On Square Roots of *M*-Matrices

G. Alefeld* and N. Schneider
FB 3 / Mathematik
Technische Universität Berlin
Strasse des 17. Juni 135
1000 Berlin 12, Germany

Submitted by Hans Schneider

ABSTRACT

The question of the existence and uniqueness of an M-matrix which is a square root of an M-matrix is discussed. The results are then used to derive some new necessary and sufficient conditions for a real matrix with nonpositive off diagonal elements to be an M-matrix.

1. INTRODUCTION

Following Ostrowski [3], a real n by n matrix $A=(a_{ij})$ is called an *M-matrix* if it can be written in the form

$$A=sI-B, \qquad s>0, \quad B\geqslant 0, \quad \rho(B)\leqslant s. \tag{1}$$

Here ρ denotes the spectral radius and I is the unit matrix. If $\rho(B)<s$, then A is called a *nonsingular M-matrix*; otherwise, a *singular M-matrix*.

In this paper we discuss the existence and uniqueness of an M-matrix which is a solution of the equation

$$X^2-A=0, \tag{2}$$

where A is a given M-matrix. A solution of (2) is called a square root of A. The following example shows that there are M-matrices which have no square root at all.

*Current address: Institut für Angewandte Mathematik, Universität Karlsruhe, Kaiserstrasse 12, 7500 Karlsruhe, Germany.

52 Vanderbilt Ave., New York, NY 10017 0024-3795/82/010119+14$02.75

EXAMPLE. Let

$$A=\begin{pmatrix}0 & -1\\ 0 & 0\end{pmatrix}=sI-B,\quad s>0,\qquad B=\begin{pmatrix}s & 1\\ 0 & s\end{pmatrix}.$$

Since $\rho(B)=s$, it follows that A is a singular M-matrix. A short discussion shows that (2) has no solution.

In this paper we prove that an M-matrix A has an M-matrix as a square root if and only if A has "property c." For a subset of the set of M-matrices with "property c" we are able to prove that there exists only one M-matrix as a square root. This subset contains M-matrices which are irreducible or nonsingular. The proof of the existence of a square root is constructive. This allows us to compute this matrix.

In Section 2 we list some definitions and well-known facts which we use in the sequel. Section 3 contains the main results. Finally, in Section 4 we give some new necessary and sufficient conditions for a real matrix with nonpositive off diagonal elements to be an M-matrix.

2. PRELIMINARIES

DEFINITION 1. An n by n matrix T is called *semiconvergent* if and only if the limit $\lim_{j\to\infty}T^j$ exists. (See [1, p. 152].)

The following result is established by use of the Jordan form for T (see [1, p. 152]).

THEOREM 1. *The n by n matrix T is semiconvergent if and only if each of the following conditions holds.*

(a) $\rho(T)\leqslant 1$.

(b) *If $\rho(T)=1$, then all elementary divisors associated with the eigenvalue 1 of T are linear.*

(c) *If $\rho(T)=1$, then $\lambda\in\sigma(T)$ ($\sigma(T)$=spectrum of T) with $|\lambda|=1$ implies $\lambda=1$.*

For the next result and in the sequel of this paper we use the definition of irreducibility of a matrix, which can be found in [4, p. 18]. Notice that a 1 by 1 matrix is irreducible iff its only element is different from zero.

THEOREM 2. *If $T=(t_{ij})\geqslant 0$, $t_{ii}>0$, $1\leqslant i\leqslant n$, then T is semiconvergent if* (a) *and* (b) *of the preceding theorem hold.*

