What Is an M-Matrix?

Nicholas J. Higham*

March 16, 2021

An M-matrix is a matrix $A \in \mathbb{R}^{n\times n}$ of the form

$$A = sI - B, \quad \text{where } B \ge 0 \text{ and } s > \rho(B). \qquad (*)$$

Here, $\rho(B)$ is the spectral radius of B, that is, the largest modulus of any eigenvalue of B, and $B \ge 0$ denotes that B has nonnegative entries. An M-matrix clearly has nonpositive off-diagonal elements. It also has positive diagonal elements, which can be shown using the result that

$$\rho(A) = \lim_{k\to\infty} \|A^k\|^{1/k} \qquad (\dagger)$$

for any consistent matrix norm:

$$s > \rho(B) = \lim_{k\to\infty} \|B^k\|_\infty^{1/k} \ge \lim_{k\to\infty} \|\mathrm{diag}(b_{ii})^k\|_\infty^{1/k} = \max_i b_{ii}.$$

Although the definition of an M-matrix does not specify s, we can set it to $\max_i a_{ii}$. Indeed let s satisfy $(*)$ and set $\widetilde{s} = \max_i a_{ii}$ and $\widetilde{B} = \widetilde{s}I - A$. Then $\widetilde{B} \ge 0$, since $\widetilde{b}_{ii} = \widetilde{s} - a_{ii} \ge 0$ and $\widetilde{b}_{ij} = -a_{ij} = b_{ij} \ge 0$ for $i \ne j$. Furthermore, for a nonnegative matrix the spectral radius is an eigenvalue, by the Perron–Frobenius theorem, so $\rho(B)$ is an eigenvalue of B and $\rho(\widetilde{B})$ is an eigenvalue of $\widetilde{B}$. Hence $\rho(\widetilde{B}) = \rho((\widetilde{s} - s)I + B) = \widetilde{s} - s + \rho(B) < \widetilde{s}$.

The concept of M-matrix was introduced by Ostrowski in 1937. M-matrices arise in a variety of scientific settings, including in finite difference methods for PDEs, input-output analysis in economics, and Markov chains in stochastic processes.

An immediate consequence of the definition is that the eigenvalues of an M-matrix lie in the open right-half plane, which means that M-matrices are special cases of positive stable matrices. Hence an M-matrix is nonsingular and the determinant, being the product of the eigenvalues, is positive. Moreover, since $C = s^{-1}B$ satisfies $\rho(C) < 1$,

$$A^{-1} = s^{-1}(I - C)^{-1} = s^{-1}(I + C + C^2 + \cdots) \ge 0.$$

In fact, nonnegativity of the inverse characterizes M-matrices. Define

$$Z_n = \{\, A \in \mathbb{R}^{n\times n} : a_{ij} \le 0,\ i \ne j \,\}.$$

Theorem 1. *A nonsingular matrix $A \in Z_n$ is an M-matrix if and only if $A^{-1} \ge 0$.*

*Department of Mathematics, University of Manchester, Manchester, M13 9PL, UK (nick.higham@manchester.ac.uk).

Sometimes an M-matrix is *defined* to be a matrix with nonpositive off-diagonal elements and a nonnegative inverse. In fact, this condition is just one of a large number of conditions equivalent to a matrix with nonpositive off-diagonal elements being an M-matrix, fifty of which are given in Berman and Plemmons (1994, Chap. 6).

It is easy to check from the definitions, or using Theorem 1, that a triangular matrix T with positive diagonal and nonpositive off-diagonal is an M-matrix. An example is

$$T_4 = \begin{bmatrix} 1 & -1 & -1 & -1 \\ & 1 & -1 & -1 \\ & & 1 & -1 \\ & & & 1 \end{bmatrix}, \quad T_4^{-1} = \begin{bmatrix} 1 & 1 & 2 & 4 \\ & 1 & 1 & 2 \\ & & 1 & 1 \\ & & & 1 \end{bmatrix}.$$

Bounding the Norm of the Inverse

An M-matrix can be constructed from any nonsingular triangular matrix by taking the comparison matrix. The comparison matrix associated with a general $B \in \mathbb{R}^{n\times n}$ is the matrix

$$M(B) = (m_{ij}), \quad m_{ij} = \begin{cases} |b_{ii}|, & i = j, \\ -|b_{ij}|, & i \neq j. \end{cases}$$

For a nonsingular triangular T, $M(T)$ is an M-matrix, and it easy to show that

$$|T^{-1}| \leq |M(T)^{-1}|,$$

where the absolute value is taken componentwise. This bound, and weaker related bounds, can be useful for cheaply bounding the norm of the inverse of a triangular matrix. For example, with e denoting the vector of ones, since $M(T)^{-1}$ is nonnegative we have

$$\|T^{-1}\|_\infty \leq \|M(T)^{-1}\|_\infty = \|M(T)^{-1}e\|_\infty,$$

and $\|M(T)^{-1}e\|_\infty$ can be computed in $O(n^2)$ flops by solving a triangular system, whereas computing T^{-1} costs $O(n^3)$ flops.

