Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Computational and Applied Mathematics 188 (2006) 89–106
JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS
www.elsevier.com/locate/cam

On optimal improvements of classical iterative schemes for Z-matrices

D. Noutsos[a,*], M. Tzoumas[b]

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

Received 9 September 2004; received in revised form 12 January 2005

Abstract

Many researchers have considered 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 unpreconditioned 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, LAA 93 (1987) 161–170], Gunawardena et al. [LAA 154–156 (1991) 123–143]. In this work we generalize the previous preconditioners to obtain optimal methods. "Good" Jacobi and Gauss–Seidel algorithms are given and preconditioners, that eliminate more than one entry per row, are also proposed and analyzed. Moreover, the behavior of the above preconditioners to the Krylov subspace methods is studied.

MSC: Primary 65F10

Keywords: Jacobi and Gauss–Seidel iterative methods; Diagonally dominant Z- and M-matrices

1. Introduction and preliminaries

Consider the linear system of algebraic equations

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

* Corresponding author.
E-mail address: dnoutsos@cc.uoi.gr (D. Noutsos).

0377-0427/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.cam.2005.03.057

where $A \in \mathbb{R}^{n,n}$ is an *irreducible dominant Z-matrix* with positive diagonal entries (see [4,18,20]), that is its off-diagonal elements are nonpositive, $(a_{ij} \leqslant 0,\ i, j = 1(1)n,\ j \neq i)$ and $b \in \mathbb{R}^n$. Since $a_{ii} > 0,\ i = 1(1)n$, we may assume for simplicity that $a_{ii} = 1,\ i = 1(1)n$. We consider the usual triangular splitting of A,

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

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 Gauss–Seidel associated with (1.1) converge and by the Stein–Rosenberg theorem [18,20,4] the Gauss–Seidel method is faster than the Jacobi one.

Many researchers have considered preconditioners, applied to system (1.1) that make the associated Jacobi and Gauss–Seidel methods converge asymptotically faster than the original ones. Milaszewicz [16], basing his idea on previous ones (see, e.g., [15,5,9]), considered as a preconditioner $P_1 \equiv I + S_1$, the matrix that eliminates the off-diagonal elements of the kth column of A. Gunawardena et al. [6] considered as a preconditioner the matrix $P_2 \equiv I + S_2$, which eliminates the elements of the first upper diagonal. Kohno et al. [10] extended the main idea in [6]. Recently Li and Sun [13] extended the class of matrices considered in [10] and very recently Hadjidimos et al. [8] extended, generalized and compared the previous works. Modifications of the above preconditioners have introduced and studied by Kotakemori et al. [11] and by Niki et al. [17] (see also Li [12]). The term "preconditioning" is used in Sections 2 and 3 as in the aforementioned works, namely, to reduce the spectral radius of the iteration matrix in order to improve the convergence of the classical iterative methods. However, the same term is often used when the goal is to improve the condition number of A and hence the convergence of the Conjugate Gradient or other Krylov subspace methods and this is done in Section 4.

This work is organized 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 A, present the convergence analysis and propose the algorithm that chooses a "good" preconditioner. In Section 3 we generalize the above preconditioners by introducing the idea of eliminating more than one entry per row and perform the corresponding convergence analysis. In Section 4 we study the behavior of the proposed preconditioners when applied to Krylov subspace methods, especially to Conjugate Gradient and to restarted GMRES methods. Finally, in Section 5, numerical examples are presented in support of our theory.

2. Extending known preconditioners

Milaszewicz's preconditioner [16] is based on the elimination of the entries of the kth column of A, $a_{ik}, i = 1(1)n,\ i \neq k$, while Gunawardena et al.'s one [6] is based on the elimination of the entries of the first upper diagonal $a_{i,i+1}, i = 1(1)n - 1$. Their common feature 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 kth 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 $A_1 = P_1 A$, which are reducible, are independent of the off-diagonal elements of the first row of A (see [8]). This does not happen in the case of Gunawardena et al.'s preconditioner, which eliminates 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 call it *cyclic* preconditioner: $P_3 \equiv I + S_3$, where

$(S_3)_{i,i+1} = -a_{i,i+1},\ i = 1(1)n-1,\ (S_3)_{n,1} = -a_{n,1}$ and 0 elsewhere. 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:

$$P \equiv I + S = \begin{bmatrix} 1 & & -a_{1k_1} & & & \\ & 1 & & & -a_{2k_2} & \\ & & 1 & -a_{3k_3} & & \\ & & & \ddots & & \\ & & & & \ddots & \\ & & & & & \ddots \\ & & -a_{nk_n} & & & 1 \end{bmatrix}, \tag{2.1}$$

where $k_i \in \{1, 2, \ldots, i-1, i+1, \ldots, n\},\ i = 1(1)n$. Obviously, there is a total number of $(n-1)^n$ choices for a preconditioner of type P in (2.1). We denote by $\widetilde{A} = PA$ the associated preconditioned matrix and 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

$$\widetilde{A}x = \widetilde{b}, \quad \text{with } \widetilde{A} = (I+S)A,\ \widetilde{b} = (I+S)b. \tag{2.2}$$

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

$$\widetilde{a}_{ij} = \begin{cases} a_{ij} - a_{ik_i}a_{k_ij} \leqslant 0, & j \neq k_i, i, \\ 0, & j = k_i, \\ 1 - a_{ik_i}a_{k_ii} > 0, & j = i. \end{cases} \tag{2.3}$$

We define the matrices

$$D_{\mathrm{s}} := \mathrm{diag}(a_{1k_1}a_{k_11}, a_{2k_2}a_{k_22}, \ldots, a_{nk_n}a_{k_nn}) \tag{2.4}$$

and

$$S(L+U-I) := L_{\mathrm{s}} + D_{\mathrm{s}} + U_{\mathrm{s}} - S_{\mathrm{L}} - S_{\mathrm{U}}, \tag{2.5}$$

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:

$$\widetilde{A} = \begin{cases} M - N = (I+S) - (I+S)(L+U), \\ M' - N' = I - (L + U + L_{\mathrm{s}} + D_{\mathrm{s}} + U_{\mathrm{s}} - S_{\mathrm{L}} - S_{\mathrm{U}}), \\ M'' - N'' = (I - D_{\mathrm{s}}) - (L + U + L_{\mathrm{s}} + U_{\mathrm{s}} - S_{\mathrm{L}} - S_{\mathrm{U}}). \end{cases} \tag{2.6}$$

