Reliable Computing **6:** 123–137, 2000.

Jacobi and Gauss-Seidel Iterations for Polytopic Systems: Convergence *via* Convex M-Matrices

DUŠAN M. STIPANOVIĆ
Department of Electrical Engineering, Santa Clara University, Santa Clara, CA 95053-0569, USA, e-mail: dstipanovic@scu.edu

and

DRAGOSLAV D. ŠILJAK
B & M Swig Professor, School of Engineering, Santa Clara University, Santa Clara, CA 95053-0569, USA, e-mail: dsiljak@scu.edu

(Received: 31 October 1998; accepted: 8 June 1999)

Abstract. A natural generalization of the Jacobi and Gauss-Seidel iterations for interval systems is to allow the matrices to reside in convex polytopes. In order to apply the standard convergence criteria involving M-matrices to iterations for polytopic systems, we derive conditions for a convex polytope of matrices to be a polytope of M-matrices in terms of its vertices. We show how the conditions are used in the convergence analysis of iterations for block and nonlinear polytopic systems.

1. Introduction

Uncertainty in solving systems of equations occurs either due to uncertain data caused by modeling errors and approximations, or due to rounding errors in the solution process. Typically, a system of equations is parametrized by a vector of uncertain parameters, which belong to a fixed hyperrectangle (box), and interval arithmetic tools are used to solve parameter dependent equations. There is a wealth of results available in convergence analysis of Jacobi and Gauss-Seidel iterations which are used in solving both linear and nonlinear systems of interval equations [10], [11], [17].

A natural generalization of the interval paradigm is to allow the uncertain parameters to reside in a convex polytope. Our objective is to derive a number of sufficient conditions for convergence of Jacobi and Gauss-Seidel iterations for polytopic systems using the standard M-matrix properties [3], [13]. The obtained conditions, which generalize those established in [21], involve only vertices of the polytope and can be readily extended to involve more flexible H-matrices. The conditions are easily tested using efficient numerical methods of linear and nonlinear convex programming.

In solving large systems of linear and nonlinear equations on parallel machines, it is common to use block iterative schemes [2], [4], [10], [22], [24], [25]. Sharp convergence results have been obtained for the block iterative methods involving interval uncertainty [10]. Our objective in this context is to derive a convergence test for the Jacobi iteration involving polytopic systems using linear programming format and the block diagonal dominance conditions of [19].

Sufficient conditions for convergence of nonlinear Gauss-Seidel and Jacobi iterations in terms of M-functions are well-known [14], [20]. By relying on the connection between M-functions and M-matrices established in [9], [15], [18], we show how our results can be applied to test convergence of the Gauss-Seidel and Jacobi iterations for polytopic nonlinear systems.

2. Polytopic Iterations

Let us consider a system of linear equations

$$Ax = b, \tag{2.1}$$

where $x \in \mathbf{R}^n$ is the vector of unknowns, $A \in \mathbf{R}^{n\times n}$ is a constant matrix, and $b \in \mathbf{R}^n$ is a constant vector. Our crucial assumption is that the matrix $A = (a_{ij})$ has uncertain elements a_{ij} and belongs to a polytope

$$\mathbf{A} = \text{conv}\{A^\ell\}, \tag{2.2}$$

which is the convex hull of known vertex matrices A^ℓ, $\ell \in \{1, 2, \dots, m\} = \mathbf{m}$. In other words, any matrix $A \in \mathbf{A}$ can be expressed as a convex combination

$$A(\alpha) = \sum_{\ell=1}^{m} \alpha_\ell A^\ell, \tag{2.3}$$

where the vector α belongs to the unit simplex

$$\mathbf{U}_m = \left\{\alpha \in \mathbf{R}^m : \sum_{\ell=1}^{m} \alpha_\ell = 1,\ \alpha_\ell \geq 0,\ \ell \in \mathbf{m}\right\}. \tag{2.4}$$

We want to formulate Gauss-Seidel and Jacobi iterations which can solve (2.1) when $A \in \mathbf{A}$. To do that we first split the matrix $A = A(\alpha)$ as

$$A(\alpha) = \Delta(\alpha) - L(\alpha) - U(\alpha), \tag{2.5}$$

where $\Delta(\alpha) = \text{diag}\{a_{11}(\alpha), a_{22}(\alpha), \dots, a_{nn}(\alpha)\}$, and $L(\alpha)$ and $U(\alpha)$ are respectively strict lower and strict upper triangular $n \times n$ matrices, whose elements are the negatives of the elements of $A(\alpha)$, and which are situated below and above the main diagonal. Then, equation (2.1) can be rewritten in the following two forms:

$$\text{Gauss-Seidel} \quad [\Delta(\alpha) - L(\alpha)]x = U(\alpha)x + b, \tag{2.6}$$

$$\text{Jacobi} \quad \Delta(\alpha)x = [L(\alpha) + U(\alpha)]x + b. \tag{2.7}$$

is positive definite. The matrix D is called the positive diagonal Liapunov solution.

(vi) A is nonsingular and $A^{-1} \geq 0$.

(vii) All eigenvalues of A have positive real parts, that is, A is positive stable, and the eigenvalue with the smallest absolute value is positive.

It is well-known (e.g., [20], [24]) that both Jacobi and Gauss-Seidel iterations are convergent if A is an M-matrix. What we want now is to derive convergence criteria for polytopic iterations by using only the vertex matrices A^ℓ of the given polytope $\mathbf{A}$, which are assumed to be M-matrices. This objective requires convexity of M-matrices which is known to be false in general. In [13], the necessary and sufficient conditions for a convex combination of two M-matrices to be an M-matrix are established, but a generalization of the conditions to a polytope of matrices is an open problem.