More generally, if we have an LU factorization $PA = LU$ of an M-matrix $A \in \mathbb{R}^{n\times n}$ then, since $A^{-1} \geq 0$,

$$\|A^{-1}\|_\infty = \|A^{-1}e\|_\infty = \|U^{-1}L^{-1}Pe\|_\infty = \|U^{-1}L^{-1}e\|_\infty.$$

Therefore the norm of the inverse can be computed in $O(n^2)$ flops with two triangular solves, instead of the $O(n^3)$ flops that would be required if A^{-1} were to be formed explicitly.

Connections with Symmetric Positive Definite Matrices

There are many analogies between M-matrices and symmetric positive definite matrices. For example, every principal submatrix of a symmetric positive definite matrix is symmetric positive definite and every principal submatrix of an M-matrix is an M-matrix. Indeed if $\widetilde{B}$ is a principal submatrix of a nonnegative B then $\rho(\widetilde{B}) \leq \rho(B)$, which follows from (†) for the ∞-norm (for example). Hence on taking principal submatrices in (∗) we have $s > \rho(\widetilde{B})$ with the same s.

A symmetric M-matrix is known as a *Stieltjes matrix*, and such a matrix is positive definite. An example of a Stieltjes matrix is minus the second difference matrix (the tridiagonal matrix arising from a central difference discretization of a second derivative), illustrated for $n = 4$ by

$$A_4 = \begin{bmatrix} 2 & -1 & & \\ -1 & 2 & -1 & \\ & -1 & 2 & -1 \\ & & -1 & 2 \end{bmatrix}, \quad A_4^{-1} = \begin{bmatrix} \frac{4}{5} & \frac{3}{5} & \frac{2}{5} & \frac{1}{5} \\ \frac{3}{5} & \frac{6}{5} & \frac{4}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{4}{5} & \frac{6}{5} & \frac{3}{5} \\ \frac{1}{5} & \frac{2}{5} & \frac{3}{5} & \frac{4}{5} \end{bmatrix}.$$

LU Factorization

Since the leading principal submatrices of an M-matrix A have positive determinant it follows that A has an LU factorization with U having positive diagonal elements. However, more is true, as the next result shows.

Theorem 2. *An M-matrix A has an LU factorization $A = LU$ in which L and U are M-matrices.*

Proof. We can write

$$A = \begin{array}{cc} \begin{array}{cc} \scriptstyle 1 & \scriptstyle n-1 \end{array} & \\ \begin{bmatrix} \alpha & b^T \\ c^T & E \end{bmatrix} & \begin{array}{c} \scriptstyle 1 \\ \scriptstyle n-1 \end{array} \end{array}, \quad \alpha > 0, \quad b \le 0, \quad c \le 0.$$

The first stage of LU factorization is

$$A = \begin{bmatrix} \alpha & b^T \\ c & E \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \alpha^{-1}c & I \end{bmatrix} \begin{bmatrix} \alpha & b^T \\ 0 & S \end{bmatrix} = L_1 U_1, \quad S = E - \alpha^{-1} c b^T,$$

where S is the Schur complement of α in A. The first column of L_1 and the first row of U_1 are of the form required for a triangular M-matrix. We have

$$A^{-1} = U_1^{-1} L_1^{-1} = \begin{bmatrix} \alpha^{-1} & -\alpha^{-1} b^T S^{-1} \\ 0 & S^{-1} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\alpha^{-1}c & I \end{bmatrix} = \begin{bmatrix} \times & \times \\ \times & S^{-1} \end{bmatrix}.$$

Since $A^{-1} \ge 0$ it follows that $S^{-1} \ge 0$. It is easy to see that $S \in Z_n$, and hence Theorem 1 shows that S is an M-matrix. The result follows by induction.

The question now arises of what can be said about the numerical stability of LU factorization of an M-matrix. To answer it we use another characterization of M-matrices, that DA is strictly diagonally dominant by columns for some diagonal matrix $D = \mathrm{diag}(d_i)$ with $d_i > 0$ for all i, that is,

$$d_j |a_{jj}| > \sum_{i \ne j} d_i |a_{ij}|, \quad j = 1\colon n.$$

(This condition can also be written as $d^T A > 0$ because of the sign pattern of A.) Matrices that are diagonally dominant by columns have the properties that an LU factorization without pivoting exists, the growth factor $\rho_n \le 2$, and partial pivoting does not require row interchanges. The effect of row scaling on LU factorization is easy to see:

$$A = LU \;\Rightarrow\; DA = DLD^{-1} \cdot DU \equiv \widetilde{L}\widetilde{U},$$

where $\widetilde{L}$ is unit lower triangular, so that $\widetilde{L}$ and $\widetilde{U}$ are the LU factors of DA. It is easy to see that the growth factor bound of 2 for a matrix diagonally dominant by columns translates into the bound

$$\rho_n \le 2\frac{\max_i d_i}{\min_i d_i} \qquad (\ddagger)$$

for an M-matrix, which was obtained by Funderlic, Neumann, and Plemmons (1982). Unfortunately, this bound can be large. Consider the matrix

$$A = \begin{bmatrix} \epsilon & 0 & -1 \\ -1 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \in Z_3, \quad 0 < \epsilon < 1.$$

