarrow \infty$. It is well known that the characteristic polynomial is a continuous function of the entries of the matrix, and so $\rho(T_\epsilon)$ is a continuous function of ϵ . Therefore, as shown by the graph of $\rho(T_\epsilon)$ (Figure 1), there exists $\epsilon_0 \geq 0$ such that $\gamma(T) \geq \rho(T_\epsilon)$ for all $\epsilon \geq \epsilon_0$ and equality holds when $\epsilon = \epsilon_0$. ■

LEMMA 5.4. *Suppose A is as defined in Lemma 5.3, and let*

$$T_\epsilon = [(1 + \epsilon)I - L]^{-1}U \quad \text{and} \quad \tilde{T}_\epsilon = [(1 + \epsilon)I - UL - L]^{-1}U^2.$$

Then $\rho(\tilde{T}_\epsilon) < \rho(T_\epsilon)$.

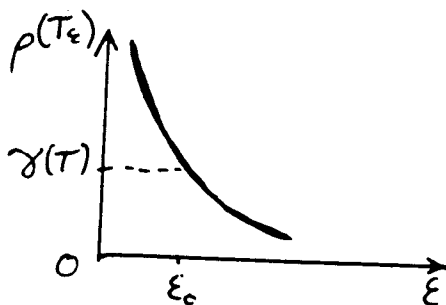


FIG. 1.

Proof. Since T_ϵ is irreducible, there exists a positive vector ω such that $T_\epsilon\omega = \rho(T_\epsilon)\omega$. That is,

$$U\omega = [(1 + \epsilon)I - L]\omega\rho(T_\epsilon). \quad (9)$$

Now consider

$$\begin{aligned} \bar{T}_\epsilon\omega - T_\epsilon\omega &= [(1 + \epsilon)I - UL - L]^{-1}U^2\omega - [(1 + \epsilon)I - L]^{-1}U\omega \\ &= [(1 + \epsilon)I - UL - L]^{-1} \\ &\quad \times \{U^2\omega - [(1 + \epsilon)I - UL - L][(1 + \epsilon)I - L]^{-1}U\omega\} \\ &= [(1 + \epsilon)I - UL - L]^{-1} \\ &\quad \times \{U[(1 + \epsilon)I - L]\rho(T_\epsilon)\omega - U\omega + UL\rho(T_\epsilon)\omega\} \quad \text{by (9)} \\ &= [(1 + \epsilon)I - UL - L]^{-1}[U(1 + \epsilon)\rho(T_\epsilon)\omega - U\omega] \\ &= [(1 + \epsilon)I - UL - L]^{-1}[(1 + \epsilon)\rho(T_\epsilon) - 1]U\omega. \end{aligned}$$

Since $\rho(T_\epsilon) < 1/(1 + \epsilon)$ by Lemma 3.9, we get $\bar{T}_\epsilon\omega < T_\epsilon\omega$. So this yields $\rho(\bar{T}_\epsilon) < \rho(T_\epsilon)$, by Theorem 2.2. ■

REMARK 5.5. By using Lemma 5.3 and Lemma 5.4 we shall establish $\rho(\bar{T}_\epsilon) < \rho(T_\epsilon) \leq \gamma(T)$ for all $\epsilon \geq \epsilon_0$. Testing of numerous examples have shown that the inequality $\gamma(\bar{T}) \leq \rho(\bar{T}_\epsilon)$, for all $\epsilon \geq \epsilon_0$, holds true in general, but a mathematical proof of this claim is still an open problem. We give an example of a tridiagonal Q -matrix below to illustrate the effectiveness of the modified method applied to Gauss-Seidel iteration matrix. All examples we tested showed a relatively large decrease in the subdominant eigenvalue, and therefore a significant improvement in the convergence rate by using the modified iteration matrix.

CONSISTENT LINEAR SYSTEMS

17

EXAMPLE 5.6. Let

$$A = \begin{pmatrix} 1 & -0.92 & 0 & 0 & 0 & 0 \\ -1 & 1 & -0.14 & 0 & 0 & 0 \\ 0 & -0.08 & 1 & -0.71 & 0 & 0 \\ 0 & 0 & -0.86 & 1 & -0.02 & 0 \\ 0 & 0 & 0 & -0.29 & 1 & -1 \\ 0 & 0 & 0 & 0 & -0.98 & 1 \end{pmatrix} = I - L - U$$

be a Q -matrix. Then $\gamma(T) = 0.9451$ and $\gamma(\bar{T}) = 0.0956$, where $T = (I - L)^{-1}U$ and $\bar{T} = (I - L - UL)^{-1}U^2$.

6. COMPARISON WITH ANOTHER METHOD

In this section we discuss miscellaneous results related to the modified iterative method. First we state the following theorem by Milaszewicz [2] and compare the modified method with his method by providing Example 6.2. Notice that we chose the modified Jacobi iteration matrix for this comparison so that it satisfies the hypothesis of Theorem 6.1.

