MATHEMATICA — REVUE D'ANALYSE NUMÉRIQUE
ET DE THÉORIE DE L'APPROXIMATION

L'ANALYSE NUMÉRIQUE ET LA THÉORIE DE L'APPROXIMATION
Tome 16, N° 2, 1987, pp. 109—115

IMPROVEMENT OF THE AREA OF CONVERGENCE OF THE AOR METHOD

LJ. CVETKOVIĆ and D. HERCEG
(Novi Sad)

1. Introduction. We consider a system of linear equations, written in the matrix form

$$Ax = b,$$

where $A \in C^{n,n}$ is a nonsingular matrix with nonzero diagonal entries, and $x, b \in C^n$ with x unknown and b known. For the numerical solution of this system we use the accelerated overrelaxation (AOR) method, which is introduced by Hadjidimos in [6], and which is a two-parameter's generalization of the SOR method. Since the AOR method had been introduced, many properties as well as numerical results concerning it have been given by several autors. Numerical examples from [1], [6] show the superiority of the AOR method. A lot of papers are referred to the linear systems with matrix which is strictly diagonally dominant (SDD), irreducible diagonally dominant (IDD), generalized diagonally dominant (GDD), an M- or an H-matrix (cf. [1], [6], [7], [9], [10], [11], [12]). In [2], [8] some new classes of linear systems have been considered. Here, we shall consider the class of H-matrices, because we had proved in [3] that all of the mentioned classes are H-matrices. By using a new technique, which is based on a generalization of Sassenfeld's criteria, we are going to get an improvement for the area of convergence of the AOR method for all of the mentioned classes of matrices.

From now on, without loss of generality, we can suppose that $a_{ii} = 1$, $i \in N$.

Let $A = E - L - U$ be the decomposition of the matrix A into its diagonal, strictly lower and strictly upper triangular parts, respectively, and let $\omega, \sigma \in R$, $\omega \neq 0$. The associated AOR method can be written as

$$x^{k+1} = M_{\sigma,\omega}\, x^k + d,\ k = 0, 1, \ldots,\ x^0 \in C^n,$$

where $M_{\sigma,\omega} = (E - \sigma L)^{-1}((1 - \omega) E + (\omega - \sigma) L + \omega U)$, $d = \omega(E - \sigma L)^{-1} b$.
Some special cases of this method are: for $\omega = \sigma$ SOR method, for $\omega = \sigma = 1$ Gauss-Seidel, for $\sigma = 0$ JOR and for $\sigma = 0$, $\omega = 1$ Jacobi method. As one can see, the AOR method is an extrapolation of either the Jacobi method (case $\sigma = 0$) or the SOR method (case $\sigma \neq 0$, where the extrapolation parameter is ω/σ).

2. Preliminaries. We shall use the following notations:

$$N = \{1, 2, \ldots, n\}, \quad N(i) = N \setminus \{i\}, \quad i \in N.$$

For any matrix $A = [a_{ij}] \in C^{n,n}$ (= set of all complex $n \times n$ matrices) and $i \in N$, we define

$$P_i(A) = \sum_{j \in N(i)} |a_{ij}|.$$

DEFINITION 1. *A real square matrix whose off-diagonal elements are all non-positive is called L-matrix.*

DEFINITION 2. *A regular L-matrix A, for which* $A^{-1} > 0$ *is called M-matrix.*

For any matrix $A = [a_{ij}] \in C^{n,n}$, we define $M(A) = [m_{ij}] \in R^{n,n}$ as follows

$$m_{ii} = |a_{ii}|, \quad i \in N, \quad m_{ij} = -|a_{ij}|, \quad i \in N, \ j \in N(i).$$

DEFINITION 3. *A matrix A is called H-matrix iff M(A) is an M-matrix.*

DEFINITION 4. *A matrix A is called generalized diagonally dominant (GDD) iff there exists a regular diagonal matrix M, so that* AM *is* SDD.

It is easy to see that the matrix A is GDD iff it is an H-matrix. By using this fact we shall conclude that it is sufficient to consider only the class of SDD matrices.

