SIAM J. MATRIX ANAL. APPL.
Vol. 25, No. 1, pp. 213–223

SOME RESULTS ON VANDERMONDE MATRICES WITH AN APPLICATION TO TIME SERIES ANALYSIS*

ANDRÉ KLEIN† AND PETER SPREIJ‡

Abstract. In this paper we study Stein equations in which the coefficient matrices are in companion form. Solutions to such equations are relatively easy to compute as soon as one knows how to invert a Vandermonde matrix (in the generic case where all eigenvalues have multiplicity one) or a confluent Vandermonde matrix (in the general case). As an application we present a way to compute the Fisher information matrix of an autoregressive moving average (ARMA) process. The computation is based on the fact that this matrix can be decomposed into blocks where each block satisfies a certain Stein equation.

Key words. ARMA process, Fisher information matrix, Stein's equation, Vandermonde matrix, confluent Vandermonde matrix

AMS subject classifications. 15A09, 15A24, 62F10, 62M10, 65F05, 65F10, 93B25

PII. S0895479802402892

1. Introduction.

In this paper we investigate some properties of (confluent) Vandermonde and related matrices aimed at and motivated by their application to a problem in time series analysis. Specifically, we show how to apply results on these matrices to obtain a simpler representation of the (asymptotic) Fisher information matrix of an autoregressive moving average (ARMA) process. The Fisher information matrix is prominently featured in the asymptotic analysis of estimators and in asymptotic testing theory, e.g., in the classical Cramér–Rao bound on the variance of unbiased estimators. See [10] for general results and see [2] for time series models. However, the Fisher information matrix has also attracted considerable attention in the signal processing literature, e.g., [6], [19], and [12]. We have previously shown (see [14]) that the Fisher information matrix of an ARMA process is the solution of a so-called Lyapunov equation. More precisely, although we don't go into detail about ARMA processes until section 5, the Fisher information matrix in this case can be decomposed into blocks that are solutions of equations such as

$$X + MXN^\top = R.$$

The coefficients M and N in this equation turn out to be in companion form in the given context of time series analysis, and the right-hand side R is another given matrix.

The plan of attack that we follow to solve such an equation is to break up the solution procedure into a number of steps that are each relatively easy to perform. First, we replace by a basis transformation the coefficient matrices with their Jordan forms, thereby also changing the variable matrix X and the right-hand side R. Since a basis of (generalized) eigenvectors of companion matrices can be represented as the columns of a (confluent) Vandermonde matrix, the basis transformation needed

*Received by the editors February 21, 2002; accepted for publication (in revised form) by S. Van Huffel November 15, 2002; published electronically May 29, 2003.

http://www.siam.org/journals/simax/25-1/40289.html

†Department of Actuarial Sciences and Econometrics, Universiteit van Amsterdam, Roetersstraat 11, 1018 WB Amsterdam, The Netherlands (aklein@fee.uva.nl).

‡Korteweg-de Vries Institute for Mathematics, Universiteit van Amsterdam, Plantage Muidergracht 24, 1018 TV Amsterdam, The Netherlands (spreij@science.uva.nl).

for this can be expressed in terms of the above-mentioned Vandermonde matrices. Performing the basis transformation requires knowing how to compute inverses of confluent Vandermonde matrices. One of the aims of our paper is to derive rather simple, but explicit, representations for these inverses. Of course this whole procedure would be meaningless if the equation in the new coordinate system were more complex than the original one. In section 4 we will see that, fortunately, the resulting equation is much easier to solve than the original one, especially in a generic case, where the solution becomes almost trivial. By applying the developed procedure to the computation of the Fisher information matrix for an ARMA process, we reach our goal of giving an alternative way to represent this Fisher information matrix. This application also motivates, from a statistical perspective, the interest of analyzing (confluent) Vandermonde matrices.

The remainder of the paper is organized as follows. In section 2 we introduce the basic notation that we use throughout the paper. Section 3 is devoted to technical results on companion matrices and confluent Vandermonde matrices, the main results concerning inversion of confluent Vandermonde matrices. In section 4 we apply these results to describe solutions to Stein equations in which the coefficient matrices are in companion form. Finally, in section 5 we investigate the special case where the solutions to certain Stein equations are given by blocks of the Fisher information matrix of an ARMA process.

2. Notation and preliminaries. Consider the matrix $A \in \mathbb{R}^{n\times n}$ in the companion form

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & 0 & 1 & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & & & 0 & 1 \\ -a_n & & & -a_2 & -a_1 \end{pmatrix}. \tag{1}$$

Let $a^\top = (a_1, \ldots, a_n)$, $u(z)^\top = (1, z, \ldots, z^{n-1})$, and $u^*(z)^\top = (z^{n-1}, \ldots, 1)$ (where $^\top$ denotes transposition). Define recursively the Hörner polynomials $a_k(\cdot)$ by $a_0(z) = 1$ and $a_k(z) = za_{k-1}(z) + a_k$. Notice that $a_n(z)$ is the characteristic polynomial of A. We will denote it by $\pi(z)$ and, occasionally, by $\pi_A(z)$ if we want to emphasize the role of the A-matrix.