The rowwise and columnwise dominance were introduced in [7] as sufficient conditions for a convex combination of a pair of M-matrices to be an M-matrix. A generalization of the rowwise condition to a matrix polytope is the following [23]:

THEOREM 3.2. *Let $A^\ell \in \mathbf{M}_n$ for all $\ell \in \mathbf{m}$. Define the augmented matrices $\overline{A}^\ell \in \mathbf{R}^{n\times(n+1)}$ such that $\overline{A}^\ell = [A^\ell, -e]$ where $e = (1, 1, \ldots, 1)^T \in \mathbf{R}^n_+$, and the augmented vector $\overline{x} = (x_1, x_2, \ldots, x_n, x_{n+1})^T = (x^T, x_{n+1})^T \in \mathbf{R}^{n+1}$. Consider the following linear programming problem:*

$$\begin{aligned} &\max x_{n+1} \\ &\text{subject to } \overline{A}^\ell \overline{x} \geq 0, \ \ \ell \in \mathbf{m}; \ \ \overline{x} \geq 0, \ \ x_{n+1} \leq 1. \end{aligned} \tag{3.4}$$

Then, $\mathbf{A} \subset \mathbf{M}_n$ if $x^0_{n+1} = \max x_{n+1} = 1$.

Proof. Observe that if $A^\ell \in \mathbf{M}_n$ for all $\ell \in \mathbf{m}$, then $A(\alpha) \in \mathbf{Z}_n$ for all $\alpha \in \mathbf{U}_m$. Now we shall prove that $x^0_{n+1} = 1$ is equivalent to the existence of a positive vector d such that $A(\alpha)d$ is positive for all $\alpha \in \mathbf{U}_m$, which by part (ii) of Theorem 3.1, further implies that $\mathbf{A}$ is a polytope of M-matrices. To prove sufficiency let us assume that $x^0_{n+1} = 1$. Then, $x \geq 0$ is such that $A^\ell x \geq e$ for all $\ell \in \mathbf{m}$. Furthermore, for a sufficiently small number $\varepsilon > 0$, for which we have $-\varepsilon A^\ell e < e$, $\ell \in \mathbf{m}$, we can set $d = x + \varepsilon e$. This means that $d > 0$ and $A^\ell d = A^\ell(x + \varepsilon e) = A^\ell x + \varepsilon A^\ell e \geq e + \varepsilon A^\ell e > 0$ for all $\ell \in \mathbf{m}$, implying

$$A(\alpha)d = \sum_{l=1}^{m} \alpha_\ell A^\ell d > 0, \quad \forall \alpha \in \mathbf{U}_m. \tag{3.5}$$

To establish necessity let us assume that there is a common $d > 0$ such that $A^\ell d > 0$ for all $\ell \in \mathbf{m}$. Notice that positivity of all vectors $A^\ell d$ is equivalent to positivity of $A(\alpha)d$ for all $\alpha \in \mathbf{U}_m$. Furthermore, we can set $\overline{x} = (d^T, 1)^T$. At this point we check if this $\overline{x}$ is a solution to (3.4) with $x^0_{n+1} = 1$. If it is not a solution, we

We assume that $a_{ii}(\alpha) \neq 0$ for all $\alpha \in \mathbf{U}_m$ and $i \in \mathbf{n} = \{1, 2, \ldots, n\}$, so that both matrices $\Delta(\alpha)$ and $\Delta(\alpha) - L(\alpha)$ are invertible for all $\alpha \in \mathbf{U}_m$. Now, the Gauss-Seidel and Jacobi iterations corresponding to polytopic system (2.1) are defined as

$$\text{Gauss-Seidel} \quad x^{k+1} = G(\alpha)x^k + [\Delta(\alpha) - L(\alpha)]^{-1}b, \tag{2.8}$$

$$\text{Jacobi} \quad x^{k+1} = J(\alpha)x^k + \Delta^{-1}(\alpha)b, \tag{2.9}$$

where $k = 0, 1, \ldots,$ and the iterations matrices are

$$\text{Gauss-Seidel} \quad G(\alpha) = [\Delta(\alpha) - L(\alpha)]^{-1}U(\alpha), \tag{2.10}$$

$$\text{Jacobi} \quad J(\alpha) = \Delta^{-1}(\alpha)[L(\alpha) + U(\alpha)]. \tag{2.11}$$

For a fixed α, the Gauss-Seidel iteration is convergent iff $\rho(G) < 1$ and, similarly, the Jacobi iteration is convergent iff $\rho(J) < 1$, where $\rho(\cdot)$ denotes the spectral radius of the indicated matrix (e.g., [24]). It is then obvious that the respective iterations for polytopic systems would be convergent iff $\rho[G(\alpha)] < 1$ and $\rho[J(\alpha)] < 1$ for all $\alpha \in \mathbf{U}_m$. These conditions are untestable as they are stated and our objective is to obtain convergence conditions in terms of M-matrix properties of the vertex matrices A^ℓ of the polytope $\mathbf{A}$.

3. Convex M-matrices

In this section, we consider the set $\mathbf{M}_n$ of M-matrices which is a subset of the set

$$\mathbf{Z}_n = \{A \in \mathbf{R}^{n \times n} : a_{ij} \leq 0, \ i \neq j, \ \forall i, j \in \mathbf{n}\} \tag{3.1}$$

of matrices with nonpositive off-diagonal elements. Before we provide a number of M-matrix properties (e.g., [3], [8]), which we need in this paper, let us define the set of positive vectors

$$\mathbf{R}_+^n = \{x \in \mathbf{R}^n : x_i > 0, \ \forall i \in \mathbf{n}\}. \tag{3.2}$$

Let us also recall the standard componentwise order: For two matrices $A, B \in \mathbf{R}^{n \times n}$, $A \geq B$ means $a_{ij} \geq b_{ij}$, and $A > B$ implies $a_{ij} > b_{ij}$. In an obvious way, the ordering carries over to vectors. Finally the superscript T denotes transposition.

