Monotonicity and Iterative Approximations Involving Rectangular Matrices

By Robert J. Plemmons*

Abstract. A new characterization of row-monotone matrices is given and is related to the Moore-Penrose generalized inverse. The M-matrix concept is extended to rectangular matrices with full column rank. A structure theorem is provided for all matrices A with full column rank for which the generalized inverse $A^+ \geq 0$. These results are then used to investigate convergent splittings of rectangular matrices in relation to iterative techniques for computing best least squares solutions to rectangular systems of linear equations.

1. Introduction. An $n \times n$ real matrix A is *monotone* if $Ax \geq 0$ implies $x \geq 0$. It was shown by Collatz [3] that A is monotone if and only if A is nonsingular and $A^{-1} \geq 0$. An important class of such matrices includes the M-matrices. Let $M \geq 0$ be an $n \times n$ matrix and α be a real number. Then a matrix A of the form

$$(1.1) \qquad\qquad A = \alpha I - M$$

is called an *M-matrix* if $\alpha > \rho(M)$, where $\rho(M)$ is the spectral radius of M. It is well known [4] that A of the form (1.1) is an M-matrix if and only if $A^{-1} \geq 0$; that is, A is monotone.

Monotone matrices have been studied extensively and the results have been applied to the study of finite difference approximations by authors such as Bramble and Hubbard [2], Collatz [4], Price [8], Varga [9], and Young [10]. Such systems frequently arise in the numerical solution of elliptic partial differential equations. An extensive theory has been developed which is important mathematically while at the same time it is of practical use. Thus far, these concepts have been applied only to the study of linear systems in which the coefficient matrix is square and nonsingular.

In [1], the study of singular matrices having monotone properties was initiated. Some conditions were given in order that a real $m \times n$ matrix A have a nonnegative Moore-Penrose generalized inverse A^+ [7]. The concept of monotonicity was extended to row-monotonicity and it was shown that if $A^+ \geq 0$ then A and A^T are row-monotone, although the converse of this statement does not hold in general.

The purpose of this paper is to extend some of the results in [1] and to apply these ideas to the study of linear systems whose coefficient matrix satisfies a generalized form of monotonicity. In Section 3, a new characterization of row-monotone matrices is given. Section 4 is concerned with an extension of the M-matrix concept

to rectangular matrices and with a characterization of matrices A of full column rank for which $A^+ \geqq 0$. It is then shown how these results yield a convergent splitting of certain rectangular matrices, leading to iterative techniques for solving rectangular linear systems.

2. Notation and Preliminaries.

R^n—the n-dimensional real space,

R^n_+—the nonnegative orthant, that is, the set $\{x = (x_i) \in R^n \mid x_i \geqq 0\}$,

$R^{m \times n}$—the $m \times n$ real matrices,

0—the zero matrix.

For $A \in R^{n \times n}$,

A^{-1}—the inverse, whenever it exists,

$\rho(A)$—the spectral radius.

For $A \in R^{m \times n}$,

A^T—the transpose,

$\Re(A)$—the range,

$\Re(A)$—the null space.

$A = (a_{i,j}) \geqq 0$ if and only if $a_{i,j} \geqq 0$,

A^+—the generalized inverse.

The generalized inverse A^+ is defined by

$A^+ y = x$ if $Ax = y$, $x \in \Re(A^T)$, and

$A^+ y = 0$ if $y \in \Re(A^T)$.

Some of the properties of A^+ that are used in the paper will now be summarized. If $A = HK$ where H has full column rank and K has full row rank, then $A^+ = K^+ H^+$. If A has full column rank, then $A^+ = (A^T A)^{-1} A^T$, and moreover $A^+ A = I_n$, the identity matrix of order n.

The following construction will be needed. For any $A \in R^{m \times n}$, with $m \geqq n$, there exist matrices $B \in R^{m \times n}$ and $G \in R^{n \times n}$ such that

$$A = BG$$

where B has full column rank, regardless of the rank, r, of A. To see this, let P be a permutation matrix of order n so that PA has r linearly independent rows among its first n rows. Suppose that

$$PA = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \\ \vdots \\ \alpha_m \end{bmatrix}, \quad \text{where } \alpha_i \text{ is the } i\text{th now of } PA,$$

and let

$$G = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}.$$

Let C denote the $(m - n) \times n$ matrix such that

$$CG = \begin{bmatrix} \alpha_{n+1} \\ \vdots \\ \alpha_m \end{bmatrix}$$

and finally take PB to be the $m \times n$ matrix

$$PB = \begin{bmatrix} I_n \\ C \end{bmatrix}.$$

Then $A = BG$ where B has rank n. A similar method for constructing B and G was described in [5]. Notice that

$$B^+ A = G, \qquad \Re(A) \subseteq \Re(B) \quad \text{and} \quad \Re(A^T) = \Re(G^T).$$

This factorization is clearly not unique. Now, in case A is of order n, B will be chosen to be the identity matrix.

