ELSEVIER

Journal of Computational and Applied Mathematics 122 (2000) 37–50

JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

www.elsevier.nl/locate/cam

Elimination techniques: from extrapolation to totally positive matrices and CAGD

M. Gasca *, G. Mühlbach

Department of Applied Mathematics, University of Zaragoza, 5009 Zaragoza, Spain

Received 20 May 1999; received in revised form 22 September 1999

Abstract

In this survey, we will show some connections between several mathematical problems such as extrapolation, linear systems, totally positive matrices and computer-aided geometric design, with elimination techniques as the common tool to deal with all of them.

1. Introduction

Matrix elimination techniques are basic tools in many mathematical problems. In this paper we will show their crucial role in some results that various authors with us have obtained in two problems apparently distant: extrapolation and computer-aided geometric design (CAGD). A brief overview of how things were developed over time will show that, once again, two results which are apparently far from each other, even obtained by different groups in different countries, are the natural consequence of a sequence of intermediate results.

Newton's interpolation formula is a classical tool for constructing an interpolating polynomial by recurrence, by using divided differences. In the 1930s, Aitken [1] and Neville [52] derived independently of each other algorithms to compute the interpolating polynomial from the solutions of two simpler interpolation problems, avoiding the explicit use of divided differences. Some papers, [38,46] among others, extended both approaches at the beginning of the 1970s, to the more general setting of Chebyshev systems. Almost simultaneously, extrapolation methods were being studied and extended by several authors, as Schneider [54], Brezinski [4,5,7], Håvie [31–33], Mühlbach [39–42,48] and Gasca and López-Carmona [19]. For a historical overview of extrapolation methods confer Brezinski's contribution [6] to this volume and the book [8]. It must be remarked that the

* Corresponding author.
E-mail address: gasca@posta.unizar.es (M. Gasca).

0377-0427/00/$ - see front matter
PII: S0377-0427(00)00356-3

techniques used by these authors were different, and that frequently the results obtained using one of these techniques induced some progress in the other ones, in a very cooperative form.

However, it is clear that the basic role in all these papers was played by elimination techniques. In [21] we studied general elimination strategies, where one strategy which we called Neville elimination proved to be well suited to work with some special classes of matrices, in particular *totally positive matrices* (that are matrices with all subdeterminants nonnegative).

This was the origin of a series of papers [24–27] where the properties of Neville elimination were carefully studied and its application to totally positive matrices allowed a much better knowledge of these matrices. Since one of the applications of totally positive matrices is CAGD, the results obtained for them have given rise in the last years to several other papers as [28,11,12]. In [11,12] Carnicer and Peña proved the optimality in their respective spaces of some well-known function bases as Bernstein polynomials and B-splines in the context of shape preserving representations. Neville elimination has appeared, once again, as a way to construct other bases with similar properties.

2. Extrapolation and Schur complement

A k-tuple $L=(\ell_1,\dots,\ell_k)$ of natural numbers, with $\ell_1<\cdots<\ell_k$, will be called an *index list* of *length* k over $\mathbb{N}$. For $I=(i_1,\dots,i_m)$ and $J=(j_1,\dots,j_n)$ two index lists over $\mathbb{N}$, we write $I\subset J$ iff every element of I is an element of J. Generally, we shall use for index lists the same notations as for sets emphasizing that $I\setminus J$, $I\cap J$, $I\cup J\dots$ always have to be ordered as above.

Let $A=(a_i^j)$ be a real matrix and $I=(i_1,\dots,i_m)$ and $J=(j_1,\dots,j_n)$ index lists contained, repectively, in the index lists of rows and columns of A. By

$$A\begin{pmatrix}J\\I\end{pmatrix}=A\begin{pmatrix}j_1,\dots,j_n\\i_1,\dots,i_m\end{pmatrix}=(a_{i_\mu}^{j_\nu})_{\mu=1,\dots,m}^{\nu=1,\dots,n}\in\mathbb{R}^{m\times n},$$

we denote the *submatrix* of A with list of rows I and list of columns J.

If $I^\circ, I^{\circ\prime}$ and $J^\circ, J^{\circ\prime}$ are partitions of I and J, respectively, i.e., $I^\circ\cup I^{\circ\prime}=I$, $I^\circ\cap I^{\circ\prime}=\emptyset$, $J^\circ\cup J^{\circ\prime}=J$, $J^\circ\cap J^{\circ\prime}=\emptyset$, we represent $A\binom{J}{I}$ in a corresponding partition

$$A\begin{pmatrix}J\\I\end{pmatrix}=\begin{pmatrix}A\begin{pmatrix}J^\circ\\I^\circ\end{pmatrix}A\begin{pmatrix}J^{\circ\prime}\\I^\circ\end{pmatrix}\\[2ex] A\begin{pmatrix}J^\circ\\I^{\circ\prime}\end{pmatrix}A\begin{pmatrix}J^{\circ\prime}\\I^{\circ\prime}\end{pmatrix}\end{pmatrix}. \tag{1}$$

If $m=n$, then by

$$A\begin{vmatrix}J\\I\end{vmatrix}:=\det A\begin{pmatrix}J\\I\end{pmatrix}=A\begin{vmatrix}j_1,\dots,j_m\\i_1,\dots,i_m\end{vmatrix},$$

we denote the determinant of $A\binom{J}{I}$ which is called a *subdeterminant* of A. Throughout we set $A\left|\begin{smallmatrix}\emptyset\\\emptyset\end{smallmatrix}\right|:=1$.

