A faster algorithm for identification of an M-Matrix

R. J. Wood*   M. J. O'Neill*

Abstract

M-matrices are important in the consideration of rates of convergence of iterative methods for solving large systems of equations and are applicable in areas such as input-output systems in economic modelling, queuing theory, and engineering. The usual definition of an M-matrix has, among other requirements, that it must be non-singular and its inverse non-negative. Following Saad (2003), Young (1971) and Berman & Shaked–Monderer (2003) two more easily checked characterisations of an M-matrix are explored. These require only the evaluation of the spectral radius of an associated non-negative matrix.

Contents

---

*School of Information Technology, Charles Sturt University, Bathurst, AUSTRALIA. mailto:rwood@csu.edu.au

1   Introduction

Some problems involving elliptic partial differential equations, when solved
by finite-difference methods, lead to a linear system where the coefficient
matrix is what is known as an M-matrix [7]. Furthermore, when such an
M-matrix is sparse, there are well-established iterative techniques for solving
the linear system [2]. M-matrices also occur in linear systems associated with
input-output analysis in economic modelling.

2   What is an M-matrix?

Definition 1 *An $n \times n$ matrix A with elements α_{ij} is said to be an M-matrix
if it satisfies the following four properties:*

1. *$\alpha_{ii} > 0$ for $i = 1, \ldots, n$;*

2. *$\alpha_{ij} \leq 0$ for $i \neq j$, $i, j = 1, \ldots, n$;*

3. *A is nonsingular;*

4. *$A^{-1} \geq 0$.*

This is the usual definition of an M-matrix. However, as most authors point out (see Saad [2]), Property 1 is redundant as Properties 2, 3 and 4 together imply Property 1. Properties 1 and 2 are easily tested, but not so Properties 3 and 4 which, for large matrices would typically involve extensive computation.

Saad [2] and Young [7] give an alternative and simpler method for determining whether a matrix is an M-matrix. This method is described in the following theorem.

Theorem 2 *Let Properties 1 and 2 hold for matrix A. Construct the matrix B where $B = I - D^{-1}A$, and D is the diagonal of A. Then A is an M-matrix if and only if $\rho(B) < 1$ (that is, the spectral radius of B is less than unity).*

Proof: See Young [7, p.42]. ♠

So Properties 3 and 4 can be replaced by the condition $\rho(B) < 1$. The following examples show the application of Theorem 2. All matrices in these examples satisfy Properties 1 and 2.

Example 3 If $A = \begin{bmatrix} 1 & -1/4 \\ -1/4 & 1 \end{bmatrix}$ then $B = \begin{bmatrix} 0 & 1/4 \\ 1/4 & 0 \end{bmatrix}$, and $\rho(B) = 1/4 < 1$, which implies that A is an M-matrix.

Example 4 If $A = \begin{bmatrix} 2 & -2 \\ -1 & 1 \end{bmatrix}$ then $B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, and $\rho(B) = 1$. So A is not an M-matrix.

Example 5 If $A = \begin{bmatrix} 2 & -3 \\ -1 & 1 \end{bmatrix}$ then $B = \begin{bmatrix} 0 & 3/2 \\ 1 & 0 \end{bmatrix}$, and $\rho(B) = \sqrt{3/2} > 1$. So A is not an M-matrix.

It may be possible to check the property $\rho(B) < 1$ without actually calculating the spectral radius of B. For example, using the fact that the spectral radius cannot exceed the row or column norm, calculation of one or other of these norms may establish immediately that $\rho(B) < 1$. This is certainly true in Example 3, where $\|B\|_\infty = 1/4$ and, since $\rho(B) \le \|B\|_\infty$, then $\rho(B) \le 1/4$. Even when this is not possible, as in Examples 4 and 5, noting that B is non-negative, an always convergent method may be used to compute $\rho(B)$. This method involves selection of an appropriate value $q > \|B\|_\infty$ and then application to B of a variant of the inverse power method, which is essentially an application of a variant of the power method to the matrix $(qI - B)^{-1}$; Wood & O'Neill [6] described the details.

This paper considers the alternative approach of Theorem 7, the proof of which uses the following easily proved Lemma.

