On Optimal Improvements of Classical Iterative Schemes for $Z-$Matrices

D. Noutsos[1] and M. Tzoumas[2]

Abstract

Many researchers have considered left preconditioners, applied to linear systems, whose matrix coefficient is a $Z-$ or an $M-$matrix, that make the associated Jacobi and Gauss-Seidel methods converge asymptotically faster than the original ones. Such preconditioners are chosen so that they eliminate the off-diagonal elements of the same column or the elements of the first upper diagonal (Milaszewicz [14], Gunawardena et al [5]). In the present work a generalization of the previous techniques is proposed in order to obtain optimal methods. The best Jacobi and Gauss-Seidel algorithms are given and preconditioners, that eliminate more than one entry per row, are also proposed and analyzed.

AMS (MOS) Subject Classifications: Primary 65F10
Keywords: Jacobi and Gauss-Seidel iterative methods, diagonally dominant $Z-$ and $M-$Matrices
Running Title: Optimally Improved Iteration Schemes

1 Introduction and Preliminaries

Consider the linear system of algebraic equations

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

where $A \in \mathbb{R}^{n,n}$ belongs to the class of *irreducibly diagonally dominant $Z-$matrices* with positive diagonal entries (see [3], [15] and [17]), that is its off-diagonal elements are nonpositive, ($a_{ij} \leq 0,\ i,j = 1(1)n,\ j \neq i$,) and $b \in \mathbb{R}^n$. Since $a_{ii} > 0,\ i = 1(1)n$, without loss of generality, we assume for simplicity that $a_{ii} = 1,\ i = 1(1)n$. We consider the usual triangular splitting of A, namely

$$A = I - L - U, \tag{1.2}$$

where I is the identity matrix and L and U are strictly lower and strictly upper triangular, respectively. Then, it is known that the iterative methods of Jacobi and of Gauss-Seidel associated with (1.1) converge and by the Stein-Rosenberg theorem ([15], [17], [3]) the Gauss-Seidel method is faster than the Jacobi one.
Many researchers have considered left preconditioners, applied to system (1.1) that make the associated Jacobi and Gauss-Seidel methods converge asymptotically faster than the original

[1]Department of Mathematics, University of Ioannina, GR-451 10 Ioannina, Greece
[2]Department of Environment and Natural Resources Management, Agrinion University School, University of Ioannina, GR-301 00 Agrinion, Greece

ones. Milaszewicz [14] who, in turn, based his idea on previous ones (see, e.g., [13], [4], [9]), considered the preconditioner $P_1 \equiv I+S_1$, which is the elementary matrix used to eliminate the elements of the first column of A below the diagonal. Gunawardena et al [5] considered as a preconditioner the matrix $P_2 \equiv I+S_2$, whose effect on A is to eliminate the elements of the first upper diagonal. Kohno et al [10] extended the main idea in [5]. Recently Li and Sun [11] extended the class of matrices considered in [10] and very recently Hadjidimos, Noutsos and Tzoumas [6] extended, generalized and compared the previous works.

In this paper we introduce a family of preconditioners, by extending the previous ones, and give the algorithms that choose a "good" preconditioner in each case. The outline of this work is as follows: In Section 2, we extend Milaszewicz's and Gunawardena et al's preconditioners by giving a family of preconditioners based on the elimination of one element in each row of the matrix A, we present the convergence analysis and propose two algorithms that choose a "good" preconditioner for the Jacobi and Gauss-Seidel iterative schemes. In Section 3 we generalize the above preconditioners by introducing the idea of eliminating more than one entry per row and study the corresponding convergence analysis. Finally, in Section 4, numerical examples are presented in support of the theory developed.

2 Extending known Preconditioners

It is known that Milaszewicz's preconditioner [14] is based on the elimination of the entries $a_{i1}, i = 2(1)n$, (the first column) while the Gunawardena et al's preconditioner [5] is based on the elimination of the entries $a_{i,i+1}, i = 1(1)n-1$, (the first upper diagonal). The common feature of the two preconditioners is that they eliminate precisely one element of A in each but one row. If we try to extend Milaszewicz's preconditioner by eliminating an off-diagonal element of the first row we obtain the same convergence results for the Jacobi and the Gauss-Seidel type schemes, since the spectral radii of the corresponding iteration matrices associated with $A_1 = P_1A$, which are reducible, are independent of the off-diagonal elements of the first row of A (see [6]). This does not happen in the case of Gunawardena et al's preconditioner by eliminating an off-diagonal element of the last row. If we choose the first element of the last row, we introduce a new preconditioner having a cyclic structure and we call it Cyclic preconditioner:

$$P_3 \equiv I + S_3 = \begin{bmatrix} 1 & -a_{12} & & & & & \\ & 1 & -a_{23} & & & & \\ & & \ddots & \ddots & & & \\ & & & 1 & -a_{i,i+1} & & \\ & & & & \ddots & \ddots & \\ & & & & & 1 & -a_{n-1,n} \\ -a_{n1} & & & & & & 1 \end{bmatrix}. \tag{2.1}$$

This observation gives us the idea of considering a family of preconditioners by eliminating exactly one element per row. So, we have a preconditioner of the following general form:

$$P \equiv I+S = \begin{bmatrix} 1 & & & -a_{1k_1} & & & \\ & 1 & & & & -a_{2k_2} & \\ & & 1 & & -a_{3k_3} & & \\ & & & \ddots & & & \\ & & & & \ddots & & \\ & & & & & \ddots & \\ & & & -a_{nk_n} & & & 1 \end{bmatrix}, \tag{2.2}$$

where $k_i \in \{1,2,\ldots,i-1,i+1,\ldots,n\}$, $i=1(1)n$. It is obvious that we have $n-1$ choices for each row, so we have a total number of $(n-1)^n$ choices for the preconditioner of type P in (2.2). We denote $\widetilde{A}=PA$ the associated preconditioned matrix. We will assume that there is at least one pair of indices i,j, such that $a_{ik_i}a_{k_ij} \neq 0$, so that at least one element of $\widetilde{A}$ is different from that of A. Applying P to (1.1) we obtain the equivalent linear system

$$\widetilde{A}x=\widetilde{b}, \quad \text{with } \widetilde{A}=(I+S)\,A,\ \widetilde{b}=(I+S)\,b. \tag{2.3}$$

The elements $\widetilde{a}_{ij}$ of $\widetilde{A}$ are given by the relationships:

$$\widetilde{a}_{ij} = \begin{cases} a_{ij}-a_{ik_i}a_{k_ij}<0, & j \neq k_i, i \\ 0, & j=k_i \\ 1-a_{ik_i}a_{k_ii}>0, & j=i \end{cases}. \tag{2.4}$$

We define the matrices

$$D_s := \operatorname{diag}(a_{1k_1}a_{k_11},\ a_{2k_2}a_{k_22},\ \ldots,a_{nk_n}a_{k_nn}) \tag{2.5}$$

and

$$S(L+U-I) := L_s+D_s+U_s-S_L-S_U, \tag{2.6}$$

where D_s, L_s and U_s are the diagonal, the strictly lower and strictly upper triangular components of $S(L+U)$ while S_L and S_U are the strictly lower and strictly upper triangular components of S, which are all nonnegative matrices. To introduce the iterative methods that will be studied the following splittings are considered:

$$\widetilde{A} = \begin{cases} M-N &= (I+S)-(I+S)\,(L+U), \\ M'-N' &= I-(L+U+L_s+D_s+U_s-S_L-S_U)\,, \\ M''-N'' &= (I-D_s)-(L+U+L_s+U_s-S_L-S_U)\,. \end{cases} \tag{2.7}$$

The corresponding Jacobi and the Jacobi type iteration matrices as well as the corresponding Gauss-Seidel and Gauss-Seidel type ones are given by

$$\begin{aligned} &B := M^{-1}N = L+U, \\ &B' := M'^{-1}N' = (L+U+L_s+D_s+U_s-S_L-S_U)\,, \\ &\widetilde{B} := M''^{-1}N'' = (I-D_s)^{-1}\,(L+U+L_s+U_s-S_L-S_U)\,, \end{aligned} \tag{2.8}$$

and

$$H := (I - L)^{-1}U,$$
$$H' := (I - L - L_s + S_L)^{-1}(U + D_s + U_s - S_U), \tag{2.9}$$
$$\widetilde{H} := (I - D_s - L - L_s + S_L)^{-1}(U + U_s - S_U).$$

The main property of this section can now be stated and proved:

Theorem 2.1 *a) Under the assumptions and the notation so far, there hold: There exist y and $z \in \mathbb{R}^n$, with $y \geq 0$ and $z \geq 0$, such that*

$$B'y \leq By \text{ and } H'z \leq Hz, \tag{2.10}$$

$$\rho\left(\widetilde{B}\right) \leq \rho(B') \leq \rho(B) < 1, \tag{2.11}$$

$$\rho\left(\widetilde{H}\right) \leq \rho(H') \leq \rho(H) < 1, \tag{2.12}$$

$$\rho\left(\widetilde{H}\right) \leq \rho\left(\widetilde{B}\right), \quad \rho(H') \leq \rho(B'), \quad \rho(H) < \rho(B) < 1. \tag{2.13}$$

(Note: Equalities in (2.13) hold if and only if $\rho(\widetilde{B}) = 0$.)
b) Suppose that A is irreducible. Then, the matrix B is also irreducible which implies that the first inequality in (2.10) and the middle inequality in (2.11) are strict.