Let $N\in\mathbb{N}$, $I:=(1,2,\dots,N+1)$ and $I^\circ:=(1,2,\dots,N)$. By a prime we denote ordered complements with respect to I. Given elements $f_1,\dots,f_N$ and $f=:f_{N+1}$ of a linear space E over $\mathbb{R}$, elements $L_1,\dots,L_N$ and $L=:L_{N+1}$ of its dual E^*, consider the problem of finding

$$\langle L, p_1^N(f)\rangle, \tag{2}$$

where $p = p_1^N(f) = c_1 \cdot f_1 + \cdots + c_N \cdot f_N$ satisfies the *interpolation conditions*

$$\langle L_i, p\rangle = \langle L_i, f\rangle \quad i \in I^\circ. \tag{3}$$

Here $\langle \cdot, \cdot \rangle$ means duality between E^* and E. If we write

$$A\begin{pmatrix} j \\ i \end{pmatrix} := \langle L_i, f_j\rangle \quad \text{for } i, j \in I, \quad (i,j) \neq (N+1, N+1),$$

and $\boldsymbol{c}$ is the vector of components c_i, this problem is equivalent to solving the bordered system (cf. [16])

$$B \cdot \boldsymbol{x} = \boldsymbol{y} \quad \text{where} \quad B = \begin{pmatrix} A\begin{pmatrix} I^\circ \\ I^\circ \end{pmatrix} & \mathbf{0} \\ A\begin{pmatrix} I^\circ \\ N+1 \end{pmatrix} & 1 \end{pmatrix}, \quad \boldsymbol{x} = \begin{pmatrix} \boldsymbol{c} \\ \xi \end{pmatrix}, \quad \boldsymbol{y} = \begin{pmatrix} A\begin{pmatrix} N+1 \\ I^\circ \end{pmatrix} \\ A\begin{pmatrix} N+1 \\ N+1 \end{pmatrix} \end{pmatrix}. \tag{4}$$

Assuming $A\binom{I^\circ}{I^\circ}$ nonsingular this can be solved by eliminating the components of $\boldsymbol{c}$ in the last equation by adding a suitable linear combination of the first N equations of (4) to the last one, yielding one equation for one unknown, namely ξ:

$$\xi = A\begin{pmatrix} N+1 \\ N+1 \end{pmatrix} - A\begin{pmatrix} I^\circ \\ N+1 \end{pmatrix} \cdot A\begin{pmatrix} I^\circ \\ I^\circ \end{pmatrix}^{-1} A\begin{pmatrix} N+1 \\ I^\circ \end{pmatrix}. \tag{5}$$

Considering the effect of this block elimination step on the matrix

$$A = \begin{pmatrix} A\begin{pmatrix} I^\circ \\ I^\circ \end{pmatrix} & A\begin{pmatrix} N+1 \\ I^\circ \end{pmatrix} \\ A\begin{pmatrix} I^\circ \\ N+1 \end{pmatrix} & A\begin{pmatrix} N+1 \\ N+1 \end{pmatrix} \end{pmatrix}, \tag{6}$$

we find it transformed to

$$\tilde{A} = \begin{pmatrix} A\begin{pmatrix} I^\circ \\ I^\circ \end{pmatrix} & A\begin{pmatrix} N+1 \\ I^\circ \end{pmatrix} \\ & \xi \end{pmatrix}.$$

If we take

$$A\begin{pmatrix} N+1 \\ N+1 \end{pmatrix} := 0, \tag{7}$$

then we have

$$\xi = -\langle L, p_1^N(f)\rangle. \tag{8}$$

On the other hand, if instead of (7) we take

$$A\begin{pmatrix} N+1 \\ N+1 \end{pmatrix} := \langle L_{N+1}, f_{N+1}\rangle, \tag{9}$$

then, in this frame, we get

$$\xi = \langle L, r_1^N(f)\rangle, \tag{10}$$

where

$$r_1^N(f) := f - p_1^N(f)$$

is the interpolation remainder.

If the systems $(f_1,\dots,f_N)$ and $(L_1,\dots,L_N)$ are independent of f and L then these problems are called *general linear extrapolation* problems, and if one or both do depend on $f=f_{N+1}$ or $L=L_{N+1}$ they are called problems of *quasilinear extrapolation*.

Observe, that with regard to determinants the block elimination step above is an elementary operation leaving the value of $\det A$ unchanged. Hence

$$\xi = \frac{\det A\begin{pmatrix} I \\ I \end{pmatrix}}{\det A\begin{pmatrix} I^\circ \\ I^\circ \end{pmatrix}},$$

which is known as the Schur complement of $A(\begin{smallmatrix} I^\circ \\ I^\circ \end{smallmatrix})$ in $A(\begin{smallmatrix} I \\ I \end{smallmatrix})$. This concept, introduced in [34,35] has found many applications in Linear Algebra and Statistics [13,53]. It may be generalized in different ways, see, for example, [21,22,44] where we used the concept of general elimination strategy which is explained in the next section.

3. Elimination strategies

In this section and the next two let $k,m,n\in\mathbb{N}$ such that $k+m=n$ and $I=(1,\dots,n)$. Given a square matrix $A=A(\begin{smallmatrix} I \\ I \end{smallmatrix})$ over $\mathbb{R}$, how can we simplify $\det A$ by elementary operations, not altering the value of $\det A$, producing zeros in prescribed columns, e.g. in columns 1 to k?. Take a permutation of all rows, $M=(m_1,\dots,m_n)$ say, then look for a linear combination of k rows from $(m_1,\dots,m_{n-1})$ which, when added to row m_n, will produce zeros in columns 1 to k. Then add to row m_{n-1} a linear combination of k of its predecessors in M, to produce zeros in columns 1 to k, etc. Finally, add to row m_{k+1} a suitable linear combination of rows $m_1,\dots,m_k$ to produce zeros in columns 1 to k. Necessarily,

$$A\begin{vmatrix} 1,\dots,k \\ j_1^r,\dots,j_k^r \end{vmatrix} \neq 0$$

is assumed when a linear combination of rows $j_1^r,\dots,j_k^r$ is added to row m_r $(r=n,n-1,\dots,k+1)$ to generate zeros in columns 1 to k, and $j_q^r < m_r$ $(q=1,\dots,k;\ r=n,n-1,\dots,k+1)$ in order that in each step an elementary operation will be performed.