Proof. It is sufficient to prove that (c) of Theorem 1 holds. There exists an n by n permutation matrix Q such that

$$QTQ^T = \begin{pmatrix} R_{11} & R_{12} & \cdots & R_{1m} \\ & R_{22} & & \vdots \\ \bigcirc & & \ddots & \\ & & & R_{mm} \end{pmatrix}, \qquad 1 \leqslant m \leqslant n,$$

where each square submatrix R_{jj}, $1 \leqslant j \leqslant m$, is irreducible. See for example [4, p. 46, (2.41)]. If T is irreducible, then $m = 1$ and we can choose $Q = I$. Since the diagonal elements of each R_{ii} are also positive, it follows from Lemma 2.8 and Theorem 2.5 in [4, p. 41] that each R_{ii} is primitive. (See Definition 2.2 in [4, p. 35]). This means that for each i there is only one eigenvalue λ of R_{ii}, namely $\rho(R_{ii})$, for which $|\lambda| = \rho(R_{ii})$ holds. From these remarks (c) of the preceding theorem follows. ■

Definition 2. An M-matrix A is said to have "*property c*" if it can be split into $A = sI - B$, $s > 0$, $B \geqslant 0$, where the matrix $T = B/s$ is semiconvergent.

All nonsingular M-matrices have "property c." There are, however, singular M-matrices which do not share this property (see [1, p. 152ff]).

Let

$$Z^{n\times n} := \{A = (a_{ij}) | a_{ij} \leqslant 0, i \neq j\}.$$

Then we have the following results.

Theorem 3.

(a) $A \in Z^{n\times n}$ *is a nonsingular M-matrix if and only if* A^{-1} *exists and* $A^{-1} \geqslant 0$ (see [1, p. 134 ff.]).

(b) $A \in Z^{n\times n}$ *is a nonsingular M-matrix if and only if the real parts of all eigenvalues are positive. The same characterization holds for the nonzero eigenvalues of a singular M-matrix* (see [1, pp. 134 ff., 147 ff.]).

(c) *If A and B are n by n M-matrices and if* $AB \in Z^{n\times n}$, *then AB is an M-matrix* (*see* [1, *p.* 159, *Exercise* 5.2]).

If we define

$$[\alpha_1,\dots,\alpha_n]:=\begin{pmatrix} \alpha_1 & \alpha_2 & \cdot & \cdot & \cdot & \cdot & \alpha_n \\ & \alpha_1 & \alpha_2 & & & & \cdot \\ & & \cdot & \cdot & & & \cdot \\ & & & \cdot & \cdot & & \cdot \\ & & & & \cdot & \cdot & \cdot \\ & & & & & \cdot & \alpha_2 \\ \bigcirc & & & & & & \alpha_1 \end{pmatrix}, \tag{3}$$

then

$$[\alpha_1,\dots,\alpha_n]\cdot[\beta_1,\dots,\beta_n]=\left[\alpha_1\beta_1,\alpha_1\beta_2+\alpha_2\beta_1,\dots,\sum_{k=1}^{n}\alpha_k\beta_{n-k+1}\right].$$

Therefore the set of matrices of the form (3) is closed under ordinary matrix multiplication.

3. SQUARE ROOTS OF *M*-MATRICES

Let there be given an n by n matrix A. We know from the definition of an M-matrix that there exists an $s>0$ such that

$$A=sI-\tilde{P},\qquad \tilde{P}\geqslant 0,\quad \rho(\tilde{P})\leqslant s.$$

If $\Delta s>0$ and $A=(s+\Delta s)I-(\Delta sI+\tilde{P})$, then using the Perron-Frobenius theorem [4, p. 46], it follows that $\rho(\Delta sI+\tilde{P})=\Delta s+\rho(\tilde{P})\leqslant\Delta s+s$. From this remark we conclude that there exists an $s_0>0$ such that for all $s\geqslant s_0$ we can represent the M-matrix A in the following manner:

$$A=sI-\tilde{P}=s(I-\tilde{P}/s)=s(I-P)$$

$$\text{with}\quad P=\tilde{P}/s,\quad \operatorname{diag}P>0,\quad \rho(P)\leqslant 1 \tag{MM}$$

(diag P denotes the diagonal part of P). Since $A/s=I-P$, one knows a square root of A if one knows a square root of $I-P$.

