A Test for Strict Sign-Regularity

M. Gasca* and J. M. Peña*
Depto. Matemática Aplicada
University of Zaragoza, Spain

Submitted by Jose A. Dias da Silva

ABSTRACT

A characterization of strict sign-regular matrices is obtained. It is given in an algorithmic form which allows the easy construction of a test to check if a matrix is strictly sign-regular. The test is based on the Neville elimination of several submatrices of A and requires less computational work than those which can be derived from other previously known characterizations of these matrices.

1. INTRODUCTION AND NOTATION

An $n \times n$ matrix A is said to be *sign-regular* if for each $1 \leqslant k \leqslant n$ all its minors of order k have the same sign (in the sense that the product of any two of them is greater than or equal to zero). The matrix is called *strictly sign-regular* (SSR for brevity) if for each $1 \leqslant k \leqslant n$ all its minors of order k are different from zero and have the same sign. *Totally positive* (*strictly totally positive*) matrices are matrices with all their minors greater than or equal to zero (greater than zero).

Sign regularity was studied in [12] as an extension of the theory of total positivity, which has important applications in many scientific fields. Sign-regular matrices were also studied for instance in [1]. One of the interesting aspects of them is that they are characterized by some variation-diminishing properties which are useful for shape-preserving representations in computer-aided geometric design (see [10, 11, 2]). In particular, a characterization of this type for SSR matrices is given by Theorem 5.3 of [1]. In the last

*Partially supported by Research grant DGICYT PS900121.

655 Avenue of the Americas, New York, NY 10010 0024-3795/94/$7.00

years [6–9] we have obtained different characterizations of totally positive and strictly totally positive (STP) matrices, which have improved several previous results (see [1, 3]). Some of our results lead to tests to check the total positivity or strict total positivity of the matrix, and in general are based on the use of the so-called Neville elimination. This elimination process was described in detail in [6] and will be briefly recalled at the end of this section. A different approach to SSR matrices in terms of matrix intervals was obtained in [5].

On the other hand, Theorem 2.5 of [1] states that only the signs of the minors with consecutive rows and columns of a matrix are needed to decide whether or not the matrix is SSR. Since this is an extension of an old result [4, 12] for STP matrices which was recently improved in Theorem 4.3 of [6], the question arises if a similar improvement can be obtained for strict sign regularity. As we shall see in Section 2, the answer is negative, but we obtain there some necessary conditions for the strict sign regularity of a matrix. Moreover, in Section 4 (Theorem 4.1) we give some conditions which are necessary and sufficient. The theorem is given in an algorithmic form that leads to the easy construction of a test to check the strict sign regularity of a matrix. In preparation, Section 3 is used to study some particular cases of SSR matrices (mainly STP and strictly totally negative matrices) which allow one to simplify the test. The computational cost of the test is considerably lower than those which can be derived from Theorem 2.5 of [1].

Our notation follows, in essence, that of [1], [6], and [9]. Given $k, n \in \mathbb{N}$, $1 \leqslant k \leqslant n$, $Q_{k,n}$ will denote the set of all increasing sequences of k natural numbers less than or equal to n.

Let A be a real square matrix of order n. For $k \leqslant n$, $l \leqslant n$, and for any $\alpha \in Q_{k,n}$ and $\beta \in Q_{l,n}$, we denote by $A[\alpha | \beta]$ the $k \times l$ submatrix of A containing rows numbered by α and columns numbered by β. For brevity we shall write $A[\alpha] := A[\alpha | \alpha]$.

The following definitions will be useful in the sequel. A *row-initial* (respectively, *column-initial*) submatrix of A is the submatrix formed by consecutive initial rows (columns) and consecutive columns (rows). The determinant of a row-initial or a column-initial submatrix is called *initial minor*. A lower (respectively upper) triangular matrix is said to be ΔSTP if all minors $\det A[\alpha | \beta]$, with $\alpha, \beta \in Q_{k,n}$ and $\alpha_i \geqslant \beta_i$ (respectively, $\alpha_i \leqslant \beta_i$) for all i, are positive (all the other minors are trivially zero).