We have

$$A^{-1} = \begin{bmatrix} \dfrac{1}{\epsilon} & 0 & \dfrac{1}{\epsilon} \\ \dfrac{1}{\epsilon} & 1 & \dfrac{1+\epsilon}{\epsilon} \\ 0 & 0 & 1 \end{bmatrix} \ge 0,$$

so A is an M-matrix. The $(2,3)$ element of the LU factor U of A is $-1 - 1/\epsilon$, which means that

$$\rho_3 \ge \frac{1}{\epsilon} + 1,$$

and this lower bound can be arbitrarily large. One can verify experimentally that numerical instability is possible when ρ_3 is large, in that the computed LU factors have a large relative residual. We conclude that pivoting is necessary for numerical stability in LU factorization of M-matrices.

Stationery Iterative Methods

A stationery iterative method for solving a linear system $Ax = b$ is based on a splitting $A = M - N$ with M nonsingular, and has the form $Mx_{k+1} = Nx_k + b$. This iteration converges for all starting vectors x_0 if $\rho(M^{-1}N) < 1$. Much interest has focused on *regular splittings*, which are defined as ones for which $M^{-1} \ge 0$ and $N \ge 0$. An M-matrix has the important property that $\rho(M^{-1}N) < 1$ for every regular splitting, and it follows that the Jacobi iteration, the Gauss-Seidel iteration, and the successive overrelaxation (SOR) iteration (with $0 < \omega \le 1$) are all convergent for M-matrices.

Matrix Square Root

The principal square root $A^{1/2}$ of an M-matrix A is an M-matrix, and it is the unique such square root. An expression for $A^{1/2}$ follows from $(*)$:

$$\begin{aligned} A^{1/2} &= s^{1/2}(I - C)^{1/2} \quad (C = s^{-1}B,\ \rho(C) < 1), \\ &= s^{1/2}\sum_{j=0}^{\infty} \binom{\frac{1}{2}}{j}(-C)^j. \end{aligned}$$

This expression does not necessarily provide the best way to compute $A^{1/2}$.

Singular M-Matrices

The theory of M-matrices extends to the case where the condition on s is relaxed to $s \geq \rho(B)$ in $(*)$, though the theory is more complicated and extra conditions such as irreducibility are needed for some results. Singular M-matrices occur in Markov chains (Berman and Plemmons, 1994, Chapter 8), for example. Let P be the transition matrix of a Markov chain. Then P is stochastic, that is, nonnegative with unit row sums, so $Pe = e$. A nonnegative vector y with $y^T e = 1$ such that $y^T P = y^T$ is called a *stationary distribution vector* and is of interest for describing the properties of the Markov chain. To compute y we can solve the singular system $Ay = (I - P^T)y = 0$. Clearly, $A \in Z_n$ and $\rho(P) = 1$, so A is a singular M-matrix.

H-Matrices

A more general concept is that of H-matrix: $A \in \mathbb{R}^{n\times n}$ is an H-matrix if the comparison matrix $M(A)$ is an M-matrix. Many results for M-matrices extend to H-matrices. For example, for an H-matrix with positive diagonal elements the principal square root exists and is the unique square root that is an H-matrix with positive diagonal elements. Also, the growth factor bound (‡) holds for any H-matrix for which DA is diagonally dominant by columns.

References

This is a minimal set of references, which contain further useful references within.

- Abraham Berman and Robert J. Plemmons, Nonnegative Matrices in the Mathematical Sciences, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1994.
- R. E. Funderlic and M. Neumann and R. J. Plemmons *LU* Decompositions of Generalized Diagonally Dominant Matrices, Numer. Math. 40, 57–69, 1982.
- Nicholas J. Higham, Accuracy and Stability of Numerical Algorithms, second edition, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2002. (Chapter 8.)
- Nicholas J. Higham, Functions of Matrices: Theory and Computation, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2008. (Section 6.8.3.)

Related Blog Posts

- What Is a Matrix Square Root? (2020)
- What Is a Symmetric Positive Definite Matrix? (2020)
- What is Numerical Stability? (2020)

This article is part of the "What Is" series, available from `https://nhigham.com/category/what-is` and in PDF form from the GitHub repository `https://github.com/higham/what-is`.