Factorizations of Cauchy–Vandermonde matrices

J.J. Martínez, J.M. Peña *

Departamento de Matemática Aplicada, Universidad de Zaragoza, Edificio de Matemáticas, 50009 Zaragoza, Spain

Received 15 September 1997; accepted 20 April 1998

Submitted by A.H. Sayed

Abstract

This paper analyzes the factorization of the inverse of a Cauchy–Vandermonde matrix as a product of bidiagonal matrices. Factorizations of Cauchy–Vandermonde matrices are also considered. The results are applied to derive a fast algorithm for solving linear systems whose coefficient matrices are Cauchy–Vandermonde matrices. © 1998 Elsevier Science Inc. All rights reserved.

1. Introduction

A Cauchy–Vandermonde matrix is an \( n \times n \) matrix \( V = (A|B) \), where the first \( l (1 \leq l \leq n) \) columns form a Cauchy matrix and the last \( n - l \) columns form a Vandermonde matrix. Cauchy Vandermonde matrices play an important role in rational interpolation approximation (cf. [6] or [5]) or numerical quadrature. They also appear in connection with the numerical solution of singular integral equations (see [12]).

Let us observe that Cauchy–Vandermonde matrices arise when computing rational interpolants with prescribed poles. A different type of unconstrained rational interpolation problem leading to different classes of structured matrices and corresponding fast algorithms is studied in [4].

* Corresponding author. E-mail: jmpena@posta.unizar.es.

Research partially supported by Spanish Research Grant DGICYT PB96-0730.

0024-3795/98/$19.00 © 1998 Elsevier Science Inc. All rights reserved.

PH: S 0 0 2 4 - 3 7 9 5 ( 9 8 ) 1 0 0 7 3 - 3
In [13] we obtained an LU factorization of a Cauchy–Vandermonde matrix which allowed us to derive a fast algorithm for solving a Cauchy–Vandermonde linear system. This algorithm generalized Björck and Pereyra algorithm [2] for solving Vandermonde linear systems, in the sense of using a Newton-type basis for the interpolation problem associated with the Cauchy–Vandermonde matrix. On the other hand, factorizations in terms of bidiagonal matrices have been useful when dealing with Vandermonde matrices (cf. Section 4.6 of [11] or [15]) and with Cauchy matrices (cf. [3]). In Theorem 2.1 we prove the existence and uniqueness of the factorization of the inverse of a Cauchy–Vandermonde matrix as a product of bidiagonal matrices. As we comment at the end of Section 2, similar results can be derived for Cauchy–Vandermonde matrices.

In Section 3 we show that the factorization of Theorem 2.1 provides a fast algorithm of $O(n^2)$ operations to solve a linear system whose coefficient matrix is an $n \times n$ Cauchy–Vandermonde matrix $V$. We also see that this method for solving a linear system $Vx = b$ also provides a method for solving the dual linear system $V^Tz = c$ ($V^T$ denotes the transpose of $V$).

The main difference between this algorithm and the algorithm proposed in [13] comes from the fact that the algorithm presented here uses the factorization in terms of bidiagonal matrices, extending to the case of Cauchy–Vandermonde matrices the results about the factorization of Cauchy matrices given in [3]. In contrast, in [13] the main tool is the construction of the triangular matrices $L$ and $U$ by using the connection between a Cauchy–Vandermonde linear system and the rational interpolation problem associated with it.

Our notation follows, in essence, that of [1,7,9]. Given $k, n \in \mathbb{N}$, $1 \leq k \leq n$, $Q_{k,n}$ 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 $k \leq n$, $m \leq n$, and for any $\alpha \in Q_{k,n}$ and $\beta \in Q_{m,n}$, we denote by $A[\alpha|\beta]$ the $k \times m$ submatrix of $A$ containing rows numbered by $\alpha$ and columns numbered by $\beta$. For brevity we shall write $A[\alpha|x] := A[\alpha|x]$. A fundamental tool for obtaining our results will be the use of an elimination method which was called Neville elimination in [7]. Neville elimination 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 $A = (a_{ij})_{1 \leq i,j \leq n}$, it consists of $n - 1$ major steps resulting in a sequence of matrices as follows:

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

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

$$ a_{ij}^{(t+1)} := \begin{cases} 
  a_{ij}^{(t)} & \text{if } i \leq t, \\
  a_{ij}^{(t)} - (a_{ii}^{(t)}/a_{i-1,i}^{(t)})a_{i-1,j}^{(t)} & \text{if } i > t + 1 \text{ and } j \geq t + 1, \\
  0 & \text{otherwise.}
\end{cases} \quad (1.1) $$
In this process the element

\[ p_{ij} := a_{ij}^{(i)} , \quad 1 \leq j \leq n , \quad j \leq i \leq n , \]

is called the \((i, j)\) pivot of the Neville elimination of \(A\). The process would break down if any of the pivots \(p_{ij} (j < i < n)\) is zero. In that case we can move the corresponding rows to the bottom and proceed with the new matrix, as described in [7]. The Neville elimination 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 \(P_{ii} = a_{ii}\) and, by Lemma 2.6 (1) of [7],

\[ p_{ij} = \frac{\det A[i-j+1, \ldots, i, 1, \ldots, j]}{\det A[i-j+1, \ldots, i, 1, \ldots, j-1]} \quad (1 < j \leq i \leq n) . \] (1.2)

The element

\[ m_{ij} = \frac{p_{ij}}{p_{i-1,j}}, \quad 1 \leq j \leq n , \quad j < i \leq n , \] (1.3)

is called the \((i, j)\) multiplier of the Neville elimination of \(A\). The matrix \(U := A\) is upper triangular and has the diagonal pivots on its main diagonal.

In some recent papers (cf. [7–10] and [14]) it has been shown that Neville elimination can be very useful when dealing with some special classes of matrices, such as totally positive matrices. A matrix is called totally positive if all its minors are nonnegative (cf. [1]). The relationship of Cauchy–Vandermonde matrices with total positivity was analyzed in Section 4 and 5 of [13]. In this paper we shall see that Neville elimination is a useful tool for studying Cauchy–Vandermonde matrices even when they are not totally positive.

2. Factorizations in terms of bidiagonal matrices

A matrix

\[ V = \begin{pmatrix} \frac{1}{c_1-d_1} & \ldots & \frac{1}{c_1-d_1} & 1 & c_1 & \ldots & c_1^{n-l} \\ \frac{1}{c_2-d_1} & \ldots & \frac{1}{c_2-d_1} & 1 & c_2 & \ldots & c_2^{n-l-1} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{1}{c_n-d_1} & \ldots & \frac{1}{c_n-d_1} & 1 & c_n & \ldots & c_n^{n-l-1} \end{pmatrix} \]

is called a Cauchy–Vandermonde matrix because if \(l = 0\) it is a classical Vandermonde matrix and if \(l = n\) it is a classical Cauchy matrix. We can interpret this matrix as the coefficient matrix of a Lagrange interpolation problem: given \(n\) different real numbers \(c_1, \ldots, c_n\) (the interpolation nodes) and \(b_1, \ldots, b_n\) (the interpolation data), find the function
\[ f(x) = \sum_{j=1}^{l} a_j \frac{1}{x - d_j} + \sum_{j=l+1}^{n} a_j x^{l-j} \]
such that \( f(c_i) = b_i \) for all \( i = 1, \ldots, n \). This problem is a rational interpolation problem of Lagrange type, whose interpolation space contains both polynomials and rational functions with \( l \) different prescribed poles \( d_1, \ldots, d_l \in \mathbb{R} \), where \( \{d_1, \ldots, d_l\} \cap \{c_1, \ldots, c_n\} = \emptyset \).