Neville elimination (NE) is a procedure to create zeros in a matrix by means of adding to a given row a suitable multiple of the previous one. For a nonsingular matrix $A = (a_{ij})_{1 \leqslant i, j \leqslant n}$, it consists of $n - 1$ major steps resulting in a sequence of matrices as follows:

$$A := A_1 \to A_2 \to \cdots \to A_n,$$

where $A_t = (a_{ij}^{(t)})_{1 \leqslant i, j \leqslant n}$ has zeros below its main diagonal in the $t-1$ first columns. The matrix A_{t+1} is obtained from A_t $(t = 1, \dots, n)$ according to the formula

$$a_{ij}^{(t+1)} := \begin{cases} a_{ij}^{(t)} & \text{if} \quad i \leqslant t, \\ a_{ij}^{(t)} - \left(a_{it}^{(t)}/a_{i-1,t}^{(t)}\right)a_{i-1,j}^{(t)} & \text{if} \quad i \geqslant t+1 \text{ and } j \geqslant t+1, \\ 0 & \text{otherwise.} \end{cases} \tag{1.1}$$

In this process the element

$$p_{ij} := a_{ij}^{(j)}, \qquad 1 \leqslant j \leqslant n, \quad j \leqslant i \leqslant n, \tag{1.2}$$

is called the (i, j) *pivot* of the NE of A. The process will break down if any of the pivots p_{ij} $(j \leqslant i < n)$ is zero. In that case we can move the corresponding rows to the bottom and proceed with the new matrix, as described in [6]. Thus we have:

REMARK 1.1. The NE of a matrix can be performed without row exchanges if all the pivots are nonzero.

The pivots p_{ii} will be referred to as *diagonal pivots*. If all the pivots p_{ij} are nonzero, then $p_{i1} = a_{i1}$ $\forall i$ and, by Lemma 2.6(1) of [6],

$$p_{ij} = \frac{\det A[i-j+1, \dots, i|1, \dots, j]}{\det A[i-j+1, \dots, i-1|1, \dots, j-1]} \qquad (1 < j \leqslant i \leqslant n). \tag{1.3}$$

The element

$$m_{ij} = p_{ij}/p_{i-1,j}, \qquad 1 \leqslant j \leqslant n, \quad j < i \leqslant n, \tag{1.4}$$

is called the (i, j) *multiplier* of the NE of A.

The matrix $U := A_n$ is upper triangular and has the diagonal pivots on its main diagonal. The *complete Neville elimination* (CNE) of a nonsingular matrix A consists in performing the NE of A until getting the upper triangular matrix U and, afterwards, proceeding with the NE of U^T (the transpose of U) until one obtains a diagonal matrix with the diagonal pivots on its main diagonal. When we say that the CNE of A is possible without row or column exchanges, we mean that there have not been any row exchanges in the NE of A or U^T. Finally, the (i, j) pivot of the CNE of A is the (i, j) pivot of the NE of A if $i \geqslant j$ and the (j, i) pivot of the NE of U^T if $i \leqslant j$. The multipliers of the CNE of A can be defined analogously.

2. NECESSARY CONDITIONS FOR STRICT SIGN REGULARITY

The following theorem characterizes a class of matrices that contains the SSR matrices. It will provide necessary conditions for the strict sign regularity of a square matrix.

THEOREM 2.1. *Let A be a nonsingular $n \times n$ matrix. Then the following conditions are equivalent*:

(i) *The initial minors of order k of A are nonzero and have the same sign for each $1 \leqslant k \leqslant n$.*

(ii) *The CNE of A can be performed without row or column exchanges, with positive multipliers and nonzero diagonal pivots.*

(iii) *A can be decomposed in the form $A = LDU$ with D a diagonal nonsingular matrix and L (respectively, U) ΔSTP and lower (upper) triangular with unit diagonal.*

Proof. First we show that (i) $\Rightarrow$ (ii). By Lemma 2.6 of [6] (see Remark 1.1 above), the NE of A can be performed without row exchanges; the pivots p_{ij} are nonzero and have the same sign for every j and $i = j, j+1, \dots, n$. Consequently, by (1.4) the multipliers are all positive.

If we denote by V the upper triangular matrix obtained from A by Neville elimination, then since the row-initial minors of A coincide with the corresponding minors of V, we can apply to the NE of V^T the same reasoning as above to obtain (ii).