THEOREM 3.1. *Let $A \in \mathbf{Z}_n$. Then, $A \in \mathbf{M}_n$ iff it satisfies one of the following equivalent conditions:*

(i) *There is a vector $x \geq 0$ such that $Ax > 0$.*

(ii) *There is a vector $x > 0$ such that $Ax > 0$.*

(iii) *There is a vector $x \geq 0$ such that $A^T x > 0$.*

(iv) *There is a vector $x > 0$ such that $A^T x > 0$.*

(v) *There is a positive diagonal matrix $D = \mathrm{diag}\{d_1, d_2, \ldots, d_n\}$ such that the matrix*

$$C = A^T D + DA \tag{3.3}$$

could find $\gamma = \min\{f(A^1 d), f(A^2 d), \ldots, f(A^m d)\}$, where the function $f : \mathbf{R}_+^n \to \mathbf{R}_+$ is defined as the smallest positive element of a positive vector. Then, the vector $\bar{x} = (\gamma^{-1} d^T, 1)^T$ is a solution of (3.4) with $x_{n+1}^0 = 1$. □

Remark 3.1. Notice that the problem (3.4) is convex and, therefore, if there is a solution to (3.4) a linear program (e.g., simplex method) would find it.

Remark 3.2. We state the columnwise version of Theorem 3.2. Let $A^\ell \in \mathbf{M}_n$ for all $\ell \in \mathbf{m}$. Define the augmented matrices $\overline{A}^\ell \in \mathbf{R}^{n\times(n+1)}$ such that $\overline{A}^\ell = [(A^\ell)^T, -e]$. From the linear programming problem (3.4) it follows that there exists a vector $\tilde{d} > 0$ such that $A^T(\alpha)\tilde{d} > 0$ iff $x_{n+1}^0 = 1$. The condition $A^T(\alpha)\tilde{d} > 0$ implies that $\mathbf{A}$ is a polytope of M-matrices.

Another way to establish $\mathbf{A}$ as a polytope of M-matrices is to apply the concept of "simultaneous Liapunov functions" introduced in [12]. We use condition (v) of Theorem 3.1 to get the following [23]:

THEOREM 3.3. *Let $A^\ell \in \mathbf{M}_n$ for all $\ell \in \mathbf{m}$. Then, the matrices A^ℓ have a common positive diagonal Liapunov solution D if the following system of linear matrix inequalities:*

$$\begin{aligned} (A^\ell)^T D + D A^\ell &\succ 0, \quad \ell \in \mathbf{m} \\ D &\succ 0 \end{aligned} \tag{3.6}$$

is feasible, where $\succ$ denotes positive definiteness.

It is obvious that the existence of D such that the inequalities (3.6) are satisfied implies that $\mathbf{A}$ is a polytope of M-matrices. Theorem 3.3 is attractive because there are efficient algorithms for solving linear matrix inequalities [6], [16].

The conditions given by Theorems 3.2, 3.3 and Remark 3.2 are mutually independent. To show this let us consider the following three pairs of M-matrices:

(i) $A^1 = \begin{bmatrix} 0.6 & -0.4 \\ -1 & 0.8 \end{bmatrix}$ and $A^2 = \begin{bmatrix} 0.6 & -0.4 \\ -0.1 & 0.1 \end{bmatrix}$ have a common $d = (1, 1.4)^T$, but do not have a common $\tilde{d}$ or D,

(ii) $A^1 = \begin{bmatrix} 0.6 & -1 \\ -0.4 & 0.8 \end{bmatrix}$ and $A^2 = \begin{bmatrix} 0.6 & -0.1 \\ -0.4 & 0.1 \end{bmatrix}$ have a common $\tilde{d} = (1, 1.4)^T$, but do not have a common d or D,

(iii) $A^1 = \begin{bmatrix} 0.4 & -0.1 \\ -0.7 & 0.5 \end{bmatrix}$ and $A^2 = \begin{bmatrix} 0.4 & -0.4 \\ -0.4 & 0.7 \end{bmatrix}$ have a common $D = \mathrm{diag}\{0.9, 0.7\}$ [5], but do not have a common d or $\tilde{d}$.

It is interesting to note that all our convexity results for M-matrices carry over to H-matrices which do not require the stringent sign structure imposed by the class

$\mathbf{Z}_n$. To show this let us recall that for a given $n \times n$ matrix $A = (a_{ij})$ the $n \times n$ *comparison matrix* $Q(A) = (q_{ij})$ is defined as

$$q_{ij} = \begin{cases} |a_{ij}|, & i = j, \\ -|a_{ij}|, & i \neq j. \end{cases} \tag{3.7}$$

Then, we have the well-known (e.g., [13])

DEFINITION 3.1. A matrix $A \in \mathbf{R}^{n\times n}$ is said to be an H-matrix if its comparison matrix $Q(A) \in \mathbf{R}^{n\times n}$ is an M-matrix.

Thus the class of H-matrices is defined as

$$\mathbf{H}_n = \{A \in \mathbf{R}^{n\times n};\ Q(A) \in \mathbf{M}_n\}. \tag{3.8}$$

To use H-matrices in proving convergence of iterations for polytopic systems, we introduce the polytope

$$\mathbf{Q} = \operatorname{conv}\{Q(A^{\ell})\}, \tag{3.9}$$

where $\ell \in \mathbf{m}$, and state the obvious.

THEOREM 3.4. *Let $A^{\ell} \in \mathbf{H}_n$ for all $\ell \in \mathbf{m}$, and let all A^{ℓ} have diagonal elements of the same sign. Then, $\mathbf{Q} \subset \mathbf{M}_n$ implies $\mathbf{A} \subset \mathbf{H}_n$.*

We also note that to test if the polytope $\mathbf{Q}$ is a polytope of M-matrices we can use Theorems 3.2, 3.3, or Remark 3.2 given above.

4. Convergence

We show now how the conditions for convexity of M-matrices, which were obtained in the preceding section, can be used to establish convergence of iterations for polytopic systems. First, the Gauss-Seidel iteration will be considered. Then, a convergence condition for the Jacobi iteration follows directly from the same derivation. We note that this derivation generalizes those in [1] to allow the use of M-matrices.