The corresponding Jacobi and the Jacobi type iteration matrices as well as their Gauss–Seidel and Gauss–Seidel counterparts are given by

$$\begin{aligned} B &:= M^{-1}N = L + U, \\ B' &:= M'^{-1}N' = (L + U + L_{\mathrm{s}} + D_{\mathrm{s}} + U_{\mathrm{s}} - S_{\mathrm{L}} - S_{\mathrm{U}}), \\ \widetilde{B} &:= M''^{-1}N'' = (I - D_{\mathrm{s}})^{-1}(L + U + L_{\mathrm{s}} + U_{\mathrm{s}} - S_{\mathrm{L}} - S_{\mathrm{U}}), \end{aligned} \tag{2.7}$$

and

$$H := (I - L)^{-1}U,$$
$$H' := (I - L - L_s + S_L)^{-1}(U + D_s + U_s - S_U),$$
$$\widetilde{H} := (I - D_s - L - L_s + S_L)^{-1}(U + U_s - S_U). \tag{2.8}$$

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

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

$$B'y \leqslant By \quad \text{and} \quad H'z \leqslant Hz, \tag{2.9}$$
$$\rho(\widetilde{B}) \leqslant \rho(B') \leqslant \rho(B) < 1, \tag{2.10}$$
$$\rho(\widetilde{H}) \leqslant \rho(H') \leqslant \rho(H) < 1, \tag{2.11}$$
$$\rho(\widetilde{H}) \leqslant \rho(\widetilde{B}), \quad \rho(H') \leqslant \rho(B'), \quad \rho(H) < \rho(B) < 1. \tag{2.12}$$

(*Note*: *Equalities in* (2.12) *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* (2.9) *and the middle inequality in* (2.10) *are strict.*

Proof. (a) (2.9): The matrices B and B' are related as follows:

$$B' = I - (I + S)A = I - (I + S)(I - B) = B - S(I - B). \tag{2.13}$$

For the nonnegative Jacobi iteration matrix B there exists a nonnegative vector y such that $By = \rho(B)y$. Postmultiplying (2.13) by y we get

$$B'y = (B - S(I - B))y = \rho(B)y - (1 - \rho(B))Sy \leqslant \rho(B)y = By, \tag{2.14}$$

which proves the first inequality of (2.9).

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 equivalently

$$\rho(H)z = \rho(H)Lz + Uz. \tag{2.15}$$

We rewrite now the matrix $\widetilde{A}$ of (2.2) as follows:

$$\begin{aligned}\widetilde{A} &= (I + S)A = (I + S_L + S_U)(I - L - U)\\ &= I - (L - S_L + S_L L + (S_L U)_L + (S_U L)_L)\\ &\quad - (U - S_U + S_U U + (S_L U)_U + (S_U L)_U) = I - \widetilde{L} - \widetilde{U},\end{aligned} \tag{2.16}$$

where by $(Q)_L$ and by $(Q)_U$ we have denoted the strictly lower and the strictly upper part of the matrix Q, respectively. So,

$$L = \widetilde{L} + S_L - S_L L - (S_L U)_L - (S_U L)_L, \quad U = \widetilde{U} + S_U - S_U U - (S_L U)_U - (S_U L)_U. \tag{2.17}$$

By substituting (2.17) in (2.15) we get

$$\rho(H)z = \rho(H)\widetilde{L}z + \widetilde{U}z + z', \tag{2.18}$$

where

$$z' = \rho(H)(S_{\mathrm{L}} - S_{\mathrm{L}}L - (S_{\mathrm{L}}U)_{\mathrm{L}} - (S_{\mathrm{U}}L)_{\mathrm{L}})z + (S_{\mathrm{U}} - S_{\mathrm{U}}U - (S_{\mathrm{L}}U)_{\mathrm{U}} - (S_{\mathrm{U}}L)_{\mathrm{U}})z. \tag{2.19}$$

If $z' \geqslant 0$ then from (2.18) we get

$$\rho(H)(I - \widetilde{L})z \geqslant \widetilde{U}z \quad \text{or} \quad \rho(H)z \geqslant (I - \widetilde{L})^{-1}\widetilde{U}z \quad \text{or} \quad \rho(H)z \geqslant H'z \tag{2.20}$$

from which the second inequality of (2.9) follows. It remains to prove that $z' \geqslant 0$:

$$\begin{aligned} z' &= \rho(H)S_{\mathrm{L}}(I - L - U)z + \rho(H)(S_{\mathrm{L}}U - (S_{\mathrm{L}}U)_{\mathrm{L}} - (S_{\mathrm{U}}L)_{\mathrm{L}})z \\ &\quad + S_{\mathrm{U}}(I - L - U)z + (S_{\mathrm{U}}L - (S_{\mathrm{L}}U)_{\mathrm{U}} - (S_{\mathrm{U}}L)_{\mathrm{U}})z \\ &= S_{\mathrm{L}}(\rho(H)(I - L) - \rho(H)U)z + \rho(H)((S_{\mathrm{L}}U)_{\mathrm{U}} - (S_{\mathrm{U}}L)_{\mathrm{L}})z \\ &\quad + \frac{1}{\rho(H)} S_{\mathrm{U}}\,(\rho(H)(I - L) - \rho(H)U)z + ((S_{\mathrm{U}}L)_{\mathrm{L}} - (S_{\mathrm{L}}U)_{\mathrm{U}})z \\ &= (1 - \rho(H))S_{\mathrm{L}}Uz + \frac{1 - \rho(H)}{\rho(H)} S_{\mathrm{U}}Uz - (1 - \rho(H))(S_{\mathrm{L}}U)_{\mathrm{U}}z + (1 - \rho(H))(S_{\mathrm{U}}L)_{\mathrm{L}}z \\ &= (1 - \rho(H))(S_{\mathrm{L}}U)_{\mathrm{L}}z + \frac{1 - \rho(H)}{\rho(H)} S_{\mathrm{U}}Uz + (1 - \rho(H))(S_{\mathrm{U}}L)_{\mathrm{L}}z \geqslant 0. \end{aligned} \tag{2.21}$$

(a) (2.10): It is known that (see [4]) a Z-matrix A is a nonsingular M-matrix iff there exists a positive vector $y(>0) \in \mathbb{R}^n$ such that $Ay > 0$. By taking such a y, the fact that $P = I + S \geqslant 0$ 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 (2.6) are regular splittings because $M'^{-1} = I^{-1} = I \geqslant 0$, $N' \geqslant 0$ and $M''^{-1} = (I - D_{\mathrm{s}})^{-1} \geqslant 0$, $N'' \geqslant 0$ and so they are convergent. Since $M''^{-1} \geqslant M'^{-1}$, it is implied (see [19]) that the left inequality in (2.10) is true. For the proof of the middle inequality in (2.10), we recall the first inequality of (2.9) which gives that $B'y \leqslant \rho(B)y$. Then, we apply Lemma 3.3 in Marek and Szyld [14] to get our assertion.