Write $a(z)$ for the n-vector $(a_0(z), \ldots, a_{n-1}(z))^\top$. Furthermore S will denote the shift matrix, so $S_{ij} = \delta_{i,j+1}$, and P will denote the *backward* or *antidiagonal* identity matrix, so $P_{ij} = \delta_{i+j,n+1}$ (assuming that $P \in \mathbb{R}^{n\times n}$). As an example we have $Pu(z) = u^*(z)$. The matrix P has the following property: If M is a Toeplitz matrix, then $PMP = M^\top$, in particular $P^2 = I$, the identity matrix.

We associate with the vector a the matrix $T_a \in \mathbb{R}^{n\times n}$ given by

$$T_a = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ a_1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \\ a_{n-1} & \cdots & a_1 & 1 \end{pmatrix}.$$

Notice that the matrices T_a and S commute and that $a(z) = T_a u(z)$.

Denoting the kth basis vector in $\mathbb{R}^n$ by e_k, we can write

$$A = -e_n a^\top P + S^\top. \tag{2}$$

If $q(\cdot)$ is a polynomial and if for some natural number k the term $(z-\alpha)^k$ is a factor of $q(z)$ (which happens if α is a zero of $q(\cdot)$ with multiplicity greater than or equal to k), then we define the polynomial $q_k(\cdot;\alpha)$ by $q_k(z;\alpha) = \frac{q(z)}{(z-\alpha)^k}$. Notice the identity $q_k(\alpha;\alpha) = q_k^{(k)}(\alpha)/k!$. In what follows we will often use D for differentiation (w.r.t. z). For instance, instead of $\frac{d}{dz}q_k(z;\alpha)$ we then write $Dq_k(z;\alpha)$, and $Dq_k(z;\alpha)$ in $z=\alpha$ is denoted by $Dq_k(\alpha;\alpha)$. Notice also the formula

$$\pi(z) - \pi(\alpha) = (z-\alpha)u^*(z)^\top a(\alpha), \tag{3}$$

which follows from the definition of the Hörner polynomials by a direct computation.

We also need some results on Lagrange and Hermite interpolation problems. Assume we are given s pairwise different complex numbers $\alpha_1,\ldots,\alpha_s$ (so $\alpha_i \neq \alpha_j$ iff $i \neq j$) and we want to find n polynomials $p_1,\ldots,p_n$ of degree at most $n-1$ such that $p_j(\alpha_i)$ take on certain given values. Notice that we have n^2 unknown parameters to determine, but only ns conditions. Therefore we add constraints by prescribing certain values of the derivatives $p_j^{(k)}(\alpha_i)$ for $k=1,\ldots,m_i-1$, where the m_i are such that $\sum_{i=1}^s m_i = n$. In this way we obtain n^2 constraints. The total set of prescribed values of the polynomials p_j and their derivatives that we consider is given by the equations

$$\frac{p_j^{(k-1)}(\alpha_i)}{(k-1)!} = \delta_{\sum_{l=1}^{i-1} m_l + k, j},$$

where $j=1,\ldots,n$, $i=1,\ldots,s$, $k=1,\ldots,m_i$, and δ denotes the Kronecker symbol. Notice that in the case where $s=n$, all m_i are equal to 1, and we only require $p_j(\alpha_i) = \delta_{ij}$.

In order to give the solution to this interpolation problem an elegant form we present the conditions as described below. We need some notation. First, we denote by $p(z)$ the column vector $(p_1(z),\ldots,p_n(z))^\top$. For each i we denote by $\Pi(i)$ the $n \times m_i$ matrix with columns $\Pi(i)_k = \frac{p^{(k-1)}(\alpha_i)}{(k-1)!}$, with $k=1,\ldots,m_i$. The constraints are now given in compact form by the equality $(\Pi(1),\ldots,\Pi(s)) = I$, where I is the $n\times n$ identity matrix.

Write $\pi(z) = \prod_{i=1}^s (z-\alpha_i)^{m_i} = \sum_{j=0}^n a_j z^{n-j}$ and let A be the associated companion matrix of (1) so that π is its characteristic polynomial. Let $U_i(z)$ be the $n \times m_i$ matrix with kth column equal to $\frac{1}{(k-1)!}u^{(k-1)}(z)$ and write $U_i = U_i(\alpha_i)$. We define the $n\times n$ matrix V (often called the *confluent* Vandermonde matrix associated with the eigenvalues of A) by $V = (U_1,\ldots,U_s)$. Similar interpolation problems involving one polynomial only are known to have a unique solution; see e.g., [17, p. 306] or [5, p. 37]. Here the situation is similar and, as an almost straightforward result from the current setup, we have the following proposition.

PROPOSITION 2.1. *The unique solution to the interpolation problem is* $p(z) = V^{-1}u(z)$.

Write $p^*(z) = z^{n-1}p(\frac{1}{z})$ and notice that we use multiplication with the same power of z for all entries of $p(\frac{1}{z})$.