Proof: a) (2.10): To prove (2.10) we give the explicit expressions of the elements of the first two Jacobi and Jacobi type iteration matrices in the splittings (2.8):

$$\begin{cases} b_{ii} = 0, \\ b_{ij} = -a_{ij}, \quad j \neq i, \end{cases} \tag{2.14}$$

$$\begin{cases} b'_{ik_i} = 0, \\ b'_{ii} = a_{ik_i}a_{k_ii} = b_{ik_i}b_{k_ii}, \\ b'_{ij} = a_{ik_i}a_{k_ij} - a_{ij} = b_{ik_i}b_{k_ij} + b_{ij}, \qquad j \neq i, k_i \end{cases} \tag{2.15}$$

For the nonnegative Jacobi iteration matrix B there exists a nonnegative vector y such that $By = \rho(B)y$. Equating the i^{th} rows, of the two vectors and replacing the elements b_{ij} of B in terms of the elements b'_{ij} of B' using (2.14) and (2.15) we successively obtain

$$\begin{aligned} \rho(B)y_i &= \textstyle\sum_{j=1,j\neq i}^n b_{ij}y_j = \sum_{j=1,j\neq i,k_i}^n b_{ij}y_j + b_{ik_i}y_{k_i} \\ &= \textstyle\sum_{j=1,j\neq i,k_i}^n b'_{ij}y_j - b_{ik_i}\sum_{j=1,j\neq i,k_i}^n b_{k_ij}y_j + b_{ik_i}y_{k_i} \\ &= \textstyle\sum_{j=1,j\neq i}^n b'_{ij}y_j - b_{ik_i}\sum_{j=1,j\neq k_i}^n b_{k_ij}y_j + b_{ik_i}b_{k_ii}y_i + b_{ik_i}y_{k_i} \\ &= \textstyle\sum_{j=1}^n b'_{ij}y_j - b_{ik_i}\sum_{j=1}^n b_{k_ij}y_j + b_{ik_i}y_{k_i}. \end{aligned} \tag{2.16}$$

Using the fact that $\rho(B)y_{k_i} = \sum_{j=1}^n b_{k_ij}y_j$ and replacing in (2.16) we have that

$$\rho(B)y_i = \sum_{j=1}^n b'_{ij}y_j + b_{ik_i}\left(\frac{1}{\rho(B)} - 1\right)\sum_{j=1}^n b_{k_ij}y_j. \tag{2.17}$$

Since the second term of the sum in (2.17) is nonnegative we have that

$$\sum_{j=1}^n b'_{ij}y_j \leq \sum_{j=1}^n b_{ij}y_j \tag{2.18}$$

from which the first of (2.10) follows.

For the nonnegative Gauss-Seidel iteration matrix H there exists a nonnegative vector z such that $Hz = \rho(H)z$. Using the fact that $H = (I-L)^{-1}U$ we have that $(I-L)^{-1}Uz = \rho(H)z$ or $Uz = \rho(H)(I-L)z$ or equivalent $\rho(H)z = \rho(H)Lz + Uz$. Equating the i^{th} rows, of the two vectors and replacing the elements b_{ij} of B in terms of the elements b'_{ij} of B' using (2.14) and (2.15) and suppose that $k_i > i$ (we have the same result in case $k_i < i$) we successively obtain

$$\begin{array}{rcl} \rho(H)z_i & = & \rho(H)\sum_{j=1}^{i-1} b_{ij}z_j + \sum_{j=i+1}^{n} b_{ij}z_j \\ & = & \rho(H)\sum_{j=1}^{i-1} b'_{ij}z_j - \rho(H)b_{ik_i}\sum_{j=1}^{i-1} b_{k_ij}z_j \\ & + & \sum_{j=i}^{n} b'_{ij}z_j - b_{ik_i}\sum_{j=i,j\neq k_i}^{n} b_{k_ij}z_j + b_{ik_i}z_{k_i}. \end{array} \tag{2.19}$$

Using the fact that $\rho(H)z_{k_i} = \rho(H)\sum_{j=1}^{k_i-1} b_{k_ij}z_j + \sum_{j=k_i+1}^{n} b_{k_ij}z_j$ or equivalently $z_{k_i} = \sum_{j=1}^{k_i-1} b_{k_ij}z_j + \frac{1}{\rho(H)}\sum_{j=k_i+1}^{n} b_{k_ij}z_j$ and replacing in (2.19) we have that

$$\begin{array}{rcl} \rho(H)z_i & = & \rho(H)\sum_{j=1}^{i-1} b'_{ij}z_j + \sum_{j=i}^{n} b'_{ij}z_j \\ & + & \underbrace{b_{ik_i}\left((1-\rho(H))\sum_{j=1}^{i-1} b_{k_ij}z_j + \left(\frac{1}{\rho(H)} - 1\right)\sum_{j=k_i+1}^{n} b_{k_ij}z_j\right)}_{\geq 0}. \end{array} \tag{2.20}$$

Since the underbraced term in (2.20) is nonnegative we have that

$$\rho(H)(z_i - \sum_{j=1}^{i-1} b'_{ij}z_j) \geq \sum_{j=i}^{n} b'_{ij}z_j\,, \quad i = 1(1)n. \tag{2.21}$$

In terms of matrices, this relation is equivalent to

$$\begin{array}{c} \rho(H)(I - L - L_s + S_L)z \geq (D_s + U + U_s - S_U)z \quad \text{or} \\ \rho(H)z \geq (I - L - L_s + S_L)^{-1}(D_s + U + U_s - S_U)z \quad \text{or} \quad \rho(H)z \geq H'z \end{array} \tag{2.22}$$

from which the second of (2.10) follows.

a) (2.11): Since a $Z-$matrix A is a nonsingular $M-$matrix iff there exists a positive vector $y(>0) \in I\!\!R^n$ such that $Ay > 0$ (see [3]), we have that $P = I + S \geq 0$, implies $\widetilde{A}y = PAy > 0$. Consequently, $\widetilde{A}$, which **is** a $Z-$matrix, is a nonsingular $M-$matrix. So, the last two splittings in (2.7) are regular splittings because $M'^{-1} = I^{-1} = I \geq 0$, $N' \geq 0$ and $M''^{-1} = (I - D_s)^{-1} \geq 0$, $N'' \geq 0$ and so they are convergent. Since $M''^{-1} \geq M'^{-1}$, it is implied (see [16]) that the left inequality in (2.11) is true. We recall now, for the proof of the middle inequality in (2.11), the first inequality of (2.10) which gives us that $B'y \leq \rho(B)y$. Then, we apply Lemma 3.3 by Marek and Szyld ([12]) to get our assertion.

a) (2.12): To prove the first inequality in (2.12) we use regular splittings of the matrix $\widetilde{A}$. Specifically, consider the following splittings that define the iteration matrices in (2.9):

$$\widetilde{A} = \left\{ \begin{array}{rcl} M - N & = & (I+S)(I-L) - (I+S)U, \\ M' - N' & = & (I - L - L_s + S_L) - (D_s + U + U_s - S_U), \\ M'' - N'' & = & (I - D_s - L - L_s + S_L) - (U + U_s - S_U) \end{array} \right. \tag{2.23}$$

where we have used the same symbols for the two matrices of each splitting as in the case of (2.7). So, the last two splittings in (2.23) are regular splittings because $M'^{-1} = (I -$

$L - L_s + S_L)^{-1} = I + (L + L_s - S_L) + \cdots + (L + L_s - S_L)^{n-1} \geq 0$, $N' \geq 0$ and $M''^{-1} = (I - D_s - L - L_s + S_L)^{-1} \geq 0$, $N'' \geq 0$ and so they are convergent. Since $M''^{-1} \geq M'^{-1}$, it is implied (see [16]) that the left inequality in (2.12) is true.

To prove the second inequality of (2.12) we consider first, that the Jacobi matrix B is irreducible. For the nonnegative Gauss-Seidel iteration matrix H there exists a nonnegative vector z such that

$$Hz = \rho(H)z \text{ or } (I - L)^{-1}Uz = \rho(H)z \text{ or } (\rho(H)L + U)z = \rho(H)z. \tag{2.24}$$

We observe here that the matrix $\rho(H)L + U$ has the same structure as the matrix B and consequently it is an irreducible matrix. So, from the Perron-Frobenius Theorem (see Varga [15]), the eigenvector z will be a positive vector. Recalling now the relation (2.22), the following property holds: There exists a positive vector z such that $\rho(H)z \geq H'z$. On this property, we can apply Lemma 3.3 by Marek and Szyld ([12]) to get the second inequality in (2.12). In the case where B is reducible, we consider a small number $\epsilon > 0$ and replace some zeros of B with ϵ such that the produced matrix $B(\epsilon)$ will be an irreducible matrix. Then, for the associated matrices $H(\epsilon)$ and $H'(\epsilon)$ there holds: $\rho(H'(\epsilon)) \leq \rho(H(\epsilon))$. Since the spectral radius is a continuous function of the elements of the matrix, the inequality above will also hold in the limit as ϵ tends to zero, which is the second inequality in (2.12).

