THE USE OF BASIC ITERATIVE METHODS FOR BOUNDING A SOLUTION OF A SYSTEM OF LINEAR EQUATIONS WITH AN M-MATRIX AND POSITIVE RIGHT-HAND SIDE

Martin Kocurek

Abstract

This article presents a simple method for bounding a solution of a system of linear equations $Ax = b$ with an M-matrix and positive right-hand side [1]. Given a suitable approximation to an exact solution, the bounds are constructed by one step in a basic iterative method.

1. Motivation

When we use iterative methods for solving sets of linear algebraic equations $Ax = b$, we guess an accuracy of the computed solution according to a residual vector. Unfortunately, small norm of the residual vector doesn't imply that we are close to the exact solution. If we could instead construct an upper and lower bound, we could guess an accuracy of the computed solution better.

2. Basic terms and definitions

Definition 2.1 *Let matrices A, B have the same dimension. We say that $A \geq B$ if $a_{ij} \geq b_{ij}$ holds for every i, j. Matrix A is called **nonnegative**, if $A \geq O$, where O is the zero matrix.*

Definition 2.2 *A real square matrix $A = (a_{ij})_{i,j=1}^{n}$ is called **M-matrix**, if*

1. $a_{ii} > 0$, $i = 1, \ldots, n$,
2. $a_{ij} \leq 0$ *for* $i \neq j$, $i, j = 1, \ldots, n$,
3. *exists* $A^{-1} \geq 0$.

Definition 2.3 *Let us split matrix A into two matrices V, W, so that $A = V - W$. If matrix V is nonsingular, then $V - W$ is called **splitting** of matrix A. The splitting of matrix A is called **regular** if V is nonsingular with $V^{-1} \geq 0$ and $W \geq 0$.*

A splitting $Ax = (V - W)x = b$ yields an iterative method

$$x^{(k+1)} = V^{-1}Wx^{(k)} + V^{-1}b,$$

which is convergent if and only if the spectral radius satisfies $\rho(V^{-1}W) < 1$.

As usual, we split matrix A into $D - L - U$, where D is the diagonal of A and L, U are strictly lower and upper triangular parts of A, respectively. The classical iterative methods are obtained by setting

- $V = I$, $W = I - A$ ... Fixed-point iterations
- $V = D$, $W = L + U$ ... Method of Jacobi
- $V = D - L$, $W = U$ ... Method of Gauss-Seidel

From now on we consider matrix A to be an M-matrix and the right-hand side to be positive. The three methods mentioned above can be written as

$$x^{(k+1)} = \mathbf{T}x^{(k)} + \mathbf{d}, \qquad \mathbf{T} := V^{-1}W, \qquad \mathbf{d} := V^{-1}b.$$

Furthermore, for all these methods (for fixed-point iterations $a_{ii} \leq 1,\ i = 1, \ldots, n$, is required) $V - W$ is a regular splitting and (see [3], Theorem 3.13)

$$\mathbf{T} \geq 0, \quad \mathbf{d} > 0, \quad \rho(\mathbf{T}) < 1.$$

3. Bounds for the solution

Lemma 3.1 *Let x be the exact solution to $Ax = b$. Let us consider an iterative process $x^{(k+1)} = \boldsymbol{T}x^{(k)} + \boldsymbol{d}$ with $\boldsymbol{T} \geq 0$ and $\rho(\boldsymbol{T}) < 1$. If*

$$x^{(l+1)} \geq x^{(l)} \tag{1}$$

for some $l \in \mathbf{N}$, then

$$x \geq x^{(l+2)} \geq x^{(l+1)}. \tag{2}$$

Similarly, if $x^{(l+1)} \leq x^{(l)}$ for some $l \in \mathbf{N}$, then

$$x \leq x^{(l+2)} \leq x^{(l+1)}.$$

Notice that condition (1) is equivalent to $Ax^{(l)} \leq b$, see [3]. Proof of this lemma is easy and can be found in [1].

If we get an approximation $x^{(k)}$ and a modificating vector v, we will try to find a vector $y^{(k)} = x^{(k)} + \delta v$ so that this $y^{(k)}$ has property (1),

$$y^{(k+1)} = \mathbf{T}y^{(k)} + \mathbf{d} \geq y^{(k)}. \tag{3}$$

Solving this inequality with variable δ we find a set of acceptable parameters δ^U. In the same way we find δ^L by solving the opposite inequality. Then we set the upper and lower bounds to be in the following form:

$$x^{(k)} + \delta^L v \leq x \leq x^{(k)} + \delta^U v.$$

Inequalities (3) have the form

$$\delta^L(I - \mathbf{T})v \leq r^{(k)} \quad \text{and} \quad \delta^U(I - \mathbf{T})v \geq r^{(k)}, \quad \text{where} \quad r^{(k)} = \mathbf{d} - (I - \mathbf{T})x^{(k)}. \tag{4}$$