3. The Convergence of AOR Method

LEMMA 1. *Let* $p_i(\sigma) = \sum_{j=1}^{i-1} |a_{ij}|(|1-\sigma| + |\sigma| p_j(\sigma)) + \sum_{j=i+1}^{n} |a_{ij}|, \ i \in N,$ $p(\sigma) = \max_i p_i(\sigma)$. *Then for the matrix* $M_{\sigma,\omega}$ *of the* AOR *method it holds that* $\|M_{\sigma,\omega}\|_\infty \leq |1 - \omega| + |\omega|\, p(\sigma)$.

Proof: From the definition of the matrix norm $\| \ \|_\infty$, there exist a vector $y \in C^n$ such that

$$\|y\|_\infty = 1, \quad \|M_{\sigma,\omega}\|_\infty = \|M_{\sigma,\omega}\, y\|_\infty.$$

We denote $z = M_{\sigma,\omega}\, y$. Hence,

(3.1) $$(E - \sigma L)\, z = ((1 - \omega)\, E + (\omega - \sigma)\, L + \omega U)\, y.$$

Now, we are going to prove that for each $i \in N$ it holds that

(3.2) $$|z_i - (1 - \omega)\, y_i| \leq |\omega|\, p_i(\sigma) \ \text{and} \ |z_i| \leq |1 - \omega| + |\omega|\, p_i(\sigma).$$

For $i = 1$, by using (3.1), we have

$$z_1 = (1 - \omega)\, y_i - \omega \sum_{j=2}^{n} a_{1j}\, y_j,$$

and (3.2) holds because of $|y_i| \leq 1$, $i \in N$.

Suppose that (3.2) holds for $i \leq k - 1$ $(k = 2, \ldots, n)$ and prove that it holds for $i = k$. From (3.1), we obtain

$$z_k - (1 - \omega)\, y_k = -\omega \sum_{j=k+1}^{n} a_{kj}\, y_j - \sum_{j=1}^{k-1} a_{kj}(\omega y_j + \sigma z_j - \sigma y_j)$$

$$= -\omega \sum_{j=k+1}^{n} a_{kj}\, y_j - \omega \sum_{j=1}^{k-1} a_{kj}[(1 - \sigma)\, y_j + \sigma(z_j - (1 - \omega)\, y_j)/\omega],$$

and

$$|z_k - (1 - \omega)\, y_k| \leq |\omega| \sum_{j=k+1}^{n} |a_{kj}| +$$

$$+ |\omega| \sum_{j=1}^{k-1} |a_{kj}|\, (|1 - \sigma| + |\sigma|\, |z_j - (1 - \omega)\, y_j| / |\omega|)\} \leq |\omega|\, p_k(\sigma).$$

Now it is easy to see that

$$|z_k| \leq |1 - \omega| + |\omega|\, p_k(\sigma).$$

The second inequality from (3.2) gives $\|z\|_\infty \leq |1 - \omega| + |\omega|\, p(\sigma)$ and proof is complete.

COROLLARY 1.1. *If* $1 - |\sigma| I_i > 0$, $i \in N$, *then*

$$\|M_{\sigma,\omega}\|_\infty \leq \max_i\, (|1 - \omega| + (|\omega|\, |1 - \sigma| - |\sigma|\, |1 - \omega|)\, I_i + |\omega|\, u_i)/(1 - |\sigma|\, I_i),$$

where $I_i = P_i(L)$, $u_i = P_i(U)$.

Proof: Obviously,

$$p(\sigma) \leq (|1 - \sigma| + |\sigma|\, p(\sigma))\, I_m + u_m,$$

for $m \in N$ for which we have $p(\sigma) = p_m(\sigma)$. Hence,

$$p(\sigma) \leq \max_i\, (|1 - \sigma|\, I_i + u_i)/(1 - |\sigma|\, I_i).$$

Now, $\|M_{\sigma,\omega}\|_\infty \leq |1 - \omega| + |\omega|\, p(\sigma) \leq$

$$\leq \max_i\, (|1 - \omega| + (|\omega|\, |1 - \sigma| - |\sigma|\, |1 - \omega|)\, I_i + |\omega|\, u_i)/(1 - |\sigma|\, I_i),$$

which completes the proof.

Corollary 1.1 gives an upper bound (let us denote it by ε) for the spectral radius of the matrix $M_{\sigma,\omega}$. So, sufficient conditions for the convergence of AOR method can be obtained from the condition $\varepsilon < 1$. It is clear that the condition

$$|1 - \omega| + |\omega|\, p(\sigma) < 1,$$

which was obtained in [4], is more general than $\varepsilon < 1$, but it does not give a possibility to say (in advance) how to choose the parameters σ and ω so that AOR method converges. By solving inequality $\varepsilon < 1$ and by the extrapolation theorem (see [7]), we obtain our area of convergence of the AOR method.