a) (2.13): Since A is a nonsingular $M-$matrix, the rightmost inequality is a straightforward implication of the Stein-Rosenberg Theorem as was mentioned before. The other two inequalities in (2.13) are implied directly by the facts that $\widetilde{A}$ is a nonsingular $M-$matrix, and the last two pairs of splittings in (2.7) and (2.23), from which the four matrices involved, $\widetilde{H}$, $\widetilde{B}$, H', B', are produced, are regular ones with $L + L_s + U + U_s - S \geq U + U_s - S_U$ and $D_s + L + L_s + U + U_s - S \geq D_s + U + U_s - S_U$. It is noted here that if $\rho(\widetilde{B}) = 0$ then $\rho(\widetilde{H}) = 0$ and the matrix $\widetilde{B}$ would be reducible with its canonical form being a strictly upper triangular matrix. On the other hand if $\rho(\widetilde{B}) = \rho(\widetilde{H})$, from the Stein-Rosenberg theorem we have either $\rho(\widetilde{B}) = \rho(\widetilde{H}) = 0$ or $\rho(\widetilde{B}) = \rho(\widetilde{H}) = 1$. Since $\rho(\widetilde{B}) < 1$, both spectral radii would be zero. For the second inequality of (2.13) we have that the matrix $\widetilde{B}$ has the same structure as the matrix B'. So, if $\rho(\widetilde{B}) = 0$, the matrix B' would be reducible with its reducible canonical form being an upper triangular matrix and its spectral radius being obtained from its diagonal elements. This means that in the directed graph $\mathcal{G}(B')$ of the matrix B', there is not any strongly connected subpath except the identity paths (loops) corresponding to the nonzero diagonal elements. For the matrix H' we have that $H' = (I - L - L_s + S_L)^{-1}(D_s + U + U_s - S_U) = (I + (L + L_s - S_L) + \cdots + (L + L_s - S_L)^{n-1})(D_s + U + U_s - S_U) = D_s + [U + U_s - S_U + ((L + L_s - S_L) + \cdots + (L + L_s - S_L)^{n-1})(D_s + U + U_s - S_U)]$. Since the matrix in the brackets is a sum of products of nonnegative parts of the matrix B', it holds that if there exists a path in the graph of this matrix, then there exists also such a path, of some order, in the graph $\mathcal{G}(B')$. So, if there exists a strongly connected subpath in the graph of the matrix in brackets, then it will exist also such a subpath in $\mathcal{G}(B')$. This means that the matrix H' has also its canonical form being a strictly upper triangular matrix, with the diagonal elements being those of B'. This proves our assertion that $\rho(B') = \rho(H')$.

b) (2.10): Since B is irreducible, The eigenvector y, corresponding to $\rho(B)$, is positive and according to the steps in the proof of (2.10) in (a) we can see that inequality (2.18) becomes a strict one.

b) (2.11): From the inequality $B'y < By$ we get $B'y < \rho(B)y$. Now we can apply Lemma 3.3 by Marek and Szyld ([12]) to get the strict inequality $\rho(B') < \rho(B)$. □

2.1 Best Jacobi Preconditioner

It was proved that the preconditioned Jacobi method converges for each choice of the matrix S and it converges faster than the initial Jacobi method. There is a question now: Is there an optimum choice of the matrix S such that the associated method will be an optimum one and if so how can one choose the matrix S? This question has not been answered yet and constitutes an open and very difficult problem. It is difficult since we have to compare the spectral radii of $(n-1)^n$ different matrices. So, instead of this we will try to answer the easier question: Is there a "good" choice of the matrix S such that the associated method will be the best among many others and maybe it will be the optimum one? We try to answer this question by working on sufficient conditions of convergence instead of sufficient and necessary ones. So, we choose the matrix S such that it will minimize the maximum norm of $\widetilde{B}$ which constitutes an upper bound for its spectral radius.

For the Jacobi iteration matrix, $\widetilde{B}$ to converge a sufficient condition is

$$\rho\left(\widetilde{B}\right) \leq \|\widetilde{B}\|_\infty \Longleftrightarrow \rho\left(\widetilde{B}\right) \leq \max_i \frac{\widetilde{l}_i + \widetilde{u}_i}{\widetilde{d}_i} < 1, \tag{2.25}$$

where

$$\begin{array}{l} \widetilde{d}_i = |\widetilde{a}_{ii}| = \widetilde{a}_{ii} = 1 - a_{ik_i}a_{k_ii}, \\ \widetilde{l}_i = \sum_{j=1}^{i-1} |\widetilde{a}_{ij}| = -\sum_{j=1}^{i-1} \widetilde{a}_{ij} = a_{ik_i}\sum_{j=1}^{i-1} a_{k_ij} - \sum_{j=1}^{i-1} a_{ij}, \\ \widetilde{u}_i = \sum_{j=i+1}^{n} |\widetilde{a}_{ij}| = -\sum_{j=i+1}^{n} \widetilde{a}_{ij} = a_{ik_i}\sum_{j=i+1}^{n} a_{k_ij} - \sum_{j=i+1}^{n} a_{ij}. \end{array} \tag{2.26}$$

We propose the following method: For each row i we choose k_i such that all the ratios $\frac{\widetilde{l}_i + \widetilde{u}_i}{\widetilde{d}_i}$ will be minimized, so, the maximum of them will also be minimized. The choice of k_i is not unique and so, the choice of S is also not unique. We conjecture here that since we minimize all the ratios for each row, which are the row sums of the nonnegative matrix $\widetilde{B}$, the new spectral radius will be as small as possible. We call this method the Best Jacobi preconditioned method and the associated preconditioner, $I+S$, the best Jacobi preconditioner. From (2.26) we get that

$$\widetilde{l}_i + \widetilde{u}_i = -\sum_{j=1,j\neq i}^{n} \widetilde{a}_{ij} = a_{ik_i}\sum_{j=1,j\neq i}^{n} a_{k_ij} - \sum_{j=1,j\neq i}^{n} a_{ij} = s_i + a_{ik_i}(1 - s_{k_i} - a_{k_ii}), \tag{2.27}$$

where $s_i = -\sum_{j=1,j\neq i}^{n} a_{ij}$ is the i^{th} row sum of B. From the nonnegativity of the Jacobi matrix B and from the diagonally dominance of A it is obvious that $0 < s_i < 1$. So, the ratios in question are given by

$$\frac{\widetilde{l}_i + \widetilde{u}_i}{\widetilde{d}_i} = \frac{s_i + a_{ik_i}(1 - s_{k_i} - a_{k_ii})}{1 - a_{ik_i}a_{k_ii}}. \tag{2.28}$$

We have to give an efficient algorithm which will choose the indices k_i and consequently the matrix S. It is well known that every iteration of the Jacobi method requires $\mathcal{O}(n^2)$ ops.

The same number of operations are required for the matrix-matrix multiplications $(I+S)A$. So, for the algorithm to be efficient regarding the cost of the choice of S, it must require at most $\mathcal{O}(n^2)$ ops. First, we have to compute all the row sums s_i which require a total number of $\mathcal{O}(n^2)$ ops. Then, we have to compute the ratios (2.28) for every i and every k_i. The total number of ratios is $(n-1)n$ and so the number of required operations for each one is $\mathcal{O}(1)$. The number of comparisons is also $\mathcal{O}(n^2)$ and so, the total cost of the algorithm is $\mathcal{O}(n^2)$ ops. This makes it an efficient one. The previous analysis of the cost is shown much clearer in the following algorithm written in a pseudocode form.

Algorithm of Best Jacobi Method

```
for i = 1(1)n
    s_i = 0
    for j = 1(1)i - 1
        s_i = s_i - a_ij
    endfor
    for j = i + 1(1)n
        s_i = s_i - a_ij
    endfor
endfor
for i = 1(1)n
    r = 1
    for j = 1(1)n
        if j ≠ i then
            t = (s_i + a_ij(1 - s_j - a_ji)) / (1 - a_ij a_ji)
            if t < r then
                r = t
                k_i = j
            endif
        endif
    endfor
endfor
End of Algorithm
```

We remark that, in case there are multiple choices of the matrix S, this algorithm chooses the one with the smallest value for each k_i.

2.2 Best Gauss-Seidel Preconditioner

The same questions, as before in the Jacobi method, are also raised in the Gauss-Seidel method. Trying to answer them, we arrive at the same conclusions, that is, the matrix S should be chosen by minimizing the maximum norm of the preconditioned Gauss-Seidel iteration matrix $\widetilde{H}$. The associated sufficient condition is ([7], [8])

$$\rho(\widetilde{H}) \leq \|\widetilde{H}\|_\infty \Longrightarrow \rho(\widetilde{H}) \leq \max_i \frac{\widetilde{u}_i}{\widetilde{d}_i - \widetilde{l}_i} < 1. \tag{2.29}$$

¿From (2.26) we have that

$$\begin{array}{l} \widetilde{d}_i - \widetilde{l}_i = \sum_{j=1}^{i} \widetilde{a}_{ij} = 1 + \sum_{j=1}^{i-1} a_{ij} - a_{ik_i} \sum_{j=1}^{i-1} a_{k_i j}, \\ \widetilde{u}_i = - \sum_{j=i+1}^{n} \widetilde{a}_{ij} = a_{ik_i} \sum_{j=i+1}^{n} a_{k_i j} - \sum_{j=i+1}^{n} a_{ij}. \end{array} \tag{2.30}$$

In this case we have one more difficulty since we have to compute all the subsums from 1 to i and from $i+1$ to n for all rows in each step, while in the Jacobi case we had to compute only the row sums. If we do it, the cost of the algorithm will become greater than $\mathcal{O}(n^2)$, and the algorithm will not be an efficient one. So, we have to be more careful in order to reduce the cost to $\mathcal{O}(n^2)$. Here we introduce the following notations:

$$f_{ji} = - \sum_{k=1, k \neq j}^{i} a_{jk} \ , \quad s_{ji} = - \sum_{k=i+1, k \neq j}^{n} a_{jk}, \quad i, j = 1(1)n. \tag{2.31}$$

In view of (2.31), the first relation of (2.30) takes the form

$$\widetilde{d}_i - \widetilde{l}_i = \left\{ \begin{array}{c} 1 - f_{ii} + a_{ik_i}(f_{k_i i} - 1), \text{ if } k_i < i \\ 1 - f_{ii} + a_{ik_i} f_{k_i i}, \text{ if } k_i > i \end{array} \right. \tag{2.32}$$

while the second relation of (2.30) takes the form

$$\widetilde{u}_i = \left\{ \begin{array}{c} s_{ii} - a_{ik_i} s_{k_i i}, \text{ if } k_i < i \\ s_{ii} - a_{ik_i}(s_{k_i i} - 1), \text{ if } k_i > i \end{array} \right. . \tag{2.33}$$