Sufficient condition for these inequalities to have a solution is $(I - \mathbf{T})v > 0$, or equivalently

$$r^{(v)} < \mathbf{d}, \tag{5}$$

where $r^{(v)} = \mathbf{d} - (I - \mathbf{T})v$. Thus, $\mathbf{d} - r^{(v)} = (I - \mathbf{T})v$ and inequalities (4) will be

$$\delta^L(\mathbf{d} - r^{(v)}) \leq r^{(k)}, \qquad \delta^U(\mathbf{d} - r^{(v)}) \geq r^{(k)}.$$

Optimal solution, which yields the highest lower bound $x^L = x^{(k)} + \delta^L v$ and the lowest upper bound $x^U = x^{(k)} + \delta^U v$, is (index i denotes i-th component of a vector)

$$\delta^L = \min_{i=1,\dots,n} \frac{r_i^{(k)}}{\mathbf{d}_i - r_i^{(v)}}, \qquad \delta^U = \max_{i=1,\dots,n} \frac{r_i^{(k)}}{\mathbf{d}_i - r_i^{(v)}}.$$

Condition $(I - \mathbf{T})v > 0$ holds for any approximation $v = x^{(k)}$, which has its residual vector $r^{(k)} < \mathbf{d}$, see (5). Here it is useful to have a positive right-hand side b (and therefore $\mathbf{d} > 0$). Therefore, if the residual vector of the approximation $x^{(k)}$ is small enough, we may take $v = x^{(k)}$, $r^{(v)} = r^{(k)}$ and the bounds will be

$$x^U = x^{(k)}(1 + \delta^U), \qquad x^L = x^{(k)}(1 + \delta^L),$$

where

$$\delta^L = \min_{i=1,\dots,n} \frac{r_i^{(k)}}{\mathbf{d}_i - r_i^{(k)}}, \qquad \delta^U = \max_{i=1,\dots,n} \frac{r_i^{(k)}}{\mathbf{d}_i - r_i^{(k)}}.$$

4. Application to irreducible Markov chains

Let us now consider a system corresponding to an automaton with n states. This automaton changes its state, switches from one state to another, in certain time steps. If a probability of switching to another state depends on the current state only, we call this system *Markov Chain*. If there exists a connection between every two states, we call this Markov chain *irreducible*.

Probability of transition from i-th state to j-th (if the system is in the i-th state) is denoted by p_{ij}. In this manner we construct a *transition probability matrix* P, which is stochastic (row sums are equal to 1).

A useful characteristic of Markov chain is its *mean first passage times matrix*, denoted M. Its elements m_{ij} are average times between leaving i-th state and reaching j-th state (it is useful when j-th state is dangerous and means some kind of failure). It is computed from the following equation, see [4],

$$M = P(M - M_D) + E,$$

where $M_D = \operatorname{diag}\{m_{11}, \dots, m_{nn}\}$ and $E = (e_{ij})_{i,j=1}^n$, $e_{ij} = 1$, $i,j = 1, \dots, n$. If we write this equation for each column separately, we get a set of linear algebraic equations

$$[I - P(I - e_i e_i^T)]M_i = e,$$

where M_i denotes the i-th column of M and $e = (1, \dots, 1)^T$. Matrix of this system is a diagonally dominant M-matrix and the method described above can be applied to find bounds for the solution.

If we use the fixed-point iterations, $M_i^{(k+1)} = P(I - e_i e_i^T)M_i^{(k)} + e$, for solving this problem with $x^{(0)} = e$, we get an approximation $x^{(k)}$, which has its residual vector $r^{(k)} < \mathbf{d} = e$ (condition (5)), after k iterations, $k \leq n$ [1]. Usually it is $k \ll n$.

5. Numerical example

We show these bounds in the following example. Let us consider a set of linear equations with the right-hand side e and matrix ([2], p. 55–56)

$$A = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1/3 & -2/3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -0.8 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1/3 & 0 & -2/3 & 0 & 0 & 0 \\ 0 & -1/7 & 0 & 0 & 1 & -2/7 & 0 & -4/7 & 0 & 0 \\ 0 & 0 & -0.2 & 0 & 0 & 1 & 0 & 0 & -0.8 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1/3 & 0 & 0 & 0 & 1 & -2/3 & 0 \\ 0 & 0 & 0 & 0 & -1/3 & 0 & 0 & 0 & 1 & -2/3 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

The method of fixed-point iterations with initial vector e is used for solving this system. The first three columns of the following tables show the vectors of the lower bounds x^L, the vector of the exact solution x, and the vectors of the upper bounds x^U. Approximate solutions $x^{(k)}$ used for creating these bounds are presented in the fourth columns and their residual vectors $r^{(k)}$ in the fifth columns. Furthermore, an error factor δ_{err} is computed as an additional criterion of convergence,

$$\delta_{\mathrm{err}} = \min_{i=1,\dots,n} \frac{x_i^L}{x_i^U}. \tag{6}$$

6. Conclusions

Systems of linear algebraic equations with an M-matrix appear in many parts of mathematics. If the right-hand side vector of the given system is positive, we may use this simple method to bound the exact solution with help of basic iterative methods.