Let us give a formal description of this general procedure. Suppose that (I_s,I_s°) $(s=1,\dots,m)$ are pairs of ordered index lists of length $k+1$ and k, respectively, over a basic index list M with $I_s^\circ\subset I_s$. Then the family

$$\Sigma := ((I_s,I_s^\circ))_{s=1,\dots,m}$$

will be called a (k,m)-*elimination strategy* over $I:=I_1\cup\cdots\cup I_m$ provided that for $s=2,\dots,m$

(i) $\operatorname{card}(I_1\cup\cdots\cup I_s)=k+s$,

(ii) $I_s^\circ\subset I_s\cap(I_1\cup\cdots\cup I_{s-1})$.

By $E(k,m,I)$ we denote the set of all (k,m)-*elimination strategies* over I. $I^\circ := I_1^\circ$ is called the *basic index list* of the strategy Σ. For each s, the zeros in the row $\alpha_\Sigma(s) := I_s \setminus I_s^\circ$ are produced with the rows of I_s°. For shortness, we shall abbreviate the phrase "elimination strategy" by e.s. Notice that, when elimination is actually performed, it is done in the reverse ordering: first in row $\alpha_\Sigma(m)$, then in row $\alpha_\Sigma(m-1)$, etc.

The simplest example of e.s. over $I = (1,\dots,m+k)$, is *Gauss elimination*:

$$\Gamma = ((G_s, G_s^\circ))_{s=1,\dots,m}, \quad G^\circ = G_s^\circ = \{1,\dots,k\}, \quad G_s = G^\circ \cup \{k+s\}. \tag{11}$$

For this strategy it is irrelevant in which order elimination is performed. This does not hold for another useful strategy over I:

$$\mathcal{N} = ((N_s, N_s^\circ))_{s=1,\dots,m} \tag{12}$$

with $N_s^\circ = (s,\dots,s+k-1), N_s = (s,\dots,s+k)$, $s = 1,\dots,m$, which we called [21,43,44] the *Neville* (k,m)–e.s. Using this strategy elimination must be performed from bottom to top. The reason for the name Neville is their relationship with Neville interpolation algorithm, based on consecutivity, see [43,23].

4. Generalized Schur complements

Suppose that $\Sigma = ((I_s, I_s^\circ))_{s=1,\dots,m} \in E(k,m,I)$ and that $\mathcal{K}^\circ \subset I$ is a fixed index list of length k. We assume that the submatrices $A\binom{\mathcal{K}^\circ}{I_s^\circ}$ of a given matrix $A = A\binom{I}{I} \in \mathbb{R}^{n\times n}$ are nonsingular for $s = 1,\dots,m$. Then the elimination strategy transforms A into the matrix $\tilde{A}$ which, partitioned with respect to $I^\circ \cup I^{\circ\prime} = I, \mathcal{K}^\circ \cup \mathcal{K}^{\circ\prime} = I$, can be written as

$$\tilde{A} = \begin{pmatrix} \tilde{A}\binom{\mathcal{K}^\circ}{I^\circ} & \tilde{A}\binom{\mathcal{K}^{\circ\prime}}{I^\circ} \\ 0 & \tilde{A}\binom{\mathcal{K}^{\circ\prime}}{I^{\circ\prime}} \end{pmatrix}$$

with

$$\tilde{A}\binom{\mathcal{K}^\circ}{I^\circ} = A\binom{\mathcal{K}^\circ}{I^\circ}, \quad \tilde{A}\binom{\mathcal{K}^{\circ\prime}}{I^\circ} = A\binom{\mathcal{K}^{\circ\prime}}{I^\circ}.$$

The submatrix $\tilde{S} := \tilde{A}\binom{\mathcal{K}^{\circ\prime}}{I^{\circ\prime}}$ of $\tilde{A}$ is called the *Schur complement* of $A\binom{\mathcal{K}^\circ}{I^\circ}$ in A with respect to the e.s. Σ and the column list $\mathcal{K}^\circ$, and is also denoted by

$$\tilde{S} = \left[A\binom{I}{I} \Big/ A\binom{\mathcal{K}^\circ}{I^\circ} \right]_\Sigma.$$

When $\Sigma = \Gamma$ as in (11) and $\mathcal{K}^\circ = \{1,\dots,k\}$, then $\tilde{S}$ is the classical Schur complement, which can also be written as

$$\tilde{A}\binom{\mathcal{K}^{\circ\prime}}{I^{\circ\prime}} = A\binom{\mathcal{K}^{\circ\prime}}{I^{\circ\prime}} - A\binom{\mathcal{K}^\circ}{I^{\circ\prime}} A\binom{\mathcal{K}^\circ}{I^\circ}^{-1} A\binom{\mathcal{K}^{\circ\prime}}{I^\circ}.$$

When $\Sigma = \mathcal{N}$ is the Neville (k,m)–e.s. (12) and $\mathcal{K}^\circ = \{1,\dots,k\}$, then the rows of the Schur complement $\tilde{S} = \tilde{A}\binom{\mathcal{K}^{\circ\prime}}{I^{\circ\prime}}$ are

$$\tilde{A}\begin{pmatrix}\mathcal{K}^{\circ\prime}\\ k+s\end{pmatrix} = A\begin{pmatrix}\mathcal{K}^{\circ\prime}\\ k+s\end{pmatrix} - A\begin{pmatrix}\mathcal{K}^{\circ}\\ k+s\end{pmatrix} A\begin{pmatrix}\mathcal{K}^{\circ}\\ s,\dots,s+k-1\end{pmatrix}^{-1} A\begin{pmatrix}\mathcal{K}^{\circ\prime}\\ s,\dots,s+k-1\end{pmatrix} \quad s=1,\dots,m.$$

Whereas, the Schur complement of a submatrix depends essentially on the elimination strategies used, its determinant does not! There holds the following generalization of Schur's classical determinantal identity [21,22,44]:

$$\det A\begin{pmatrix}I\\ I\end{pmatrix} = (-1)^\beta \det A\begin{pmatrix}\mathcal{K}^{\circ}\\ I^{\circ}\end{pmatrix} \det\left[A\begin{pmatrix}I\\ I\end{pmatrix}\Big/ A\begin{pmatrix}\mathcal{K}^{\circ}\\ I^{\circ}\end{pmatrix}\right]_\Sigma$$

for all e.s. $\Sigma \in E(k,m,I)$, where β is an integer depending only on Σ and $\mathcal{K}^\circ$.