It is obvious that A has an M-matrix as a square root if and only if $I-P$ has an M-matrix as a square root. Therefore we can restrict the discussion in the following lemma to M-matrices which have the special form

$$A=I-P, \qquad P\geqslant 0, \quad \operatorname{diag} P>0, \quad \rho(P)\leqslant 1.$$

LEMMA 1. *Let there be given an n by n matrix $P=(p_{ij})\geqslant 0$ where $\rho(P)\leqslant 1$ and $\operatorname{diag} P>0$. Let $\alpha\geqslant 1$. Then the following three statements hold:*

(a) *There exists an n by n matrix $B\geqslant 0$ where $\rho(B)\leqslant\alpha$ and $I-P=(\alpha I-B)^2$ if and only if the iteration method*

$$X_{i+1}=\frac{1}{2\alpha}\left[P+(\alpha^2-1)I+X_i^2\right], \qquad X_0=0, \tag{4}$$

is convergent. In this case $B\geqslant X^*=\lim_{i\to\infty}X_i$, $X^*\geqslant 0$, $\rho(X^*)\leqslant\alpha$, $\operatorname{diag} X^*>0$, and $(\alpha I-X^*)^2=I-P$.

(b) *If* (4) *is convergent, it follows that P and X^*/α are semiconvergent.*

(c) *If P is semiconvergent, then* (4) *is convergent for all $\alpha\geqslant 1$. Denoting in this case the limit of the iteration method*

$$Y_{i+1}=\tfrac{1}{2}\left(P+Y_i^2\right), \qquad Y_0=0,$$

by Y^*, the equation

$$\alpha I-X^*=I-Y^*$$

holds.

Proof. (a), $\Rightarrow$: We prove by induction that the sequence which is computed using (4) is increasing and bounded. Assuming $X_{i+1}\geqslant X_i$, $X_i\leqslant B$, which is obviously true for $i=0$, it follows that

$$X_{i+2}=\frac{1}{2\alpha}\left[P+(\alpha^2-1)I+X_{i+1}^2\right]\geqslant\frac{1}{2\alpha}\left[P+(\alpha^2-1)I+X_i^2\right]=X_{i+1}$$

and

$$\begin{aligned}B-X_{i+1}&=B-\frac{1}{2\alpha}\left[P+(\alpha^2-1)I+X_i^2\right]\\&=\frac{1}{2\alpha}\left[I-P-\left(\alpha^2I-2\alpha B+X_i^2\right)\right]\geqslant\frac{1}{2\alpha}\left[I-P-(\alpha I-B)^2\right]=0.\end{aligned}$$

Therefore the iteration method (4) is convergent. For its limit X^* we have $\lim_{i\to\infty} X_i = X^* \leqslant B$, $(\alpha I - X^*)^2 = I - P$, diag $X^* > 0$, $X^* \geqslant 0$. Because of $0 \leqslant X^* \leqslant B$, it follows that $\rho(X^*) \leqslant \rho(B) \leqslant \alpha$ from the Perron-Frobenius theorem [4, p. 46, Theorem 2.7].

$\Leftarrow$: Assuming the convergence of the method (4), we have $X^* = \lim_{i\to\infty} X_i \geqslant 0$ and $I - P = (\alpha I - X^*)^2$. Let U be a matrix which transforms P into Jordan form:

$$U^{-1}PU = \begin{pmatrix} J_1 & & \bigcirc \\ & \ddots & \\ \bigcirc & & J_k \end{pmatrix},$$

where

$$J_j = \begin{pmatrix} \lambda_j & 1 & & & \bigcirc \\ & \cdot & \cdot & & \\ & & \cdot & \cdot & \\ & & & \cdot & 1 \\ \bigcirc & & & & \lambda_j \end{pmatrix}, \qquad 1 \leqslant j \leqslant k.$$

Setting $Z_i := U^{-1} X_i U$, we get from (4)

$$Z_{i+1} = \frac{1}{2\alpha}\left[U^{-1}PU + (\alpha^2 - 1)I + Z_i^2\right], \qquad Z_0 = 0, \tag{5}$$

and therefore each diagonal element μ of $U^{-1}PU$ is related to a diagonal element $\lambda^{(i+1)}$ of Z_{i+1} by the equation

$$\lambda^{(i+1)} = \frac{1}{2\alpha}\left[\mu + (\alpha^2 - 1) + (\lambda^{(i)})^2\right]. \tag{6}$$