THEOREM 4.1. *The Gauss-Seidel iteration is convergent for all $A \in \mathbf{A}$ if $\mathbf{A} \subset \mathbf{M}_n$.*

Proof. From $\mathbf{A} \subset \mathbf{M}_n$ and Theorem 3.1 it follows that for any $\alpha \in \mathbf{U}_m$, there exists a $d(\alpha) > 0$ such that $A(\alpha)d(\alpha) > 0$. This is equivalent to saying that the matrix $A(\alpha)$ is quasidiagonally dominant, that is, there exist positive numbers $d_i(\alpha)$ so that

$$d_i(\alpha)a_{ii}(\alpha) = d_i(\alpha)|a_{ii}(\alpha)| > \sum_{j\neq i}^{n} d_j(\alpha)|a_{ij}(\alpha)|, \quad \forall i \in \mathbf{n}. \tag{4.1}$$

Let us define for each $\alpha \in \mathbf{U}_m$ a matrix

$$B(\alpha, \lambda) = \lambda[\Delta(\alpha) - L(\alpha)] - U(\alpha), \tag{4.2}$$

where $\lambda \in \mathbf{C}$. If $|\lambda| \geq 1$, then from (4.1) we obtain for all $i \in \mathbf{n}$,

$$\begin{aligned} d_i(\alpha)|\lambda a_{ii}(\alpha)| &> \sum_{j \neq i}^{n} d_j(\alpha)|a_{ij}(\alpha)| \\ &= \sum_{j=1}^{i-1} d_j(\alpha)|\lambda a_{ij}(\alpha)| + |\lambda| \sum_{j=i+1}^{n} d_j(\alpha)|a_{ij}(\alpha)| \\ &\geq \sum_{j=1}^{i-1} d_j(\alpha)|\lambda a_{ij}(\alpha)| + \sum_{j=i+1}^{n} d_j(\alpha)|a_{ij}(\alpha)|. \end{aligned} \tag{4.3}$$

The last inequality implies that the matrix $\tilde{B} = (\tilde{b}_{ij})$ defined as

$$\tilde{B}(\alpha, \lambda) = D^{-1}(\alpha)B(\alpha, \lambda)D(\alpha), \tag{4.4}$$

where $D(\alpha) = \operatorname{diag}\{d_1(\alpha), d_2(\alpha), \dots, d_n(\alpha)\}$ is strictly diagonally dominant, that is,

$$\tilde{b}_{ii}(\alpha, \lambda) > \sum_{j \neq i}^{n} |\tilde{b}_{ij}(\alpha, \lambda)|, \quad \forall i \in \mathbf{n}. \tag{4.5}$$

These inequalities further imply (e.g., Theorem 6.1.10 in [13]) that $\tilde{B}(\alpha, \lambda)$ and, thus, $B(\alpha, \lambda)$, are nonsingular for $|\lambda| \geq 1$.

Now, we obtain

$$\begin{aligned} \det[\lambda I - G(\alpha)] &= \det\{\lambda I - [\Delta(\alpha) - L(\alpha)]^{-1} U(\alpha)\} \\ &= \det\big([\Delta(\alpha) - L(\alpha)]^{-1}\{\lambda[\Delta(\alpha) - L(\alpha)] - U(\alpha)\}\big) \\ &= \det[\Delta(\alpha) - L(\alpha)]^{-1} \det B(\alpha, \lambda). \end{aligned} \tag{4.6}$$

Since $\Delta(\alpha) - L(\alpha)$ is nonsingular by assumption, that is, $[\Delta(\alpha) - L(\alpha)]^{-1}$ exists for all $\alpha \in \mathbf{U}_m$, and since $B(\alpha, \lambda)$ is nonsingular for all $\alpha \in \mathbf{U}_m$ and $|\lambda| \geq 1$, we conclude from (4.6) that $\rho[G(\alpha)] < 1$, and the Gauss-Seidel iteration is convergent for all $\alpha \in \mathbf{U}_m$. □

To show that the same theorem holds for Jacobi iterations, we first define a matrix

$$C(\alpha, \lambda) = \lambda\Delta(\alpha) - L(\alpha) - U(\alpha). \tag{4.7}$$

Then, we reproduce (4.6) as in [1] to get

$$\begin{aligned} \det[\lambda I - J(\alpha)] &= \det\{\lambda I - \Delta^{-1}(\alpha)[L(\alpha) + U(\alpha)]\} \\ &= \det\{\lambda\Delta^{-1}(\alpha)\Delta(\alpha) - \Delta^{-1}(\alpha)[L(\alpha) + U(\alpha)]\} \\ &= \det\{\Delta^{-1}(\alpha)[\lambda\Delta(\alpha) - L(\alpha) - U(\alpha)]\} \\ &= \det[\Delta^{-1}(\alpha)]\det[\lambda\Delta(\alpha) - L(\alpha) - U(\alpha)] \\ &= \det[\Delta^{-1}(\alpha)]\det C(\alpha, \lambda). \end{aligned} \tag{4.8}$$

To repeat the same argument used in the proof of Theorem 4.1, we need to carry the majorization (4.3) one step further and get

$$d_i(\alpha)|\lambda a_{ii}(\alpha)| > \sum_{j=1}^{i-1} |a_{ij}(\alpha)| + \sum_{j=i+1}^{n} |a_{ij}(\alpha)|, \tag{4.9}$$

which is again valid for all $\alpha \in \mathbf{U}_m$, and $|\lambda| \geq 1$. Now, from (4.8) we conclude that $\det[\lambda I - J(\alpha)] = 0$ only if $\det C(\alpha, \lambda) = 0$, which cannot take place for any $\alpha \in \mathbf{U}_m$ and $|\lambda| \geq 1$ because of (4.9). Therefore, $\rho[J(\alpha)] < 1$ and the Jacobi iteration is convergent for all $\alpha \in \mathbf{U}_m$, as well.