So, the ratio $\frac{\widetilde{u}_i}{\widetilde{d}_i - \widetilde{l}_i}$ is given by

$$\frac{\widetilde{u}_i}{\widetilde{d}_i - \widetilde{l}_i} = \left\{ \begin{array}{l} \frac{s_{ii} - a_{ik_i} s_{k_i i}}{1 - f_{ii} + a_{ik_i}(f_{k_i i} - 1)}, \text{ if } k_i < i \\ \frac{s_{ii} - a_{ik_i}(s_{k_i i} - 1)}{1 - f_{ii} + a_{ik_i} f_{k_i i}}, \text{ if } k_i > i \end{array} \right. . \tag{2.34}$$

We can observe here that in the search of k_1 we need only the row sums f_{i1} and s_{i1} for all rows $i = 1(1)n$, in the search of k_2 we need only the row sums f_{i2} and s_{i2} and so on. So, to construct an efficient algorithm we use the following technique: First, we compute all the off-diagonal row sums denoted by s_{i0} ($s_{i0} = - \sum_{k=1, k \neq i}^{n} a_{ik}$), $i = 1(1)n$, and set $f_{i0} = 0$, $i = 1(1)n$. Then, we can see that the following relations hold:

$$\begin{array}{c} f_{ji} = f_{j,i-1} - a_{ji} \ , \quad s_{ji} = s_{j,i-1} + a_{ji} \ , \quad i, j = 1(1)n \ , \ i \neq j, \\ f_{ii} = f_{i,i-1} \ , \quad s_{ii} = s_{i,i-1} \ , \quad i = 1(1)n \ , \end{array} \tag{2.35}$$

which means that form one step to the next only one addition is required to compute the f_{ji}'s and s_{ji}'s. This observation reduces the cost of the algorithm to $\mathcal{O}(n^2)$ ops. Since $\widetilde{u}_n = 0$, the ratio $\frac{\widetilde{u}_n}{\widetilde{d}_n - \widetilde{l}_n}$ is also equal to zero. In this case, the sufficient condition does not give us any information for the choice of k_n. We can choose it randomly or by using the best Jacobi algorithm, since $\rho(\widetilde{H}) \leq \rho(\widetilde{B}) \leq \max_i \frac{\widetilde{l}_i + \widetilde{u}_i}{\widetilde{d}_i}$. For $i = n-1$ we can see that we have the unique choice $k_{n-1} = n$ in which case the ratio in question takes the value zero. So we have $n-2$ searching steps. In pseudocode, the Best Gauss-Seidel Algorithm is as follows.

Algorithm of Best Gauss-Seidel Method

```
for i = 1(1)n
    s_i = 0
    f_i = 0
    for j = 1(1)i - 1
        s_i = s_i - a_ij
    endfor
    for j = i + 1(1)n
        s_i = s_i - a_ij
    endfor
endfor
for i = 1(1)n - 2
    r = s_i / (1 - f_i)
    k_i = 0
    for j = 1(1)i - 1
        s_j = s_j + a_ji
        f_j = f_j - a_ji
        t = (s_i - a_ij s_j) / (1 - f_i + a_ij (f_j - 1))
        if t < r then
            r = t
            k_i = j
        endif
    endfor
    for j = i + 1(1)n
        s_j = s_j + a_ji
        f_j = f_j - a_ji
        t = (s_i - a_ij (s_j - 1)) / (1 - f_i + a_ij f_j)
        if t < r then
            r = t
            k_i = j
        endif
    endfor
endfor
k_{n-1} = n
End of Algorithm
```

The same remark, as in the Best Jacobi algorithm can also be made here. Moreover, we can see that the parameters f_j and s_j were used here instead of f_{ji} and s_{ji}, respectively. This is because, only two positions of memory are needed to store the quantities f_{ji} and s_{ji} for each row. Furthermore, it is seen in the algorithm that, before the beginning of the search, in each row, we give $\frac{u_i}{d_i - l_i}$ as the initial value to the ratio (the value of the ratio without any improvement) and $k_i = 0$. This is done in order to cover the case where the ratio is already the smallest one, meaning that no improvement on the row i is required. In this case we have as output $k_i = 0$.

3 Generalized Preconditioners based on multiple elimination

In this section we will generalize and extend our improved method by eliminating two or more off-diagonal elements in each row. So, the matrix S, introduced in (2.2), will have more than one elements in each row, at exactly the same positions as the elements we want to eliminate. First, we consider that in the i^{th} row we have to eliminate the elements k_i and l_i, where, without loss of generality, $k_i < l_i$. For this we have to compute the elements s_{ik_i} and s_{il_i} of the matrix S. Denoting $\widetilde{A} = (I + S)A$ we have the equations:

$$\begin{array}{rcll} \widetilde{a}_{ik_i} = & 0 & = a_{ik_i} + s_{ik_i} + s_{il_i}a_{l_ik_i} \\ \widetilde{a}_{il_i} = & 0 & = a_{il_i} + s_{ik_i}a_{k_il_i} + s_{il_i} \end{array} \Leftrightarrow \begin{array}{rcl} s_{ik_i} + s_{il_i}a_{l_ik_i} & = & -a_{ik_i} \\ s_{ik_i}a_{k_il_i} + s_{il_i} & = & -a_{il_i} \end{array} \tag{3.1}$$

or

$$\begin{gathered} (s_{ik_i} \;\; s_{il_i}) \begin{pmatrix} 1 & a_{k_il_i} \\ a_{l_ik_i} & 1 \end{pmatrix} = -(a_{ik_i} \;\; a_{il_i}) \quad \Leftrightarrow \\ (s_{ik_i} \;\; s_{il_i}) = -(a_{ik_i} \;\; a_{il_i}) \begin{pmatrix} 1 & a_{k_il_i} \\ a_{l_ik_i} & 1 \end{pmatrix}^{-1}. \end{gathered} \tag{3.2}$$

We will try to generalize the above relations by considering that m elements of the i^{th} row are to be eliminated. For this we give the following definitions. Let $\hat{k}_i^T = (k_{i_1} k_{i_2} \dots k_{i_m})$ be a multiindex, where the indices $k_{i_1} < k_{i_2} < \dots < k_{i_m}$ denote the positions of the elements of row i to be eliminated. Then, we define by $s^T_{i\hat{k}_i} = (s_{ik_{i_1}} s_{ik_{i_2}} \dots s_{ik_{i_m}})$ the vector of nonzero off-diagonal elements of the i^{th} row of S, $a^T_{i\hat{k}_i} = (a_{ik_{i_1}} a_{ik_{i_2}} \dots a_{ik_{i_m}})$ the vector of the elements of the i^{th} row of A to be eliminated and the matrix

$$A_{\hat{k}_i} = \begin{bmatrix} 1 & a_{k_{i_1}k_{i_2}} & \cdots & a_{k_{i_1}k_{i_m}} \\ a_{k_{i_2}k_{i_1}} & 1 & \cdots & a_{k_{i_2}k_{i_m}} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k_{i_m}k_{i_1}} & a_{k_{i_m}k_{i_2}} & \cdots & 1 \end{bmatrix}, \tag{3.3}$$

which consists of the rows and columns of A indexed by the multiindex $\hat{k}_i$. The matrix $A_{\hat{k}_i}$ is a principal submatrix of the $M-$matrix A. So, $A_{\hat{k}_i}$ is also an $M-$matrix and its inverse $A^{-1}_{\hat{k}_i}$ is a positive matrix if $A_{\hat{k}_i}$ is irreducible or a nonnegative one if $A_{\hat{k}_i}$ is reducible. Therefore, relation (3.2) takes the following generalized form:

$$s^T_{i\hat{k}_i} = -a^T_{i\hat{k}_i} A^{-1}_{\hat{k}_i}. \tag{3.4}$$

After the previous notations and considerations the elements $\widetilde{a}_{ij}$ of $\widetilde{A}$ are as follows:

$$\widetilde{a}_{ij} = \begin{cases} a_{ij} - a^T_{i\hat{k}_i} A^{-1}_{\hat{k}_i} a_{\hat{k}_i j} < 0, & j \neq i, j \notin \hat{k}_i, \\ 0, & j \in \hat{k}_i, \\ 1 - a^T_{i\hat{k}_i} A^{-1}_{\hat{k}_i} a_{\hat{k}_i i} > 0, & j = i. \end{cases} \tag{3.5}$$

The first inequality in (3.5) is obvious while the last one is to be proved. For this we denote by $A_{ij;kl}$ the submatrix of A after the deletion of its i^{th} and j^{th} rows as well as its k^{th} and l^{th} columns. We give now the next lemma which is needed for the proof.