3. Monotonicity. A matrix $A \in R^{m \times n}$ is said to be *row-monotone* if

$$(3.1) \qquad Ax \geqq 0, \qquad x \in \Re(A^T), \quad \text{implies } x \geqq 0.$$

If $m = n$ and A is a nonsingular matrix satisfying (3.1), then A is monotone in the usual sense and $A^{-1} \geqq 0$. Moreover, if A is rectangular with full column rank, then A is monotone in the sense discussed by Mangasarian [6], who showed that in this case (3.1) is equivalent to the existence of a nonnegative left inverse of A. The general case was considered in [1], where it was shown that (3.1) is equivalent to

$$(3.2) \qquad YA = A^+ A \quad \text{for some } Y \geqq 0.$$

The purpose of this section is to provide an alternate characterization of (3.1). As opposed to the nonsingular case, (3.1) is not equivalent to $A^+ \geqq 0$, although A satisfies (3.1) whenever $A^+ \geqq 0$.

Let E be any orthogonal projection in $R^{m \times n}$. That is $E^2 = E = E^T$. Then, $E^+ = E$ can be expressed as $E^+ = I + (E - I)$, where, of course, $I \geqq 0$ and $(E - I)E = 0$. The following theorem shows that all row-monotone matrices share this property.

THEOREM 1. *The matrix $A \in R^{m \times n}$ is row-monotone if and only if there exist matrices B and C such that*

$$(3.3) \qquad A^+ = B + C, \qquad B \geqq 0 \quad and \quad CA = 0.$$

Proof. If A is row-monotone, then from (3.2) there is some $Y \geqq 0$ such that $YA = A^+ A$. Then $A^+ = Y + (A^+ - Y)$ and taking $B = Y$ and $C = A^+ - Y$, $B \geqq 0$ and $CA = 0$.

Conversely, if A satisfies (3.3), then $A^+ A = BA + CA = BA$ so that A is row-monotone from (3.2), since $B \geqq 0$. $\square$

COROLLARY 1. *Let $B, C \in R^{m \times n}$ where $B \geqq 0$ and $BC^T = -CC^T$. Then, $(B + C)^+$ is row-monotone. Conversely, every row-monotone matrix is obtainable in this way.*

Proof. If $BC^T = -CC^T$ then $C(B^T + C^T) = 0$ so that $C(B + C)^+ = 0$, since

$$\Re(B^T + C^T) = \Re[(B + C)^T] = \Re[(B + C)^+].$$

The proof then follows from Theorem 1 by taking $A = (B + C)^+$. □

The section is concluded by considering the special case where A is *EPr*, that is, where $m = n$ and $\Re(A) = \Re(A^T)$.

COROLLARY 2. *If* $A \in R^{n \times n}$ *is EPr and* A *is row-monotone, then there exists a matrix* Y *such that* $A^+ = YA$, $Y \geqq 0$.

Proof. Premultiplying (3.3) by B and postmultiplying by A yields

$$A^+ = A^+AA^+ = BAA^+ = BA^+A = B^2A + BCA = B^2A,$$

since $AA^+ = A^+A$. Then $Y = B^2 \geqq 0$. □

4. Rectangular Splittings. Monotone matrices arise very naturally as coefficient matrices for linear systems in connection with finite difference methods for solving certain boundary value problems. In particular, the *M*-matrix concept and regular splittings [9], [10] have played a fundamental role in the theory. Research papers thus far have considered only the case where resulting linear systems are square and nonsingular. In this section, some of these concepts are extended to the rectangular case by considering conditions for which the generalized inverse is nonnegative.

In attempting to extend the concept of an *M*-matrix to the rectangular case, one might try relaxing the requirement that $\alpha > \rho(M)$ in (1.1). However, this approach is somewhat unsatisfactory since it can be shown that if M is irreducible and $\alpha \leqq \rho(M)$, then each row and each column of A^+, where $A = \alpha I - M$, contains negative entries. In what follows, the factorization $M = BG$, where $M \in R^{m \times n}$ and where $B \in R^{m \times n}$ has full column rank, is used to consider a rectangular form of an *M*-matrix. The treatment follows that of Varga [9, Section 3.5].