THEOREM 2. *Let A be a strictly diagonally dominant matrix and let $l_i + P_i(L)$, $u_i = P_i(U)$, $i \in N$. Then* AOR *method converges for*:

(*i*) $0 < \sigma < 2/(1+p(M_{0,1}(M(A)))) = : s$, $0 < \omega < 2\sigma/(i+p(M_{\sigma,\sigma})) = : r$ *or*

(*ii*) $0 < \omega \leqslant 1$, $-\min_i (1 - l_i - u_i)/2l_i < \sigma < \min_i (1 + l_i - u_i)/2l_i$ *or*

(*iii*) $1 < \omega < 2 - \max_i 2u_i/(1 + u_i - l_i) = : q$,

$$\max\{0, \max_i ((\omega(1 + l_i + u_i) - 2)/(2(\omega - 1) l_i))\} < \sigma <$$

$$< \min_i (2 - \omega(1 - l_i + u_i))/2l_i \text{ or}$$

(*iv*) $1 < \omega < 2/(1 + \max_i (l_i + u_i)) = : t$,

$$\max_i (\omega(1 + l_i + u_i) - 2)/2l_i < \sigma < 0.$$

Proof: It is easy to verify that for each σ, which satisfies one of the conditions (*ii*)—(*iv*), we have

$$1 - |\sigma| l_i > 0, \; i \in N.$$

(*i*) Since A is SDD matrix, then $M(A)$ is an M-matrix, and from [16] it follows that for $0 < \sigma < s$ it holds that

$$p(M_{\sigma,\sigma}) < 1.$$

It is known that for $\sigma \neq 0$, $M_{\sigma,\omega} = \left(1 - \dfrac{\omega}{\sigma}\right) E + \dfrac{\omega}{\sigma} M_{\sigma,\sigma}$.

If $0 < \omega/\sigma < r$, by using the Extrapolation theorem, [7], we conclude that $p(M_{\sigma,\omega}) < 1$.

(*ii*) If $0 < \sigma < 1$, it holds $|\omega| |1 - \sigma| - |\sigma| |1 - \omega| = \omega - \sigma$ and

$$1 - \omega + (\omega - \sigma) l_i + \omega u_i < 1 - \sigma l_i, \; i \in N \text{ because of}$$

$$-\omega(1 - l_i - u_i) < 0, \; i \in N.$$

If $\sigma > 1$, we have $|\omega| |1 - \sigma| - |\sigma| |1 - \omega| = 2\sigma\omega - \sigma - \omega$ and

$$\sigma < (1 - u_i + l_i)/2l_i$$

$$\Rightarrow 2\sigma l_i < 1 - u_i + l_i$$

$$\Rightarrow 2\sigma\omega l_i - \omega + \omega u_i - \omega l_i < 0$$

$$\Rightarrow 1 - \omega + (2\sigma\omega - \sigma - \omega) l_i + \omega u_i < 1 - \sigma l_i, \; i \in N,$$

and from Corollary 1.1 we obtain $p(M_{\sigma,\omega}) < 1$.

If $\sigma < 0$, we have $|\omega| |1 - \sigma| - |\sigma| |1 - \omega| = \sigma + \omega - 2\sigma\omega$ and

$$\sigma > -(1 - l_i - u_i)/2l_i$$

$$\Rightarrow -2\sigma\omega l_i < \omega - \omega l_i - \omega u_i$$

$$\Rightarrow 1 - \omega + (\sigma + \omega - 2\sigma\omega) l_i + \omega u_i < 1 + \sigma l_i, \; i \in N.$$

From Corollary 1.1 it holds that $p(M_{\sigma,\omega}) < 1$.

(*iii*) and (*iv*) can be proved similarly, by using the same Corollary.

Detailed analysis shows that the area of convergence for the class of SDD matrices, given in [12], is always smaller than this one. Here we illustrate this fact by the following example.