Also, Sylvester's classical determinantal identity [55,56] has a corresponding generalization, see [18,21,22,43,44] for details. In the case of Gauss elimination we get Sylvester's classical identity [9,10,55,56]

$$\det\left(A\begin{vmatrix}1,\dots,k,\ k+t\\ 1,\dots,k,\ k+s\end{vmatrix}\right)_{s=1,\dots,m}^{t=1,\dots,m} = \det A\ \left(A\begin{vmatrix}1,\dots,k\\ 1,\dots,k\end{vmatrix}\right)^{m-1}.$$

In the case of Neville elimination one has

$$\det\left(A\begin{vmatrix}1,\dots,k, & k+t\\ s,\dots,s+k-1, & s+k\end{vmatrix}\right)_{s=1,\dots,m}^{t=1,\dots,m} = \det A\prod_{s=2}^{m} A\begin{vmatrix}1,\dots,k\\ s,\dots,s+k-1\end{vmatrix}.$$

Another identity of Sylvester's type has been derived in [3]. Also some applications to the E-algorithm [5] are given there.

As we have seen, the technique of e.s. has led us in particular to general determinantal identities of Sylvester's type. It can also be used to extend determinantal identities in the sense of Muir [51], see [47].

5. Application to quasilinear extrapolation problems

Suppose we are given elements $f_1,\dots,f_N$ of a linear space E and elements $L_1,\dots,L_N$ of its dual E^*. Consider furthermore elements $f =: f_{N+1}$ of E and $L =: L_{N+1}$ of E^*. Setting $I = (1,\dots,N+1)$, by A we denote the *generalized Vandermonde matrix*

$$A = A\begin{pmatrix}I\\ I\end{pmatrix} = V\begin{pmatrix}f_1,\dots,f_N,f_{N+1}\\ L_1,\dots,L_N,L_{N+1}\end{pmatrix} := \left(\langle L_i, f_j\rangle\right)_{i=1,\dots,N+1}^{j=1,\dots,N+1}. \tag{13}$$

Assume now that $k,m \in \mathbb{N}, m \leqslant N+1-k$ and that

$$\Sigma = ((I_s, I_s^\circ))_{s=1,\dots,m} \tag{14}$$

is a $(k-1,m)$–e.s. over $\bigcup_{s=1}^m I_s \subset (1,\dots,N)$. Let $G := (1,\dots,k)$. If the submatrices

$$A\begin{pmatrix}G\\ I_s\end{pmatrix} \quad \text{are nonsingular for } s = 1,\dots,m, \tag{15}$$

then for $s = 1, \ldots, m$ the interpolants

$$p_s^k(f) := \sum_{j=1}^{k} c_{s,j}^k(f) \cdot f_j, \tag{16}$$

satisfying the interpolation conditions

$$\langle L_i, p_s^k(f)\rangle = \langle L_i, f\rangle \quad \text{for } i \in I_s$$

are well defined as well as

$$\tau_s^k(f) := \langle L, p_s^k(f)\rangle.$$

Clearly, in case of general linear extrapolation the mapping

$$E \ni f \overset{p_s^k}{\to} p_s^k(f)$$

is a linear projection onto $\operatorname{span}\{f_1, \ldots, f_N\}$ and

$$E \ni f \overset{c_{s,j}^k}{\to} c_{s,j}^k(f)$$

is a linear functional. In case of quasilinear extrapolation we assume that, as a function of $f \in E$, p_s^k remains idempotent. Then, as a function of $f \in E$, in general the coefficients $c_{s,j}^k(f)$ are not linear. We assume that, as functions of $f \in \operatorname{span}\{f_1, \ldots, f_N\}$, $c_{s,j}^k(f)$ remain linear.

The task is

(i) to find conditions, such that $p_1^N(f), \tau_1^N(f)$ are well defined, and
(ii) to find methods to compute these quantities from $p_s^k(f), \tau_s^k(f) (s = 1, \ldots, m)$, respectively.

When translated into pure terms of Linear Algebra these questions mean: Consider matrix (13) and assume (15),

(i) under which conditions can we ensure that $A\binom{1,\ldots,N}{1,\ldots,N}$ is nonsingular?

The coefficient problem reads:

(ii′) Suppose that we do know the solutions

$$\boldsymbol{c}_s^k(f) = (c_{s,j}^k(f))_{j=1,\ldots,k}$$

of the linear systems

$$A\begin{pmatrix} G \\ I_s \end{pmatrix} \cdot \boldsymbol{c}_s^k(f) = A\begin{pmatrix} N+1 \\ I_s \end{pmatrix}, \quad s = 1, \ldots, m.$$

How to get from these the solution $\boldsymbol{c}_1^N(f) = (c_{1,j}^N(f))_{j=1,\ldots,N}$ of

$$A\begin{pmatrix} 1,\ldots,N \\ 1,\ldots,N \end{pmatrix} \cdot \boldsymbol{c}_1^N(f) = A\begin{pmatrix} N+1 \\ 1,\ldots,N \end{pmatrix}?$$

The value problem reads:

(iii) Suppose that we do know the values

$$\tau_s^k(f) = \langle L, p_s^k(f)\rangle, \quad s = 1, \ldots, m.$$

How to get from these the value $\tau_1^N(f) = \langle L, p_1^N(f)\rangle$?

A *dual coefficient problem* can be also considered interchanging the roles of the spaces E and E^*. These problems were considered and solved in [20,7,19,31,40–42,45,48,50].

6. Applications to special classes of matrices

General elimination strategies, in particular the Neville e.s. and generalized Schur complements have found other applications in matrix theory and related problems.