We shall also use the following well-known formula (cf. [6] or [13]) for the determinant of a Cauchy-Vandermonde matrix:

\[
\det V = \frac{\prod_{1 \leq i < j < n} (c_j - c_i) \prod_{1 \leq i < j < l} (d_i - d_j)}{\prod_{1 \leq i \leq n, 1 \leq j < l} (c_i - d_j)}.
\]  

**Theorem 2.1.** Let \( A = (a_{ij})_{1 \leq i, j \leq n} \) be a Cauchy-Vandermonde matrix whose corresponding interpolation nodes and poles satisfy \( \{d_1, \ldots, d_l\} \cap \{c_1, \ldots, c_n\} = \emptyset \) and \( c_i \neq 0 \) \( \forall i \). Then

\[
A^{-1} = G_1 G_2 \cdots G_{n-1} D^{-1} F_{n-1} F_{n-2} \cdots F_1,
\]

where Eqs. (2.4)–(2.6), Eqs. (2.8) and (2.11) hold. Moreover, this factorization is unique among factorizations of that type.

**Proof.** Let us observe that the minors of \( A \) using initial consecutive columns are Cauchy-Vandermonde determinants and so formula (2.1) shows that they are nonzero. In consequence, by Lemma 2.6 (2) of [7] all pivots of the Neville elimination of \( A \) are nonzero and this elimination can be performed without row exchanges (i.e. \( A \) satisfies the property called WR-condition in [9]). By Eq. (2.12) of [9] we can write

\[
F_{n-1} F_{n-2} \cdots F_1 A = R
\]

with \( R \) an upper triangular matrix and the matrices \( F_i \) \( (1 \leq i \leq n - 1) \) are bidiagonal matrices of the form

\[
F_i = \begin{bmatrix}
1 & 0 & 0 & \cdots & 0 \\
0 & 1 & 0 & \cdots & 0 \\
& & \ddots & \ddots & \ddots \\
& & & \ddots & \ddots \\
& & & & \ddots & 1 \\
& & & & & -m_{i+1,i} & 1 \\
& & & & & & -m_{i+2,i} & 1 \\
& & & & & & & \ddots & \ddots \\
& & & & & & & & \ddots & \ddots \\
& & & & & & & & & -m_{ni} & 1
\end{bmatrix}.
\]
Since the entries $m_{ij}$ are the multipliers of the Neville elimination of $A$, they have the form

$$m_{ij} = \frac{p_{ij}}{p_{i-1,j}} = \frac{a_{ij}^{(j)}}{a_{i-1,j}^{(j)}}.$$  

(2.5)

By Eqs. (1.2) and (2.1) we derive for $j \geq 2$

$$a_{ij}^{(j)} = \begin{cases} \frac{\prod_{k=j}^{l-1} (c_i - c_k) \prod_{k=j}^{l-1} (-d_k + d_i)}{\prod_{k=1}^{l-1} (c_k - d_k) \prod_{k=j}^{l-1} (c_i - d_k)}, & j \leq l; \\ \frac{\prod_{k=j}^{l-1} (c_i - c_k)}{\prod_{k=1}^{l-1} (c_k - d_k)}, & j > l. \end{cases} \quad (2.6)$$

Let us now see that the minors of $A^T$ using initial consecutive columns and consecutive rows are nonzero. In fact, these minors are either Cauchy–Vandermonde determinants (and we can apply to them formula (2.1)) or determinants of the form

$$\det \begin{pmatrix} c_0^k & c_1^k & \cdots & c_l-k+1 \\ c_0^{k+1} & c_1^{k+1} & \cdots & c_l-k+1 \\ \vdots & \vdots & \ddots & \vdots \\ c_0^{l-k} & c_1^{l-k} & \cdots & c_l^{l-k+1} \end{pmatrix} = c_0^k c_1^k \cdots c_l-k+1 \det M,$$

where $M^T$ is a nonsingular Vandermonde matrix. Then by Lemma 2.6 (2) of [7] all pivots of the Neville elimination of $A^T$ are nonzero and this elimination can be performed without row exchanges (i.e. $A^T$ satisfies the WR-condition in [9]). By Eq. (2.12) of [9] we can write

$$G_{n-1}^T G_{n-2}^T \cdots G_1^T A^T = V^T, \quad (2.7)$$

where $V^T$ is an upper triangular matrix and the matrices $G_i^T (1 \leq i \leq n - 1)$ are bidiagonal matrices of the form

$$G_i^T = \begin{bmatrix} 1 & & & & \\ 0 & 1 & & & \\ & \ddots & \ddots & \ddots & \\ & & 0 & 1 \\ -\tilde{m}_{i+1,i} & -\tilde{m}_{i+2,i} & \ddots & \ddots & -\tilde{m}_{ni} \end{bmatrix}. \quad (2.8)$$
Using Eqs. (2.3) and (2.7) and the uniqueness of the LDU-factorization of a nonsingular matrix we can deduce that

\[ A = F_1^{-1} \cdots F_{n-1}^{-1} DG_{n-1}^{-1} \cdots G_1^{-1}, \]

where \( D \) is a nonsingular diagonal matrix. Therefore Eq. (2.2) follows.