Let Π^* be defined by $\Pi^* = V^{-1}PV$. Then the matrix Π^* is involutive, i.e., $(\Pi^*)^2 = I$.

Proposition 2.2. *The polynomials p and p^* are related by*

$$p^*(z) = V^{-1}PVp(z) = \Pi^* p(z). \tag{4}$$

In particular, $p^*(0) = V^{-1}e_n$.

Proof. This follows from

$$p^*(z) = z^{n-1}V^{-1}u\left(\frac{1}{z}\right) = V^{-1}Pu(z) = V^{-1}PVp(z). \qquad \square$$

3. Confluent Vandermonde matrices. The main point of this section is to give some formulas for the inverse of a confluent Vandermonde matrix. We need some auxiliary results. First we give an expression for $\operatorname{adj}(z - A)$, where A is a companion matrix of the form (2). The next proposition is an alternative to formula (31) in [7, p. 84].

Proposition 3.1. *Let A be a companion matrix with π as its characteristic polynomial. The following equation holds true:*

$$\operatorname{adj}(z - A) = u(z)a(z)^\top P - \pi(z)\sum_{j=0}^{n-1} z^j S^{j+1}. \tag{5}$$

Proof. First we show that

$$a(z)^\top P(z - A) = \pi(z)e_1^\top. \tag{6}$$

Using (2), we have

$$\begin{aligned} a(z)^\top P(z - A) &= a(z)^\top P(z - S^\top + e_n a^\top P) \\ &= a(z)^\top (z - S + Pe_n a^\top)P \\ &= (\pi(z)e_n^\top - a^\top + a(z)^\top Pe_n a^\top)P \\ &= \pi(z)e_n^\top P, \end{aligned}$$

which gives (6). Multiply the right-hand side of (5) by $(z - A)$. First we consider $a(z)^\top P(z - A)$. In view of (6), this is just

$$\pi(z)e_1^\top. \tag{7}$$

Then we consider $\sum_{j=0}^{n-1} z^j S^{j+1}(z - A) = \sum_{j=0}^{n-1} z^{j+1}S^{j+1} + \sum_{j=0}^{n-1} z^j S^{j+1}(-S^\top + e_n a^\top P)$. Since $Se_n = 0$, this reduces to $\sum_{j=0}^{n-1} z^{j+1}S^{j+1} - \sum_{j=0}^{n-1} z^j S^{j+1}S^\top$. Now use the equality $SS^\top = I - e_1e_1^\top$ to rewrite this as $\sum_{j=0}^{n-1} z^j S^j(zS - I + e_1e_1^\top)$, which equals $\sum_{j=0}^{n-1} z^j S^j(zS - I) + \sum_{j=0}^{n-1} z^j e_{j+1}e_1^\top$. However, this is equal to $-I + u(z)e_1^\top$ because the first summation is just $-I$ and the latter one equals $u(z)e_1^\top$. Hence

$$\sum_{j=0}^{n-1} z^j S^{j+1}(z - A) = -I + u(z)e_1^\top. \tag{8}$$

So we obtain from (7) and (8) that the right-hand side of (5) multiplied by $z - A$ is equal to

$$u(z)\pi(z)e_1^\top + \pi(z)(I - u(z)e_1^\top),$$

which is $\pi(z)I$, precisely what we have to prove. ◻

For the application to time series that we have in mind, as explained in the introduction, we need the inverse of a (confluent) Vandermonde matrix. In the 1970s this was an especially popular topic and many papers appeared on the subject. Quite often attention has been paid to the finding of efficient procedures to carry out the inversion numerically. Recently, there has been a renewed interest in a related subject, the inversion of Cauchy–Vandermonde matrices. These matrices appear in rational interpolation problems and are beyond the scope of this paper.

Below we provide inversion formulas for confluent Vandermonde matrices. Some of these can be found in the older literature, but the derivation below is different. Of the many possible references we mention [11] and [4], which give results for the relatively simple case of a genuine Vandermonde matrix or, in the spirit of our Proposition 3.3 (but obtained by different methods), for a confluent Vandermonde matrix, and mention [20] which has elementwise expressions. Related results of a different nature include [9], [3], and [18].

We need the Jordan decomposition of A. We use the notation S_{m_i} to denote the shift matrix of size $m_i \times m_i$. Recall that the confluent Vandermonde matrix as we defined it is such that the columns are independent eigenvectors of A. The Jordan form of A is determined by the relation $V^{-1}AV = J_A$, and J_A is block diagonal with the ith block given by $\alpha_i I_{m_i} + S_{m_i}^\top$. As a first step toward expressions for the inverse of a Vandermonde matrix we will use the next proposition.