Lemma 3.1 *Let the matrix $A \in \mathbb{R}^{n,n}$ and the multiindex $\hat{\imath} = (2, 3, \ldots, n)$, then there holds*

$$\det(A) = a_{11}\det(A_{1;1}) - a_{1\hat{\imath}}^T adj(A_{1;1}) a_{\hat{\imath}1} \tag{3.6}$$

Proof: It is easy to see that the $(j,i)^{th}$ element of the adjoint matrix $(adj(A_{1;1}))$ of $A_{1;1}$ is equal to $(-1)^{i+j}\det(A_{1j;1i})$. So, by expanding $\det(A)$, first along its first row and then along its first column we get:

$$\begin{aligned} \det(A) &= \textstyle\sum_{i=1}^n (-1)^{i+1} a_{1i}\det(A_{1;i}) \\ &= a_{11}\det(A_{1;1}) + \textstyle\sum_{i=2}^n (-1)^{i+1} a_{1i} \sum_{j=2}^n (-1)^j a_{j1}\det(A_{1j;1i}) \\ &= a_{11}\det(A_{1;1}) - \textstyle\sum_{i=2}^n \sum_{j=2}^n (-1)^{i+j} a_{1i} a_{j1}\det(A_{1j;1i}) \\ &= a_{11}\det(A_{1;1}) - a_{1\hat{\imath}} adj(A_{1;1}) a_{\hat{\imath}1} \end{aligned} \tag{3.7}$$

□

We consider the matrix

$$\widetilde{A}_{\hat{k}_i} = \left(\begin{array}{c|c} 1 & a_{i\hat{k}_i}^T \\ \hline a_{\hat{k}_i i} & A_{\hat{k}_i} \end{array} \right) \tag{3.8}$$

which is an extension of $A_{\hat{k}_i}$ by adding the i^{th} row and column as the first row and column, respectively. Since the matrices $\widetilde{A}_{\hat{k}_i}$ and $A_{\hat{k}_i}$ are two principal submatrices of the $M-$matrix A they are also $M-$matrices and so they have positive determinants. We use now the well known relation which connects the inverse of a matrix with respect to its adjoint matrix and its determinant and the previous Lemma to get

$$1 - a_{i\hat{k}_i}^T A_{\hat{k}_i}^{-1} a_{\hat{k}_i i} = 1 - a_{i\hat{k}_i}^T \frac{adj(A_{\hat{k}_i})}{\det(A_{\hat{k}_i})} a_{\hat{k}_i i} = \frac{\det(A_{\hat{k}_i}) - a_{i\hat{k}_i}^T adj(A_{\hat{k}_i}) a_{\hat{k}_i i}}{\det(A_{\hat{k}_i})} = \frac{\det(\widetilde{A}_{\hat{k}_i})}{\det(A_{\hat{k}_i})} > 0. \tag{3.9}$$

We define the matrix

$$D_s := \mathrm{diag}(a_{1\hat{k}_1}^T A_{\hat{k}_1}^{-1} a_{\hat{k}_1 1},\ a_{2\hat{k}_2}^T A_{\hat{k}_2}^{-1} a_{\hat{k}_2 2},\ \ldots, a_{n\hat{k}_n}^T A_{\hat{k}_n}^{-1} a_{\hat{k}_n n}), \tag{3.10}$$

which is the diagonal part of the matrix $S(L+U)$. We use the same notations (2.5) and (2.6), we consider the same splittings (2.7), and the associated Jacobi (2.8) and the Gauss-Seidel schemes (2.9). So, we can give and prove the theorem below, which is the generalization of Theorem 2.1.

Theorem 3.1 *a) Under the assumptions and the notation so far, there hold:*
There exist y and $z \in \mathbb{R}^n$, with $y \geq 0$ and $z \geq 0$, such that

$$B'y \leq By \text{ and } H'z \leq Hz, \tag{3.11}$$

$$\rho\left(\widetilde{B}\right) \leq \rho(B') \leq \rho(B) < 1, \tag{3.12}$$

$$\rho\left(\widetilde{H}\right) \leq \rho(H') \leq \rho(H) < 1, \tag{3.13}$$

$$\rho\left(\widetilde{H}\right) \leq \rho\left(\widetilde{B}\right),\quad \rho(H') \leq \rho(B'),\quad \rho(H) < \rho(B) < 1. \tag{3.14}$$

(Note: Equalities in (3.14) hold if and only if $\rho(\widetilde{B}) = 0$.)
b) Suppose that A is irreducible. Then, the matrix B is also irreducible which implies that the first inequality in (3.11) and the middle inequality in (3.12) are strict.

Proof: a) (3.11): To prove (3.11) we use the explicit expressions of the elements for the two Jacobi and Jacobi type iteration matrices. The elements of the Jacobi matrix B are the same as those given in (2.14) while those of type Jacobi matrix B' are modified as follows:

$$\begin{cases} b'_{ij} = 0, & j \in \hat{k}_i, \\ b'_{ii} = a^T_{i\hat{k}_i} A^{-1}_{\hat{k}_i} a_{\hat{k}_i j} = b^T_{i\hat{k}_i} A^{-1}_{\hat{k}_i} b_{\hat{k}_i j}, & \\ b'_{ij} = a^T_{i\hat{k}_i} A^{-1}_{\hat{k}_i} a_{\hat{k}_i j} - a_{ij} = b^T_{i\hat{k}_i} A^{-1}_{\hat{k}_i} b_{\hat{k}_i j} + b_{ij}, & j \neq i, j \notin \hat{k}_i \end{cases} \tag{3.15}$$

where $b^T_{i\hat{k}_i}$ and $b_{\hat{k}_i j}$ denote the row vector with elements the ones corresponding to the multiindex $\hat{k}_i$ of the i^{th} row of B and the column vector with elements the ones corresponding to the multiindex $\hat{k}_i$ of the j^{th} column of B.

We follow the same steps as in the proof of Theorem 2.1. So,

$$\begin{aligned} \rho(B)y_i &= \textstyle\sum_{j=1, j \notin \{i,\hat{k}_i\}}^n b_{ij}y_j + b^T_{i\hat{k}_i} y_{\hat{k}_i} \\ &= \textstyle\sum_{j=1, j \notin \{i,\hat{k}_i\}}^n b'_{ij}y_j - b_{i\hat{k}_i} A^{-1}_{\hat{k}_i} \sum_{j=1, j \notin \{i,\hat{k}_i\}}^n b_{\hat{k}_i j}y_j + b^T_{i\hat{k}_i} A^{-1}_{\hat{k}_i} A_{\hat{k}_i} y_{\hat{k}_i} \\ &= \textstyle\sum_{j=1, j \neq i}^n b'_{ij}y_j - b_{i\hat{k}_i} A^{-1}_{\hat{k}_i} \sum_{j=1, j \notin \{i,\hat{k}_i\}}^n b_{\hat{k}_i j}y_j + b^T_{i\hat{k}_i} A^{-1}_{\hat{k}_i} y_{\hat{k}_i} - b^T_{i\hat{k}_i} A^{-1}_{\hat{k}_i} B_{\hat{k}_i} y_{\hat{k}_i} \\ &= \textstyle\sum_{j=1}^n b'_{ij}y_j - b^T_{i\hat{k}_i} A^{-1}_{\hat{k}_i} \sum_{j=1}^n b_{\hat{k}_i j}y_j + b^T_{i\hat{k}_i} A^{-1}_{\hat{k}_i} y_{\hat{k}_i}. \end{aligned} \tag{3.16}$$

Using the fact that $\rho(B)y_{\hat{k}_i} = \sum_{j=1}^n b_{\hat{k}_i j}y_j$, pre-multiplying it b $b^T_{i\hat{k}_i} A^{-1}_{\hat{k}_i}$ and replacing in (3.16) we have that

$$\rho(B)y_i = \sum_{j=1}^n b'_{ij}y_j + \left(\frac{1}{\rho(B)} - 1\right) b^T_{i\hat{k}_i} A^{-1}_{\hat{k}_i} \sum_{j=1}^n b_{\hat{k}_i j}y_j. \tag{3.17}$$

Since the second term of the sum in (3.17) is nonnegative we have that

$$\sum_{j=1}^n b'_{ij}y_j \leq \sum_{j=1}^n b_{ij}y_j \tag{3.18}$$

from which the first of (3.11) follows.

For the nonnegative Gauss-Seidel iteration matrix H we use analogous considerations to get the generalized version of (2.19) and analogous notations as in the nonnegative Jacobi case. Moreover we define the multiindex $\hat{k}'_i$ which consists of the indices of $\hat{k}_i$ that are greater than i. Then, the inner product $b^T_{i\hat{k}'_i} z_{\hat{k}'_i}$, which will appear in the sequel, can be replaced by $b^T_{i\hat{k}_i} \tilde{z}_{\hat{k}_i}$. As $\tilde{z}_{\hat{k}_i}$ we define the multiindexed by $\hat{k}_i$ vector which consists of zeros in the places where the index is less than i and of the elements of $z_{\hat{k}'_i}$ in the places where the index is greater than i. For simplicity we use ρ instead of $\rho(H)$. After these notations we can get:

$$\begin{aligned} \rho z_i &= \rho \textstyle\sum_{j=1, j \notin \hat{k}_i}^n b_{ij}z_j + (1-\rho) \sum_{j=i+1, j \notin \hat{k}_i}^n b_{ij}z_j + \rho b^T_{i\hat{k}_i} z_{\hat{k}_i} + (1-\rho) b^T_{i\hat{k}'_i} z_{\hat{k}'_i} \\ &= \rho \textstyle\sum_{j=1}^n b'_{ij}z_j + (1-\rho) \sum_{j=i}^n b'_{ij}z_j - \rho b^T_{i\hat{k}_i} A^{-1}_{\hat{k}_i} \sum_{j=1, j \notin \hat{k}_i}^n b_{\hat{k}_i j}z_j \\ &- (1-\rho) b^T_{i\hat{k}_i} A^{-1}_{\hat{k}_i} \textstyle\sum_{j=i, j \notin \hat{k}_i}^n b_{\hat{k}_i j}z_j + \rho b^T_{i\hat{k}_i} A^{-1}_{\hat{k}_i} A_{\hat{k}_i} z_{\hat{k}_i} + (1-\rho) b^T_{i\hat{k}_i} A^{-1}_{\hat{k}_i} A_{\hat{k}_i} \tilde{z}_{\hat{k}_i} \\ &= \rho \textstyle\sum_{j=1}^n b'_{ij}z_j + (1-\rho) \sum_{j=i}^n b'_{ij}z_j - \rho b^T_{i\hat{k}_i} A^{-1}_{\hat{k}_i} \sum_{j=1}^n b_{\hat{k}_i j}z_j \\ &- (1-\rho) b^T_{i\hat{k}_i} A^{-1}_{\hat{k}_i} \textstyle\sum_{j=i}^n b_{\hat{k}_i j}z_j + \rho b^T_{i\hat{k}_i} A^{-1}_{\hat{k}_i} z_{\hat{k}_i} + (1-\rho) b^T_{i\hat{k}_i} A^{-1}_{\hat{k}_i} \tilde{z}_{\hat{k}_i}. \end{aligned} \tag{3.19}$$