Applying Theorem 2.2 of [9] to \( A \) and \( A^T \) we obtain the uniqueness of the factorizations Eqs. (2.3) and (2.7), and so the uniqueness of Eq. (2.2).

Since the elements \( \tilde{m}_{ij} \) are the multipliers of the Neville elimination of \( A^T \), they have the form

\[
\tilde{m}_{ij} = \frac{(A^T)_{ij}^{(j)}}{(A^T)_{i-1,j}^{(j)}}. \tag{2.9}
\]

By Eqs. (1.2) and (2.1) we can obtain for \( j \geq 2 \)

\[
(A^T)_{ij}^{(j)} = \begin{cases}
\prod_{k=j}^{i-1} (c_j - c_k) \prod_{k=j}^{i-1} (c_j - d_k), & i \leq l, \\
\prod_{k=j}^{i-1} (c_j - c_k) \prod_{k=i-j}^{i-1} (c_j - d_k), & i - j + 1 \leq l < i, \\
\prod_{k=i-j}^{i-1} (c_j - c_k), & i - j = l, \\
c_j^{i-j-l} \prod_{k=i-j}^{i-1} (c_j - c_k), & l < i - j.
\end{cases} \tag{2.10}
\]

Finally, the diagonal matrix \( D \) in Eq. (2.2) has as \( i \)th diagonal element \((1 \leq i \leq n)\) the \( i \)th pivot \( p_{ii} = a_{ii}^{(i)} \) of the Neville elimination of \( A \):

\[
D = \text{diag}\{a_{11}^{(1)}, a_{22}^{(2)}, \ldots, a_{nn}^{(n)}\}. \tag{2.11}
\]

The expression of the elements \( a_{ii}^{(i)} \) for \( i \geq 2 \) can be found in Eq. (2.6).

In Section 3 we shall also be interested in solving the dual linear system

\[ A^T x = b, \]

where \( A \) is a Cauchy–Vandermonde matrix and \( b \in \mathbb{R}^n \). These linear systems arise in the construction of formulae of numerical integration and numerical differentiation of interpolatory type. A key tool will be the following factorization of \( (A^T)^{-1} \), which is a consequence of Theorem 2.1.

**Corollary 2.2.** Let \( A = (a_{ij})_{1 \leq i, j \leq n} \) be a Cauchy–Vandermonde matrix whose corresponding interpolation nodes and poles satisfy \( \{d_1, \ldots, d_l\} \cap \{c_1, \ldots, c_n\} = \emptyset \) and \( c_i \neq 0 \ \forall i \). Then

\[
(A^T)^{-1} = F_1 F_2^T \cdots F_{n-1}^T (D^T)^{-1} G_{n-1}^T G_{n-2}^T \cdots G_1^T, \tag{2.12}
\]

where Eqs. (2.4), (2.6), (2.8) and (2.11) hold. Moreover, this factorization is unique among factorizations of that type.
Due to our purposes for the following section, we have been interested in a factorization of $A^{-1}$ as a product of bidiagonal matrices, where $A$ is a Cauchy–Vandermonde matrix. The results of Sections 2 and 3 of [9] show that a similar factorization of $A$ could be obtained from that of $A^{-1}$. In particular, using Theorem 2.6 of [9] one can deduce that the corresponding factors of $A$ contain the numbers opposite in sign to the multipliers $m_{ij}$'s and $\tilde{m}_{ij}$'s, but occurring in a different order.

3. Fast algorithms for solving Cauchy–Vandermonde linear systems

This section is devoted to construct fast algorithms for solving Cauchy–Vandermonde linear systems and their dual systems. As we shall see, both problems are closely related.