We prove by mathematical induction that $\rho(X_i) \leqslant \alpha$ and therefore that $\rho(X^*) \leqslant \alpha$ holds. For $i = 0$ this is trivially the case. Using $\rho(P) \leqslant 1$ and the induction hypothesis, we have

$$|\lambda^{(i+1)}| \leqslant \frac{1}{2\alpha}\left[|\mu| + \alpha^2 - 1 + |\lambda^{(i)}|^2\right] \leqslant \frac{1}{2\alpha}\left(|\mu| - 1 + 2\alpha^2\right) \leqslant \alpha. \tag{7}$$

Therefore the assertion holds with $B:=X^*$.

(b): Let (4) be convergent, and assume that P is not semiconvergent. Again let U be a matrix which transforms P into Jordan form. Since $\rho(P)\leqslant 1$ and $\operatorname{diag} P>0$, it follows, using Theorem 2, that at least one of the submatrices J_j, $1\leqslant j\leqslant k$, has the form

$$J_j=\begin{pmatrix} 1 & 1 & & & \bigcirc \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ \bigcirc & & & & 1 \end{pmatrix},$$

where the order of J_j is greater than one. Let this be the case for $j:=\bar{j}$. Setting again $Z_i=U^{-1}X_iU$, we see from (5) that Z_i has the form

$$Z_i=\begin{pmatrix} J_1^{(i)} & & \bigcirc \\ & \ddots & \\ \bigcirc & & J_k^{(i)} \end{pmatrix}.$$

$J_{\bar{j}}^{(i)}$ has the form

$$J_{\bar{j}}^{(i)}=\left[\beta_1^{(i)},\dots,\beta_{\nu_{\bar{j}}}^{(i)}\right].$$

The sequence $(\beta_1^{(i)})$ is computed by the recursion

$$\beta_1^{(i+1)}=\frac{1}{2\alpha}\left[1+(\alpha^2-1)+\left(\beta_1^{(i)}\right)^2\right],\qquad \beta_1^{(0)}=0.$$

Since this sequence is increasing and bounded, it is convergent and $\lim \beta_1^{(i)}=\alpha$. Assume now that the sequence $(\beta_2^{(i)})$ which is computed using the iteration method

$$\beta_2^{(i+1)}=\frac{1}{2\alpha}\left(1+2\beta_1^{(i)}\beta_2^{(i)}\right),\qquad \beta_2^{(0)}=0 \tag{8}$$

is convergent. Setting $\lim \beta_2^{(i)}=\beta$, we get from (8)

$$\beta=\frac{1}{2\alpha}(1+2\alpha\beta),$$

which is a contradiction. Therefore the sequence (Z_i) and hence the sequence (X_i) cannot be convergent. Contradiction! This means that P is semiconvergent. It remains to be shown that also X^*/α is semiconvergent. Since by Lemma 1(a), diag $X^*>0$ and therefore also diag $(X^*/\alpha)>0$, it is sufficient by Theorem 2 to show that (a) and (b) of Theorem 1 hold. In part (a) of Lemma 1 we have shown that $\rho(X^*)\leqslant\alpha$, and therefore it follows that $\rho(X^*/\alpha)\leqslant 1$. This is (a) of Theorem 1. Let now $\rho(X^*/\alpha)=1$. In order to prove (b) of Theorem 1, we have to show that all elementary divisors associated with the eigenvalue 1 of X^*/α are linear. Passing to the limit in (6), it follows that every eigenvalue μ of P is associated with an eigenvalue λ of X^* by the equation $1-\mu=(\alpha-\lambda)^2$. If now $\lambda/\alpha=1$, then $\mu=1$ follows. Passing to the limit in (5), it follows that if μ is an eigenvalue of P the elementary divisors of which are linear, then the elementary divisors of the related eigenvalue λ are also linear. Since P is semiconvergent, (b) of Theorem 1 holds for P and therefore also for X^*/α.