PROPOSITION 3.2. *Let J_A be the Jordan form of the companion matrix A. Then*

$$\mathrm{adj}(z - J_A) = p(z)a(z)^\top PV - \pi(z)V^{-1}\sum_{j=0}^{n-1} z^j S^{j+1}V. \tag{9}$$

In particular

$$\mathrm{adj}(\alpha_k - J_A) = \pi(\alpha_k)a(\alpha_k)^\top PV. \tag{10}$$

Proof. This follows from Propositions 3.1 and 2.1. ◻

Next we proceed with some results of a general nature. Let M be the block diagonal matrix with s blocks $M(i)$ of size $m_i \times m_i$ specified by

$$M(i)_{kl} = \begin{cases} \frac{1}{(k+l-m_i-1)!}D^{k+l-m_i-1}\pi_{m_i}(\alpha_i;\alpha_i) & \text{if } k+l-m_i-1 \geq 0, \\ 0 & \text{else.} \end{cases} \tag{11}$$

Notice that the $M(i)$ are symmetric Hankel matrices and that the $M(i)_{kl}$ are zero for $k + l \leq m_i$. We have for the matrices $M(i)$ the alternative expression

$$M(i) = \sum_{l=0}^{m_i-1} \delta_l S^l P, \text{ where } \delta_l = \frac{1}{l!}D^l \pi_{m_i}(\alpha_i;\alpha_i).$$

Here we denoted by S the $m_i \times m_i$ shift matrix and by P the $m_i \times m_i$ backward identity matrix.

The computation of the inverse of an $M(i)$ is simple because of its triangular structure and the fact that it is Hankel. Indeed, it is sufficient to know the first row of $M(i)^{-1}$, call it r_1, since all rows r_j are of the form $r_1 S^{j-1}$. As a matter of fact, the inverses of the matrices $M(i)$ have a particular simple structure. To clarify this we introduce, for a given $m-1$ times continuously differentiable real function f, the matrix valued function $L^f(z)$ of size $m \times m$ defined by

$$L^f_{kl}(z) = \begin{cases} \frac{1}{(k-l)!} D^{k-l} f(z) & \text{if } k \geq l, \\ 0 & \text{else.} \end{cases}$$

Notice that the matrices $L^f(z)$ are lower triangular and Toeplitz. One readily verifies that $(L^f(z))^{-1} = L^{\frac{1}{f}}(z)$ in the points z where f doesn't vanish. In particular, the last row of $(L^f(z))^{-1}$ is given by

$$\left(\frac{1}{f(z)}, \dots, \frac{1}{(m-1)!} D^{m-1} \left(\frac{1}{f(z)} \right) \right) P,$$

where P is, as above, of size $m \times m$.

Now we apply this result to $f(z) = \pi_{m_i}(z; \alpha_i)$ and $m = m_i$ to get the inverse of $M(i)$. We then have for this choice of f that $M(i) = L^f(\alpha_i)P$. The first row of $M(i)^{-1}$ is then seen to be

$$\left(\frac{1}{\pi_{m_i}(\alpha_i; \alpha_i)}, \dots, \frac{1}{(m_i-1)!} D^{m_i-1} \left(\frac{1}{\pi_{m_i}(\alpha_i; \alpha_i)} \right) \right) P. \tag{12}$$

Next we define a matrix N consisting of blocks $N(ij)$ of size $m_i \times m_j$. To do so we need some additional notation. We write $\pi^*(z) = z^n \pi(\frac{1}{z})$ and $\pi^*_k(z; \alpha) = z^{n-1} \pi_k(\frac{1}{z}; \alpha)$. Then we define the entries of the $N(ij)$ by

$$N(ij)_{kl} = \frac{1}{(k-1)!} D^{k-1} \pi^*_l(\alpha_i; \alpha_j).$$

Unfortunately, the matrix N doesn't share the nice properties (block diagonal, block Hankel, block symmetric) with the matrix M above.

PROPOSITION 3.3. *The following equalities hold:*

$$u^*(z)^\top T_a = a(z)^\top P, \tag{13}$$

$$u^*(z)^\top T_a V = (\pi_1(z; \alpha_1), \dots, \pi_{m_1}(z; \alpha_1), \dots, \pi_1(z; \alpha_s), \dots, \pi_{m_s}(z; \alpha_s)), \tag{14}$$

$$V^\top P T_a V = M, \tag{15}$$

$$u(z)^\top T_a V = (\pi^*_1(z; \alpha_1), \dots, \pi^*_{m_1}(z; \alpha_1), \dots, \pi^*_1(z; \alpha_s), \dots, \pi^*_{m_s}(z; \alpha_s)), \tag{16}$$

$$V^\top T_a V = N, \tag{17}$$

$$V^{-1} = M^{-1} V^\top P T_a = M^{-1} (T_a V)^\top P. \tag{18}$$

Proof. The equality (13) is the result of the string $u^*(z)^\top T_a = u(z)^\top P T_a = u(z)^\top T_a^\top P = a(z)^\top P$.

We continue with showing (14). Consider (3) and differentiate k times w.r.t. α. We obtain $-D^k \pi(\alpha) = u^*(z)^\top ((z-\alpha) D^k a(\alpha) - k D^{k-1} a(\alpha))$.