Lemma 6 *Let A be any real $n \times n$ matrix with elements α_{ij}. If A has non-positive off-diagonal elements then it can be decomposed into the form $A = \lambda I - C$ with $\lambda \ge \max_i \alpha_{ii}$ and $C \ge 0$.*

Theorem 7 *Let Properties 1 and 2 hold for matrix A. Decompose the matrix A into the form $A = \lambda I - C$, where $\lambda > \max_i \alpha_{ii}$. Then A is an M-matrix if and only if $\rho(C) < \lambda$. (See Berman & Shaked–Monderer [1], where Theorem 7 is used as the definition of an M-matrix.)*

Proof: This now follows from Theorem 3.8 in Varga [5] which states that if $C \ge 0$ is an $n \times n$ matrix, then the following two statements are equivalent:

1. $\lambda > \rho(C)$;

2. $(\lambda I - C)$ is non-singular and $(\lambda I - C)^{-1} \ge 0$.

♠

As with the method described in Theorem 2, the Properties 3 and 4 can be replaced by the condition $\rho(C) < \lambda$. For the matrix A in Example 3, choosing $\lambda = 2$ gives

$$C = \left[\begin{array}{cc} 1 & 1/4 \\ 1/4 & 1 \end{array} \right],$$

and, since $\rho(C) = 5/4 < 2$, then A is an M-matrix. For the matrix A in Example 4, choosing $\lambda = 3$ gives

$$C = \left[\begin{array}{cc} 1 & 2 \\ 1 & 2 \end{array} \right]$$

with $\rho(C) = 3 = \lambda$. So the matrix A is not an M-matrix. For the matrix A in Example 5, choosing $\lambda = 3$ gives

$$C = \left[\begin{array}{cc} 1 & 3 \\ 1 & 2 \end{array} \right]$$

with $\rho(C) = \left(3 + \sqrt{13} \right)/2 > \lambda$. Therefore the matrix A is not an M-matrix.

As previously, it may be possible to check the property $\rho(C) < \lambda$ using row or column norms, without actually calculating the spectral radius of C. This is certainly true for the matrix C from Example 3. Failing this, the fact that C is non-negative allows the always convergent method mentioned previously to be used to compute $\rho(C)$.

Note: there is considerable flexibility in choosing the λ of Theorem 7. However, if the λ chosen does not produce a matrix C such that $\rho(C) < \lambda$, then increasing the value of λ will not remedy the situation, since increasing λ by k increases all eigenvalues of C by k, and hence leaves $\rho(C) - \lambda$ unchanged.

In both the methods of Theorem 2 and 7 it may be required, as was seen in Examples 4 and 5, to compute the spectral radius of the non-negative matrix. The procedure for this is greatly simplified if it is known that we are

dealing with an irreducible matrix for it is in that case it can be guaranteed that the method recommended is always convergent. Wood & O'Neill [6] justified this. The following theorem describes conditions under which it is known that the matrices B and C in Theorem 2 and 7 are irreducible.

Theorem 8 *If Ω is a non-singular diagonal matrix and λ a scalar, then $G = \lambda I - \Omega^{-1}A$ is irreducible if and only if A is irreducible.*

Proof: This follows from the definition of "irreducible", since, if A and G have (ij)th elements α_{ij} and η_{ij} respectively, then, for all $j \neq i$, $\eta_{ij} = 0$ if and only if $\alpha_{ij} = 0$. Therefore A and G have the same zero elements.   ♠

So knowing that A is irreducible, guarantees that both B and C are irreducible.

3   Computational aspects

Both the method of Theorem 2 and that of Theorem 7 can be quite robust, in the sense of not being unduly affected by an ill-conditioned matrix. The next example shows this.

Example 9 If
$$A = \begin{bmatrix} 10^p & -10^p + \varepsilon \\ -10^p - \varepsilon & 10^p \end{bmatrix},$$
with p moderately large and ε small, then A is ill-conditioned—the condition number $\kappa(A) = \|A\|_\infty \|A^{-1}\|_\infty$ is approximately $4 \times 10^{2p}/\varepsilon^2$. For this matrix A,

$$B = \begin{bmatrix} 0 & 1 - 10^{-p}\varepsilon \\ 1 + 10^p\varepsilon & 0 \end{bmatrix} \quad \text{and} \quad \rho(B) = \sqrt{1 - 10^{-2p}\varepsilon^2} < 1.$$

Theorem 2 then gives the correct conclusion that A is an M-matrix.