To see that (ii) $\Rightarrow$ (iii) we can just apply Theorem 3.1 of [9] to obtain the LDU factorization of A, in which L and U are ΔSTP by Theorems 2.2 and 4.3 of [9].

Finally let us prove that (iii) $\Rightarrow$ (i). It is easy to check that

$$\begin{aligned} \det A[i, i+1, \dots, i+k-1 | 1, 2, \dots, k] \\ = \det L[i, i+1, \dots, i+k-1 | 1, 2, \dots, k] \\ \times \det D[1, \dots, k] \det U[1, \dots, k]. \quad (2.1) \end{aligned}$$

Since $\det L[i, i+1, \ldots, i+k-1 | 1, 2, \ldots, k] > 0$ because L is ΔSTP and $\det U[1, \ldots, k] = 1$, we deduce from (2.1) that all the column-initial minors of A are nonzero and for each k have the same sign (the sign of $\det D[1, \ldots, k] = p_{11} \cdots p_{kk}$).

An analogous conclusion can be obtained for the row-initial minors of A from

$$\begin{aligned} &\det A[1, 2, \ldots, k | i, i+1, \ldots, i+k-1] \\ &\quad = \det L[1, \ldots, k] \det D[1, \ldots, k] \\ &\qquad \times \det U[1, 2, \ldots, k | i, i+1, \ldots, i+k-1]. \quad \blacksquare \end{aligned}$$

REMARK 2.2. Let us observe that, as we have seen in this proof, the sign of the minors of order k in condition (i) coincides with the sign of the product of the first k diagonal pivots of A.

REMARK 2.3. Condition (i) of Theorem 2.1 [and consequently the equivalent conditions (ii) or (iii)] is a necessary condition for a square matrix A to be SSR, but it is not sufficient. For example, the matrices

$$A = \begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 4 & 3 \end{pmatrix}$$

satisfy condition (i) and are not SSR.

3. SPECIAL CLASSES OF SSR MATRICES

Theorem 2.1 gives a necessary (not sufficient) condition for strict sign regularity. However, for some classes of matrices those conditions are also sufficient, as we shall see in this section. STP matrices play an important role in many fields (approximation theory, economics, mechanics, etc.). Taking into account Theorem 4.3 of [6], the positivity of all its initial minors is a necessary and sufficient condition for a matrix to be STP. Hence, for this type of matrices we have the following proposition.

PROPOSITION 3.1. *Let A be a nonsingular matrix. Then the following conditions are equivalent*:

(i) *A is STP.*

(ii) *The initial minors of A are positive.*

(iii) *The CNE of A can be performed without row or column exchanges with positive multipliers and positive diagonal pivots.*

(iv) *A can be decomposed as $A = LDU$ with L (U) a lower (upper) ΔSTP matrix and D a diagonal matrix with positive entries on its main diagonal.*

REMARK 3.2. The proposition can be used to characterize in a similar way the negatives of STP matrices. For a nonsingular $n \times n$ matrix A the following properties are equivalent:

(i) $-A$ is STP.

(ii) For each $1 \leqslant r \leqslant n$, the initial minors of order r of A have the sign of $(-1)^r$.

(iii) The CNE of A can be performed without row or column exchanges with positive multipliers and negative diagonal pivots.

(iv) A can be decomposed as $A = LDU$ with L (U) a lower (upper) ΔSTP matrix and D a diagonal matrix with negative entries on its main diagonal.

Let us now consider matrices with all their minors negative. They will be referred to as *strictly totally negative* (STN) matrices. The main reason to consider them here is that STP and STN matrices play an interesting role in simplifying the test for strict sign regularity which will be obtained in the next section.

First we need some auxiliary results and notation.

LEMMA 3.3. *Let J be the diagonal matrix $J := \operatorname{diag}\{1, -1, 1, \dots, (-1)^{n-1}\}$, and A a nonsingular $n \times n$ matrix. The matrix A is SSR (or ΔSTP) if and only if $JA^{-1}J$ is.*

The lemma is a direct consequence of (1.32) of [1], which gives the value of $\det(JA^{-1}J)[\alpha|\beta]$ for any $\alpha, \beta \in Q_{k,n}$.