In [21,22,44] we have considered some classes $\mathscr{L}_n$ of real $n \times n$-matrices A including the classes

(i) $\mathscr{C}_n$ of matrices satisfying $\det A\binom{J}{J} > 0$ for all $J \subset (1,\dots,n)$, $\det A\binom{K}{J} \cdot \det A\binom{J}{K} > 0$ for all $J, K \subset (1,\dots,n)$ of the same cardinality, which was considered in [36];
(ii) of symmetric positive-definite matrices;
(iii) of *strictly totally positive matrices* (STP), which are defined by the property that all square submatrices have positive determinants [36];
(iv) of Minkowski matrices, defined by

$$A\begin{pmatrix} j \\ i \end{pmatrix} < 0 \quad \text{for all } i \neq j, \quad \det A\begin{pmatrix} 1,\dots,k \\ 1,\dots,k \end{pmatrix} > 0 \quad \text{for all } 1 \leqslant k \leqslant n.$$

In [21] we have proved that

$$A \in \mathscr{L}_n \Rightarrow \tilde{S} \in \mathscr{L}_m,$$

where $m = n - k$ and $\tilde{S}$ denotes the classical Schur complement of $A\binom{1,\dots,k}{1,\dots,k}$ in A. For STP matrices also generalized Schur complements with respect to the Neville e.s. are STP. Using the Neville e.s. in [21,49] tests of algorithmic complexity $O(N^4)$ for matrices being STP were derived for the first time. Neville elimination, based on consecutivity, proved to be especially well suited for STP matrices, because these matrices were characterized in [36] by the property of having all subdeterminants with *consecutive* rows and columns positive.

Elimination by consecutive rows is not at all new in matrix theory. It has been used to prove some properties of special classes of matrices, for example, totally positive (TP) matrices, which, as it has already been said, are matrices with all subdeterminants nonnegative. However, motivated by the above mentioned algorithm for testing STP matrices, Gasca and Peña [24] initiated an exhaustive study of Neville elimination in an algorithmic way, of the pivots and multipliers used in the proccess to obtain new properties of totally positive matrices and to improve and simplify the known characterizations of these matrices.

Totally positive matrices have interesting applications in many fields, as, for example, vibrations of mechanical systems, combinatorics, probability, spline functions, computer-aided geometric design, etc., see [36,37]. For this reason, remarkable papers on total positivity due to specialists on these fields have appeared, see for example the ones collected in [29].

The important survey [2] presents a complete list of references on totally positive matrices before 1987. One of the main points in the recent study of this class of matrices has been that of characterizing them in practical terms, by factorizations or by the nonnegativity of some minors (instead of all of them, as claimed in the definition).

In [24] for example, it was proved that a matrix is STP if and only if all subdeterminants with lists of consecutive rows and consecutive columns, starting at least one of these lists by 1, are positive. Necessarily, one of the lists must start with 1. Observe, that the new characterization considerably decreases the number of subdeterminants to be checked, compared with the classical characterization, due to Fekete and Pólya [17], which used all subdeterminants with consecutive rows and columns.

This result means that the set of all subdeterminants of a matrix A with consecutive rows and columns, of the form

$$A\begin{vmatrix} 1,\dots,j \\ i,\dots,i+j-1 \end{vmatrix}, \quad A\begin{vmatrix} i,\dots,i+j-1 \\ 1,\dots,j \end{vmatrix},$$

called in [24] column- and row-initial minors, play in total positivity a similar role to that of the leading principal minors

$$A\begin{vmatrix} 1,\dots,j \\ 1,\dots,j \end{vmatrix}$$

in positive definiteness of symmetric real matrices. An algorithm based on Neville elimination was given in [24] with a complexity $O(N^3)$ for a matrix of order N, instead of the one with $O(N^4)$ previously obtained in [21,49]. Other similar simplifications were obtained in [24] for the characterization of totally positive matrices (not strictly).

Concerning factorizations, in [26] Neville elimination was described in terms of a product by bidiagonal unit-diagonal matrices. Some of the most well-known characterizations of TP and STP matrices are related to their LU factorization. Cryer [14,15], in the 1970s, extended to TP matrices what was previously known for STP matrices, thus obtaining the following result.

A square matrix A is TP (resp. STP) iff it has an LU factorization such that L and U are TP (ΔSTP).

Here, as usual, L (resp. U) denotes a lower (upper) triangular matrix and ΔSTP means triangular nonnegative matrices with all the nontrivial subdeterminants of any order strictly positive.

Also Cryer pointed out that the matrix A is STP iff it can be written in the form

$$A = \prod_{r=1}^{N} L_r \prod_{s=1}^{M} U_s$$

where each L_r (resp. U_s) is a lower (upper) ΔSTP matrix. Observe that this result does not mention the relation of N or M with the order n of the matrix A.

The matricial description of Neville elimination obtained in [26] produced in the same paper the following result.

Let A be a nonsingular matrix of order n. Then A is STP iff it can be expressed in the form:

$$A = F_{n-1}\cdots F_1 D G_1 \cdots G_{n-1},$$

where, for each $i=1,2,\dots,n-1$, F_i is a bidiagonal, lower triangular, unit diagonal matrix, with zeros in positions $(2,1),\dots,(i,i-1)$ and positive entries in $(i+1,i),\dots,(n,n-1)$, G_i has the transposed form of F_i and D is a diagonal matrix with positive diagonal.

Similar results were obtained in [26] for TP matrices. In that paper all these new characterizations were collected in three classes: characterizations in terms of determinants, in terms of algorithms and in terms of factorizations.

7. Variation diminution and computer-aided geometric design

An $n \times n$ matrix A is said to be *sign-regular* (SR) if for each $1 \leqslant k \leqslant n$ all its minors of order k have the same (non strict) sign (in the sense that the product of any two of them is greater than or

equal to zero). The matrix is *strictly sign-regular* (SSR) if for each $1 \leqslant k \leqslant n$ all its minors of order k are different from zero and have the same sign. In [27] a test for strict sign regularity is given.