THEOREM 6.1 (Milaszewicz [2]). *Let T be a nonnegative irreducible matrix such that $t_{ii} = 0$ for all i , $1 \leq i \leq n$. Let k be an arbitrary integer between 1 and n , and S be the matrix whose only nonvanishing terms belong to its k th column and coincide with the corresponding ones in T . Set $T_m = ST + T - S$. If $\rho(T) < 1$, then $\rho(T_m) < \rho(T)$.*

EXAMPLE 6.2. Let

$$A = \begin{pmatrix} 1 & -0.1 & -0.2 & 0 & -0.3 & -0.5 \\ -0.2 & 1 & -0.3 & 0 & -0.4 & -0.1 \\ 0 & -0.3 & 1 & -0.6 & -0.2 & 0 \\ -0.2 & -0.3 & 0 & 1 & -0.1 & -0.3 \\ 0 & -0.3 & -0.2 & -0.1 & 1 & -0.2 \\ -0.2 & -0.3 & 0 & -0.3 & -0.1 & 1 \end{pmatrix} = I - L - U.$$

Then $\rho(T_j) = 0.9530$, $\rho(\bar{T}_j) = 0.9371$, and $\rho(T_m) = 0.9451$, where

$$T_j = L + U, \quad \bar{T}_j = (I - SL)^{-1}(L + U - S + SU)$$

are as defined in Remark 3.3, and T_m is as defined in Theorem 6.1. Note that $\rho(\tilde{T}_j) < \rho(T_m)$.

Next we present an interesting theorem which contains some results on the modified method applied to a special type of matrix:

$$\begin{pmatrix} 1 & a & 0 & 0 & 0 & \cdots & 0 & 0 \\ -b & 1 & a & 0 & 0 & \cdots & 0 & 0 \\ 0 & -b & 1 & a & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -b & 1 & a \\ 0 & 0 & 0 & 0 & \cdots & 0 & -b & 1 \end{pmatrix}. \tag{10}$$

THEOREM 6.3. *Let $A = I - L - U$ be a matrix of the form (10) such that a, b are positive scalars with $a < 1$ and $b < 1$. Assume $\rho(T) \leq a(1 + b)$, where $T = (I - L)^{-1}U$. Then for $\tilde{T} = (I - L - UL)^{-1}U^2$, $\rho(\tilde{T}) < \rho(T)$. (Indeed, it suffices to assume that $a < 1$ and $ab < 1$.)*

Proof. Using the series expansion of $(I - L)^{-1}$ and $(I - L - UL)^{-1}$, one can show that

$$T = (I - L)^{-1}U = -a \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & b & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & b^2 & b & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & b^{n-2} & b^{n-3} & b^{n-4} & \cdots & b^2 & b & 1 \\ 0 & b^{n-1} & b^{n-2} & b^{n-3} & \cdots & b^3 & b^2 & b \end{pmatrix}$$

and

$$\begin{aligned} \tilde{T} &= (I - L - UL)^{-1}U^2 \\ &= \frac{a^2}{c} \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & d & 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & d^2 & d & 1 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & d^{n-3} & d^{n-4} & \cdot & \cdot & \cdot & d & 1 \\ 0 & 0 & d^{n-2} & d^{n-3} & \cdot & \cdot & \cdot & d^2 & d \\ 0 & 0 & bd^{n-2} & bd^{n-3} & \cdot & \cdot & \cdot & bd^2 & bd \end{pmatrix}, \end{aligned}$$

CONSISTENT LINEAR SYSTEMS

19

where $c = 1 + ab$ and $d = b/c$. Now we partition T and \bar{T} so that

$$T = -a \begin{pmatrix} 0 & T_0 \\ 0 & T_1 \end{pmatrix} \quad \text{and} \quad \bar{T} = \frac{a^2}{c} \begin{pmatrix} 0 & \bar{T}_0 \\ 0 & \bar{T}_1 \end{pmatrix},$$

where T_1 and \bar{T}_1 are $(n-1) \times (n-1)$ matrices. Since T_1 is irreducible by Theorem 2.1, there exists a positive vector x such that $T_1 x = \lambda x$, where $\lambda = \rho(T_1)$. Now $\rho(T) = a\rho(T_1) \leq a(1+b)$ implies $\lambda \leq 1+b$. Thus we get

$$\begin{pmatrix} b & 1 & 0 & 0 & \cdots & 0 & 0 \\ b^2 & b & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ b^{n-2} & b^{n-3} & b^{n-4} & \cdots & b^2 & b & 1 \\ b^{n-1} & b^{n-2} & b^{n-3} & \cdots & b^3 & b^2 & b \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} \leq (1+b) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix}.$$