Example 1. The area of convergence for σ and ω obtained by Theorem 2 in case when

$$A = \begin{bmatrix} 1 & -0.0625 \\ -0.25 & 1 \end{bmatrix},$$

is:

(*i*) $0 < \sigma < 16/9$, $0 < \omega < 2\sigma/(1 + p(M_{\sigma,\omega}))$ or

(*ii*) $0 < \omega \leqslant 1$, $-1.5 < \sigma < 2.5$ or

(*iii*) $1 < \omega < 32/17$, $2.5 - 1.5/(\omega - 1) < \sigma < (8 - 3\omega)/2$ or

(*iv*) $1 < \omega < 1.6$, $(5\omega - 8)/2 < \sigma < 0$.

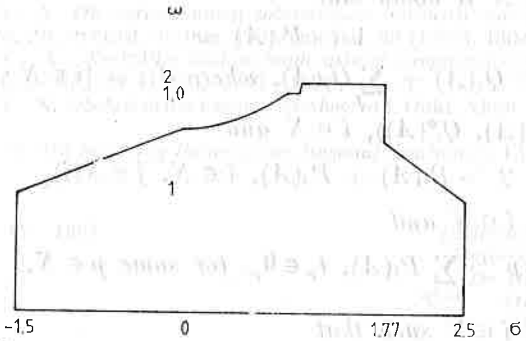

Fig. 1

We give a geometric interpretation of Theorem 2 for this example (fig. 1). We can see that the area of convergence obtained here is larger that the one from Theorem 4 from [12] (fig. 2).

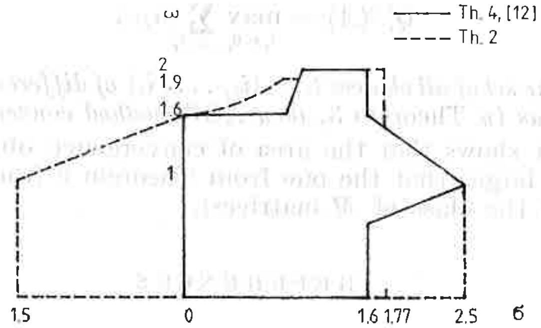

Fig. 2

Now, we can use the result of Theorem 2 in order to improve the area of convergence for the parameters σ and ω in case when A is an H-matrix, i.e. GDD matrix.

Since $p(M_{0,1}(M(A))) = p(M_{0,1}(M(AW)))$ and $p(M_{\sigma,\omega}(A)) = p(M_{\sigma,\omega}(AW))$ for a regular matrix W, we obtain the following theorem.

THEOREM 3. *If A is an H-matrix (i.e. GDD) and the parameters σ and ω are chosen as in* Theorem 2, *where $l_i = P_i(LW)$ and $u_i = P_i(UW)$, $i \in N$, then $p(M_{\sigma,\omega}(A)) < 1$.*

COROLLARY 3.1. *Let A be an* IDD *or an M-matrix or a matrix whose elements satisfy at least one of the following conditions:*

(*i*) $1 > P_i(A)$, $i \in N$ (SDD),

(*ii*) $1 > P_{i,\alpha}(A)$, $i \in N$, *for some* $\alpha \in [0, 1]$,

(*iii*) $1 > P_i^{\alpha}(A)\, Q_i^{1-\alpha}(A)$, $i \in N$, *for some* $\alpha \in [0, 1]$,

(*iv*) $1 > P_i(A)\, P_j(A)$, $i \in N$, $j \in N(i)$,

(*v*) $1 > P_i^{\alpha}(A)\, Q_i^{1-\alpha}\, P_j^{\alpha}(A)\, Q_i^{1-\alpha}(A)$, $i \in N$, $j \in N(i)$, *for some* $\alpha \in [0, 1]$,

(*vi*) *For each $i \in N$ it holds that*

$$1 > P_i(A) \text{ or}$$

$$1 + \operatorname{card}(J) > Q_i(A) + \sum_{j \in J} Q_j(A), \text{ where } J := \{i \in N : 1 \leqslant Q_i(A)\},$$