The importance of these types of matrices comes from their *variation diminishing properties*. By a *sign sequence* of a vector $x = (x_1, \ldots, x_n)^{\mathrm{T}} \in \mathbb{R}^n$ we understand any signature sequence ε for which $\varepsilon_i x_i = |x_i|,\ i = 1, 2, \ldots, n$. The number of sign changes of x associated to ε, denoted by $\mathscr{C}(\varepsilon)$, is the number of indices i such that $\varepsilon_i \varepsilon_{i+1} < 0,\ 1 \leqslant i \leqslant n-1$. The *maximum* (resp. *minimum*) *variation of signs*, $V_+(x)$ (resp. $V_-(x)$), is by definition the maximum (resp. minimum) of $\mathscr{C}(\varepsilon)$ when ε runs over all sign sequences of x. Let us observe that if $x_i \neq 0$ for all i, then $V_+(x) = V_-(x)$ and this value is usually called the exact variation of signs. The next result (see [2, Theorems 5.3 and 5.6]) characterizes sign-regular and strictly sign-regular matrices in terms of their variation diminishing properties.

Let A be an $n \times n$ nonsingular matrix. Then:

(i) *A is SR* $\Leftrightarrow V_-(Ax) \leqslant V_-(x)\ \forall x \in \mathbb{R}^n$.
(ii) *A is SR* $\Leftrightarrow V_+(Ax) \leqslant V_+(x)\ \forall x \in \mathbb{R}^n$.
(iii) *A is SSR* $\Leftrightarrow V_+(Ax) \leqslant V_-(x)\ \forall x \in \mathbb{R}^n \setminus \{0\}$.

The above matricial definitions lead to the corresponding definitions for systems of functions. A system of functions $(u_0, \ldots, u_n)$ is *sign-regular* if all its collocation matrices are sign-regular of the same kind. The system is *strictly sign-regular* if all its collocation matrices are strictly sign-regular of the same kind. Here a *collocation matrix* is defined to be a matrix whose (i, j)-entry is of the form $u_i(x_j)$ with any system of strictly increasing points x_j.

Sign-regular systems have important applications in CAGD. Given $u_0, \ldots, u_n$, functions defined on $[a, b]$, and $P_0, \ldots, P_n \in \mathbb{R}^k$, we may define a curve $\gamma(t)$ by

$$\gamma(t) = \sum_{i=0}^{n} u_i(t) P_i.$$

The points $P_0, \ldots, P_n$ are called *control points*, because we expect to modify the shape of the curve by changing these points adequately. The polygon with vertices $P_0, \ldots, P_n$ is called *control polygon* of γ.

In CAGD the functions $u_0, \ldots, u_n$ are usually nonnegative and normalized ($\sum_{i=0}^{n} u_i(t) = 1\ \forall t \in [a, b]$). In this case they are called *blending functions*. These requirements imply that the curve lies in the convex hull of the control polygon (*convex hull property*). Clearly, $(u_0, \ldots, u_n)$ is a system of blending functions if and only if all the collocation matrices are stochastic (that is, they are nonnegative matrices such that the elements of each row sum up to 1). For design purposes, it is desirable that the curve imitates the control polygon and that the control polygon even "exaggerates" the shape of the curve, and this holds when the system satisfies variation diminishing properties. If $(u_0, \ldots, u_n)$ is a sign-regular system of blending functions then the curve γ preserves many shape properties of the control polygon, due to the variation diminishing properties of $(u_0, \ldots, u_n)$. For instance, any line intersects the curve no more often than it intersects the control polygon.

A characterization of SSR matrices A by the Neville elimination of A and of some submatrices of A is obtained in [26, Theorem 4.1].

A system of functions $(u_0, \ldots, u_n)$ is said to be *totally positive* if all its collocation matrices are totally positive. The system is *normalized totally positive* (NTP) if it is totally positive and $\sum_{i=0}^{n} u_i = 1$.

Normalized totally positive systems satisfy an interesting shape-preserving property, which is very convenient for design purposes and which we call *endpoint interpolation property*: the initial and final endpoints of the curve and the initial and final endpoints (respectively) of the control polygon coincide. In summary, these systems are characterized by the fact that they always generate curves γ satisfying simultaneously the convex hull, variation diminishing and endpoint interpolation properties.

Now the following question arises. Given a system of functions used in CAGD to generate curves, does there exist a basis of the space generated by that system with optimal shape preserving properties? Or equivalently, is there a basis such that the generated curves γ imitate better the form of the corresponding control polygon than the form of the corresponding control polygon for any other basis?

In the space of polynomials of degree less than or equal to n on a compact interval, the Bernstein basis is optimal. This was conjectured by Goodman and Said in [30], and it was proved in [11]. In [12], there is also an affirmative answer to the above questions for any space with TP basis. Moreover, Neville elimination provides a constructive way to obtain optimal bases. In the space of polynomial splines, B-splines form the optimal basis.

Since the product of TP matrices is a TP matrix, if $(u_0,\dots,u_n)$ is a TP system of functions and A is a TP matrix of order $n+1$, then the new system $(u_0,\dots,u_n)A$ is again a TP system (which satisfies a "stronger" variation diminishing property than $(u_0,\dots,u_n)$). If we obtain from a basis $(u_0,\dots,u_n)$, in this way, all the totally positive bases of the space, then $(u_0,\dots,u_n)$ will be the "least variation diminishing" basis of the space. In consequence, the control polygons with respect to $(u_0,\dots,u_n)$ will imitate the form of the curve better than the control polygons with respect to other bases of the space. Therefore, we may reformulate the problem of finding an optimal basis $(b_0,\dots,b_n)$ in the following way:

Given a vector space $\mathscr{U}$ with a TP basis, is there a TP basis $(b_0,\dots,b_n)$ of $\mathscr{U}$ such that, for any TP basis $(v_0,\dots,v_n)$ of $\mathscr{U}$ there exists a TP matrix K satisfying $(v_0,\dots,v_n)=(b_0,\dots,b_n)K$?.

The existence of such *optimal basis* $(b_0,\dots,b_n)$ was proved in [12], where it was called *B-basis*. In the same paper, a method of construction, inspired by the Neville elimination process, was given. As mentioned above, Bernstein polynomials and B-splines are examples of B-bases.

Another point of view for B-bases is closely related to corner cutting algorithms, which play an important role in CAGD.