From the convergence proof of Theorem 4.1 we see that it is insensitive to the signs of the off-diagonal elements of the matrix A. We can capitalize on this fact and prove the following:

THEOREM 4.2. *Let $A^\ell \in \mathbf{H}_n$ for all $\ell \in \mathbf{m}$, and let all A^ℓ have diagonal elements of the same sign. If the polytope $\mathbf{Q}$ is a polytope of M-matrices, then both Gauss-Seidel and Jacobi iterations are convergent for all $A \in \mathbf{A}$.*

Proof. Let us first define the $n \times n$ matrix $\overline{Q} = (\overline{q}_{ij})$ as a convex combination

$$\overline{Q}(\alpha) = \sum_{\ell=1}^{m} \alpha_\ell Q(A^\ell) \tag{4.10}$$

where $\alpha \in \mathbf{U}_m$. Since by assumption $\overline{Q}(\alpha)$ is an M-matrix, then for any $\alpha \in \mathbf{U}_m$ there exists a positive vector $d(\alpha)$ such that $\overline{Q}(\alpha)d(\alpha)$ is positive. This fact implies that

$$d_i(\alpha)\overline{q}_{ii}(\alpha) > -\sum_{j\neq i}^{n} d_j(\alpha)\overline{q}_{ij}(\alpha), \quad \forall i \in \mathbf{n}. \tag{4.11}$$

Since all A^ℓ have diagonal elements of the same sign, we have

$$\overline{q}_{ii}(\alpha) = \sum_{\ell=1}^{m} \alpha_\ell q_{ii}^\ell = \sum_{\ell=1}^{m} \alpha_\ell |a_{ii}^\ell| = |a_{ii}(\alpha)|, \quad \forall i \in \mathbf{n} \tag{4.12}$$

where the individual matrices are defined as

$$\begin{aligned} A^\ell &= (a_{ij}^\ell), & A(\alpha) &= [a_{ij}(\alpha)], \\ Q(A^\ell) &= (q_{ij}^\ell), & \overline{Q}(\alpha) &= [q_{ij}(\alpha)]. \end{aligned} \tag{4.13}$$

For off-diagonal elements we write

$$\overline{q}_{ij}(\alpha) = \sum_{\ell=1}^{m} \alpha_\ell q_{ij}^\ell = -\sum_{\ell=1}^{m} \alpha_\ell |a_{ij}^\ell| \leq -\left|\sum_{\ell=1}^{m} \alpha_\ell a_{ij}^\ell\right| = -|a_{ij}(\alpha)|, \tag{4.14}$$

which implies

$$-\overline{q}_{ij}(\alpha) \geq |a_{ij}(\alpha)|, \quad \forall i,j \in \mathbf{n}, \quad i \neq j. \tag{4.15}$$

From (4.11), (4.12) and (4.15) we obtain the inequalities (4.1) which establish $A(\alpha)$ as a quasidominant diagonal matrix for all $\alpha \in \mathbf{U}_m$. This means that the proof from this point on can follow the proof of Theorem 4.1 reaching the same conclusion regarding convergence of the Gauss-Seidel iterations. The same reasoning applies to Jacobi iteration as well. □

5. Block Jacobi Iteration

When parallel processing is considered to speed up the convergence, a block Jacobi iteration becomes attractive because of low communication overhead in multiprocessor architectures. To define this kind of iteration, let us partition equation (2.1) as

$$\begin{bmatrix} A_{11} & A_{12} & \dots & A_{1N} \\ A_{21} & A_{22} & \dots & A_{2N} \\ \vdots & \vdots & & \vdots \\ A_{N1} & A_{N2} & \dots & A_{NN} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{bmatrix}. \tag{5.1}$$

Where $A_{ij} \in \mathbf{R}^{n_i \times n_j}$ are submatrices and $x_i \in \mathbf{R}^{n_i}$, $b_i \in \mathbf{R}^{n_i}$ are subvectors such that $n = \sum_{i=1}^{N} n_i$ and $i, j \in \mathbf{N} = \{1, 2, \dots, N\}$. Next let us define the following block matrices

$$A_D = \text{diag}\{A_{11}, A_{22}, \dots, A_{NN}\}, \quad A_C = A - A_D. \tag{5.2}$$

We assume that A_D is nonsingular and introduce the normalized block matrix $\hat{A} = (\hat{A}_{ij})$ where

$$\hat{A}_{ij} = \begin{cases} 0, & i = j, \\ A_{ii}^{-1} A_{ij}, & i \neq j. \end{cases} \tag{5.3}$$

Then, the block Jacobi iteration is defined as

$$x^{k+1} = -\hat{A} x^k + \hat{b}, \quad k = 0, 1, 2, \dots \tag{5.4}$$

where $\hat{b} = A_D^{-1} b$.

The block Jacobi iteration is obviously convergent iff $\rho(\hat{A}) < 1$. To consider iterations in the block matrix context we need to interpret this convergence condition in terms of M-matrices. The necessary notion is that of the monotone matrix norm.

Let $F : \mathbf{R}^{n \times n} \to \mathbf{R}^{N \times N}$ be the matrix function $F(A) = [F_{ij}(A_{ij})]$ of a block matrix $A = (A_{ij})$, where $F_{ij}(A_{ij}) = \|A_{ij}\|$ and $\|\cdot\|$ denotes a matrix norm which may be different for each F_{ij}. Let $\|\cdot\|_\mu : \mathbf{R}^{N \times N} \to \mathbf{R}_+$ be a *monotone matrix norm*, that is, for any two matrices A and $B, A \geq B \geq 0$ implies $\|A\|_\mu \geq \|B\|_\mu$. Then, the function $\|\cdot\|_* : \mathbf{R}^{n \times n} \to \mathbf{R}_+$ defined by $\|A\|_* = \|F(A)\|_\mu$ is a matrix norm. We need

THEOREM 5.1. *The following conditions are equivalent [19]:*

(i) *There exists a monotone norm* $\|\cdot\|_\mu$ *such that* $\|F(A)\|_\mu < 1$.

(ii) *There exists a matrix norm* $\|\cdot\|$ *such that* $\|F(A)\| < 1$.

(iii) $\rho[F(A)] < 1$.