(*vii*) $1 > \min(P_i(A), Q_i^*(A))$, $i \in N$ *and*

$$2 > P_i(A) + P_j(A), \ i \in N, \ j \in N(i),$$

(*viii*) $1 > Q_i^{(p)}(B)$, $i \in N$ *and*

$$p > \sum_{j \in t_p} P_i(A), \ t_p \in \theta_p, \text{ for some } p \in N,$$

(*ix*) *There exists $i \in N$ such that*

$$1 - P_j(A) + |a_{ji}|) > P_i(A) |a_{ji}|, \ j \in N(i),$$

where
$$Q_i(A) = \sum_{j \in N(i)} |a_{ji}|,$$

$$P_{i,\alpha}(A) = \alpha P_i(A) + (1 - \alpha)\, Q_i(A), \quad Q_i^*(A) = \max_{j \in N(i)} |a_{ji}|,$$

$$Q_i^{(r)}(A) = \max_{t_r \in \theta_r} \sum_{j \in t_r} |a_{ji}|,$$

$r \in N$ and θ_r is the set of all choices $t_r = \{i_1, \ldots, i_r\}$ of different indices from N. If σ and ω are as in Theorem 3, *then* AOR *method converges.*

Example 1 shows that the area of convergence obtained by Corollary 3.1 is still larger that the one from Theorem 8 from [12] (which is related only to the class of M-matrices).

REFERENCES

[1] Avdelas, G., Hadjidimos, A., Yeyios, A., *Some theoretical and computational results concerning the accelerated overrelaxation* (AOR) *method*, Anal. Numer. Theor. Approx., **9** (1980), No. 1, 5–10.

[2] Cvetković, Lj., Herceg, D. *Some sufficient conditions for convergence of* AOR-*method*, Numerical Methods and Approximation Theory (G. V. Milovanović, ed.), Niš, 1984, 143–148.

[3] Cvetković, Lj., Herceg, D. *Some Results on M-and H-matrices*, Zb. Rad. Prir. Mat. Fak. Univ. Novom Sadu Ser. Mat., **16**, No. 2 (1986) (to appear).

[4] Chernyak, V. Ya. *O vliyanii nelineinosti reshaemoĭ sistemy na shodimost' metodov Yakobi*, PVR- *i* UVR-*N'yutona*, Chislennye metody analiza i ih prilozheniya, Akad. Nauk SSSR Sibirsk. Otdel., Energet. Inst., Irkutsk, 1983, 139–150.

[5] Dashnic, M. S., Zusmanovich, M. S. *O nekotoryh kriteriyah regulyarnosti matric i lokalizacii ih spektra*, Zh. vychisl. matem. i matem. fiz., **5** (1970), 1092–1097.

[6] Hadjidimos, A. *Accelerated Overrelaxation Method*, Math., Comp., **32** (1978), 149–157.

[7] Hadjidimos, A., Yeyios, A. *The principle of extrapolation in connection with the accelerated overrelaxation method*, Linear Algebra Appl., **30** (1980), 115–128.

[8] Herceg, D. *Über die Konvergenz des* AOR-*Verfahrens*, Z. Angew. Math. Mech., **65** (1985), 378–379.

[9] Martins, M. *On an accelerated overrelaxation iterative method for linear systems with strictly diagonally dominant matrix*, Math. Comp., **35** (1980), 1269–1273.

[10] Martins, M. *Note on irreducible diagonally dominant matrices and the convergence of the* AOR *iterative method*, Math. Comp., **37** (1981), 101–103.

[11] Martins, M. *Generalized diagonal dominance in connection with the accelerated overrelaxation* (AOR) *method*, BIT, **22** (1982), 73–78.

[12] Martins, M. *An improvement for the area of convergence of the accelerated overrelaxation iterative method*, Anal. Numer. Theor. Approx., **12** (1983), No. 1, 65–76.

[13] Pupkov, V. A. *Ob izolirovannom sobstvennom znachenii matricy i strukture ego sobstvennogo vektora*, Zh. vychisl. matem. i matem. fiz., **23** (1983), 1304–1313.

[14] Pupkov, V. A. *Neskol'ko dostatochnyh uslovii nevyrozhdenosti matric*, Zh. vychisl. matem. i matem. fiz., **24** (1984), 1733–1737.

[15] Solov'ev, V. N. *Obobshchenie teoremy Gershgorina*, Dokl. Akad. Nauk SSSR, **47** (1983), 1285–1301.

[16] Varga, R. S. *On recurring theorems on diagonal dominance*, Linear Algebra Appl., **13** (1976), 1–9.

Received 2. IV. 1987

Institute of Mathematics
University of Novi Sad
Dr Ilije Djuričića 4
21000 Novi Sad
Yugoslavia