Given two NTP bases, $(p_0,\dots,p_n)$, $(b_0,\dots,b_n)$, let K be the nonsingular matrix such that

$$(p_0,\dots,p_n)=(b_0,\dots,b_n)K.$$

Since both bases are normalized, if K is a nonnegative matrix, it is clearly stochastic.

A curve γ can be expressed in terms of both bases

$$\gamma(t)=\sum_{i=0}^{n} B_i b_i(t)=\sum_{i=0}^{n} P_i p_i(t), \quad t\in[a,b],$$

and the matrix K gives the relationship between both control polygons

$$(B_0,\dots,B_n)^{\mathrm{T}}=K(P_0,\dots,P_n)^{\mathrm{T}}.$$

An *elementary corner cutting* is a transformation which maps any polygon $P_0 \cdots P_n$ into another polygon $B_0 \cdots B_n$ defined by:

$$\begin{aligned} &B_j = P_j, \quad j \neq i, \\ &B_i = (1-\lambda)P_i + \lambda P_{i+1}, \quad \text{for one} \quad i \in \{0, \ldots, n-1\} \end{aligned} \tag{17}$$

or

$$\begin{aligned} &B_j = P_j, \quad j \neq i, \\ &B_i = (1-\lambda)P_i + \lambda P_{i-1}, \quad \text{for one} \quad i \in \{1, \ldots, n\}. \end{aligned} \tag{18}$$

Here $\lambda \in (0,1)$.

A *corner-cutting algorithm* is the algorithmic description of a corner cutting transformation, which is any composition of elementary corner cutting transformations.

Let us assume now that the matrix K above is TP. Since it is stochastic, nonsingular and TP, it can be factorized as a product of bidiagonal nonnegative matrices, (as we have mentioned in Section 6), which can be interpreted as a corner cutting transformation. Such factorizations are closely related to the Neville elimination of the matrix [28]. From the variation diminution produced by the totally positive matrices of the process, it can be deduced that the curve γ imitates better the form of the control polygon $B_0 \cdots B_n$ than that of the control polygon $P_0 \cdots P_n$. Therefore, we see again that an NTP basis $(b_0, \ldots, b_n)$ of a space $\mathscr{U}$ has *optimal shape-preserving properties* if for any other NTP basis $(p_0, \ldots, p_n)$ of $\mathscr{U}$ there exists a (stochastic) TP matrix K such that

$$(p_0, \ldots, p_n) = (b_0, \ldots, b_n)K. \tag{19}$$

Hence, a basis has optimal shape preserving properties if and only if it is a normalized B-basis. Neville elimination has also inspired the construction of B-bases in [11,12]. Many of these results and other important properties and applications of totally positive matrices have been collected, as we have already said in [28, Section 6].

References

[1] A.G. Aitken, On interpolation by iteration of proportional parts without the use of differences, Proc. Edinburgh Math. Soc. 3 (1932) 56–76.
[2] T. Ando, Totally positive matrices, Linear Algebra Appl. 90 (1987) 165–219.
[3] B. Beckermann, G. Mühlbach, A general determinantal identity of Sylvester type and some applications, Linear Algebra Appl. 197,198 (1994) 93–112.
[4] Cl. Brezinski, The Mühlbach–Neville–Aitken-algorithm and some extensions, BIT 20 (1980) 444–451.
[5] Cl. Brezinski, A general extrapolation algorithm, Numer. Math. 35 (1980) 175–187.
[6] Cl. Brezinski, Convergence acceleration during the 20th century, this volume, J. Comput. Appl. Math. 122 (2000) 1–21.
[7] Cl. Brezinski, Recursive interpolation, extrapolation and projection, J. Comput. Appl. Math. 9 (1983) 369–376.
[8] Cl. Brezinski, M. Redivo Zaglia, Extrapolation methods, theory and practice, North-Holland, Amsterdam, 1991.
[9] R.A. Brualdi, H. Schneider, Determinantal identities: Gauss, Schur, Cauchy, Sylvester, Kronecker, Jacobi, Binet, Laplace, Muir and Cayley, Linear Algebra Appl. 52/53 (1983) 769–791.
[10] R.A. Brualdi, H. Schneider, Determinantal identities revisited, Linear Algebra Appl. 59 (1984) 183–211.
[11] J.M. Carnicer, J.M. Peña, Shape preserving representations and optimality of the Bernstein basis, Adv. Comput. Math. 1 (1993) 173–196.