(iv) $I - F(A)$ *is an* M*-matrix.*

(v) *There exists a vector* $d > 0$ *such that* $[I - F(A)]d > 0$.

(vi) *There exists a vector* $\tilde{d} > 0$ *such that* $[I - F(A)]^T\tilde{d} > 0$.

(vii) *There exists a vector* $d \geq 0$ *such that* $[I - F(A)]d \geq 0$.

(viii) *There exists a vector* $\tilde{d} \geq 0$ *such that* $[I - F(A)]^T\tilde{d} \geq 0$.

Proof. Obviously, (i) implies (ii). Since $\rho(A) \leq \|A\|$ for any square matrix A, (ii) implies (iii). From the properties of M-matrices in Theorem 3.1, the conditions (iii)–(viii) are equivalent. Finally, we show that (v) implies (i). Define $\|\cdot\|_\mu : \mathbf{R}^{N\times N} \to \mathbf{R}_+$ as

$$\|A\|_\mu = \max_i \left\{ \sum_{j=1}^{N} d_i^{-1} d_j |a_{ij}| \right\}. \tag{5.5}$$

Then, $\|\cdot\|_\mu$ is a monotone matrix norm, and (v) implies $\|F(A)\|_\mu < 1$. □

To define iterations in the block context, let us assume that matrix A belongs to a polytope

$$\mathbf{A} = \mathrm{conv}\{A^\ell\}, \tag{5.6}$$

where $A_D^\ell = A_D$; the diagonal blocks are the same for all vertex matrices of $\mathbf{A}$. Then, $A = A(\alpha)$ can be expressed as

$$A(\alpha) = A_D + \sum_{\ell=1}^{m} \alpha_\ell A_C^\ell = \sum_{\ell=1}^{m} \alpha_\ell A^\ell, \quad \forall \alpha \in \mathbf{U}_m. \tag{5.7}$$

We prove the following:

THEOREM 5.2. *Let* $A^\ell \in \mathbf{R}^{n\times n}$ *be given for all* $\ell \in \mathbf{m}$. *Define the augmented matrices* $\overline{A}^\ell \in \mathbf{R}^{N\times(N+1)}$ *such that* $\overline{A}^\ell = [I - F(\hat{A}^\ell), -e]$ *where* $e = (1, 1, \ldots, 1)^T \in \mathbf{R}_+^N$, *and the augmented vector* $\bar{x} = (x_1, x_2, \ldots, x_N, x_{N+1})^T \in \mathbf{R}_+^{N+1}$. *Consider the following linear programming problem:*

$$\begin{aligned} &\max x_{N+1} \\ &\text{subject to } \ \overline{A}^\ell \bar{x} \geq 0, \ \ \ell \in \mathbf{m}; \ \ \bar{x} \geq 0, \ \ x_{N+1} \leq 1. \end{aligned} \tag{5.8}$$

Then, a Jacobi iteration is convergent for all $A \in \mathbf{A}$ *if* $x_{N+1}^0 = \max x_{N+1} = 1$.

Proof. Let $x_{N+1} = 1$. Then, for all $\ell \in \mathbf{m}$,

$$\overline{A}^\ell \bar{x} = [I - F(\hat{A}^\ell)]x - e \geq 0 \tag{5.9}$$

or

$$[I - F(\hat{A}^{\ell})]x \geq e. \tag{5.10}$$

For any $A(\alpha) \in \mathbf{A}$ we write

$$F_{ij}[A_{ij}(\alpha)] = \begin{cases} 0, & i = j, \\ \|A_{ii}^{-1}(\alpha)A_{ij}(\alpha)\|, & i \neq j. \end{cases} \tag{5.11}$$

Since $A_D^{\ell} = A_D$ for all $\ell \in \mathbf{m}$, we have

$$A_{ii}(\alpha) = \sum_{\ell=1}^{m} \alpha_{\ell} A_{ii}^{\ell} = A_{ii} \sum_{\ell=1}^{m} \alpha_{\ell} = A_{ii}, \tag{5.12}$$

and for $i \neq j$ we obtain

$$\begin{aligned} \|A_{ii}^{-1}(\alpha)A_{ii}(\alpha)\| &= \left\| A_{ii}^{-1} \sum_{\ell=1}^{m} \alpha_{\ell} A_{ij}^{\ell} \right\| \\ &= \left\| \sum_{\ell=1}^{m} \alpha_{\ell} A_{ii}^{-1} A_{ij}^{\ell} \right\| \\ &\leq \sum_{\ell=1}^{m} \alpha_{\ell} \|A_{ii}^{-1} A_{ij}^{\ell}\|. \end{aligned} \tag{5.13}$$

The last inequality implies

$$F[\overline{A}(\alpha)] \leq \sum_{\ell=1}^{m} \alpha_{\ell} F[A^{\ell}(\alpha)] \tag{5.14}$$

and we obtain

$$\begin{aligned} \left(I - F[\overline{A}(\alpha)]\right)x &\geq \left(I - \sum_{\ell=1}^{m} \alpha_{\ell} F[\overline{A}^{\ell}(\alpha)]\right)x \\ &= \sum_{\ell=1}^{m} \alpha_{\ell}\left(I - F[\overline{A}^{\ell}(\alpha)]\right)x \\ &\geq \sum_{\ell=1}^{m} \alpha_{\ell} e \\ &= e. \end{aligned} \tag{5.15}$$

We recall that the vector e is positive and by the help of Theorem 5.1 conclude from (5.14) that

$$\left\|F[\overline{A}(\alpha)]\right\|_{\mu} < 1, \quad \forall \alpha \in \mathbf{U}_m. \tag{5.16}$$