(c): Let P be semiconvergent. Then P belongs to the class M (see [2, p. 47]). Hence a natural matrix norm exists for which $\|P\|\leqslant 1$ holds. Using mathematical induction, we prove that the sequence computed using (4) is norm-bounded. If $\|X_i\|\leqslant\alpha$, which is the case for $i=0$, then it follows that

$$\|X_{i+1}\|\leqslant\frac{1}{2\alpha}\left(\|P\|+\alpha^2-1+\|X_i\|^2\right)\leqslant\frac{1}{2\alpha}(1+\alpha^2-1+\alpha^2)=\alpha.$$

The sequence is also increasing, and therefore it is convergent. Setting again $Z_i=U^{-1}X_iU$ and in addition $\tilde{Z}_i=U^{-1}Y_iU$, we have

$$Z_i=\begin{pmatrix} J_1^{(i)} & & \bigcirc \\ & \ddots & \\ \bigcirc & & J_k^{(i)}\end{pmatrix},\qquad \tilde{Z}_i=\begin{pmatrix} \tilde{J}_1^{(i)} & & \bigcirc \\ & \ddots & \\ \bigcirc & & \tilde{J}_k^{(i)}\end{pmatrix}.$$

Here $J_j^{(i)}$ and $\tilde{J}_j^{(i)}$, $1\leqslant j\leqslant k$, are matrices of the form

$$J_j^{(i)}=\left[\beta_1^{(i)},\ldots,\beta_{\nu_j}^{(i)}\right],\qquad \tilde{J}_j^{(i)}=\left[\tilde{\beta}_1^{(i)},\ldots,\tilde{\beta}_{\nu_j}^{(i)}\right],$$

where the elements are computed in the following way:

$$\beta_1^{(i+1)} = \frac{1}{2\alpha}\left[\lambda_j + (\alpha^2 - 1) + \left(\beta_1^{(i)}\right)^2\right], \qquad \beta_1^{(0)} = 0,$$

$$\tilde{\beta}_1^{(i+1)} = \tfrac{1}{2}\left[\lambda_j + \left(\tilde{\beta}_1^{(i)}\right)^2\right], \qquad \tilde{\beta}_1^{(0)} = 0,$$

$$\beta_2^{(i+1)} = \frac{1}{2\alpha}\left[1 + 2\beta_2^{(i)}\beta_1^{(i)}\right], \qquad \beta_2^{(0)} = 0,$$

$$\tilde{\beta}_2^{(i+1)} = \tfrac{1}{2}\left[1 + 2\tilde{\beta}_2^{(i)}\tilde{\beta}_1^{(i)}\right], \qquad \tilde{\beta}_2^{(0)} = 0,$$

and in general

$$\left.\begin{aligned} \beta_r^{(i+1)} &= \frac{1}{2\alpha}\sum_{l=1}^{r} \beta_l^{(i)}\beta_{r-l+1}^{(i)}, \quad \beta_r^{(0)} = 0 \\ \tilde{\beta}_r^{(i+1)} &= \frac{1}{2}\sum_{l=1}^{r} \tilde{\beta}_l^{(i)}\tilde{\beta}_{r-l+1}^{(i)}, \quad \tilde{\beta}_r^{(0)} = 0 \end{aligned}\right\} \qquad 3 \leqslant r \leqslant \nu_j.$$

(In order to simplify the notation we suppress the fact that the elements are actually dependent also on j.) Setting $\beta_r = \lim_{i\to\infty}\beta_r^{(i)}$, $\tilde{\beta}_r = \lim_{i\to\infty}\tilde{\beta}_r^{(i)}$, $1 \leqslant r \leqslant \nu_j$, we have $(\alpha - \beta_1)^2 = 1 - \lambda_j$ and therefore $\beta_1 = \alpha - \sqrt{1-\lambda_j}$. [Since $A = I - P$ is an M-matrix, all eigenvalues of A have nonnegative real parts. See Theorem 3, part (b). Since $\alpha - \beta_1$ is an eigenvalue of $\alpha I - X^*$, which is an M-matrix by (a), we have to choose the unique square root of $1 - \lambda_j$ which has a nonnegative real part.] Since also $\tilde{\beta}_1 = 1 - \sqrt{1-\lambda_j}$, we have $\beta_1 = \alpha - 1 + \tilde{\beta}_1$. Using mathematical induction, we are able to prove that $\beta_r = \tilde{\beta}_r$, $r = 1(1)\nu_j$. This can be done for all j, $1 \leqslant j \leqslant k$. This means