(a) (2.11): To prove the first inequality in (2.11) we use regular splittings of the matrix $\widetilde{A}$. Specifically, consider the following splittings that define the iteration matrices in (2.8):

$$\widetilde{A} = \begin{cases} M - N = (I + S)(I - L) - (I + S)U, \\ M' - N' = (I - L - L_{\mathrm{s}} + S_{\mathrm{L}}) - (D_{\mathrm{s}} + U + U_{\mathrm{s}} - S_{\mathrm{U}}), \\ M'' - N'' = (I - D_{\mathrm{s}} - L - L_{\mathrm{s}} + S_{\mathrm{L}}) - (U + U_{\mathrm{s}} - S_{\mathrm{U}}), \end{cases} \tag{2.22}$$

where we have used the same symbols for the two matrices of each splitting as in the case of (2.6). So, the last two splittings in (2.22) are regular splittings because $M'^{-1} = (I - L - L_{\mathrm{s}} + S_{\mathrm{L}})^{-1} = I + (L + L_{\mathrm{s}} - S_{\mathrm{L}}) + \cdots + (L + L_{\mathrm{s}} - S_{\mathrm{L}})^{n-1} \geqslant 0$, $N' \geqslant 0$ and $M''^{-1} = (I - D_{\mathrm{s}} - L - L_{\mathrm{s}} + S_{\mathrm{L}})^{-1} \geqslant 0$, $N'' \geqslant 0$ and so they are convergent. Since $M''^{-1} \geqslant M'^{-1}$, it is implied (see [19]) that the left inequality in (2.11) is true.

To prove the second inequality of (2.11) 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

$$Hz = \rho(H)z \quad \text{or} \quad (I - L)^{-1}Uz = \rho(H)z \quad \text{or} \quad (\rho(H)L + U)z = \rho(H)z. \tag{2.23}$$

We observe here that the matrix $\rho(H)L + U$ has the same structure as the matrix B and consequently it is also irreducible. So, from the Perron–Frobenius Theorem (see Varga [18]), the eigenvector z will be a

positive vector. Recalling relation (2.20), the following property holds: There exists a positive vector z such that $\rho(H)z \geqslant H'z$. Based on this, we can apply Lemma 3.3 in Marek and Szyld [14] to get the second inequality in (2.11). In the case where B is reducible, we consider a small number $\varepsilon > 0$ and replace some zeros of B with ε so that the produced matrix $B(\varepsilon)$ becomes irreducible. Then, for the associated matrices $H(\varepsilon)$ and $H'(\varepsilon)$ there holds: $\rho(H'(\varepsilon)) \leqslant \rho(H(\varepsilon))$. 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 (2.11).

(a) (2.12): 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 (2.12) are implied directly by the facts that $\widetilde{A}$ is a nonsingular M-matrix, and the last two pairs of splittings in (2.6) and (2.22), from which the four matrices involved, $\widetilde{H}$, $\widetilde{B}$, H', B', are produced, are regular ones with $L + L_s + U + U_s - S \geqslant U + U_s - S_U$ and $D_s + L + L_s + U + U_s - S \geqslant D_s + U + U_s - S_U$. It is noted that if $\rho(\widetilde{B}) = 0$ then $\rho(\widetilde{H}) = 0$ and the matrix $\widetilde{B}$ would be reducible with its *normal form* being a strictly upper triangular matrix. By the Stein–Rosenberg theorem, if $\rho(\widetilde{B}) = \rho(\widetilde{H})$ it is either $\rho(\widetilde{B}) = \rho(\widetilde{H}) = 0$ or $\rho(\widetilde{B}) = \rho(\widetilde{H}) = 1$. Since $\rho(\widetilde{B}) = 0$, both spectral radii would be zero. For the second inequality of (2.12) 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 *normal form* being an upper triangular matrix. This means that in the directed graph $\mathcal{G}(B')$ of the matrix B', there is no strongly connected subpath except for identity paths (loops) corresponding to the nonzero diagonal elements. For the matrix H' we have that $H' = (I - L - L_s + S_L)^{-1}(D_s + U + U_s - S_U) = (I + (L + L_s - S_L) + \cdots + (L + L_s - S_L)^{n-1})(D_s + U + U_s - S_U) = D_s + [U + U_s - S_U + ((L + L_s - S_L) + \cdots + (L + L_s - S_L)^{n-1})(D_s + U + U_s - S_U)]$. The matrix in the brackets is a sum of products of nonnegative parts of B'. Therefore, 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 also exist such a subpath in $\mathcal{G}(B')$. This means that the matrix H' has also its *normal form* an upper triangular matrix, with its diagonal elements those of B'. This proves our assertion that $\rho(B') = \rho(H')$.

(b) (2.9): Since B is irreducible, The eigenvector y, corresponding to $\rho(B)$, is positive and according to the steps in the proof of (2.9) in (a) we can see that inequality (2.14) becomes a strict one.

(b) (2.10): From the inequality $B'y < By$ we get $B'y < \rho(B)y$. Now we can apply Lemma 3.3 in Marek and Szyld [14] to get the strict inequality $\rho(B') < \rho(B)$. □

2.1. "Good" Jacobi and Gauss–Seidel preconditioners

It was proved that the preconditioned Jacobi and Gauss–Seidel methods converge for each choice of the matrix S and converge faster than their unpreconditioned counterparts. A question, then, arises: Is there an optimal choice for the matrix S so that the associated method will be an optimal one and if so how can one choose such a matrix S? This question cannot be answered yet and constitutes an open problem. It seems to be difficult since we have to compare the spectral radii of $(n-1)^n$ different matrices. Instead, we will try to answer a simpler related question: Is there a "good" choice of the matrix S such that the associated method will be the best among many others and possibly the optimal one? To find a "good" Jacobi preconditioner or a "good" Gauss–Seidel one we will work using sufficient conditions of convergence rather than necessary and sufficient ones. So, we choose the matrix S such that to minimize the maximum norm of $\widetilde{B}$ (or of $\widetilde{H}$) which constitutes an upper bound for its spectral radius. In the following we give the analysis and the associated algorithm for the Jacobi method only. The analysis for

the Gauss–Seidel method is analogous and straightforward and its associated algorithm is of the same order of complexity. In the last section, we will show by numerical experiments, the efficiency of the "good" Gauss–Seidel method as well.