If α is a zero with multiplicity m, then $D^k \pi(\alpha) = 0$ for $k \leq m-1$. So we get the system of equations $0 = u^*(z)^\top ((z-\alpha) D^k a(\alpha) - k D^{k-1} a(\alpha))$ for $1 \leq k \leq m-1$ and $\pi(z) = (z-\alpha) u^*(z)^\top a(\alpha)$. Now write $q_k(z) = u^*(z)^\top D^k a(\alpha)$, then $q_0(z) = \frac{\pi(z)}{z-\alpha}$, and we have the recursive system of equations $0 = (z-\alpha) q_k(z) - k q_{k-1}(z)$ for $k = 1, \dots, m-1$. Solving this system yields $q_k(z) = k! \frac{\pi(z)}{(z-\alpha)^{k+1}} = k! \pi_{k+1}(z; \alpha)$. In other words, we find

$$u^*(z)^\top D^k a(\alpha) = k! \pi_{k+1}(z; \alpha). \tag{19}$$

Consider now $a(w) = T_a u(w) = T_a V p(w)$, where p is the interpolation polynomial. Then we also have $u^*(z)^\top a(w) = u^*(z)^\top T_a V p(w)$. Take in this equation derivatives w.r.t. w, substitute α_i for w, and use the definition of the interpolation polynomial to get

$$u^*(z)^\top T_a V = \left(a(\alpha_1), \ldots, \frac{D^{m_1-1}a(\alpha_1)}{(m_1-1)!}, \ldots, a(\alpha_s), \ldots, \frac{D^{m_s-1}a(\alpha_s)}{(m_s-1)!}\right). \tag{20}$$

Combining (19) and (20) yields (14).

To prove (15) we start from (14). Take the appropriate jth order derivatives, divide by $j!$, and substitute the α_i in the resulting expression. Doing so results in a block diagonal matrix, with the $M(i)$ on the diagonal.

Equation (16) immediately follows from (14) by definition of the polynomials $\pi_k^*(z;\alpha)$.

The proof of (17) completely parallels that of (15) and is therefore omitted. Now we turn to (18). First we observe that all the matrices $M(i)$ are invertible because of their triangular structure and the nonzero elements $\pi_{m_i}(\alpha_i;\alpha_i)$ (α_i had multiplicity m_i) on the antidiagonal. Therefore M also is invertible and, taking inverses in (15), yields the first equality of (18). The second then follows from $PT_a = T_a^\top P$. ∎

Remark 3.4. The most important formula of Proposition 3.3 is (18), which gives an expression for the inverse of the confluent Vandermonde matrix. We see that the only inversion that has to be carried out is that of M. For that we have (12) at our disposal.

COROLLARY 3.5. *The matrices M and N are related through the identities*

$$M = N^\top \Pi^*, \tag{21}$$

$$N = (\Pi^*)^\top M. \tag{22}$$

Moreover $NM^{-1} = MN^{-1}$, and thus NM^{-1} is involutive.

Proof. From (17) we get $V^{-\top} = T_a V N^{-1}$, and hence $V^\top P V^{-\top} N = V^\top P T_a V$ and, in view of (15), this equals M. Now Π^* was defined as $\Pi^* = V^{-1}PV$, so we get $(\Pi^*)^\top N = M$ and, since M is symmetric, we obtain (21). However, we also have $N = (\Pi^*)^{-\top} M = (\Pi^*)^\top M$ since Π^* is involutive. For the same reason the final assertion of the corollary follows. ∎

In the next proposition we present integral representations for the matrices M and M^{-1}. Below we use the notation $u_{m_i}(z)^\top = (1, z, \ldots, z^{m_i-1})$ and $u_{m_i}^*(z)^\top = (z^{m_i-1}, \ldots, z, 1)$, and the Γ_{α_i} are sufficiently small contours around α_i.

PROPOSITION 3.6. *The following integral representations for the matrices $M(i)$ and $M(i)^{-1}$ are valid:*

$$M(i) = \frac{1}{2\pi i}\oint_{\Gamma_{\alpha_i}} u_{m_i}^*(z-\alpha_i) u_{m_i}^*(z-\alpha_i)^\top \frac{\pi(z)}{(z-\alpha_i)^{2m_i}} dz, \tag{23}$$

$$M(i)^{-1} = \frac{1}{2\pi i}\oint_{\Gamma_{\alpha_i}} u_{m_i}(z-\alpha_i) u_{m_i}(z-\alpha_i)^\top \frac{1}{\pi(z)} dz. \tag{24}$$

As we have previously noticed, $M(i)^{-1}$ is completely determined by its first row (or column). From Proposition 3.6 we get, using Cauchy's theorem, that this first row is given by

$$\frac{1}{2\pi i}\oint_{\Gamma_{\alpha_i}} u_{m_i}(z-\alpha_i)^\top \frac{1}{\pi(z)} dz = \left(\frac{1}{\pi_{m_i}(\alpha_i;\alpha_i)}, \ldots, \frac{1}{(m_i-1)!} D^{m_i-1} \frac{1}{\pi_{m_i}(\alpha_i;\alpha_i)}\right) P,$$

in agreement with what we already found in (12).

4. Application to Stein equations. The goal of this section is to obtain a way to compute the solution of Stein's equation, where the coefficients are matrices in companion form. Apart from its interest this is chiefly motivated by the computation of Fisher's information matrix of an ARMA process. As we stated in the introduction, the blocks of Fisher's information matrix are solutions to such a Stein equation; see [14]. We postpone the application to ARMA processes until section 5.