$$U^{-1}X^*U = (\alpha - 1)I + U^{-1}Y^*U,$$

and therefore we have proved the assertion. ∎

Using the preceding lemma, we can establish the following result.

THEOREM 4. *Let A be an n by n M-matrix, and let $A = s(I-P)$ be a representation of A of the form* (MM). *A has an M-matrix as a square root if and only if A has "property c." In this case let Y^* denote the limit of the* *sequence generated by*

$$Y_{i+1}=\tfrac{1}{2}\left(P+Y_i^2\right),\qquad Y_0=0.$$

Then $\sqrt{s}\,(I-Y^*)$ is an M-matrix with "property c" which is a square root of A. For any other M-matrix Z^* which is a square root of A the relation $Z^*\leqslant\sqrt{s}\,(I-Y^*)$ holds.

Proof. If A has an M-matrix as a square root, then we get from Lemma 1, parts (a) and (b), that P is semiconvergent and therefore that A has "property c." The other direction follows from Lemma 1, parts (a) and (c). From Lemma 1, part (b), we also obtain that the matrix $\sqrt{s}\,(I-Y^*)$ has "property c."

Finally let Z^* be an M-matrix which is a square root of A, and let $Z^*=\beta I-B$, $\beta\geqslant\sqrt{s}$, be a representation of this matrix of the form (MM). The matrix $I-P$ can be written in the form

$$I-P=\frac{A}{s}=\left(\frac{\beta}{\sqrt{s}}I-\frac{B}{\sqrt{s}}\right)^2,$$

where $\beta/\sqrt{s}\geqslant 1$. Setting $\alpha:=\beta/\sqrt{s}$ and denoting by X^* the limit of the sequence generated by

$$X_{i+1}=\frac{1}{2\alpha}\left[P+(\alpha^2-1)I+X_i^2\right],\qquad X_0=0,$$

we get from Lemma 1, parts (a) and (c), that $X^*\leqslant B/\sqrt{s}$ and $\alpha I-B/\sqrt{s}\leqslant\alpha I-X^*=\alpha I-[(\alpha-1)I+Y^*]=I-Y^*$, that is,

$$\sqrt{s}\,(I-Y^*)\geqslant\beta I-B=Z^*.$$

This shows that the matrix $\sqrt{s}\,(I-Y^*)$ is the largest M-matrix which is a square root of A. ∎

We consider now the 2 by 2 zero matrix $A=0$. Each M-matrix of the form

$$B=\begin{pmatrix}0 & b\\ 0 & 0\end{pmatrix},\qquad b\leqslant 0,$$

is a square root of A which has "property c." Therefore the problem of finding an M-matrix which is a square root of an M-matrix with "property c" is in general not uniquely solvable. However, for a certain subset of the set of M-matrices with "property c" the following theorem shows the uniqueness of the solution of this problem. This subset contains the set of nonsingular M-matrices and the set of irreducible M-matrices.

THEOREM 5. *An M-matrix $A=(a_{ij})$ has exactly one M-matrix as a square root if 0 is at most a simple zero of the characteristic polynomial of A.*

Proof. If 0 is at most a simple zero of the characteristic polynomial of A, then A has "property c" and Theorem 4 guarantees the existence of an M-matrix which is a square root of A. Assume again that A is expressed in the form (MM), that is, $A=s(I-P)$. We know from Theorem 4 that if Y^* is the limit of the sequence (Y_i) generated by

$$Y_{i+1}=\tfrac{1}{2}\left(P+Y_i^2\right), \qquad Y_0=0,$$

then $I-Y^*$ is an M-matrix which is a square root of $I-P$. We assume that there exists another M-matrix Z^* which is a square root of $I-P$, and that $Z^*=\alpha I-B$, $\alpha\geqslant 1$, denotes a representation of the form (MM).