Using the facts that the l^{th} component of $\rho z_{\hat{k}_i}$ is given by $\rho z_{k_{i_l}} = \rho \sum_{j=1}^{n} b_{k_{i_l}j} z_j + (1-\rho)\sum_{j=k_{i_l}+1}^{n} b_{k_{i_l}j} z_j$ we have that $\rho z_{\hat{k}_i} = \rho \sum_{j=1}^{n} b_{\hat{k}_ij} z_j + (1-\rho) y_{\hat{k}_i}$ where $y_{k_{i_l}} = \sum_{j=k_{i_l}+1}^{n} b_{k_{i_l}j} z_j$. The vector $\rho \tilde{z}_{\hat{k}_i}$ has the components of $\rho z_{\hat{k}_i}$ in the places where the index is greater than i and zeros elsewhere. So, pre-multiplying these vectors by $b_{i\hat{k}_i}^T A_{\hat{k}_i}^{-1}$ and replacing into (3.20) we have that

$$\begin{array}{rcl} \rho z_i & = & \rho \sum_{j=1}^{n} b'_{ij} z_j + (1-\rho) \sum_{j=i}^{n} b'_{ij} z_j \\ & + & b_{i\hat{k}_i}^T A_{\hat{k}_i}^{-1} (-\rho \sum_{j=1}^{n} b_{\hat{k}_ij} z_j - (1-\rho) \sum_{j=i}^{n} b_{\hat{k}_ij} z_j + \rho \sum_{j=1}^{n} b_{\hat{k}_ij} z_j \\ & + & (1-\rho) y_{\hat{k}_i} + (1-\rho) \sum_{j=1}^{n} \tilde{b}_{\hat{k}_ij} z_j + (\frac{1}{\rho} - 1) \tilde{y}_{\hat{k}_i}) \\ & = & \rho \sum_{j=1}^{n} b'_{ij} z_j + (1-\rho) \sum_{j=i}^{n} b'_{ij} z_j \\ & + & b_{i\hat{k}_i}^T A_{\hat{k}_i}^{-1} \underbrace{\left((1-\rho) \left(-\sum_{j=i}^{n} b_{\hat{k}_ij} z_j + y_{\hat{k}_i} + \sum_{j=1}^{n} \tilde{b}_{\hat{k}_ij} z_j \right) + \left(\frac{1}{\rho} - 1 \right) \tilde{y}_{\hat{k}_i} \right)}, \end{array} \tag{3.20}$$

where the vectors $\tilde{b}_{\hat{k}_i}$ and $\tilde{y}_{\hat{k}_i}$ have the components of $b_{\hat{k}_i}$ and $y_{\hat{k}_i}$, respectively, in the places where the index is greater than i and zeros elsewhere. The l^{th} component of the underbraced term in (3.20) is given by:

$$\begin{array}{l} (1-\rho)(-\sum_{j=i}^{n} b_{k_{i_l}j} z_j + \sum_{j=k_{i_l}+1}^{n} b_{k_{i_l}j} z_j) = (1-\rho) \sum_{j=k_{i_l}+1}^{i-1} b_{k_{i_l}j} z_j \geq 0, \quad k_{i_l} < i \\ \\ (1-\rho)(-\sum_{j=i}^{n} b_{k_{i_l}j} z_j + \sum_{j=k_{i_l}+1}^{n} b_{k_{i_l}j} z_j + \sum_{j=1}^{n} b_{k_{i_l}j} z_j + (\frac{1}{\rho} - 1) \sum_{j=k_{i_l}+1}^{n} b_{k_{i_l}j} z_j) \\ = (1-\rho)(\sum_{j=1}^{i-1} b_{k_{i_l}j} z_j + \sum_{j=k_{i_l}+1}^{n} b_{k_{i_l}j} z_j) + (\frac{1}{\rho} - 1) \sum_{j=k_{i_l}+1}^{n} b_{k_{i_l}j} z_j \geq 0, \quad k_{i_l} > i \end{array} \tag{3.21}$$

So, the unterbraced term of (3.20) is a nonnegative vector. Hence, we get that

$$\rho(z_i - \sum_{j=1}^{i-1} b'_{ij} z_j) \geq \sum_{j=i}^{n} b'_{ij} z_j \,, \quad i = 1(1)n. \tag{3.22}$$

This relation is equivalent, in terms of matrices, to

$$\begin{array}{ll} \rho(I - L - L_s + S_L) z \geq (D_s + U + U_s - S_U) z & \text{or} \\ \rho z \geq (I - L - L_s + S_L)^{-1} (D_s + U + U_s - S_U) z & \text{or} \quad \rho z \geq H' z \end{array} \tag{3.23}$$

from which the second of (3.11) follows.
If we use the same notations, the proof of Theorem 3.1 for all the other inequalities follows step by step that of Theorem 2.1, and so, the proof of the present statement is complete. □

¿From now on we have to prove convergence theorems of the proposed improved methods when compared to each other. For the multiindexed improved method, it remains to study the best choice of the multiindex $\hat{k}_i$. This is a very difficult problem, more difficult than that of the "best" choice corresponding to one elimination per row. However, it is possible to compare the multiindexed improved methods, in the case where the one is a subcase of the other as this is given in the following comparison theorem.

Theorem 3.2 *a) Let A be a nonsingular $M-$matrix with $a_{ii}=1$. We consider the multiindexed by $\hat{k}_i$ and $\hat{k'}_i$, $i=1(1)n$ preconditioning matrices $I+S_{\hat{k}}$ and $I+S_{\hat{k'}}$, respectively, where $\hat{k'}_i\subset\hat{k}_i$, $i=1(1)n$. Then for the associated improved type Jacobi matrices $B'_{\hat{k}}$, $B'_{\hat{k'}}$, improved type Gauss-Seidel matrices $H'_{\hat{k}}$, $H'_{\hat{k'}}$, improved Jacobi matrices $\widetilde{B}_{\hat{k}}$, $\widetilde{B}_{\hat{k'}}$ and improved Gauss-Seidel matrices $\widetilde{H}_{\hat{k}}$, $\widetilde{H}_{\hat{k'}}$, there hold:*

$$\rho(\widetilde{B}_{\hat{k}})\leq\rho(B'_{\hat{k}})\leq\rho(\widetilde{B}_{\hat{k'}})\leq\rho(B'_{\hat{k'}})<1, \tag{3.24}$$

$$\rho(\widetilde{H}_{\hat{k}})\leq\rho(H'_{\hat{k}})\leq\rho(\widetilde{H}_{\hat{k'}})\leq\rho(H'_{\hat{k'}})<1, \tag{3.25}$$

b) If the matrix $B'_{\hat{k'}}$ is an irreducible matrix then the second from the left inequality of (3.24) becomes strict.

Proof: a) The fact that all the spectral radii are less than 1 is obvious from Theorem 3.1. We will prove all the other inequalities by proving the following property:

Suppose we apply successively the multiindexed by $\hat{k}'$ preconditioned technique to the matrix A and then the multiindexed by $\hat{k}$ preconditioned technique to the matrix $I-\widetilde{B}_{\hat{k'}}$. We denote by $\widetilde{B}_{\hat{k'},\hat{k}}$, $B'_{\hat{k'},\hat{k}}$, $\widetilde{H}_{\hat{k'},\hat{k}}$ and $H'_{\hat{k'},\hat{k}}$ the associated improved Jacobi, improved type Jacobi, improved Gauss-Seidel and improved type Gauss Seidel matrices, respectively, produced in the second part of the successively improved technique. Then, there holds:

$$\widetilde{B}_{\hat{k'},\hat{k}}=\widetilde{B}_{\hat{k'}},\quad B'_{\hat{k'},\hat{k}}\geq B'_{\hat{k}} \tag{3.26}$$

$$\widetilde{H}_{\hat{k'},\hat{k}}=\widetilde{H}_{\hat{k'}},\quad H'_{\hat{k'},\hat{k}}\geq H'_{\hat{k}}. \tag{3.27}$$

If this property holds, then from (3.12) of Theorem 3.1 and from the nonnegativity of the above matrices, we get that

$$\rho(\widetilde{B}_{\hat{k}})\leq\rho(B'_{\hat{k'},\hat{k}})\leq\rho(\widetilde{B}_{\hat{k'}})\ ,\quad \rho(\widetilde{B}_{\hat{k}})\leq\rho(B'_{\hat{k}})\leq\rho(B)\ \text{ and }\ \rho(B'_{\hat{k}})\leq\rho(B'_{\hat{k'},\hat{k}}), \tag{3.28}$$

since the matrix $\widetilde{B}_{\hat{k'}}$ plays the role of the Jacobi matrix corresponding to the initial matrix $I-\widetilde{B}_{\hat{k'}}$ in the second part of the successively improving technique. The above three relations give us

$$\rho(\widetilde{B}_{\hat{k}})\leq\rho(B'_{\hat{k}})\leq\rho(\widetilde{B}_{\hat{k'}}). \tag{3.29}$$

So, relation (3.24) is proved. The inequalities (3.25) are proved in the same way from (3.13) of Theorem 3.1 and (3.27), while the assertion (b) is proved from the assertion (b) of Theorem 3.1. It remains to prove the above property.
For simplicity we define by C the matrix $I-\widetilde{B}_{\hat{k'}}$ and we use the same notations applied to the matrix C as in the second part of the successively improved technique. So,

$$\widetilde{c}_{ij}=\begin{cases} c_{ij}-c^T_{i\hat{k}_i}C^{-1}_{\hat{k}_i}c_{\hat{k}_ij}<0, & j\neq i,\ j\notin\hat{k}_i,\\ 0, & j\in\hat{k}_i,\\ 1-c^T_{i\hat{k}_i}C^{-1}_{\hat{k}_i}c_{\hat{k}_ii}>0, & j=i.\end{cases} \tag{3.30}$$