The *conversion* of the matrix $A = (a_{ij})_{1 \leqslant i,j \leqslant n}$ [1, 9] is the matrix $A^{\#}$ of order n whose (i, j) entry is $a_{n-i+1, n-j+1}$. It is straightforward to see that $(AB)^{\#} = A^{\#}B^{\#}$ and $(A^{\#})^{-1} = (A^{-1})^{\#}$. The following theorem characterizes the class of STN matrices.

THEOREM 3.4. *Let A be a nonsingular $n \times n$ matrix. Then A is STN if and only if its (n, n) entry is negative and the CNE of A can be performed without row or column exchanges, with positive multipliers, and with diago-*

nal pivots p_{ii} *verifying*

$$p_{11} < 0, \qquad p_{ii} > 0 \quad \forall i > 1. \tag{3.1}$$

Proof. The necessity of the condition is a consequence of Theorem 2.1 and (1.3). Let us prove that it is also sufficient. By Theorem 2.1 the initial minors of A (one of them is det A) are negative. Therefore we have only to see that A is SSR. By Lemma 3.3, A is SSR if and only if so is $JA^{-1}J$. Since obviously a matrix A is SSR if and only if $A^{\#}$ is, the strict sign regularity of A follows from that of

$$B := (JA^{-1}J)^{\#} = J^{\#}(A^{-1})^{\#}J^{\#} = J(A^{-1})^{\#}J. \tag{3.2}$$

In fact, we shall prove that any minor of the submatrices $B[1, 2, \dots, n-1]$, $B[2, 3, \dots, n|1, 2, \dots, n-1]$, $B[1,2, \dots, n-1|2, 3, \dots, n]$, $B[2, 3, \dots, n]$ (that is, any minor of order not greater than $n-1$ of B) is positive. The only minor of order n, det $B = 1/\det A$, is negative.

Again by Theorem 2.1, A can be written in the form $A = LDU$ with L, U ΔSTP matrices and D the diagonal matrix $D = \operatorname{diag}\{p_{11}, p_{22}, \dots, p_{nn}\}$. Since $(A^{-1})^{\#} = (U^{-1})^{\#}(D^{-1})^{\#}(L^{-1})^{\#}$ and $J = J^{-1}$, one has

$$B = \{J(U^{-1})^{\#}J\}\{J(D^{-1})^{\#}J\}\{J(L^{-1})^{\#}J\}.$$

The matrix $\overline{L} := J(U^{-1})^{\#}J = J(U^{\#})^{-1}J$ $[\overline{U} := J(L^{-1})^{\#}J]$ is lower [upper] triangular with unit diagonal and, by Lemma 3.3, ΔSTP. The matrix $\overline{D} := J(D^{-1})^{\#}J$ is diagonal, with negative (n, n) entry and all the other diagonal entries positive. Now we have

$$B[1,2,\dots,n-1] = \overline{L}[1,2,\dots,n-1]\overline{D}[1,2,\dots,n-1] \times \overline{U}[1,2,\dots,n-1],$$

which, by Proposition 3.1, is STP.

In consequence, all the row-initial minors of order not greater than $n-2$ of $B[2, 3, \dots, n|1, 2, \dots, n-1]$ are positive. The column-initial minors of any order of this matrix (one of them being the whole determinant) are also column-initial minors of $B = \overline{L}\overline{D}\overline{U}$, which, by Theorem 2.1, are positive. Hence, by Proposition 3.1, the matrix $B[2, 3, \dots, n|1, 2, \dots, n-1]$ is STP. By similar reasoning, so is the matrix $B[1, 2, \dots, n-1|2, 3, \dots, n]$.