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

$$\rho(\widetilde{B}) \leqslant \|\widetilde{B}\|_\infty \iff \rho(\widetilde{B}) \leqslant \max_i \frac{\widetilde{l}_i + \widetilde{u}_i}{\widetilde{d}_i} < 1, \tag{2.24}$$

where

$$\begin{aligned}
\widetilde{d}_i &= |\widetilde{a}_{ii}| = \widetilde{a}_{ii} = 1 - a_{ik_i}a_{k_i i},\\
\widetilde{l}_i &= \sum_{j=1}^{i-1} |\widetilde{a}_{ij}| = -\sum_{j=1}^{i-1} \widetilde{a}_{ij} = a_{ik_i} \sum_{j=1}^{i-1} a_{k_i j} - \sum_{j=1}^{i-1} a_{ij},\\
\widetilde{u}_i &= \sum_{j=i+1}^{n} |\widetilde{a}_{ij}| = -\sum_{j=i+1}^{n} \widetilde{a}_{ij} = a_{ik_i} \sum_{j=i+1}^{n} a_{k_i j} - \sum_{j=i+1}^{n} a_{ij}.
\end{aligned} \tag{2.25}$$

For each row i, of the method we propose, we choose k_i such that all the ratios $\frac{\widetilde{l}_i+\widetilde{u}_i}{\widetilde{d}_i}$ are minimized and so their maximum will also be minimized. Since the choice of k_i is not unique, S is also not unique. We conjecture that since we minimize all the ratios for each row (the row sums of the nonnegative matrix $\widetilde{B}$), the new spectral radius will be as small as possible. We call this method the "Good" Jacobi preconditioned method and the associated preconditioner, $I + S$, the "Good" Jacobi preconditioner. From (2.25) we get that

$$\widetilde{l}_i + \widetilde{u}_i = -\sum_{j=1, j\neq i}^{n} \widetilde{a}_{ij} = a_{ik_i} \sum_{j=1, j\neq i}^{n} a_{k_i j} - \sum_{j=1, j\neq i}^{n} a_{ij} = s_i + a_{ik_i}(1 - s_{k_i} - a_{k_i i}), \tag{2.26}$$

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

$$\frac{\widetilde{l}_i + \widetilde{u}_i}{\widetilde{d}_i} = \frac{s_i + a_{ik_i}(1 - s_{k_i} - a_{k_i i})}{1 - a_{ik_i}a_{k_i i}}. \tag{2.27}$$

To give an efficient algorithm that will choose the indices k_i and consequently the matrix S we observe that the Jacobi method requires $\mathcal{O}(n^2)$ ops per iteration. The same number of operations is required for the multiplication $(I + S)A$. So, the cost of the choice of S must require at most $\mathcal{O}(n^2)$ ops. First, we compute all the row sums s_i that require a total number of $\mathcal{O}(n^2)$ ops. Then, we compute the ratios (2.27) for every i and every k_i. The 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 the total cost of the algorithm is $\mathcal{O}(n^2)$ ops per iteration. This makes it an efficient one. The previous analysis for the cost is illustrated more clearly in the following algorithm written in a pseudocode form.

Algorithm of "*Good*" *Jacobi Preconditioner*
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.

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 (2.1), 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 ith row we have to eliminate the elements k_i and l_i, where $k_i < l_i$. 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:

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

or

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

We generalize the above relations by considering that m elements of the ith row are to be eliminated. For this we give the following definitions. Let $\hat{k}_i^{\mathrm{T}} = (k_{i_1}k_{i_2}\ldots k_{i_m})$ be a multiindex, where the indices $k_{i_1} < k_{i_2} < \cdots < k_{i_m}$ denote the positions of the elements of row i to be eliminated. Then, we define

by $s_{i\hat{k}_i}^{\mathrm{T}} = (s_{ik_{i_1}} s_{ik_{i_2}} \dots s_{ik_{i_m}})$ the vector of nonzero off-diagonal elements of the ith row of S, $a_{i\hat{k}_i}^{\mathrm{T}} = (a_{ik_{i_1}} a_{ik_{i_2}} \dots a_{ik_{i_m}})$ the vector of the elements of the ith 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 $A_{\hat{k}_i}^{-1}$ 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:

$$s_{i\hat{k}_i}^{\mathrm{T}} = -a_{i\hat{k}_i}^{\mathrm{T}} A_{\hat{k}_i}^{-1}. \tag{3.4}$$

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

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

The first inequality in (3.5) is obvious while the last one is to be proved.

We consider the matrix

$$\widetilde{A}_{\hat{k}_i} = \begin{pmatrix} 1 & a_{i\hat{k}_i}^{\mathrm{T}} \\ a_{\hat{k}_i i} & A_{\hat{k}_i} \end{pmatrix}. \tag{3.6}$$

The quantity $1 - a_{i\hat{k}_i}^{\mathrm{T}} A_{\hat{k}_i}^{-1} a_{\hat{k}_i i}$ is the Schur complement of the above matrix $\widetilde{A}_{\hat{k}_i}$. Since the Schur complement of a nonsingular M-matrix is also a nonsingular M-matrix (e.g., see [7, p. 128]), we get that

$$1 - a_{i\hat{k}_i}^{\mathrm{T}} A_{\hat{k}_i}^{-1} a_{\hat{k}_i i} > 0. \tag{3.7}$$