THEOREM 2. *Let* $M \in R^{m \times n}$, $m \geqq n$, *and let* $M = BG$ *where* $B \in R^{m \times n}$ *has rank* n. *If* $\rho(G) < 1$, *then* $B - M$ *has full column rank and*

$$(4.1) \qquad (B - M)^+ = B^+ + GB^+ + G^2B^+ + \cdots ,$$

the series on the right converging. Conversely, if the series on the right converges, then $\rho(G) < 1$.

Proof. Assume $\rho(G) < 1$. Then by [9, Theorem 3.7], $I - G$ is nonsingular and

$$(4.2) \qquad (I - G)^{-1} = I + G + G^2 + \cdots .$$

Then $B - M = B(I - G)$ has full column rank and since B has full column rank, $(B - M)^+ = (I - G)^{-1}B^+$, yielding (4.1).

For the converse, note that since B^+ has full row rank, the convergence of the series on the right in (4.1) implies the convergence of the series on the right in (4.2). Thus, $\rho(G) < 1$. □

THEOREM 3. *Let* $M \in R^{m \times n}$, $m \geqq n$, *and let* $M = BG$, *where* $B \in R^{m \times n}$ *has rank* n. *Assume that* $M \geqq 0$ *and* $B^+ \geqq 0$ *and let* α *be a real number. Then the following statements are equivalent.*

(1) $\alpha > \rho(G)$.
(2) $(\alpha B - M)^+ = \sum_{i=0}^{\infty} (1/\alpha)G^iB^+ \geqq 0$.

Proof. Notice that $\alpha B - M = B(\alpha I - G)$ and that $G = B^+M \geqq 0$.
If (1) holds, then, by [9, Theorem 3.8], $\alpha I - G$ is nonsingular and

$$(\alpha I - G)^{-1} = \sum_{i=0}^{\infty} \frac{1}{\alpha} G^i \geqq 0.$$

Consequently, since B has full column rank and $B^+ \geqq 0$,

$$(\alpha B - M)^+ = (\alpha I - G)^{-1} B^+ = \sum_{i=0}^{\infty} \frac{1}{\alpha} G^i B^+ \geqq 0.$$

If (2) holds, then, since B^+ has full row rank, $\sum_{i=0}^{\infty} (1/\alpha) G^i$ converges to $(\alpha I - G)^{-1}$. Then since $G \geqq 0$, $(\alpha I - G)^{-1} \geqq 0$ so that $\alpha > \rho(G)$ by [9, Theorem 3.8]. $\square$

In view of Theorem 3, a matrix $A \in R^{m \times n}$, $m \geqq n$, that can be expressed in the form

(4.3) $A = \alpha B - M, \qquad M = BG, \qquad B \in R^{m \times n}$ of rank n and $B^+ \geqq 0$

will be called a *rectangular M-matrix* whenever $\alpha > \rho(G)$. Then a rectangular M-matrix A has full column rank and moreover $A^+ \geqq 0$. Now if $m = n$ and B is chosen to be I, then the definition reduces to the usual one for square matrices.

The preceding theorem determines a class of matrices A having full column rank for which $A^+ \geqq 0$. In [1, Theorem 2] it was shown that in general $A^+ \geqq 0$ if and only if

(4.4) $Ax \in [R_+^m + \mathfrak{N}(A^T)], \qquad x \in \mathfrak{R}(A^T), \quad \text{implies } x \geqq 0.$

Then if A has full column rank, (4.4) reduces to

(4.5) $Ax \in [R_+^m + \mathfrak{N}(A^T)] \quad \text{implies } x \geqq 0.$

However, the class of matrices satisfying (4.5) is not easily recognized and no general structure theorem for them has previously been given. Thus the following extension of a result by H. S. Price [8, Theorem 2.2] could prove useful.

THEOREM 4. *Let $A \in R^{m \times n}$ and suppose A has full column rank. Then $A^+ \geqq 0$ if and only if there exists a matrix N such that*
  1. $M = A + N$ has $\mathfrak{R}(N) \subseteq \mathfrak{R}(M)$ and $M^+ \geqq 0$,
  2. $M^+ N \geqq 0$,
  3. $\rho(M^+ N) < 1$.

Proof. The proof parallels that given for Theorem 2.2 in [8]. If $A^+ \geqq 0$ then N can be chosen to be the null matrix 0 so that the above properties are trivially satisfied.