Let A be a complex matrix of size $n \times n$ (not necessarily in companion form). If f is a $\mathbb{C}^{n\times l}$ valued analytic function, then we define $f(A)$ as $\sum_{k=0}^{\infty} \frac{1}{k!} A^k f^{(k)}(0)$. We use the following known result (see, for instance, [17, section 9.9, Theorem 2]).

LEMMA 4.1. *Let A be a complex matrix $(n \times n)$ whose eigenvalues lie strictly inside the unit disk Γ. Then for a $\mathbb{C}^{n\times l}$ valued analytic function f one has*

$$\frac{1}{2\pi i} \oint_{\Gamma} (z - A)^{-1} f(z) dz = f(A).$$

As an application of Lemma 4.1 we solve the Stein equation. Given matrices A, C, and H of appropriate dimensions (we also assume that the eigenvalues of both A and C lie inside the unit disk), we are looking for the solution for S of

$$S - ASC^{\top} = H. \tag{25}$$

This equation is of interest in matrix and operator theory (e.g., the operator that takes S to $S - ASC$ is called a displacement operator; see [8]). In [15] we study this equation further and relate solutions of various Stein equations to a certain Fisher information matrix.

The solution to (25) (see [16]) is given by $\frac{1}{2\pi i} \oint_{\Gamma} (z - A)^{-1} f(z) dz$, with $f(z) = H(I - zC)^{-\top}$, and hence is equal to $\sum_{k=0}^{\infty} A^k H (C^{\top})^k$.

We continue with presenting an alternative way to obtain a solution for the special case where both the matrices A and C are in companion form. Let V_A be the Vandermonde matrix associated with A and let V_C be associated with C. Let $\hat{S} = V_A^{-1} S V_C^{-\top}$ and $\hat{H} = V_A^{-1} H V_C^{-\top}$. The results of section 3 on inverses of confluent Vandermonde matrices enable us to compute $\hat{H}$.

Premultiplication of (25) with V_A^{-1}, together with postmultiplication with $V_C^{-\top}$, results in

$$\hat{S} - J_A \hat{S} J_C^{\top} = \hat{H}, \tag{26}$$

where J_A and J_C are the Jordan forms of A and B, respectively.

Let $v = \mathrm{vec}\,(\hat{S})$ and $h = \mathrm{vec}\,(\hat{H})$. Then it is known (see [16]) that v is given by $v = (I - J_C \otimes J_A)^{-1} h$ under the assumption that no product of an eigenvalue of A and an eigenvalue of C equals 1. This assumption is typically fulfilled in the context of stationary and invertible ARMA processes, where these eigenvalues are the zeros of both AR- and MA-polynomials and thus lie inside the unit circle; see section 5.

The computation of the inverse of the matrix $I - J_C \otimes J_A$ can now be done in an efficient way. Let $J_{A,i}$ be the Jordan block of J_A associated with the eigenvalue α_i and let $J_{C,j}$ be the Jordan block of J_C associated with the eigenvalue γ_j. Then $I - J_C \otimes J_A$ is block diagonal with diagonal blocks $I - J_{C,j} \otimes J_{A,i}$. Moreover, these blocks are upper triangular and even almost block diagonal. On the diagonal we find the blocks $I - \gamma_j J_{A,i}$ and on the subdiagonal just above it find the blocks $-J_{A,i}$. Therefore, $(I - J_{C,j} \otimes J_{A,i})^{-1}$ is again upper triangular with, on the diagonal, the blocks $(I - \gamma_j J_{A,i})^{-1}$ and, on the kth subdiagonal above it ($k \leq m_j - 1$ with m_j the multiplicity of γ_j), one finds the blocks $(I - \gamma_j J_{A,i})^{-k-1} J_{A,i}^k$. Finally, the inverses of the $I - \gamma_j J_{A,i}$ are upper triangular Toeplitz matrices with kl-element given by $\gamma_j^{k-l}(1-\alpha_i\gamma_j)^{-k+l-1}$ for $k \geq l$.

The generic case is that in which all the eigenvalues of A and all the eigenvalues of C have multiplicity 1. Consequently the matrices J_A and J_C are diagonal. In this case (26) has a very simple solution: $\hat{S}$ has elements $\hat{S}_{ij} = \frac{1}{1-\alpha_i\gamma_j}\hat{H}_{ij}$.

5. Application to ARMA processes. Consider an ARMA(p,q) process y, a stationary discrete time stochastic process that satisfies

$$y_t + a_1 y_{t-1} + \cdots + a_p y_{t-p} = \varepsilon_t + c_1 \varepsilon_{t-1} + \cdots + c_q \varepsilon_{t-q}, \tag{27}$$

where ε is a Gaussian white noise sequence with unit variance. The real constants $a_1, \ldots, a_p$ and $c_1, \ldots, c_q$ will be fixed throughout the rest of this section.