[12] J.M. Carnicer, J.M. Peña, Totally positive bases for shape preserving curve design and optimality of B-splines, Comput. Aided Geom. Design 11 (1994) 633–654.
[13] R.W. Cottle, Manifestations of the Schur complement, Linear Algebra Appl. 8 (1974) 189–211.
[14] C. Cryer, The LU-factorization of totally positive matrices, Linear Algebra Appl. 7 (1973) 83–92.
[15] C. Cryer, Some poperties of totally positive matrices, Linear algebra Appl. 15 (1976) 1–25.
[16] D.R. Faddeev, U.N. Faddeva, Computational Methods of Linear Algebra, Freeman, San Francisco, 1963.
[17] M. Fekete, G. Pólya, Über ein Problem von Laguerre, Rend. C.M. Palermo 34 (1912) 89–120.
[18] M. Gasca, A. López-Carmona, V. Ramírez, A generalized Sylvester's identity on determinants and 1st application to interpolation problems, in: W. Schempp, K. Zeller (Eds.), Multivariate Approximation Theory II, ISNM, Vol. 61, Biskhäuser, Basel, 1982, pp. 171–184.
[19] M. Gasca, A. López-Carmona, A general interpolation formula and its application to multivariate interpolation, J. Approx. Theory 34 (1982) 361–374.
[20] M. Gasca, E. Lebrón, Elimination techniques and interpolation, J. Comput. Appl. Math. 19 (1987) 125–132.
[21] M. Gasca, G. Mühlbach, Generalized Schur-complements and a test for total positivity, Appl. Numer. Math. 3 (1987) 215–232.
[22] M. Gasca, G. Mühlbach, Generalized Schur-complements, Publicacciones del Seminario Matematico Garcia de Galdeano, Serie II, Seccion 1, No. 17, Universidad de Zaragoza, 1984.
[23] M. Gasca, J.M. Peña, Neville elimination and approximation theory, in: S.P. Singh (Ed.), Approximation Theory, Wavelets and Applications, Kluwer Academic Publishers, Dordrecht, 1995, pp. 131–151.
[24] M. Gasca, J.M. Peña, Total positivity and Neville elimination, Linear Algebra Appl. 165 (1992) 25–44.
[25] M. Gasca, J.M. Peña, On the characterization of TP and STP matrices, in: S.P. Singh (Ed.), Aproximation Theory, Spline Functions and Applications, Kluwer Academic Publishers, Dordrecht, 1992, pp. 357–364.
[26] M. Gasca, J.M. Peña, A matricial description of Neville elimination with applications to total positivity, Linear Algebra Appl. 202 (1994) 33–54.
[27] M. Gasca, J.M. Peña, A test for strict sign-regularity, Linear Algebra Appl. 197–198 (1994) 133–142.
[28] M. Gasca, J.M. Peña, Corner cutting algorithms and totally positive matrices, in: P.J. Laurent, A. Le Méhauté, L.L. Schumaker (Eds.), Curves and Surfaces II, 177–184, A.K. Peters, Wellesley, MA, 1994.
[29] M. Gasca, C.A. Micchelli (Eds.), Total Positivity and its Applications, Kluwer Academic Publishers, Dordrecht, 1996.
[30] T.N.T. Goodman, H.B. Said, Shape preserving properties of the generalized ball basis, Comput. Aided Geom. Design 8 (115–121) 1991.
[31] T. Håvie, Generalized Neville type extrapolation schemes, BIT 19 (1979) 204–213.
[32] T. Håvie, Remarks on a unified theory of classical and generalized interpolation and extrapolation, BIT 21 (1981) 465–474.
[33] T. Håvie, Remarks on the Mühlbach–Neville–Aitken-algorithm, Math. a. Comp. Nr. 2/80, Department of Numerical Mathematics, The University of Trondheim, 1980.
[34] E. Haynsworth, Determination of the inertia of a partitioned Hermitian matrix, Linear Algebra Appl. 1 (1968) 73–81.
[35] E. Haynsworth, On the Schur Complement, Basel Mathemathics Notes No. 20, June 1968.
[36] S. Karlin, Total Positivity, Stanford University Press, Standford, 1968.
[37] S. Karlin, W.J. Studden, Tchebycheff Systems: with Applications in Analysis and Statistics, Interscience, New York, 1966.
[38] G. Mühlbach, Neville–Aitken Algorithms for interpolation by functions of Čebyšev-systems in the sense of Newton and in a generalized sense of hermite, in: A.G. Law, B.N. Sahney (Eds.), Theory of Approximation, with Applications, Proceedings of the International Congress on Approximation Theory in Calgary, 1975, Academic Press, New York, 1976, pp. 200–212.
[39] G. Mühlbach, The general Neville–Aitken-algorithm and some applications, Numer. Math. 31 (1978) 97–110.
[40] G. Mühlbach, On two general algorithms for extrapolation with applications to numerical differentiation and integration, in: M.G. de Bruin, H. van Rossum (Eds.), Padé Approximation and its Applications, Lecture Notes in Mathematics, Vol. 888, Springer, Berlin, 1981, pp. 326–340.
[41] G. Mühlbach, Extrapolation algorithms as elimination techniques with applications to systems of linear equations, Report 152, Institut für Mathematik der Universität Hannover, 1982, pp. 1–47.
[42] G. Mühlbach, Algorithmes d'extrapolation, Publication ANO 118, Université de Lille 1, January 1984.

[43] G. Mühlbach, Sur une identité généralisée de Sylvester, Publication ANO 119, Université de Lille 1, January 1984.
[44] G. Mühlbach, M. Gasca, A generalization of Sylvester's identity on determinants and some applications, Linear Algebra Appl. 66 (1985) 221–234.
[45] G. Mühlbach, Two composition methods for solving certain systems of linear equations, Numer. Math. 46 (1985) 339–349.
[46] G. Mühlbach, A recurrence formula for generalized divided differences and some applications, J. Approx. Theory. 9 (1973) 165–172.
[47] G. Mühlbach, On extending determinantal identities, Publicaciones del Seminario Matematico Garcia de Galdeano, Serie II, Seccion 1, No. 139, Universidad de Zaragoza, 1987.
[48] G. Mühlbach, Linear and quasilinear extrapolation algorithms, in: R. Vichnevetsky, J. Vignes (Eds.), Numerical Mathematics and Applications, Elsevier, North-Holland, Amsterdam, IMACS, 1986, pp. 65–71.
[49] G. Mühlbach, M. Gasca, A test for strict total positivity via Neville elimination, in: F. Uhlig, R. Grone (Eds.), Current Trends in Matrix Theory, North-Holland, Amsterdam, 1987, pp. 225–232.
[50] G. Mühlbach, Recursive triangles, in: D. Beinov, V. Covachev (Eds.), Proceedings of the third International Colloquium on Numerical Analysis, Utrecht, VSP, 1995, pp. 123–134.
[51] T. Muir, The law of extensible minors in determinants, Trans. Roy. Soc. Edinburgh 30 (1883) 1–4.
[52] E.H. Neville, Iterative interpolation, J. Indian Math. Soc. 20 (1934) 87–120.
[53] D.V. Ouellette, Schur complements and statistics, Linear Algebra Appl. 36 (1981) 186–295.
[54] C. Schneider, Vereinfachte rekursionen zur Richardson-extrapolation in spezialfällen, Numer. Math. 24 (1975) 177–184.
[55] J.J. Sylvester, On the relation between the minor determinants of linearly equivalent quadratic functions, Philos. Mag. (4) (1851) 295–305.
[56] J.J. Sylvester, Collected Mathematical Papers, Vol. 1, Cambridge University Press, Cambridge, 1904, pp. 241–250.