We define the matrix

$$D_{\mathrm{s}} := \mathrm{diag}(a_{1\hat{k}_1}^{\mathrm{T}} A_{\hat{k}_1}^{-1} a_{\hat{k}_1 1}, a_{2\hat{k}_2}^{\mathrm{T}} A_{\hat{k}_2}^{-1} a_{\hat{k}_2 2}, \dots, a_{n\hat{k}_n}^{\mathrm{T}} A_{\hat{k}_n}^{-1} a_{\hat{k}_n n}), \tag{3.8}$$

which is the diagonal part of the matrix $S(L+U)$. Using the notation in (2.4) and (2.5), and considering the splittings in (2.6), as well as the associated Jacobi (2.7) and Gauss–Seidel (2.8), we can 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* $y \geqslant 0$ *and* $z \geqslant 0$, *such that*

$$B'y \leqslant By \quad \text{and} \quad H'z \leqslant Hz, \tag{3.9}$$

$$\rho(\widetilde{B}) \leqslant \rho(B') \leqslant \rho(B) < 1, \tag{3.10}$$

$$\rho(\widetilde{H}) \leqslant \rho(H') \leqslant \rho(H) < 1, \tag{3.11}$$

$$\rho(\widetilde{H}) \leqslant \rho(\widetilde{B}), \quad \rho(H') \leqslant \rho(B'), \quad \rho(H) < \rho(B) < 1. \tag{3.12}$$

(*Note*: *Equalities in* (3.12) *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* (3.9) *and the middle inequality in* (3.10) *are strict.*

Proof. Using the same notation, the proof of Theorem 3.1 follows step by step that of Theorem 2.1, and so, the proof of the present statement 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 has 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 has a cost of $\mathcal{O}(n^2)$ ops per iteration. So, to obtain an efficient algorithm, the number m must be chosen very small and independent of the dimension n. A question which arises is: 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. Therefore there must be a golden section, which also depends on the matrix A. Another question which arises is: How can one choose the multiindices $\hat{k}_i$'s? Since this is difficult to answer we follow the idea we did in the "Good" Jacobi or Gauss–Seidel algorithms, where we provided search algorithms by using sufficient criteria instead of sufficient and necessary ones. In the case of $m = 1$, the cost of the searching algorithm is $\mathcal{O}(n^2)$. 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 increases 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, keeping the others fixed. As we will see in the numerical examples, in many cases, by keeping the multiindices fixed, the multiindexed preconditioner is better than the "Good" algorithm.

4. Eliminated preconditioners for Krylov subspace methods

In this section we study the behavior of the proposed preconditioners when they are applied to Krylov subspace methods. First we study the conjugate gradient (CG) method when the matrix A is a real symmetric positive definite matrix. It is well known that the convergence theory of the preconditioned conjugate gradient (PCG) method holds if the preconditioner is also a symmetric and positive definite matrix. This means that, to construct our preconditioner $I + S$, we have to choose the entries to be eliminated in symmetric positions. We will study here the behavior of the multiindexed preconditioner, where we have fixed the multiindices by taking $\hat{k}_1 = (2), \hat{k}_i = (i-1\ \ i+1)^{\mathrm{T}},\ i = 2(1)n-1$ and $\hat{k}_n = (n-1)$. In other words we eliminate the elements of the first upper and lower diagonals. So, the eliminated entries

are in symmetric positions. The study of other multiindexed preconditioners having entries in symmetric positions, could be straightforwardly generalized. First, we state and prove the following lemma which is useful for the convergence properties below:

Lemma 4.1. *Let A be a nonsingular M-matrix with $a_{ii}=1$. Let also $\widetilde{A}=(I+S)A$ be the preconditioned matrix with $I+S$ being the multiindexed preconditioner defined by $\hat{k}_1=(2), \hat{k}_i=(i-1\ i+1)^{\mathrm{T}},\ i=2(1)n-1$ and $\hat{k}_n=(n-1)$. Then all the principal minors of $\widetilde{A}$ are positive and less than or equal to the associated principal minors of A.*

$$0<\det(\widetilde{A}_{i_1 i_2 \ldots i_k})\leqslant \det(A_{i_1 i_2 \ldots i_k}), \tag{4.1}$$

where by the indices $i_1 i_2 \ldots i_k$ we denote the rows and the associated columns which form the principal submatrix.

Proof. The first inequality holds since all the principal submatrices of $\widetilde{A}$ and A are M-matrices. We will prove the second inequality by induction. First we prove that for all the principal submatrices of the preconditioner $I+S$, there holds

$$0<\det((I+S)_{i_1 i_2 \ldots i_k})\leqslant 1.$$

Since each principal submatrix of order k can be put in the first k rows and columns by a permutation transformation of the matrix $I+S$, we will prove, without loss of generality, the above inequalities only for the upper left principal minors, i.e.,

$$0<\det((I+S)_k)\leqslant 1,$$

where $(I+S)_k$ is the principal submatrix, consisting of the first k rows and columns, of the matrix $(I+S)$. We denote also by A_k and by $\widetilde{A}_k$ the principal submatrices consisting of the first k rows and columns of the matrices A and $\widetilde{A}$, respectively. We define now the $k\times k$ matrix

$$A'_k=A_k+S_{k,n-k}A_{n-k,k},$$

where $S_{k,n-k}$ is the $k\times(n-k)$ submatrix of $I+S$ consisting of the first k rows and the last $n-k$ columns, while $A_{n-k,k}$ is the $n-k\times k$ submatrix of A consisting from the last $n-k$ rows and the first k columns. From the above definition it is obvious that A'_k is the principal submatrix consisting from the first k rows and columns of the matrix $(I+S'_{k,n-k})A$, where $S'_{k,n-k}=\begin{pmatrix}0 & S_{k,n-k}\\ 0 & 0\end{pmatrix}$. Since $I+S'_{k,n-k}$ is a nonnegative matrix, A is an M-matrix and $(I+S'_{k,n-k})A$ is a Z-matrix we get that $(I+S'_{k,n-k})A$ is also an M-matrix. Therefore its principal submatrix A'_k is an M-matrix.

From the symmetric structure of the matrix $I+S$ and since $s_{ij}\geqslant 0,\ i,j=1,2,\ldots n,\ i\neq j$, it is easily checked that if we use the Gauss elimination process step by step, the diagonal elements do not increase. So, all the first principal minors would be less than or equal to 1. We will prove in the sequel that these principal minors could not be less than 0. We will prove inequality (4.13) for the first principal minors, i.e.