¿From (3.9) we get that

$$\widetilde{c}_{ii} = \frac{\det(\widetilde{C}_{\hat{k}_i})}{\det(C_{\hat{k}_i})} > 0, \quad i = 1(1)n. \tag{3.31}$$

Now we define by $C^{ij}_{\hat{k}_i}$ the matrix which is the extension of $C_{\hat{k}_i}$ by adding as its first row the associated elements of the i^{th} row of C and as its first column the associated elements of the j^{th} column of C. So,

$$C^{ij}_{\hat{k}_i} = \left(\begin{array}{c|c} c_{ij} & c^T_{i\hat{k}_i} \\ \hline c_{\hat{k}_i j} & C_{\hat{k}_i} \end{array} \right). \tag{3.32}$$

We use now the relation which connects the inverse of a matrix with respect to its adjoint matrix and its determinant and Lemma 3.1 to get

$$\widetilde{c}_{ij} = c_{ij} - c^T_{i\hat{k}_i} \frac{adj(C_{\hat{k}_i})}{\det(C_{\hat{k}_i})} c_{\hat{k}_i j} = \frac{c_{ij}\det(C_{\hat{k}_i}) - c^T_{i\hat{k}_i} adj(C_{\hat{k}_i}) c_{\hat{k}_i j}}{\det(C_{\hat{k}_i})} = \frac{\det(C^{ij}_{\hat{k}_i})}{\det(C_{\hat{k}_i})}. \tag{3.33}$$

The matrix C is produced from the matrix A after some elimination process corresponding to the first step of improved, and then by dividing each row with the associated diagonal elements. It is well known that the determinants as well as all the minor determinants are invariant under an elimination process. So, $\det(C^{ij}_{\hat{k}_i})$ is equal to $\det(A^{ij}_{\hat{k}_i})$ divided by the product of the associated diagonal elements of $\widetilde{A}$, corresponding to the first step, while the denominator $\det(C_{\hat{k}_i})$ is equal to $\det(A_{\hat{k}_i})$ divided by the product of the associated diagonal elements of $\widetilde{A}$. The diagonal elements corresponding to the indices of $\hat{k}_i$ coincide and are eliminated, while the element $\widetilde{a}_{ii}$ which divides the numerator remains unchanged. So,

$$\widetilde{c}_{ij} = \frac{\det(C^{ij}_{\hat{k}_i})}{\det(C_{\hat{k}_i})} = \frac{\det(A^{ij}_{\hat{k}_i})}{\det(A_{\hat{k}_i})\widetilde{a}_{ii}} = \frac{\det(A^{ij}_{\hat{k}_i})\det(A_{\hat{k}'_i})}{\det(A_{\hat{k}_i})\det(\widetilde{A}_{\hat{k}'_i})}, \quad j = 1(1)n, j \neq i, \tag{3.34}$$

while

$$\widetilde{c}_{ii} = \frac{\det(\widetilde{C}_{\hat{k}_i})}{\det(C_{\hat{k}_i})} = \frac{\det(\widetilde{A}_{\hat{k}_i})}{\det(A_{\hat{k}_i})\widetilde{a}_{ii}} = \frac{\det(\widetilde{A}_{\hat{k}_i})\det(A_{\hat{k}'_i})}{\det(A_{\hat{k}_i})\det(\widetilde{A}_{\hat{k}'_i})}. \tag{3.35}$$

The elements of $\widetilde{B}_{\hat{k}',\hat{k}}$, $B'_{\hat{k}',\hat{k}}$, as well as of $\widetilde{B}_{\hat{k}}$ and $B'_{\hat{k}}$, are given by

$$(\widetilde{B}_{\hat{k}',\hat{k}})_{ij} = \frac{-\widetilde{c}_{ij}}{\widetilde{c}_{ii}} = -\frac{\det(A^{ij}_{\hat{k}_i})\det(A_{\hat{k}'_i})\det(A_{\hat{k}_i})\det(\widetilde{A}_{\hat{k}'_i})}{\det(A_{\hat{k}_i})\det(\widetilde{A}_{\hat{k}'_i})\det(\widetilde{A}_{\hat{k}_i})\det(A_{\hat{k}'_i})} = -\frac{\det(A^{ij}_{\hat{k}_i})}{\det(\widetilde{A}_{\hat{k}_i})}, \quad j \neq i, \tag{3.36}$$

$$(B'_{\hat{k}',\hat{k}})_{ij} = \begin{cases} -\widetilde{c}_{ij} = -\frac{\det(A^{ij}_{\hat{k}_i})\det(A_{\hat{k}'_i})}{\det(A_{\hat{k}_i})\det(\widetilde{A}_{\hat{k}'_i})}, & j \neq i, \\ 1 - \widetilde{c}_{ii} = 1 - \frac{\det(\widetilde{A}_{\hat{k}_i})\det(A_{\hat{k}'_i})}{\det(A_{\hat{k}_i})\det(\widetilde{A}_{\hat{k}'_i})}, & j = i, \end{cases} \tag{3.37}$$

$$(\widetilde{B}_{\hat{k}})_{ij} = \frac{-\widetilde{a}_{ij}}{\widetilde{a}_{ii}} = -\frac{\det(A^{ij}_{\hat{k}_i})\det(A_{\hat{k}_i})}{\det(A_{\hat{k}_i})\det(\widetilde{A}_{\hat{k}_i})} = -\frac{\det(A^{ij}_{\hat{k}_i})}{\det(\widetilde{A}_{\hat{k}_i})}, \quad j \neq i, \quad \text{and} \tag{3.38}$$

$$(B'_{\hat{k}})_{ij} = \begin{cases} -\widetilde{a}_{ij} = -\frac{\det(A^{ij}_{\hat{k}_i})}{\det(A_{\hat{k}_i})}, & j \neq i, \\ 1 - \widetilde{a}_{ii} = 1 - \frac{\det(\widetilde{A}_{\hat{k}_i})}{\det(A_{\hat{k}_i})}, & j = i. \end{cases} \tag{3.39}$$

¿From relations (3.36) and (3.38) we get the equality of (3.26). From (3.9) we have that

$$\frac{\det(\widetilde{A}_{\hat{k}'_i})}{\det(A_{\hat{k}'_i})} = 1 - a^T_{i\hat{k}'_i} A^{-1}_{\hat{k}'_i} a_{\hat{k}'_i i} < 1, \tag{3.40}$$

which, in connection with (3.37) and (3.39), proves the inequality of (3.26). The equality of (3.27) is a direct consequence of the equality (3.26), while the inequality (3.27) is easily proved from the inequality of (3.26), by writing the associated Gauss-Seidel matrices in terms of the nonnegative lower and upper triangular parts of the Jacobi ones and the proof of our assertion is complete. □

We remark here that our proposed multiindexed method is the most general method, among many other improved methods, based on elimination techniques. It is a generalization of the block elimination improved method proposed by Alanelli and Hadjidimos [1], [2]. They have studied the block Milaszewicz's improved method, which eliminates the elements of the first k_1 columns of A below the diagonal. This is precisely the multiindexed method with $\hat{k}_i = (1\ 2\ 3\ \cdots\ i-1)$, $i = 1(1)k_1$ and $\hat{k}_i = (1\ 2\ 3\ \cdots\ k_1)$, otherwise.

As regards the cost of the present method we can observe that, for the construction of the matrix S, we have to solve an $m \times m$ linear system for each row which requires a cost of $\mathcal{O}(m^3 n)$ ops. The cost of the matrix-matrix product $(I+S)A$ is $\mathcal{O}(mn^2)$ ops. Then follows the standard iterative process of Jacobi or Gauss-Seidel method which requires a cost of $\mathcal{O}(n^2)$ ops per iteration. We can remark here that, to obtain an efficient algorithm, the number m must be chosen very small and independent of the dimension n. Thus a question can be raised here which sets an open problem: Which is the best choice of m? Observe that by increasing m what is gained in number of iterations is lost in the construction of S and in the matrix-matrix product. Which is then the golden section? This also depends on the matrix A. Another question that can be raised is: How can one choose the multiindices $\hat{k}_i$'s? This is very difficult to anwer. As we worked in the "best" Jacobi or Gauss-Seidel algorithm, we can also give searching algorithms by using sufficient criteria instead of sufficient and necessary ones. In the "best" case, where we had $m = 1$, the cost of the searching algorithm was $\mathcal{O}(n^2)$. If we take $m = 2$, we have to do a double searching per row, so the cost increases to $\mathcal{O}(n^3)$ and the algorithm becomes non-efficient. If the value of m is increased further, the power of n in the cost increases too. So, the only efficient searching algorithm is the one where we search along one component of the multiindex, taking the others fixed. As we will see in the numerical examples, in many cases, by taking the multiindices fixed, the multiindexed algorithm is better than that of the "best" preconditioned algorithm.