From Lemma 1, part (a), we know that if X^* denotes the limit of the sequence generated by

$$X_{i+1}=\frac{1}{2\alpha}\left[P+(\alpha^2-1)I+X_i^2\right], \qquad X_0=0,$$

then

$$B\geqslant X^*, \qquad (\alpha I-X^*)^2=I-P=(\alpha I-B)^2.$$

Let Q be a permutation matrix which transforms the matrix B into the reducible normal form

$$\tilde{B}=QBQ^T=\begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1m} \\ & B_{22} & & B_{2m} \\ \bigcirc & & \ddots & \vdots \\ & & & B_{mm} \end{pmatrix}, \qquad 1\leqslant m\leqslant n,$$

where each square submatrix B_{ii} is irreducible, since we have diag $B>0$. If B

is irreducible, then $m=1$ and we can choose $Q=I$. Because $B\geq X^*\geq 0$ we get

$$\tilde{X}=QX^*Q^T=\begin{pmatrix} X_{11} & X_{12} & \cdots & X_{1m} \\ & X_{22} & & X_{2m} \\ \bigcirc & & \ddots & \vdots \\ & & & X_{mm} \end{pmatrix},$$

where B_{ij} and X_{ij} have the same size.

We first show that $X_{jj}=B_{jj}$, $1\leq j\leq m$. We know that $X_{jj}\leq B_{jj}$, and we assume $X_{jj}\neq B_{jj}$ for some j. From the Perron-Frobenius theorem for irreducible matrices it follows that $\rho(B_{jj})>\rho(X_{jj})$. From $(\alpha I-X^*)^2=(\alpha I-B)^2$ it follows $(\alpha I-\tilde{B})^2=(\alpha I-\tilde{X})^2$, and therefore $(\alpha I-B_{jj})^2=(\alpha I-X_{jj})^2$. Setting$\mu:=\rho(B_{jj})\leq\alpha$, there must exist an eigenvalue $\tilde{\mu}$ of $\tilde{X}_{jj}$ such that

$$(\alpha-\mu)^2=(\alpha-\tilde{\mu})^2,$$

or

$$(\mu-\tilde{\mu})(\mu+\tilde{\mu})=2\alpha(\mu-\tilde{\mu}).$$

Since $|\tilde{\mu}|<\mu$, we get $\mu+\tilde{\mu}=2\alpha$, which is a contradiction. Therefore we have $X_{jj}=B_{jj}$, $1\leq j\leq m$.

Since 1 is at most a simple zero of the characteristic polynomial of P, we know from (6) that α is at most a simple zero of the characteristic polynomial of $\tilde{X}$. Therefore we know that at most one of the diagonal blocks X_{jj} of $\tilde{X}$ has the eigenvalue α. We prove by mathematical induction that

$$X_{lk}=B_{lk},\qquad 0\leq k-l\leq m-1,$$

holds.

For $k-l=0$ we have already shown the assertion. From $(\alpha I-\tilde{B})^2=(\alpha I-\tilde{X})^2$ it follows that

$$\tilde{B}^2-2\alpha\tilde{B}=\tilde{X}^2-2\alpha\tilde{X},$$

and therefore

$$\sum_{j=l}^{k}B_{lj}B_{jk}-2\alpha B_{lk}=\sum_{j=l}^{k}X_{lj}X_{jk}-2\alpha X_{lk}.$$