$$0<\det(\widetilde{A}_k)\leqslant \det(A_k),\quad k=1,2,\ldots,n.$$

The proof for any other determinant is given directly by a similar analysis which becomes a little more complicated. From the notation above, we have that

$$\widetilde{A}_k = (I+S)_k A_k + S_{k,n-k} A_{n-k,k} = (I+S)_k A_k + \begin{pmatrix} 0_{k-1,k} \\ s_{k,k+1} a^{\mathrm{T}}_{k+1;1:k} \end{pmatrix},$$

since the matrix $S_{k,n-k}$ has only one entry different from zero in the position $(k, k+1)$. We have denoted by $a^{\mathrm{T}}_{k+1;1:k}$ the row vector consisting of the first k entries of the $(k+1)$st row of the matrix A and by $0_{k-1,k}$ the $(k-1)\times k$ zero matrix. The above equality is written as

$$(I+S)_k A_k = \widetilde{A}_k - \begin{pmatrix} 0_{k-1,k} \\ s_{k,k+1} a^{\mathrm{T}}_{k+1;1:k} \end{pmatrix} = \left(I_k - \begin{pmatrix} 0_{k-1,k} \\ s_{k,k+1} a^{\mathrm{T}}_{k+1;1:k} \end{pmatrix} \widetilde{A}_k^{-1} \right) \widetilde{A}_k. \tag{4.2}$$

Since $\widetilde{A}$ is an M-matrix, $\widetilde{A}_k$ is also an M-matrix and consequently $\widetilde{A}_k^{-1}$ is a nonnegative matrix. So, $-s_{k,k+1} a^{\mathrm{T}}_{k+1;1:k} \widetilde{A}_k^{-1}$ is a nonnegative row vector. Therefore, the matrix

$$\left(I_k - \begin{pmatrix} 0_{k-1,k} \\ s_{k,k+1} a^{\mathrm{T}}_{k+1;1:k} \end{pmatrix} \widetilde{A}_k^{-1} \right)$$

is a lower triangular matrix with all the diagonal entries equal to one except the last one which is greater than or equal to 1. This means that

$$\det\left(\left(I_k - \begin{pmatrix} 0_{k-1,k} \\ s_{k,k+1} a^{\mathrm{T}}_{k+1;1:k} \end{pmatrix} \widetilde{A}_k^{-1} \right) \right) \geqslant 1.$$

By taking determinants of the matrices in (4.2) we get

$$\det((I+S)_k)\det(A_k) = \det\left(I_k - \begin{pmatrix} 0_{k-1,k} \\ s_{k,k+1} a^{\mathrm{T}}_{k+1;1:k} \end{pmatrix} \widetilde{A}_k^{-1} \right) \det(\widetilde{A}_k).$$

Since A_k and $\widetilde{A}_k$ are both M-matrices, the associated determinants are both positive. So, from the above equality it has been proved that $\det((I+S)_k) > 0$. From the fact that $\det((I+S)_k) \leqslant 1$ and

$$\det\left(I_k - \begin{pmatrix} 0_{k-1,k} \\ s_{k,k+1} a^{\mathrm{T}}_{k+1;1:k} \end{pmatrix} \widetilde{A}_k^{-1} \right) \geqslant 1$$

the above equality gives us the inequality we had to prove:

$$\det(\widetilde{A}_k) \leqslant \det(A_k). \qquad \square$$

4.1. Preconditioned conjugate gradient method

Theorem 4.1. *Let A be an irreducible and symmetric positive definite M-matrix with $a_{ii} = 1$ and with its eigenvalues $\lambda_i \in [a, b]$, $i = 1, 2, \ldots, n$. Let also $\widetilde{A} = (I+S)A$ be the preconditioned matrix with $I+S$ being the multiindexed preconditioner defined by $\hat{k}_1 = (2)$, $\hat{k}_i = (i-1\ i+1)^{\mathrm{T}}$, $i = 2(1)n-1$*

and $\hat{k}_n = n-1$, which is symmetric and positive definite. Then the eigenvalues $\widetilde{\lambda}_i$, $i = 1, 2, \ldots, n$, of the preconditioned matrix $\widetilde{A}$ are clustered in the interval $[\widetilde{a}, \widetilde{b}]$ except for a possible small number of them that are greater than $\widetilde{b}$, where $\widetilde{a} > a$ and $\widetilde{b} < b$. Consequently, the Preconditioned Conjugate Gradient method converges faster than the unpreconditioned one.

Proof. It is well known that the characteristic polynomial of a matrix A is given by

$$P_n(\lambda) = \lambda^n - \sum_{i=1}^{n} a_{ii}\lambda^{n-1} + \cdots + (-1)^k \sum_{i_1 i_2 \cdots i_k} \det(A_{i_1 i_2 \cdots i_k})\lambda^{n-k} + \cdots + (-1)^n \det(A). \quad (4.3)$$

In terms of the eigenvalues the same polynomial is given by

$$P_n(\lambda) = \lambda^n - \left(\sum_{i=1}^{n} \lambda_i\right)\lambda^{n-1} + \cdots + (-1)^k \left(\sum_{i_1 i_2 \cdots i_k} \lambda_{i_1}\lambda_{i_2}\cdots\lambda_{i_k}\right)\lambda^{n-k} + \cdots + (-1)^n \lambda_1\lambda_2\cdots\lambda_n. \quad (4.4)$$

So,

$$\sum_{i_1 i_2 \cdots i_k} \lambda_{i_1}\lambda_{i_2}\cdots\lambda_{i_k} = \sum_{i_1 i_2 \cdots i_k} \det(A_{i_1 i_2 \cdots i_k}), \quad k = 1, 2, \ldots, n, \quad (4.5)$$

where the sums have been taken over all the combinations of n over k. By taking the same arguments of the preconditioned matrix $\widetilde{A}$ we take

$$\sum_{i_1 i_2 \cdots i_k} \widetilde{\lambda}_{i_1}\widetilde{\lambda}_{i_2}\cdots\widetilde{\lambda}_{i_k} = \sum_{i_1 i_2 \cdots i_k} \det(\widetilde{A}_{i_1 i_2 \cdots i_k}), \quad k = 1, 2, \ldots, n. \quad (4.6)$$

From Lemma 4.1 we have that

$$0 < \det(\widetilde{A}_{i_1 i_2 \ldots i_k}) \leqslant \det(A_{i_1 i_2 \ldots i_k}), \quad \forall i_1 i_2 \ldots i_k,\ k = 1, 2, \ldots, n. \quad (4.7)$$

This inequality and equalities (4.5), (4.6) give us the main inequality which compares the eigenvalues of the preconditioned matrix with those of the unpreconditioned one:

$$\sum_{i_1 i_2 \cdots i_k} \widetilde{\lambda}_{i_1}\widetilde{\lambda}_{i_2}\cdots\widetilde{\lambda}_{i_k} \leqslant \sum_{i_1 i_2 \cdots i_k} \lambda_{i_1}\lambda_{i_2}\cdots\lambda_{i_k}, \quad k = 1, 2, \ldots, n. \quad (4.8)$$

We observe that if for at least one combination $i_1 i_2 \cdots i_k$ inequality (4.7) is strict, then so is inequality (4.20). It is easily seen that there are values of k such that an inequality exists. Obvious values of such k's are $k = 1$, $k = 2$ or $k = n$.