4 Numerical Examples

Below we present four representative example matrices (taken from [6]) for which the spectral radii of the corresponding iteration matrices considered are given in the subsequent two Tables for the Jacobi and Gauss-Seidel method, respectively. In the Tables, M, G, C, B and M_2 denote the Milaszewicz's, the Gunawardena et al's, the Cyclic, the "Best" and the multiindexed, with $m = 2$, preconditioner, respectively. For the multiindexed preconditioner we have fixed the multiindices by taking $\hat{k}_1 = (2 \quad n)^T$, $\hat{k}_i = (i-1 \quad i+1)^T, \quad i = 2(1)n-1$ and $\hat{k}_n = (1 \quad n-1)^T$. In other words we have eliminated the elements corresponding to the Cyclic preconditioner and their transposed ones.

$$A_1 = \begin{bmatrix} 1.00000 & -0.00580 & -0.19350 & -0.25471 & -0.03885 \\ -0.28424 & 1.00000 & -0.16748 & -0.21780 & -0.21577 \\ -0.24764 & -0.26973 & 1.00000 & -0.18723 & -0.08949 \\ -0.13880 & -0.01165 & -0.25120 & 1.00000 & -0.13236 \\ -0.25809 & -0.08162 & -0.13940 & -0.04890 & 1.00000 \end{bmatrix},$$

$$A_2 = \begin{bmatrix} 1.00000 & -0.15359 & -0.24342 & -0.02303 & -0.03363 \\ -0.01756 & 1.00000 & -0.00630 & -0.14703 & -0.18174 \\ -0.01087 & -0.03714 & 1.00000 & -0.25258 & -0.17673 \\ -0.12507 & -0.01414 & -0.07603 & 1.00000 & -0.14130 \\ -0.00515 & -0.24496 & -0.23477 & -0.27707 & 1.00000 \end{bmatrix},$$

$$A_3 = \begin{bmatrix} 1.00000 & -0.27149 & -0.20650 & -0.02972 & -0.12557 \\ -0.12416 & 1.00000 & -0.18328 & -0.07729 & -0.25528 \\ -0.31163 & -0.02827 & 1.00000 & -0.15184 & -0.39463 \\ -0.12292 & -0.00477 & -0.23299 & 1.00000 & -0.20115 \\ -0.37067 & -0.09086 & -0.20368 & -0.30835 & 1.00000 \end{bmatrix},$$

$$A_4 = \begin{bmatrix} 1.00000 & -0.23661 & -0.37369 & -0.25833 & -0.05480 \\ -0.13602 & 1.00000 & -0.10578 & -0.38675 & -0.32750 \\ -0.12569 & -0.01525 & 1.00000 & -0.26597 & -0.17207 \\ -0.14603 & -0.18344 & -0.34914 & 1.00000 & -0.35613 \\ -0.15730 & -0.34795 & -0.09515 & -0.00397 & 1.00000 \end{bmatrix}.$$

In the first column of the following two Tables we give the spectral radii of the non-preconditioned method and, in the other columns, the associated spectral radii of the corresponding improved methods, indexed in the head of the Tables.

Jacobi method

Matrix		M	G	C	B	M_2
A_1	0.629054	0.553502	0.584773	0.572500	0.553502	0.463763
A_2	0.484223	0.460575	0.418960	0.418438	0.378143	0.362226
A_3	0.758521	0.693935	0.715067	0.692129	0.690212	0.624807
A_4	0.806792	0.767901	0.763008	0.756508	0.729308	0.708140

Gauss-Seidel method

Matrix		M	G	C	B	M_2
A_1	0.384958	0.295976	0.285946	0.247030	0.258751	0.215618
A_2	0.266686	0.232881	0.160474	0.159189	0.144649	0.141635
A_3	0.603046	0.480367	0.497869	0.428684	0.405759	0.394486
A_4	0.684691	0.622791	0.568660	0.546671	0.557928	0.511027

We can remark here that, in both methods, the multiindexed preconditioner is better than all the others, the "best" preconditioner is better than the remaining ones, the cyclic preconditioner is better than the rest except in one case for the Jacobi method, the Gunawardena et al's preconditioner is better than the Milaszewicz's one for the Gauss-Seidel method, while the score of the last two preconditioners is two-to-two for the Jacobi method. Also, we can see that in the matrix A_1 and the Jacobi method, the Milaszewicz's and the "best" preconditioners are equivalent. We have checked this case and we have seen that both preconditioners coincide except that the "best" one eliminates one more element in the first row. As was noticed in the beginning of Section 2, the Milaszewicz's preconditioning matrix is reducible, the elimination of one element of the first row does not reduce the spectral radius and thus we have the equivalence observed.

For 10000 randomly generated nonsingular $M-$matrices for $n = 10$, $n = 20$ and $n = 50$ we have determined the spectral radii of the iteration matrices of all the methods mentioned previously. Below, we present two Tables with results of percentages, for the Jacobi and Gauss-Seidel methods, respectively. The numbers in the Tables represent the percentages in which the method indexed in the first column is better than the one indexed in the head of the Table.

Jacobi method

	$n=10$				$n=20$				$n=50$			
	M	G	C	B	M	G	C	B	M	G	C	B
M		60.88	48.8	6.19		58.01	48.94	1.35		54.61	48.98	0
G	39.12		0	1.05	41.99		0	0.06	45.39		0	0
C	51.2	100		2.88	51.06	100		0.19	51.02	100		0
B	93.81	98.95	97.12		98.65	99.94	99.81		100	100	100	
M_2	99.83	100	100	97.23	99.99	100	100	99.42	100	100	100	99.96

Gauss-Seidel method

	$n=10$				$n=20$				$n=50$			
	M	G	C	B	M	G	C	B	M	G	C	B
M		2.54	0.44	1.08		0.1	0.02	0		0	0	0
G	97.46		0	36.62	99.9		0	25.49	100		0	5.11
C	99.56	100		56.63	99.98	100		36.34	100	100		7.13
B	98.92	63.38	43.37		100	74.51	63.66		100	94.89	92.87	
M_2	99.75	100	100	68.39	100	100	100	44.8	100	100	100	8.95

We have to remark here that for the Jacobi method and for n large enough, the multiindexed improved method is 100% better than all the others. The "best" preconditioner is 100% better than the remaining ones, the cyclic preconditioner is 100% better than that of Gunawardena et al's, while the last two preconditioners tend to be equivalent, regarding their performance, to the one of Milaszewicz's, as n increases. For the Gauss-Seidel method we can see that the "best" preconditioner is better than all the others, then follow the multi-indexed, the cyclic, the Gunawardena et al's and finally the the Milaszewicz's preconditioner. At this point, another question is raised: Is the "best" preconditioner indeed better than the multiindexed one? The answer is **no**! It depends on the choice of the multiindices. In this example we have chosen one element over the diagonal and one under it. We observe here that the elimination of the over-diagonal elements play the most important role for the Gauss-Seidel method than that of the under-diagonal one. Since we have chosen one fixed over-diagonal element per row, for the multiindexed preconditioner, while for the "best" one we have searched one element per row (we think that the most of them are over-diagonal elements), the last preconditioner becomes better than the first one. It has been checked that the multiindexed preconditioner ($m = 2$) is 100% better than all the others for the Gauss-Seidel method if we choose both elements to be over-diagonal for each row. So, the numerical examples confirm the theoretical results for the proposed improved techniques.

References

[1] M. Alanelli, *Block Elementary Block Elimination Preconditioners for the Numerical Solution of Non-singular and Singular Linear Systems.* Master Dissertation (in Greek), Department of Mathematics, University of Crete, Heraklion, Greece, 2001.

[2] M. Alanelli and H. Hadjidimos, *Block Gauss Elimination Followed by a Classical Iterative Method for the Solution of Linear Systems.* Preprint no. 02-1/2002, Department of Mathematics, University of Crete, Heraklion, Greece.

[3] A. Berman and R.J. Plemmons *Nonnegative Matrices in the Mathematical Sciences. Classics in Applied Mathematics.* SIAM, Philadelphia, 1994.

[4] R.E. Funderlic and R.J. Plemmons, *LU Decomposition of M-Matrices by Elimination Without Pivoting.* Linear Algebra Appl. 41 (1981), 99–110.

[5] A.D. Gunawardena, S.K. Jain and L. Snyder, *Modified Iterative Methods for Consistent Linear Systems.* Linear Algebra Appl. 154–156 (1991), 123–143.

[6] A. Hadjidimos, D. Noutsos and M. Tzoumas, *More on Modifications and Improvements of Classical Iterative Schemes for $Z-$Matrices.* Preprint no. 01-1/2001, Department of Mathematics, University of Crete, GR-714 09 Heraklion, Greece, 2001.

[7] K.R. James, *Convergence of Matrix Iterations Subject to Diagonal Dominance.* SIAM J. Numer. Anal. 10 (1973), 478–484.

[8] K.R. James and W. Riha, *Convergence Criteria for Successive Overrelaxation.* SIAM J. Numer. Anal. 12 (1975), 137–143.

[9] M.L. Juncosa and T.W. Mulliken, *On the Increase of Convergence Rates of Relaxation Procedures for Elliptic Partial Differential Equations.* J. Assoc. Comput. Mach. 7 (1960), 29–36.

[10] T. Kohno, H. Kotakemori, H. Niki and M. Usui, *Improving the Gauss-Seidel Method for Z-Matrices.* Linear Algebra Appl. 267 (1997), 113–123.

[11] W. Li and W. Sun, *Modified Gauss-Seidel type methods and Jacobi type methods for Z-matrices.* Linear Algebra Appl. 317 (2000), 227–240.

[12] I. Marek and D. B. Szyld, *Comparison theorems for weak splittings of bounded operators.* Numerische Mathematik 58 (1990), 387–397.

[13] J.P. Milaszewicz, *On Modified Jacobi Linear Operators*, Linear Algebra Appl. 51 (1983), 127–136.

[14] J.P. Milaszewicz, *Improving Jacobi and Gauss-Seidel Iterations*, Linear Algebra Appl. 93 (1987), 161–170.

[15] R.S. Varga, *Matrix Iterative Analysis.* Prentice-Hall, Englewood Cliffs, NJ, 1962. (Also: 2nd Edition, Revised and Expanded, Springer, Berlin, 2000.)

[16] Z. Woźnicki, *Nonnegative Splitting Theory,* Japan Journal of Industrial and Applied Mathematics 11 (1994), 289–342.

[17] D.M. Young, *Iterative Solution of Large Linear Systems.* Academic Press, New York, 1971.