Using the induction hypothesis it follows that

$$X_{ll}B_{lk}+B_{lk}X_{kk}-2\alpha B_{lk}=X_{ll}X_{lk}+X_{lk}X_{kk}-2\alpha X_{lk},$$

and therefore

$$(X_{ll}-\alpha I)(B_{lk}-X_{lk})+(B_{lk}-X_{lk})(X_{kk}-\alpha I)=0. \tag{9}$$

Since the matrices $\alpha I-X_{ll}$ and $\alpha I-X_{kk}$ are M-matrices and since at most one of these two is singular, it follows that Equation (9) has only the trivial solution, that is, $B_{lk}=X_{lk}$. (See [5, p. 262, Theorem 8.5.1].) Therefore we have $B=X^*$, and using Lemma 1, part (c), we obtain

$$\alpha I-B=\alpha I-X^*=I-Y^*.$$

Therefore $\sqrt{s}\,(I-Y^*)$ is the only M-matrix which is a square root of the given matrix A. ∎

4. SOME CHARACTERIZATIONS OF M-MATRICES

In [1, p. 134 ff.] there is listed a series of conditions which characterize matrices $A\in Z^{n\times n}$ that are nonsingular M-matrices or singular M-matrices with "property c." Using the Theorem 4 of the preceding section we can establish the following results.

COROLLARY 1. *$A\in Z^{n\times n}$ is a nonsingular M-matrix if and only if there exists a nonsingular M-matrix Y^* for which $A=(Y^*)^2$ holds.*

Proof. If $A\in Z^{n\times n}$ is a nonsingular M-matrix, then the assertion follows from Theorem 4, since A has "property c." If on the other hand Y^* is a nonsingular M-matrix, then using Theorem 3, part (a), it follows from $A=(Y^*)^2$ that A is a nonsingular M-matrix. ∎

COROLLARY 2. *$A\in Z^{n\times n}$ is a nonsingular M-matrix if and only if there exists a nonnegative matrix Z^* for which $A(Z^*)^2=I$ holds.*

Proof. Let $A\in Z^{n\times n}$ and let $Z^*\geqslant 0$. Then it follows from $A(Z^*)^2=I$ that A^{-1} exists and $A^{-1}\geqslant 0$ holds. Using Theorem 3 this means that A is a nonsingular M-matrix. If on the other hand A is a nonsingular M-matrix, then

Theorem 4 guarantees the existence of a nonsingular M-matrix Y^* for which $A=(Y^*)^2$ holds. Therefore the equation $A(Z^*)^2=I$, $Z=(Y^*)^{-1}\geqslant 0$, follows. ∎

The next result deals with the singular case.

COROLLARY 3. *$A\in Z^{n\times n}$ is a singular M-matrix with "property c" if and only if there exists a singular M-matrix Y^* with "property c" for which $A=(Y^*)^2$ holds.*

Proof. If $A\in Z^{n\times n}$ is a singular M-matrix with "property c," then the assertion follows from Theorem 4. If on the other hand Y^* is a singular M-matrix with $A=(Y^*)^2$, then using Theorem 3, part (c), it follows that A is a singular M-matrix. We have to show that A has "property c." Since Y^* has "property c," it follows from Lemma (4.11) in [1, p. 153] that $\operatorname{rank} Y^*=\operatorname{rank}(Y^*)^2$. This implies the equation $\operatorname{rank}(Y^*)^2=\operatorname{rank}(Y^*)^4$ or $\operatorname{rank} A=\operatorname{rank} A^2$. Applying again Lemma (4.11) from [1], we have the result that A has "property c." ∎

The authors are very much indebted to Professor Hans Schneider for his many suggestions for improving an earlier version of this paper.

REFERENCES

1 A. Berman and R. J. Plemmons, *Non-negative Matrices in the Mathematical Sciences*, Academic, 1979.
2 A. S. Householder, *The Theory of Matrices in Numerical Analysis*, Ginn (Blaisdell), Waltham, Mass. 1964.
3 A. M. Ostrowski, Über die Determinanten mit überwiegender Hauptdiagonale, *Comment. Math. Helv.* 10:69–96 (1937).
4 R. S. Varga, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, N.J. (1962).
5 P. Lancaster, *Theory of Matrices*, Academic, 1969.

Received 18 November 1980; revised 28 April 1981