Since $\rho(F) \leq \|F\|$ for any square matrix F, the inequality (5.15) implies

$$\rho[\bar{A}(\alpha)] < 1, \quad \forall \alpha \in \mathbf{U}_m \tag{5.17}$$

that is, the Jacobi iteration is convergent for all $A \in \mathbf{A}$. □

6. Nonlinear Polytopic Systems

Our interest now is to consider a nonlinear algebraic equation

$$f(x, p) = b, \tag{6.1}$$

where $x \in \mathbf{R}^n$ is again the vector of unknowns x_i and $b \in \mathbf{R}^n$ is a known vector. The parameter vector $p \in \mathbf{R}^r$ is assumed to belong to a convex polytope

$$\mathbf{P} = \operatorname{conv}\{p^\ell\}, \tag{6.2}$$

where $\ell \in \mathbf{m}$. We want to derive convergence conditions for solving a polytopic system (6.1) by the Gauss-Seidel iteration:

$$\begin{aligned} &\text{Solve } f_i(x_1^{k+1}, \ldots, x_{i-1}^{k+1}, x_i, x_{i+1}^k, \ldots, x_n^k, p) = b_i \text{ for } x_i \\ &\text{Set } x_i^{k+1} = x_i, \ \ i = 1, 2, \ldots, n; \ \ k = 0, 1, 2, \ldots \end{aligned} \tag{6.3}$$

or by the Jacobi iteration:

$$\begin{aligned} &\text{Solve } f_i(x_1^k, \ldots, x_{i-1}^k, x_i, x_{i+1}^k, \ldots, x_n^k, p) = b_i \text{ for } x_i \\ &\text{Set } x_i^{k+1} = x_i, \ \ i = 1, 2, \ldots, n; \ \ k = 0, 1, 2, \ldots \end{aligned} \tag{6.4}$$

In the absence of the uncertain parameter p the two iteration schemes were considered by many authors, most notably in [14], [20]. These two papers contain a wide variety of convergence conditions for the nonlinear iterations, which are based on the notion of M-functions introduced earlier by Ortega (see [20]). In order to derive convergence conditions along the same lines for polytopic systems, we need to review briefly those properties of M-functions which are relevant to our development.

M-functions represent a natural generalization of M-matrices. Similarly, as in case of M-matrices, they are assumed to have certain monotone properties. A function $f : \mathbf{R}^n \to \mathbf{R}^n$ is said to be off-diagonally decreasing if for any $x \in \mathbf{R}^n$ the functions $\varphi_{ij} : \{\mu \in \mathbf{R} : x + \mu e^j\} \to \mathbf{R}$ defined as $\varphi_{ij}(\mu) = f_i(x + \mu e^j)$, $i \neq j$, $i, j \in \mathbf{n}$, are decreasing functions of μ, where $e^j \in \mathbf{R}^n$, $j \in \mathbf{n}$, are the unit basis vectors with j-th component one and all others zero. As the inverse of an M-matrix is a nonnegative matrix, so an M-function is required to be inverse increasing. This means that a function $f(x)$ is inverse increasing if $f(x) \leq f(y)$ for any $x, y \in \mathbf{R}^n$ implies $x \leq y$. It turns out that the two properties completely define the class of M-functions, and we say that a function $f(\cdot)$ is an M-function if it is off-diagonally decreasing and inverse increasing.

To use our conditions for convexity of M-matrices in the context of polytopic nonlinear systems, we shall interpret the properties of M-functions in terms of their Jacobians. To do this we start with the result from [18] which states that a function $f : \mathbf{R}^n \to \mathbf{R}^n$ is a differentiable M-function if and only if the Jacobian matrix $A = (a_{ij}) \in \mathbf{R}^{n\times n}$, where $a_{ij} = \partial f_i(x) / \partial x_j$, is an M-matrix everywhere in $\mathbf{R}^n$. From [20] we know that if $f(x)$ is continuous and surjective M-function, then for any fixed $b \in \mathbf{R}^n$ and any starting point $x^0 \in \mathbf{R}^n$ both nonlinear Gauss-Seidel and Jacobi iterations converge.

We now turn our attention to polytopic system (6.1), and start with the assumption that $f(x,p)$ is a differentiable function of $x \in \mathbf{R}^n$, that is, the corresponding Jacobian matrix $A(x,p) = [a_{ij}(x,p)]$ with coefficients

$$a_{ij}(x,p) = \frac{\partial f_i(x,p)}{\partial x_j} \tag{6.5}$$

is well defined. We also assume that $a_{ij}(x,p)$ are concave functions of $p \in \mathbf{P}$, that is,

$$a_{ij}(x,p) = a_{ij}\left(x, \sum_{\ell=1}^{m} \alpha_\ell p^\ell\right) \geq \sum_{\ell=1}^{m} \alpha_\ell a_{ij}(x,p^\ell), \quad \forall x \in \mathbf{R}^n \tag{6.6}$$

such that there exist constant matrices $\overline{\mathbf{A}}^\ell = (a_{ij}^\ell)$ providing the bounds for the Jacobian matrix as

$$\begin{aligned} a_{ii}(x,p^\ell) &\geq \overline{a}_{ii}^\ell, \quad i = j, \\ 0 \geq a_{ij}(x,p^\ell) &\geq \overline{a}_{ij}^\ell, \quad i,j \in \mathbf{n}, \ i \neq j, \ \forall x \in \mathbf{R}^n. \end{aligned} \tag{6.7}$$

From (6.6) we conclude that the matrices $\overline{A}^\ell$ define a matrix polytope

$$\overline{\mathbf{A}} = \text{conv}\{\overline{A}^\ell\} \tag{6.8}$$

to which our results on convex M-matrices of Section 3 can be applied via

THEOREM 6.1. *Let $f(x,p)$ be a differentiable, surjective M-function with respect to x for each $p \in \mathbf{P}$. Then, both nonlinear Gauss-Seidel and Jacobi iterations for polytopic system (6.1) are convergent if $\overline{\mathbf{A}} \subset \mathbf{M}_n$.*