On the other hand, from the inequalities (3.10) of Theorem 3.1 and from the irreducibility of A we have that $\rho(B') < \rho(B)$ or $\rho(I - \widetilde{A}) < \rho(I - A)$ which means that

$$\widetilde{\lambda}_{\min} > \lambda_{\min}. \quad (4.9)$$

Inequalities (4.20) and (4.21) imply that, when we apply the improved process to the matrix A, although the minimum eigenvalue increases, some sums of products of them do not increase while some others strictly decrease. This observation and the strong numerical evidence suggest that, although the new largest eigenvalue might be outside the interval $[a, b]$, the majority of the remaining eigenvalues must lie (are clustered) in a smaller interval $[\widetilde{a}, \widetilde{b}] \subset [a, b]$. By the known Axelsson's Theorem [3] concerning the convergence of the Conjugate Gradient method, we have that the Preconditioned Conjugate Gradient method converges faster than the unpreconditioned one. □

We remark that we could have had the same result as in Theorem 4.1 if we had proved the above inequalities only for exact values of k ($k = 1, 2, n$) for which the proof could have been very easy. We think that this general proof makes our result stronger. It is also remarked that the same result could be proved, in the more general case of the choice of the multiindex. This will be shown by experiments, in the next section, when we solve the Helmholz equation.

4.2. Preconditioned GMRES method

Theorem 4.2. *Let A be an irreducible M-matrix with $a_{ii}=1$ and let its eigenvalues $\lambda_i \in \mathscr{S}$, $i=1, 2, \ldots, n$, where $\mathscr{S}$ is a bounded convex region in the complex plane with $a=\min_{x\in\mathscr{S}} Re(x) > 0$. Let also $\widetilde{A}=(I+S)A$ be the preconditioned matrix with $I + S$ being the multiindexed preconditioner defined by $\hat{k}_1 = (2)$, $\hat{k}_i = (i-1\ i+1)^{\mathrm{T}}$, $i=2(1)n-1$ and $\hat{k}_n = n-1$. Then, under the assumption that the matrices V and $\widetilde{V}$ of the eigenvectors of A and of $\widetilde{A}$, respectively have both small enough euclidian condition number, the eigenvalues $\widetilde{\lambda}_i$, $i = 1, 2, \ldots, n$, of the preconditioned matrix $\widetilde{A}$ are clustered in a convex region $\widetilde{\mathscr{S}} \subset \mathscr{S}$ with $\widetilde{a} = \min_{x\in\widetilde{\mathscr{S}}} Re(x) > a$, except for a possible small number of them that are outside $\widetilde{\mathscr{S}}$, but with real parts no less than $\widetilde{a}$. Consequently, the preconditioned GMRES method converges faster than the unpreconditioned one.*

Proof. The proof is straightforward and the same as the one of Theorem 4.1. All relations (4.2)–(4.21) hold with the exception that now $\lambda_i, \widetilde{\lambda}_i, i = 1, 2, \ldots, n$, are complex numbers with positive real parts. Since B' and B are nonnegative matrices, $\lambda_{\min}$ and $\widetilde{\lambda}_{\min}$ are real positive numbers with (4.21) holding. It is obvious that the sums of products in (4.20) decrease when the moduli of the eigenvalues decrease. Since the smaller of them increase, most of the larger moduli should decrease. So, the eigenvalues are clustered in a shorter region $\widetilde{\mathscr{S}}$ which belongs in $\mathscr{S}$, except for a small number of eigenvalues that lie outside $\widetilde{\mathscr{S}}$. Let c be the modulus of the center of the region in which most of the eigenvalues lie and s be the radius of the same region. It is known that the convergence rate of GMRES method depends on the ratio $\frac{s}{c}$, in the case where the euclidian condition number of the matrix of the eigenvectors is small enough. For the eigenvalues that lie outside, a constant number of GMRES additional iterations are performed. From the analysis above, we have no information for the centers of the regions. Since the smallest moduli increase and the largest ones decrease, the centers should be almost the same. For the radii, it is obvious that the radius of $\widetilde{\mathscr{S}}$ is smaller than the one of S, due to the better clustering. So, the preconditioned GMRES method converges faster than the unpreconditioned one, and the proof is complete. □

The same remarks, as before, could be stated here. In the next section we show the validity of the above result by solving the Convection–Diffusion equation.

5. Numerical examples

For 10,000 randomly generated nonsingular M-matrices for $n = 10$, 20 and 50 we have determined the spectral radii of the iteration matrices of all the classical methods mentioned previously. Below, we present two tables for the Jacobi and Gauss–Seidel methods, respectively (Tables 1 and 2). Each number in the tables represents the percentage of the case worked out, the method in the first column is better than the one in the head of the table, where M, G, C, "G" and M_2 denote a method with the Milaszewicz's, the Gunawardena et al.'s, the Cyclic, the "Good" and the multiindexed, with $m = 2$, preconditioners, respectively. For the multiindexed preconditioner we have fixed the multiindices by taking $\hat{k}_1 = (2\ n)^{\mathrm{T}}$, $\hat{k}_i = (i-1\ i+1)^{\mathrm{T}}$, $i = 2(1)n-1$ and $\hat{k}_n = (1\ n-1)^{\mathrm{T}}$. In other words we have eliminated the elements corresponding to the Cyclic preconditioner and the symmetrically placed elements.

We remark that for the Jacobi method and for n large enough, the multiindexed improved method is 100% better than all the others. The "Good" 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 "Good" preconditioner is better than all the others, then follow the multiindexed, the cyclic, the Gunawardena et al.'s and finally the Milaszewicz's preconditioner. At this point, another question is raised: Is the "Good" 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 that the elimination of the over-diagonal

Table 1
Jacobi method

	$n = 10$				$n = 20$				$n = 50$			
	M	G	C	"G"	M	G	C	"G"	M	G	C	"G"
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
"G"	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

Table 2
Gauss–Seidel method

	$n = 10$				$n = 20$				$n = 50$			
	M	G	C	"G"	M	G	C	"G"	M	G	C	"G"
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
"G"	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

Table 3
Preconditioned conjugate gradient method, for two-dimension Helmholtz equation

n	I	S_2	S_4
16	51	45	27
32	100	86	52
64	191	163	100
128	374	306	194

elements play the most important role for the Gauss–Seidel method than that of the under-diagonal ones. Since we have chosen one fixed over-diagonal element per row, for the multiindexed preconditioner, while for the "Good" one we have searched one element per row (we guess that 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 ones for each row. So, the numerical examples confirm the theoretical results for the proposed improved techniques.