Using $\lambda = 10^p + 1$ gives

$$C = \begin{bmatrix} 1 & 10^p - \varepsilon \\ 10^p + \varepsilon & 1 \end{bmatrix},$$

and $\rho(C) = 1 + \sqrt{10^{2p} - \varepsilon^2} = 1 + 10^p\sqrt{1 - 10^{-2p}\varepsilon^2} < 1 + 10^p.$

Theorem 7 also gives the correct conclusion that A is an M-matrix. When $p = \varepsilon = 1$, for example, a single iteration of the always convergent method gives an upper bound for the spectral radius which is less than the required value for an M-matrix in each case. However, if $10^{-2p}\varepsilon^2$ underflows to zero in approximate arithmetic, a wrong conclusion will result in each case.

In the method of Theorem 2, if D has very small elements, the computation of $D^{-1}A$ has the potential to compound any errors in A. This is not a problem with the method of Theorem 7. Furthermore, when D has very small elements, it is possible that B will be a non-symmetric matrix with elements which differ widely in magnitude. In such circumstances, B is likely to have ill-conditioned eigenvalues. Example 10 is a case in point.

Example 10 If

$$A = \begin{bmatrix} 0.6909 & 0 & -0.0059 \\ -0.8166 & 10^{-6} & 0 \\ -0.9810 & -0.0697 & 0.6909 \end{bmatrix}$$

then

$$B = 10^5 \begin{bmatrix} 0 & 0 & 8.54 \times 10^{-8} \\ 8.166 & 0 & 0 \\ 1.42 \times 10^{-5} & 1.009 \times 10^{-6} & 0 \end{bmatrix}$$

.

The eigenvalues of B are 8.8943, and $-4.4471 \pm 7.7019\,i$, which are almost equal in magnitude. This results in very slow convergence of the always convergent method described above. The dominant eigenvalue of B is ill-conditioned and this causes convergence to a slightly incorrect approximation. However, balancing the matrix remedies this situation. The method of Theorem 7, with $\lambda = 1.6909$, produces none of the above difficulties. Another advantage of the method of Theorem 7 when A is symmetric is that it preserves symmetry, since $\lambda I - A$ is symmetric if and only if A is symmetric. When A is symmetric, $I - D^{-1}A$ is symmetric if and only if $AD = DA$.

Compare the number of operations of each method. The method of Theorem 2 requires $(n^2 + n)$ operations to calculate B whereas, with the method of Theorem 7, only n operations are required to calculate C. Hence when n is large there is a possibility of greater rounding errors in the method of Theorem 2.

The always convergent method operates most effectively when the nonnegative matrix is irreducible. If the matrix is reducible, then there are at least two ways to proceed

1. Convert it to normal form. This form is block upper-triangular with diagonal blocks which are either null or irreducible. Theorem 2 or Theorem 7 may then be applied to the irreducible blocks on the diagonal. Normal form is achieved using pre-multiplication of the matrix, C, say, by a particular permutation matrix P and post-multiplication by its transpose P^T. If C is reducible the diagonal blocks of PCP^T will be either null or irreducible. For details of this process, see Senata [3]. The following example shows how to apply this technique.

Example 11 If

$$
A = \begin{bmatrix}
8 & 0 & -2 & 0 & 0 \\
0 & 10 & 0 & 0 & 0 \\
-2 & 0 & 10 & 0 & 0 \\
-2 & -2 & -2 & 10 & -2 \\
0 & 0 & 0 & 0 & 8
\end{bmatrix}
$$

This matrix is obviously reducible. It satisfies Properties 1 and 2.

Converting C with $\lambda = 11$ to normal form, and noting that the normal form will have the same eigenvalues as C, we obtain

$$
PCP^T = \begin{bmatrix}
[1] & 2 & 2 & 2 & 2 \\
0 & \begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix} & & 0 & 0 \\
0 & & & 0 & 0 \\
0 & 0 & 0 & [1] & 0 \\
0 & 0 & 0 & 0 & [3]
\end{bmatrix},
$$

where the four irreducible diagonal blocks have been indicated. Applying the always convergent method to the $(2,2)$ block gives $[4, 4.333]$ as the interval of uncertainty for the spectral radius of this block and ultimately then for $\rho(C)$. Since all values in this interval are less than $\lambda = 11$, A is confirmed as an M-matrix. This example illustrates the technique, but observe that the column norm of C reveals immediately that $\rho(C) \leq 7 \leq 11$.