The initial minors of order not greater than $n-2$ of $B[2,3,\ldots,n]$ (which are minors of $B[2,3,\ldots,n|1,2,\ldots,n-1]$ or $B[1,2,\ldots,n-1|2,3,\ldots,n]$) are also positive. For the only initial minor of order $n-1$, which is $\det B[2,3,\ldots,n]$, we have

$$\det B[2,3,\ldots,n] = \det (JA^{-1}J)^{\#}[2,3,\ldots,n]$$
$$= \det (JA^{-1}J)[1,2,\ldots,n-1]$$

and, by (1.32) of [1],

$$\det (JA^{-1}J)[1,2,\ldots,n-1] = \frac{\det A[n]}{\det A} > 0.$$

Therefore, the four submatrices $B[1,2,\ldots,n-1]$, $B[2,3,\ldots,n|1,2,\ldots,n-1]$, $B[1,2,\ldots,n-1|2,3,\ldots,n]$, and $B[2,3,\ldots,n]$ are STP, B is SSR, and A is STN. ■

REMARK 3.5. The matrix

$$\begin{pmatrix} -1 & -1 \\ -2 & 3 \end{pmatrix}$$

shows that the condition $a_{nn} < 0$ cannot be suppressed in Theorem 3.4.

REMARK 3.6. By Theorem 2.1 and (1.3), Theorem 3.4 can be reformulated in a similar way to Proposition 3.1:

Let $A = (a_{ij})_{1 \leqslant i,j \leqslant n}$ be a nonsingular matrix with $a_{nn} < 0$. Then the following properties are equivalent:

(i) A is STN.

(ii) The initial minors of A are negative.

(iii) The CNE of A can be performed without row or column exchanges with positive multipliers and with diagonal pivots verifying

$$p_{11} < 0, \qquad p_{ii} > 0 \quad \forall i > 1. \tag{3.3}$$

(iv) A can be decomposed $A = LDU$ with L (U) a lower (upper) ΔSTP matrix and D a diagonal matrix whose entries p_{ii} on its main diagonal satisfy (3.3).

Now, as in Remark 3.2, a similar characterization can be obtained for the negatives of STN matrices. The sign of their initial minors of order r will be that of $(-1)^{r+1}$, and the diagonal pivots will satisfy

$$p_{11} > 0, \qquad p_{ii} < 0 \quad \forall i > 1.$$

4. A TEST FOR STRICT SIGN REGULARITY

In this section it will be useful to denote $A^{(l)} := A[l, l+1, \dots, n]$. Let us also denote by $p_{11}^{(l)}, \dots, p_{n-l+1,n-l+1}^{(l)}$ the diagonal pivots of $A^{(l)}$, and by sg(a) the sign of a nonzero number a. The following theorem provides a test to check the strict sign regularity of a matrix A.

THEOREM 4.1. *Let A be a nonsingular $n \times n$ matrix. Then A is SSR if and only if the following conditions hold*:

(1) *The CNE of $A^{(1)} = A$ can be performed without row or column exchanges, with positive multipliers, and with nonzero diagonal pivots* $p_{11}, \dots, p_{nn}$.

(2) $\text{sg}(a_{nn}) = \text{sg}(p_{11})$.

(3) *Let r be the integer such that* $\text{sg}(p_{11}) = \cdots = \text{sg}(p_{rr}) \neq \text{sg}(p_{r+1,r+1})$ *($r := n$ if $\text{sg}(p_{ii}) = \text{sg}(p_{11})\ \forall i$). If $r = 1$, let s be the integer such that* $\text{sg}(p_{11}) \neq \text{sg}(p_{22}) = \cdots = \text{sg}(p_{s+1,s+1}) \neq \text{sg}(p_{s+2,s+2})$ *($s := n - 1$ if $\text{sg}(p_{ii}) = \text{sg}(p_{22})\ \forall i \geqslant 2$). If $r > 1$ ($r = 1$), then for each $2 \leqslant l \leqslant n - r + 1$ ($2 \leqslant l \leqslant n - s$) the CNE of $A^{(l)}$ can be performed without row or column exchanges, with positive multipliers, and with nonzero diagonal pivots $p_{ii}^{(l)}$ ($1 \leqslant i \leqslant n - l + 1$) having signs*

$$\text{sg}(p_{11}^{(l)}) = \text{sg}(p_{11}), \dots, \text{sg}(p_{n-l+1,n-l+1}^{(l)}) = \text{sg}(p_{n-l+1,n-l+1}). \quad (4.1)$$

Proof. If A is SSR, then (1) holds by Theorem 2.1, and (2) is obvious. On the other hand, the submatrices $A^{(l)}$ are also SSR and, again by Theorem 2.1 [taking into account (1.3) to derive (4.1)], (3) follows.