Introduce the monic polynomials $a(z) = \sum_{i=0}^p a_{p-i} z^i$ and $c(z) = \sum_{i=0}^q c_{q-i} z^i$ and let a^* and c^* be the corresponding reciprocal polynomials so that $a^*(z) = \sum_{i=0}^n a_i z^i$ and $c^*(z) = \sum_{i=0}^q c_i z^i$. We make the common assumption that the ARMA process is causal and invertible, meaning that a and c have their zeros strictly inside the unit circle [2, Chapter 3].

Write $\theta = (a_1, \ldots, a_p, c_1, \ldots, c_q)^\top$. Notice that the observations y (given random variables or their realized values) of course don't depend on the parameter θ, but then the noise sequence ε does. The Fisher information matrix $F_n(\theta)$ for n observations is defined (see [1]) as the covariance matrix of the score function and, because of the assumed Gaussian distribution of ε, it is asymptotically equal to n times the stationary Fisher information matrix

$$F(\theta) = \mathbb{E}_\theta \frac{\partial \varepsilon}{\partial \theta} \frac{\partial \varepsilon}{\partial \theta}^\top,$$

where $\mathbb{E}_\theta$ denotes expectation under the parameter θ. Knowledge of the Fisher information matrix is crucial for asymptotic statistical analysis. For instance, it is known (see, e.g., [2]) that maximum likelihood estimators of the parameters of an ARMA process are consistent and have (using n observations) an asymptotic covariance matrix that is n^{-1} times the inverse (provided that it exists) of the stationary Fisher information matrix. The inverse exists if the polynomials a and c have no common zeros; see [13].

The matrix $F(\theta)$ has a representation in the spectral domain given by the block decomposition

$$F(\theta) = \begin{pmatrix} F_{aa} & F_{ac} \\ F_{ac}^\top & F_{cc} \end{pmatrix}, \tag{28}$$

where the matrices appearing here have the elements

$$F_{aa}^{jk} = \frac{1}{2\pi i} \oint_{|z|=1} \frac{z^{j-k+p-1}}{a(z)a^*(z)} dz, (j,k = 1, \ldots, p),$$
$$F_{ac}^{jk} = \frac{1}{2\pi i} \oint_{|z|=1} \frac{z^{j-k+q-1}}{c(z)a^*(z)} dz, (j = 1, \ldots, p, k = 1, \ldots, q),$$
$$F_{cc}^{jk} = \frac{1}{2\pi i} \oint_{|z|=1} \frac{z^{j-k+q-1}}{c(z)c^*(z)} dz, (j,k = 1, \ldots, q).$$

With $k(z) = a(z)a^*(z)c(z)c^*(z)$, $u_p(z) = (1, \ldots, z^{p-1})^\top$, $u_q(z)$ likewise, and u_p^* and u_q^* their reciprocal polynomials, we have the following compact expression for the whole Fisher information matrix:

$$F(\theta) = \frac{1}{2\pi i} \oint_{|z|=1} \frac{1}{k(z)} \begin{pmatrix} c^*(z)u_p(z) \\ -a^*(z)u_q(z) \end{pmatrix} \begin{pmatrix} c(z)u_p^*(z)^\top & -a(z)u_q^*(z)^\top \end{pmatrix} dz. \tag{29}$$

As in section 2 we let $A \in \mathbb{R}^{p\times p}$ be the companion matrix associated with the polynomial $a(\cdot)$ (its precise form is given by (1) for $n = p$). The matrix $C \in \mathbb{R}^{q\times q}$ associated with the polynomial $c(\cdot)$ has an analogous form.

Let the matrix $\tilde{A} \in \mathbb{R}^{(p+q)\times(p+q)}$ be given by

$$\tilde{A} = \begin{pmatrix} A & 0 \\ 0 & C \end{pmatrix}.$$

In [14] we showed that the Fisher information matrix $F(\theta)$ is the solution of the Stein equation

$$F(\theta) - \tilde{A}F(\theta)\tilde{A}^\top = ee^\top, \tag{30}$$

where $e^\top = (e_{pp}^\top, e_{qq}^\top)$ with e_{pp} the pth standard basis vector in $\mathbb{R}^p$ and e_{qq} the qth standard basis vector in $\mathbb{R}^q$. Using for $F(\theta)$ the block decomposition (28), we see that each of the blocks involved satisfies a Stein equation with appropriate coefficients. For instance, for $F_{ac} \in \mathbb{R}^{p\times q}$ we have

$$F_{ac} - AF_{ac}C^\top = H_{ac}, \tag{31}$$

with $H_{ac} = e_{pp}e_{qq}^\top$. As we already announced in the introduction, (31) as well as the analogous equation for the other blocks of Fisher's information matrix motivated the study of solutions to Stein's equation, in which the coefficient matrices are in companion form.