This yields $n-1$ inequalities:

$$(i_1) \quad bx_1 + x_2 \leq (1+b)x_1,$$

$$(i_2) \quad b^2x_1 + bx_2 + x_3 \leq (1+b)x_2,$$

$$\vdots$$

$$(i_{n-1}) \quad b^{n-2}x_1 + b^{n-3}x_2 + \cdots + bx_{n-1} + x_n \leq (1+b)x_{n-1}.$$

Hence it follows that

$$x_1 \geq x_2 > \cdots > x_{n-1} > x_n.$$

That is, the components of the vector x are in decreasing order. Now consider

$$\frac{a^2}{c} \bar{T}_1 x - a\lambda x = \frac{a}{c} (a\bar{T}_1 x - cT_1 x).$$

We will show that $a\bar{T}_1x - cT_1x < 0$. Let $y = a\bar{T}_1x - cT_1x$. That is,

$$y = a \begin{pmatrix} 0 & d & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & d^2 & d & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & d^{n-3} & d^{n-4} & \cdot & \cdot & \cdot & d & 1 \\ 0 & d^{n-2} & d^{n-3} & \cdot & \cdot & \cdot & d^2 & d \\ 0 & bd^{n-2} & bd^{n-3} & \cdot & \cdot & \cdot & bd^2 & bd \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-2} \\ x_{n-1} \end{pmatrix}$$

$$- c \begin{pmatrix} b & 1 & 0 & 0 & \cdots & 0 & 0 \\ b^2 & b & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ b^{n-2} & b^{n-3} & b^{n-4} & \cdots & b^2 & b & 1 \\ b^{n-1} & b^{n-2} & b^{n-3} & \cdots & b^3 & b^2 & b \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-2} \\ x_{n-1} \end{pmatrix}$$

This yields $n-1$ equations,

$$y_1 = (adx_2 - cbx_1) + (ax_3 - cx_2),$$

$$y_2 = (ad^2x_2 - cb^2x_1) + (adx_3 - cbx_2) + (ax_4 - cx_3),$$

$$y_3 = (ad^3x_2 - cb^3x_1) + (ad^2x_3 - cb^2x_2) + (adx_4 - cbx_3) + (ax_5 - cx_4),$$

\vdots

$$y_{n-2} = (ad^{n-2}x_2 - cb^{n-2}x_1) + (ad^{n-3}x_3 - cb^{n-3}x_2) + \cdots$$

$$+ (adx_n - cbx_{n-1}) - cx_n,$$

$$y_{n-1} = by_{n-2}.$$

Since $a < 1$ and $c > 1$, we have $a < c^{r+1}$ for all $r = 1, 2, \dots$. This implies $a(b/c)^r < cb^r$ and hence $ad^r < cb^r$ for all $r = 1, 2, \dots$. Therefore $y_i < 0$ for all i , since $x_1 > x_2 > \cdots > x_n$. This yields

$$a\bar{T}_1x < cT_1x = c\lambda x.$$

By Theorem 2.2, we have $(a^2/c)\rho(\bar{T}_1) < a\lambda$. This implies that $\rho(\bar{T}) < \rho(T)$. ■

COROLLARY 6.4. *If $b < 1$, then a necessary condition for the hypothesis of Theorem 6.3 to hold is that $b < (\sqrt{5} - 1)/2$.*

Proof. Consider the series

$$b^{n-1} + b^{n-2} + b^{n-3} + \cdots + b^2 + b = \frac{b(1-b^{n-1})}{1-b}. \quad (11)$$

If $b > (\sqrt{5} - 1)/2$, then $b^2 - b - 1 > 0$ and hence $b/(1-b) > 1 + b$. Now for $b < 1$,

$$\frac{b - b^n}{1 - b} \approx \frac{b}{1 - b}$$

and so by (11)

$$b^{n-1} + b^{n-2} + b^{n-3} + \cdots + b^2 + b > b + 1.$$

Therefore the minimum row sum for T_1 is $b + 1$, and by Theorem 2.3, $\rho(T_1) > 1 + b$. ■

We have found some examples of matrices of the type (10) where the modified method converges quite fast even though the standard method diverges. Such examples are very encouraging, because the modified method might be applied to solve linear systems even when the standard Gauss-Seidel and Jacobi methods fail to give convergence.

REFERENCES

- 1 A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, Academic, 1979.
- 2 J. P. Milaszewicz, Improving Jacobi and Gauss-Seidel iterations, *Linear Algebra Appl.* 93:161-170 (1987).
- 3 M. E. Mokari-Bolhassan and T. N. Trick, A new iterative algorithm for the solutions of large scale systems, presented at 28th Midwest Symposium on Circuits and Systems Louisville, Ky. 19-20 Aug. 1985.
- 4 M. E. Mokari-Bolhassan, D. Smart, and T. N. Trick, A new robust relaxation technique for VLSI circuit simulation, preprint.
- 5 H. Schneider, Theorems on M -splitting of a singular M -matrix, *Linear Algebra Appl.* 58:407-424 (1984).

Received 6 March 1990; final manuscript accepted 17 July 1990.