Conversely, assume that (1), (2), and (3) hold. By Theorem 2.1 all the *initial* minors of order k ($1 \leqslant k \leqslant n - l + 1$) of $A^{(l)}$ ($1 \leqslant l \leqslant n - r + 1$ if $r > 1$, $1 \leqslant l \leqslant n - s$ otherwise) have the same sign [$\text{sg}(p_{11}^{(l)} \cdots p_{kk}^{(l)})$ by Remark 2.2], which by (4.1) is $\text{sg}(p_{11} \cdots p_{kk})$. In the case $r > 1$ this implies, for $l + n - r + 1$, that either $A^{(n-r+1)}$ or $-A^{(n-r+1)}$ is STP, by Proposition 3.1 or Remark 3.2. Therefore, the sign of any minor of order k ($1 \leqslant k \leqslant r$) of $A^{(n-r+1)}$ coincides with $\text{sg}(p_{11} \cdots p_{kk})$. On the other hand, any minor of order k of A with consecutive rows and columns is a row-initial or a column-initial minor when it is considered as a minor of some matrix $A^{(h)}$ with $1 \leqslant h \leqslant n$. If $h \geqslant n - r + 1$, then the sign of that minor is $\text{sg}(p_{11} \cdots p_{kk})$, since it is a minor of order k of $A^{(n-r+1)}$. If $h \leqslant n - r$, its sign is also $\text{sg}(p_{11} \cdots p_{kk})$ by the beginning of this reasoning.

The case $r = 1$ is quite similar, with $A^{(n-r+1)}$ replaced by $A^{(n-s)}$, and using Remark 3.6 to prove that either $A^{(n-s)}$ or $-A^{(n-s)}$ is STN. For it, we can take into account that the $(s+1) \times (s+1)$ matrix $A^{(n-s)}$ has a_{nn} as $(s+1, s+1)$ entry and that, by (4.1), $\text{sg}(a_{nn}) = \text{sg}(p_{11}) = \text{sg}(P_{11}^{(n-s)})$.

Now the strict sign regularity of A follows from Theorem 2.5 of [1]. ■

REMARK 4.2. The sequence of the signs of the minors, that is, the signature sequence of the SSR matrix A (see [1]), becomes apparent as soon as we obtain the diagonal pivots of A and apply (1.3).

REFERENCES

1 T. Ando, Totally positive matrices, *Linear Algebra Appl.* 90.165–219 (1987).

2 J. M. Carnicer and J. M. Peña, Shape preserving representations and optimality of the Bernstein basis, *Adv. Comput. Math.*, 1:173–196 (1993).

3 C. Cryer, *LU*-factorization of totally positive matrices, *Linear Algebra Appl.* 7:83–92 (1973).

4 M. Fekete, Über ein Problem von Laguerre, *Rend. Circ. Mat. Palermo* 34:89–100, 110–120 (1913).

5 J. Garloff, Criteria for sign-regularity of sets of matrices, *Linear Algebra Appl.* 15:1–25 (1982).

6 M. Gasca and J. M. Peña, Total positivity and Neville elimination, *Linear Algebra Appl.* 165:25–44 (1992).

7 M. Gasca and J. M. Peña, On the characterization of totally positive matrices, in *Approximation Theory, Spline Functions and Applications* (S. P. Singh, Ed.), Kluwer Academic, Dordrecht, 1992, pp. 357–364.

8 M. Gasca and J. M. Peña, Total positivity, *QR*-factorization and Neville elimination, *SIAM J. Matrix Anal.*, to appear.

9 M. Gasca and J. M. Peña, A matricial description of Neville elimination, with applications to total positivity, *Linear Algebra Appl.*, to appear.

10 T. N. T. Goodman, Shape preserving representations, in *Mathematical Methods in CAGD* (T. Lyche and L. L. Schumaker, Eds.), Academic, Boston, 1989, pp. 333–357.

11 T. N. T. Goodman and H. B. Said, Shape preserving properties of the generalized Ball basis, *Comput. Aided Geom. Design* 8:115–121 (1991).

12 S. Karlin, *Total Positivity*, Vol. I, Stanford U.P., 1968.

Received 26 October 1992; final manuscript accepted 3 March 1993