Proof. By relying on the results on M-matrices, which were reviewed above, a sketch of the proof would suffice. Since $\overline{\mathbf{A}} \subset \mathbf{M}_n$, then from (6.7) it follows that $A(x,p) \in \mathbf{Z}_n$ and there is an $\overline{A} \in \overline{\mathbf{A}}$ such that $A(x,p) \geq \overline{A}$ for all $x \in \mathbf{R}^n$. Since $\overline{A}$ is an M-matrix, then $A(x,p)$ is an M-matrix for all $x \in \mathbf{R}^n$ and all $p \in \mathbf{P}$, which implies that $f(x,p)$ is an M-function in $\mathbf{R}^n$ for all $p \in \mathbf{P}$. Furthermore, $f(x,p)$ is surjective in $x \in \mathbf{R}^n$ for each $p \in \mathbf{P}$ implying that nonlinear Gauss-Seidel and Jacobi iterations for system (6.1) are convergent. □

7. Conclusion

In this paper, we have formulated linear and nonlinear Jacobi and Gauss-Seidel iterations as well as block Jacobi iteration for polytopic systems. A number of sufficient conditions have been derived, which can be used to determine if a matrix polytope is contained in the set of M-matrices by testing only the vertices of the polytope. These conditions have been used to establish the convergence of the new type of Jacobi and Gauss-Seidel iterations. From an application point of view, it is interesting to note that the convergence can be checked using efficient convex programming methods.

Finding necessary and sufficient conditions for an arbitrary dimensional polytope of matrices to be a polytope of M-matrices is still an open problem. These conditions would considerably strengthen our results in application. This is a subject for future research.

Acknowledgement

The research reported herein has been supported by the NSF Grant ECS-9526142.

References

1. Bagnara, R.: A Unified Proof for the Convergence of Jacobi and Gauss-Seidel Methods, *SIAM Review* **37** (1995), pp. 93–97.
2. Barán, B., Kaszkurewicz, E. and Bhaya, A.: Parallel Asynchronous Team Algorithms: Convergence and Performance Analysis, *IEEE Transactions on Parallel and Distributed Systems* **7** (1996), pp. 677–688.
3. Berman, A. and Plemmons, R. J.: *Nonnegative Matrices in the Mathematical Sciences*, Academic Press, New York, 1979.
4. Bertsekas, D. P. and Tsitsiklis, J. N.: *Parallel and Distributed Computation: Numerical Methods*, Prentice-Hall, Englewood Cliffs, NJ, 1989.
5. Bhaya, A. and Kaszkurewicz, E.: Comments Regarding "On Stability of Interval Matrices," *IEEE Transactions on Automatic Control* **41** (1996), pp. 762–766.
6. Boyd, S., Ghaoui, L. E., Feron, E., and Balakrishnan, V.: *Linear Matrix Inequalities in System and Control Theory*, SIAM, Philadelphia, PA, 1994.
7. Fan, K.: Inequalities for the Sum of Two M-Matrices, in: Shisha, O. (ed.), *Inequalities*, Academic Press, New York, 1967, pp. 105–117.
8. Fiedler, M. and Pták, V.: On Matrices with Non-Positive Off-Diagonal Elements and Positive Minors, *Czechoslovakian Mathematical Journal* **12** (1962), pp. 382–400.
9. Gale, D. and Nikaido, H.: The Jacobian Matrix and Global Univalence of Mappings, *Mathematische Annalen* **159** (1965), pp. 81–93.
10. Garloff, J.: Block Methods for the Solution of Linear Interval Equations, *SIAM Journal of Matrix Analysis and Applications* **11** (1990), pp. 89–106.
11. Hansen, E.: *Global Optimization Using Interval Analysis*, Marcel Dekker, New York, 1992.
12. Horisberger, H. P. and Belanger, P. R.: Regulators for Linear, Time Invariant Plants with Uncertain Parameters, *IEEE Transactions on Automatic Control* **AC-21** (1976), pp. 705–708.
13. Horn, R. A. and Johnson, C. R.: *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.
14. Moré, J. J.: Nonlinear Generalizations of Matrix Diagonal Dominance with Application to Gauss-Seidel Iterations, *SIAM Journal on Numerical Analysis* **9** (1972), pp. 357–378.
15. Moré, J. J. and Rheinboldt, W.: On P- and S-Functions and Related Classes of n-dimensional Nonlinear Mappings, *Linear Algebra and Its Applications* **6** (1973), pp. 45–68.
16. Nesterov, Yu. and Nemirovsky, A.: An Interior-Point Polynomial Methods in Convex Programming, *Studies in Applied Mathematics*, volume 13, SIAM, Philadelphia, PA, 1994.
17. Neumaier, A.: *Interval Methods for Systems of Equations*, Cambridge University Press, Cambridge, 1990.

18. Nikaido, H.: *Convex Structures and Economic Theory*, Academic Press, New York, 1968.
19. Ohta, Y. and Šiljak, D. D.: Overlapping Block Diagonal Dominance and Existence of Liapunov Functions, *Journal of Mathematical Analysis and Applications* **112** (1985), pp. 396–410.
20. Rheinboldt, W. C.: On *M*-functions and Their Applications to Nonlinear Gauss-Seidel Iterations and Network Flows, *Journal of Mathematical Analysis and Applications* **32** (1970), pp. 275–307.
21. Sezer, M. E. and Šiljak, D. D.: On Stability of Interval Matrices, *IEEE Transactions on Automatic Control* **39** (1994), pp. 368–371.
22. Šiljak, D. D. and Zečević, A. I.: Overlapping Block-Iterative Methods for Solving Algebraic Equations, *Journal of Difference Equations and Applications* **1** (1995), pp. 125–136.
23. Stipanović, D. M. and Šiljak, D. D.: Convex *M*-Matrices and Polytopic Dynamic Systems, in: *Proceedings of the 36th Conference on Decision and Control*, San Diego, CA, 1997, pp. 4366–4368.
24. Varga, R. S.: *Matrix Iterative Analysis*, Prentice Hall, Englewood Cliffs, NJ, 1962.
25. Zečević, A. I. and Šiljak, D. D.: A Block-Parallel Newton Method via Overlapping Epsilon Decomposition, *SIAM Journal of Matrix Analysis and Applications* **15** (1994), pp. 824–844.