We apply the results of the previous sections as follows. Let V_A be a matrix whose columns are the generalized eigenvectors of A, and let V_C be the corresponding matrix for C. As we have seen, these matrices are confluent Vandermonde matrices. By J_A and J_C we denote the Jordan forms of A and C, respectively. Let also $\hat{F}_{ac} = V_A^{-1}F_{ac}V_C^{-\top}$ and $\hat{H}_{ac} = V_A^{-1}H_{ac}V_C^{-\top}$. Then we can replace (31) with the equivalent equation

$$\hat{F}_{ac} - J_A\hat{F}_{ac}J_C^\top = \hat{H}_{ac}. \tag{32}$$

A little more can be said. The matrix $\hat{H}_{ac}$ here becomes $V_A^{-1}e_{pp}(V_C^{-1}e_{qq})^\top$ and we observe that both $V_A^{-1}e_{pp}$ and $V_C^{-1}e_{qq}$ are the last columns of the inverse of a Vandermonde matrix. We have already seen in section 2 how these columns are related to interpolation polynomials. We have, for instance, that $V_A^{-1}e_{pp}$ is equal to $p_A^*(0)$, where $p_A^*(z) = z^{p-1}p_A(\frac{1}{z})$ and p_A is the interpolation polynomial related to the eigenvalues of A as described in Proposition 2.2. Likewise $V_C^{-1}e_{qq} = p_C^*(0)$.

Let us finish by considering the generic case of Fisher's information matrix; i.e., we assume that A and C only have eigenvalues of multiplicity 1. It then follows that $\hat{F}_{ac}$ has as its ijth element

$$\frac{p_A^*(0)_i p_C^*(0)_j}{1 - \alpha_i\gamma_j}.$$

Now it is easy to compute $F_{ac} = V_A\hat{F}_{ac}V_C^\top$. To the other blocks of the Fisher information matrix the same procedure applies.

Acknowledgment. We thank a referee for many detailed suggestions that helped us in writing the revision.

REFERENCES

[1] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time Series Analysis*, 3rd ed., Prentice Hall Inc., Englewood Cliffs, NJ, 1994.

[2] P. J. Brockwell and R. A. Davis, *Time Series: Theory and Methods*, 2nd ed., Springer-Verlag, Berlin, New York, 1991.

[3] F. C. Chang, *The inverse of the generalized Vandermonde matrix through the partial fraction expansion*, IEEE Trans. Automat. Control, AC-19 (1974), pp. 151–152.

[4] F. G. Csáki, *Some notes on the inversion of confluent Vandermonde matrices*, IEEE Trans. Automat. Control, AC-20 (1975), pp. 154–157.

[5] P. J. Davis, *Interpolation and Approximation*, Dover Publications Inc., New York, 1975.

[6] B. Friedlander, *On the computation of the Cramér-Rao bound for ARMA parameter estimation*, IEEE Trans. Acoust. Speech Signal Process., 32 (1984), pp. 721–727.

[7] F. R. Gantmacher, *The Theory of Matrices*, Vol. 1, Chelsea Publishing, New York, 1959.

[8] I. Gohberg and V. Olshevsky, *The fast generalized Parker-Traub algorithm for inversion of Vandermonde and related matrices*, J. Complexity, 13 (1997), pp. 208–234.

[9] I. C. Göknar, *Obtaining the inverse of the generalized Vandermonde matrix of the most general type*, IEEE Trans. Automat. Control, AC-18 (1973), pp. 530–532.

[10] I. A. Ibragimov and R. Z. Has'minskiĭ, *Statistical Estimation. Asymptotic Theory*, Springer-Verlag, New York, 1981.

[11] I. Kaufman, *The inversion of the Vandermonde matrix and the transformation to the Jordan canonical form*, IEEE Trans. Automat. Control, AC-14 (1969), pp. 774–777.

[12] A. Klein and G. Mélard, *Computation of the Fisher information matrix for SISO models*, IEEE Trans. Signal Process., 42 (1994), pp. 684–688.

[13] A. Klein and P. J. C. Spreij, *On Fisher's information matrix of an ARMAX process and Sylvester's resultant matrices*, Linear Algebra Appl., 237/238 (1996), pp. 579–590.

[14] A. Klein and P. J. C. Spreij, *On Fisher's information matrix of an ARMA process*, in Stochastic Differential and Difference Equations, Birkhäuser Boston, Boston, MA, 1997, pp. 273–284.

[15] A. Klein and P. J. C. Spreij, *On Stein's equation, Vandermonde matrices and Fisher's information matrix of time series processes.* I. *The autoregressive moving average process*, Linear Algebra Appl., 329 (2001), pp. 9–47.

[16] P. Lancaster and L. Rodman, *Algebraic Riccati Equations*, Oxford Science Publications, The Clarendon Press, Oxford University Press, New York, 1995.

[17] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*, 2nd ed., Academic Press, Orlando, FL, 1985.

[18] L. Lupaş, *On the computation of the generalized Vandermonde matrix inverse*, IEEE Trans. Automat. Control, AC-20 (1975), pp. 559–561.

[19] B. Porat and B. Friedlander, *Computation of the exact information matrix of Gaussian time series with stationary random components*, IEEE Trans. Acoust. Speech Signal Process., 34 (1986), pp. 118–130.

[20] R. H. Schappelle, *The inverse of the confluent Vandermonde matrix*, IEEE Trans. Automat. Control, AC-17 (1972), pp. 724–725.