The obtained bounds may be used to verify the accuracy of the computed solution. The approximate solutions x^L, x^U computed in Table 1 can be used to restart the iterative process [1].

Lower bnd. x^L	Exact sol. x	Upper bnd. x^U	Approx. $x^{(k)}$	Residual $r^{(k)}$
99.269406	105.000000	106.412140	79.876550	0.236850
98.320974	104.000000	105.395466	79.113400	0.236850
82.846169	87.579104	88.807202	66.661688	0.195356
104.635728	110.710448	112.164585	84.194530	0.247642
102.307712	108.223881	109.669062	82.321305	0.244195
98.700281	104.376119	105.802065	79.418607	0.237525
104.397207	110.453731	111.908902	84.002605	0.249366
103.464330	109.453731	110.908902	83.251972	0.249366
101.479317	107.325373	108.781061	81.654742	0.239748
99.647874	105.376119	106.817840	80.181082	0.237525

Tab. 1: *Solution after $k = 150$ iterations, $\|r^{(k)}\|$= 0.752329, δ_{err}=0.932877.*

Lower bnd. x^L	Exact sol. x	Upper bnd. x^U	Approx. $x^{(k)}$	Residual $r^{(k)}$
104.727984	105.000000	105.062731	103.534808	0.013813
103.730431	104.000000	104.061990	102.548621	0.013813
87.354444	87.579104	87.633660	86.359207	0.011393
110.422097	110.710448	110.775045	109.164048	0.014442
107.943055	108.223881	108.288080	106.713250	0.014241
104.106702	104.376119	104.439464	102.920605	0.013852
110.166244	110.453731	110.518374	108.911110	0.014543
109.169430	109.453731	109.518374	107.925653	0.014543
107.047876	107.325373	107.390039	105.828270	0.013982
105.104214	105.376119	105.440165	103.906753	0.013852

Tab. 2: *Solution after $k = 450$ iterations, $\|r^{(k)}\|$= 0.043876, δ_{err}=0.996814.*

Lower bnd. x^L	Exact sol. x	Upper bnd. x^U	Approx. $x^{(k)}$	Residual $r^{(k)}$
104.999085	105.000000	105.000210	104.995017	0.000047
103.999094	104.000000	104.000208	103.995064	0.000047
87.578349	87.579104	87.579287	87.574955	0.000039
110.709478	110.710448	110.710664	110.705188	0.000049
108.222936	108.223881	108.224096	108.218743	0.000048
104.375213	104.376119	104.376332	104.371169	0.000047
110.452765	110.453731	110.453948	110.448485	0.000049
109.452775	109.453731	109.453948	109.448534	0.000049
107.324440	107.325373	107.325590	107.320281	0.000048
105.375205	105.376119	105.376334	105.371122	0.000047

Tab. 3: *Solution after k=1050 iterations, $\|r^{(k)}\|$= 0.000149, δ_{err}=0.999989.*

Disadvantages of this approach are given by strict conditions that need to be fulfilled. Most restrictive conditions are the positive right-hand side and the need for a modificating vector. The positive right-hand side appears in some problems arising in modelling of Markov chains. The modificating vector is obtained either by computing a sufficient approximation, which is sometimes very difficult, or by using an extremely slow iterative method. On the other hand, having the modificating vector, one matrix-vector multiplication is enough to construct these bounds.

References

[1] M. Kocurek: *Acceleration of iterative methods for computing the mean first passage times matrix*. Diploma Thesis, MFF UK Praha, 2005 (in Czech).

[2] W.J. Stewart: *Introduction to the numerical solution of Markov chains.* Princeton University Press, Princeton, New Jersey, 1994.

[3] R.S.Varga: *Matrix iterative analysis.* Prentice-Hall, Englewood Cliffs, New Jersey, 1962.

[4] Š. Klapka: *Markov models in signalling systems.* Disertation Thesis, MFF UK Praha, 2002 (in Czech).

[5] Y. Saad: *Iterative methods for sparse linear systems.* SIAM, Philadelphia, 1996.