If A, N and M satisfy 1, 2 and 3 then $MM^+ N = N$ since $\mathfrak{R}(N) \subseteq \mathfrak{R}(M)$ and since MM^+ is a projection on $\mathfrak{R}(M)$. Moreover $\mathfrak{R}(A) \subseteq \mathfrak{R}(M)$, for if $y = Ax$ then $y = Mx - Nx \in \mathfrak{R}(M)$, so that $\mathfrak{R}(A) = \mathfrak{R}(M)$ since A has full column rank. Then

$$A = M - N = M(I - M^+ N).$$

Now since $M^+ N \geqq 0$ and $\rho(M^+ N) < 1$, $I - M^+ N$ is an M-matrix so that $(I - M^+ N)^{-1} \geqq 0$. But M has full column rank and $M^+ \geqq 0$ and thus

$$A^+ = [M(I - M^+ N)]^+ = (I - M^+ N)^{-1} M^+ \geqq 0. \quad \square$$

The paper is concluded by showing how these concepts can be applied to the study of iterative solutions of systems of linear equations. Consider the problem of finding an n-vector x which satisfies

(4.6) $Ax = b,$

where $A \in R^{m \times n}$ has full column rank n. If $b \in \mathfrak{R}(A)$, then (4.6) has a unique solution, otherwise the system is inconsistent. In any case, $A^+ b$ provides the best least squares

approximate solution to (4.6). That is, of all vectors x which minimize $||Ax - b||$, A^+b has the smallest $||x||^2$.

For the case where $m = n$ so that A is nonsingular, many iterative techniques for solving this problem such as the point Jacobi, Gauss-Seidel and successive over-relaxation methods [9], [10] can be obtained by splitting the matrix into the difference of two $n \times n$ real matrices M and N where M is nonsingular. If $A = M - N$, then the iteration

$$x_{i+1} = M^{-1}Nx_i + M^{-1}b$$

converges whenever $\rho(M^{-1}N) < 1$.

Now the splittings discussed in Theorems 3 and 4 hold for rectangular matrices and thus can be applied to iterative solutions to (4.6) in the general case. Notice that if A is a rectangular M-matrix then since A has full column rank and $A^+ \geq 0$, the splitting for A given in (4.3) is a special case of the splitting provided by Theorem 4.

THEOREM 5. *Consider the rectangular system (4.6). If $A^+ \geq 0$ so that A can be expressed in the form $A = M - N$ where A, M and N satisfy 1, 2 and 3 of Theorem 4, then the iteration*

(4.7) $$x_{i+1} = M^+Nx_i + M^+b$$

converges to the best least squares solution A^+b to (4.6), for any x_0.

Proof. By the proof of Theorem 4, M has full column rank and since $\rho(M^+N) < 1$ by 2, the iteration (4.7) converges to some vector y. Then $(I - M^+N)y = M^+b$ so that $(M - MM^+N)y = MM^+b$. But $\Re(N) \subseteq \Re(M)$ by 1, so that $MM^+N = N$. Also since $\Re(M) = \Re(A)$, $MM^+ = AA^+$ and thus

$$Ay = (M - N)y = (M - MM^+N)y = MM^+b = AA^+b.$$

Then $y = A^+Ay = A^+AA^+b = A^+b$. □

Mathematics and Computer Science Departments
The University of Tennessee
Knoxville, Tennessee 37916

1. A. BERMAN & R. PLEMMONS, "Monotonicity and the generalized inverse," *SIAM J. Appl. Math.*, v. 22, 1972, pp. 155–161.

2. J. H. BRAMBLE & B. E. HUBBARD, "New monotone type approximations for elliptic problems," *Math. Comp.*, v. 18, 1964, pp. 349–367. MR **29** #2982.

3. L. COLLATZ, "Aufgaben monotoner Art," *Arch. Math.*, v. 3, 1952, pp. 366–376. MR **14**, 799.

4. L. COLLATZ, *Funktionalanalysis und numerische Mathematik*, Die Grundlehren der math. Wissenschaften, Band 120, Springer-Verlag, Berlin, 1964; English transl., Academic Press, New York, 1966. MR **29** #2931; MR **34** #4961.

5. V. N. JOSHI, "A note on the solution of rectangular linear systems by iteration," *SIAM Rev.*, v. 12, 1970, pp. 463–466. MR **41** #7831.

6. O. L. MANGASARIAN, "Characterizations of real matrices of monotone kind," *SIAM Rev.*, v. 10, 1968, pp. 439–441. MR **38** #5818.

7. R. PENROSE, "A generalized inverse for matrices," *Proc. Cambridge Philos. Soc.*, v. 51, 1955, pp. 406–413. MR **16**, 1082.

8. H. S. PRICE, "Monotone and oscillation matrices applied to finite difference approximations," *Math. Comp.*, v. 22, 1968, pp. 489–516. MR **38** #875.

9. R. S. VARGA, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, N. J., 1962. MR **28** #1725.

10. D. M. YOUNG, *Iterative Solutions to Large Linear Systems*, Academic Press, New York, 1971.