2. Perturb slightly the reducible matrix.

   If the matrix B, say, is reducible then converging bounds do not necessarily occur, when the always convergent method is applied. However, this can be overcome by adding to B the matrix E, where

$$
E = \begin{bmatrix}
0 & \xi & 0 & \cdots & 0 \\
\vdots & 0 & \ddots & & 0 \\
0 & & \ddots & & \xi \\
\xi & 0 & & \cdots & 0
\end{bmatrix},
\quad \text{with small } \xi > 0.
$$

This ensures that $B + E$ is irreducible, and the always convergent method can be applied effectively. An attendant difficulty is then whether the perturbation in matrix B significantly affects the spectral radius. The stated perturbation effectively ensures, by the Perron–Frobenius Theorem, that $B + E$ has a unique dominant eigenvalue. To determine the impact of the perturbation a result from Stewart [4] is helpful. This result states that if λ is a simple eigenvalue of B with right eigenvector x and left eigenvector y, with $\|x\|_2 = 1$ and $y^T x = 1$, and B is deflated using an orthogonal matrix R such that $R^T B R = \begin{bmatrix} \lambda & h^T \\ 0 & F \end{bmatrix}$, then

$$|\lambda - \lambda'| \leq \varepsilon \|y\|_2 + \frac{\varepsilon^2}{\delta} + \eta \mathcal{O}(\varepsilon^2),$$

where λ' is the corresponding eigenvalue of the perturbed matrix $B+E$. Also, $\varepsilon = \|E\|_2$, $\delta = \|(\lambda I - F)^{-1}\|_2^{-1}$ and $\eta = \|h\|_2$. So the numbers $\|y\|_2$, δ and η give a measure of the condition of the simple eigenvalue λ.

For the matrix B which can be computed from A in Example 11, choose $\varepsilon = 10^{-6}$. Then $\|y\|_2 = 1.6210$, $\delta = 0.1523$, and $\eta = 0.3617$. This gives $|\lambda - \lambda'| < 1.6210 \times 10^{-6}$, indicating that the difference between the spectral radius of B and $B + E$ is probably suspect in the sixth decimal place.

The always convergent method requires solution of a system of linear equations. If this is computationally expensive, an alternative method is possible. This method was described by Wood & O'Neill [6] and is closely related to the ordinary power method. It is also an always convergent method provided that the non-negative matrix is primitive. If A is irreducible, then it can be easily shown that C is primitive, but B may be cyclic.

4 Conclusion

This paper discusses two methods for determining whether a matrix is an M-matrix. Both methods overcome the difficulty of showing that A^{-1} exists and is non-negative, by calculating or bounding the spectral radius of an associated non-negative matrix. If the non-negative matrix is irreducible, an always convergent method is available for calculating its spectral radius when this is necessary. This always convergent method makes use of the inverse power method but is a special variation of this method, appropriate to non-negative matrices. If the matrix is reducible, then two ways are suggested to handle this situation.

The method of Theorem 7 requires fewer operations than that of Theorem 2 and is thus potentially faster and more accurate. Also, it avoids the potential problem of amplification of errors in A, which can result in the calculation of $B = I - D^{-1}A$, when some elements of D are very small. However, take care to choose λ large enough to avoid drastic loss of significant digits in the calculation of $C = \lambda I - A$. The method of Theorem 7 preserves the symmetry of matrix A, but the method of Theorem 2 does not, in general. This has implications for the condition of the eigenvalues of B and C. An alternative method is described when there is an imperative to avoid the solution of a large system of linear equations. In this context, B may be cyclic, which results in non-convergence of the method. This problem does not occur with matrix C. It is reasonable then to conclude that the method of Theorem 7 is computationally superior to that of Theorem 2.

References

[1] A. Berman and N. Shaked–Monderer. *Completely Positive Matrices.* World Scientific. New Jersey, 2003. C735

[2] Y. Saad. *Iterative Methods for Sparse Linear Systems.* SIAM, Philadelphia, 2003. C733, C734

[3] E. Senata. *Non-Negative Matrices*, George Allen and Unwin, London, 1973. C739

[4] G. W. Stewart. *Introduction to Matrix Computations,* Academic Press, New York, 1973. C741

[5] R. Varga. *Matrix Iterative Analysis*, Prentice-Hall Inc, Englewood Cliffs, New Jersey, 1962. C735

[6] R. J. Wood and M. J. O'Neill. An always convergent method for finding the spectral radius of a non-negative matrix, *ANZIAM J.*, 45(E): C474–C485, 2004.
http://anziamj.austms.org.au/V45/CTAC2003/Wood C735, C737, C741

[7] D. M. Young. *Iterative Solutions of Large Linear Systems*, Academic Press, New York, 1971. C733, C734