In order to solve the linear system $Ax = b$, where $A$ is an $n \times n$ Cauchy–Vandermonde matrix, we use Theorem 2.1 obtaining

$$x = G_1 G_2 \cdots G_{n-1} D^{-1} F_{n-1} F_{n-2} \cdots F_1 b.$$ (3.1)

Since matrices $G_i$'s and $F_i$'s are bidiagonal and $D^{-1}$ is diagonal, it is clear that the whole product can be carried out with computational complexity $O(n^2)$. It remains to see that the construction of those matrices can also be carried out with computational complexity $O(n^2)$. Equivalently, by Eqs. (2.4), (2.5), (2.8), (2.9) and (2.11), it is sufficient to show that the entries $a_{ij}^{(j)}$ and $(A^T)_{ij}^{(j)}$ can be obtained in $O(n^2)$ arithmetic operations.

Let us start with the construction of the entries $a_{ij}^{(j)}$ given by Eq. (2.6). First of all, let us observe that for $j = 1$ we have $a_{11}^{(1)} = a_{11}$, i.e. the entries in the first column of $A$. So we must compute $a_{ij}^{(j)}$ for $j = 2, \ldots, n$, $i = j, j + 1, \ldots, n$.

Defining

$$B_{ij} = \prod_{k=i-j+1}^{i} (c_k - d_j), \quad j = 2, \ldots, l; \quad i = j, j + 1, \ldots, n,$$ (3.2)

$$G_j = \prod_{r=1}^{j-1} (-d_j + d_r), \quad j = 2, \ldots, l,$$ (3.3)

we have

$$a_{ij}^{(j+1)} = \begin{cases} a_{ij}^{(j)} \frac{G_{i+1}}{G_j} \frac{B_{ij} (c_i - c_{i-j})}{B_{ij} (c_i - d_j)}, & j < l, \\ a_{il}^{(l)} \frac{B_{ij} (c_i - c_{i-j})}{G_j (c_l - d_h)}, & j = l, \\ a_{ij}^{(j)} (c_i - c_{i-j}), & j > l. \end{cases}$$ (3.4)

Clearly, the numbers $G_j$ of Eq. (3.3) can be computed in $O(n^2)$ elementary operations. As for the numbers $B_{ij}$, if we begin by computing $B_{jj}$ for $j = 1, \ldots, l$, and we take into account that
\[ B_{ij} = B_{i-1,j}(c_i - d_j)/(c_{i-j} - d_j), \quad i = j + 1, \ldots, n, \]

then we conclude that the computational complexity of the construction of all \( B_j \)'s is again \( O(n^2) \).

Finally, taking into account Eq. (3.4) it can be easily seen that the construction of \( d^{(j)}_{ij} \) \((j = 2, \ldots, n, i = j, \ldots, n)\) can be carried out with computational complexity \( O(n^2) \).

A similar algorithm with computational complexity \( O(n^2) \) can be derived for obtaining the entries \( (A^T)^{(j)}_{ij} \) given by Eq. (2.10).

In order to solve the dual linear system \( A^Tx = b \) where \( A \) is an \( n \times n \) Cauchy-Vandermonde matrix, we use Corollary 2.2 and then we have

\[ x = F_1^T F_2^T \cdots F_{n-1}^T (D^T)^{-1} G_{n-1}^T G_{n-2}^T \cdots G_1^T b. \] (3.5)

Therefore, since the matrices appearing in Eq. (3.5) are simply the transpose matrices of those in Eq. (3.1), the dual system can also be solved with computational complexity \( O(n^2) \).

Let us remark that these algorithms extend to the case of Cauchy-Vandermonde matrices a bidiagonal algorithm for Cauchy matrices given in [3]. In [3] the authors show that their algorithm is stable when the Cauchy matrix is totally positive. In Proposition 4.1 of [13] we proved that a Cauchy–Vandermonde matrix \( V \) (see Eq. (2.1) for an explicit expression of its determinant) is totally positive when

\[ 0 < c_1 < c_2 < \cdots < c_n \quad \text{and} \quad 0 < -d_1 < -d_2 < \cdots < -d_n. \]

References