For the Conjugate Gradient method we solved the two-dimensional Helmholtz equation

$$-\Delta u + u = f$$

in the unit square Ω with Dirichlet boundary conditions. We have discretized the unit square by taking the $n \times n$ uniform grid, then we approximated the above problem by finite differences. We solved the $n^2 \times n^2$ irreducible positive definite system yielded by the CG method, by the PCG method, with improved two-indexed preconditioner S_2, as was described in Theorem 4.1, and by the PCG method, with improved four-indexed preconditioner S_4 which is based on the elimination of the first and nth diagonals over and under the main diagonal. As a stopping criterion we chose $\|r\|_2 \leqslant 0.5 \times 10^{-6}$, where r is the residual vector. In Table 3 the numbers of required iterations, for the CG and for the two PCG methods, are given for various values of n. The efficiency of the proposed method is clearly shown.

For the GMRES method we solved the two-dimensional convection–diffusion equation

$$-\Delta u + \frac{\partial u}{\partial x} + 2\frac{\partial u}{\partial y} = f$$

in the unit square Ω with Dirichlet boundary conditions. We followed the same technique to approximate the above problem. Then, we solved the $n^2 \times n^2$ irreducible M-system yielded by the GMRES method, by the PGMRES(S_2) method and by the PGMRES(S_4) one. The last two PGMRES methods are based on the improved two-indexed S_2 and four-indexed S_4 preconditioners, respectively, as they were described previously. The same stopping criterion was chosen and the restarted technique of GMRES was run. In the first part of Table 4 the number of $m = n$ iterations was chosen to restart the method while in the second, $m = \frac{n}{2}$. For each method, a pair of integers is given, where the first one corresponds to the number of completed outer iterations while the second, to the number of the inner iterations, required in the last uncompleted outer one. We also counted the number of the required operations in each case. So, in Table 4 we give them, for PGMRES methods, in percentages with respect to the GMRES method, for each case.

Table 4
Restarted preconditioned GMRES method, for two-dimension convection–diffusion equation

$m = n$								
n	*GMRES*		*PGMRES*(S_2)		ops%	*PGMRES*(S_4)		ops%
4	6	0	4	0	93.95	2	0	55.71
8	5	3	3	7	97.42	2	0	59.37
16	4	9	3	5	92.10	1	14	59.75
32	4	1	2	25	78.67	1	26	56.15
64	3	42	2	28	71.96	1	45	53.07
128	3	83	2	35	64.77	1	65	41.75
$m = n/2$								
4	17	0	10	0	86.62	6	2	70.61
8	14	3	9	3	94.14	5	3	66.23
16	12	5	7	6	83.08	4	6	59.91
32	12	4	7	11	79.03	4	2	48.96
64	11	20	7	17	75.89	3	29	43.78
128	10	51	7	1	72.63	3	54	41.55

Acknowledgements

Warm thanks are due to the referees who made useful suggestions and provided some more references that helped us to improve the quality of the paper.

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, A. Hadjidimos, Block gauss elimination followed by a classical iterative method for the solution of linear systems, JCAM 163 (2004) 381–400.
[3] O. Axelsson, G. Lindskog, On the rate of convergence of the preconditioned conjugate gradient method, Numer. Math. 48 (1986) 499–523.
[4] A. Berman, R.J. Plemmons, Nonnegative Matrices in the Mathematical Sciences, Classics in Applied Mathematics, SIAM, Philadelphia, 1994.
[5] R.E. Funderlic, R.J. Plemmons, LU decomposition of M-matrices by elimination without pivoting, LAA 41 (1981) 99–110.
[6] A.D. Gunawardena, S.K. Jain, L. Snyder, Modified iterative methods for consistent linear systems, LAA 154–156 (1991) 123–143.
[7] R.A. Horn, C.R. Jonson, Topics in Matrix Analysis, Cambridge University Press, Cambridge, 1991 (Also: second ed. Revised and Expanded, Springer, Berlin, 2000.).
[8] A. Hadjidimos, D. Noutsos, M. Tzoumas, More on modifications and improvements of classical iterative schemes for Z-matrices, LAA 364 (2003) 253–279.
[9] M.L. Juncosa, T.W. Mulliken, On the increase of convergence rates of relaxation procedures for elliptic partial differential equations, J. Assoc. Comput. Math. 7 (1960) 29–36.

[10] T. Kohno, H. Kotakemori, H. Niki, M. Usui, Improving the Gauss–Seidel Method for Z-Matrices, LAA 267 (1997) 113–123.
[11] H. Kotakemori, K. Harada, M. Morimoto, H. Niki, A comparison theorem for the iterative method with the preconditioner (I+Smax), JCAM 145 (2002) 373–378.
[12] W. Li, Comparison results for solving preconditioned linear systems, JCAM 176 (2005) 319–329.
[13] W. Li, W. Sun, Modified Gauss–Seidel type methods and Jacobi type methods for Z-matrices, LAA 317 (2000) 227–240.
[14] I. Marek, D.B. Szyld, Comparison theorems for weak splittings of bounded operators, Numer. Math. 58 (1990) 387–397.
[15] J.P. Milaszewicz, On modified Jacobi linear operators, LAA 51 (1983) 127–136.
[16] J.P. Milaszewicz, Improving Jacobi and Gauss–Seidel iterations, LAA 93 (1987) 161–170.
[17] H. Niki, K. Harada, M. Morimoto, M. Sakakihara, The survey of preconditioners used for accelerating the rate of convergence in the Gauss–Seidel method, JCAM 164–165 (2004) 587–600.
[18] R.S. Varga, Matrix Iterative Analysis, Prentice-Hall, Englewood Cliffs, NJ, 1962 (Also: second ed., Revised and Expanded, Springer, Berlin, 2000.).
[19] Z. Woźnicki, Nonnegative splitting theory, Jap. J. Ind. Appl. Math. 11 (1994) 289–342.
[20] D.M. Young, Iterative Solution of Large Linear Systems, Academic Press, New York, 1971.