On the application of Vandermonde matrices to time series analysis

André Klein
Department of Actuarial Sciences and Econometrics
Universiteit van Amsterdam
Roetersstraat 11
1018 WB Amsterdam

Peter Spreij
Korteweg-de Vries Institute for Mathematics
Universiteit van Amsterdam
Plantage Muidergracht 24
1018 TV Amsterdam

November 16, 2000

Abstract

In this paper we present a way to compute the Fisher information matrix of an ARMA process. The computation is based on the fact that this matrix satisfies a Stein equation. Solutions of the Stein equation are relatively easy to compute as soon as one knows how to invert a Vandermonde matrix (in the generic case where all zeros and poles of the transfer function have multiplicity one) or a confluent Vandermonde matrix (in the general case).

Keywords ARMA process, Fisher information matrix, Stein's equation, Vandermonde matrix, confluent Vandermonde matrix

AMS classification 15A09, 15A24, 62F10, 62M10, 65F05, 65F10, 93B25

1 Introduction

In this paper we investigate some properties of Vandermonde and some related matrices aimed at application of them to a problem in time series analysis. Specifically we show how to apply results on these matrices to the computation of the (asymptotic) Fisher matrix of an ARMA process. This matrix prominently features in the asymptotic analysis of estimators and in asymptotic testing theory.
We have previously shown, see [13], that this Fisher matrix is the solution of a so-called Stein equation. The coefficients in this equation turn out to be matrices in companion form and a basis of eigenvectors of companion matrices can be represented as the columns of a (confluent) Vandermonde matrix. Therefore, also from a statistical perspective there is an interest in analyzing (confluent) Vandermonde matrices.

The remainder of the paper is organised as follows. In section 2 we introduce the basic notations 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 Vandermonde matrices. Finally in section 4 we apply these results to describe solutions to Stein equations and investigate the special case where the solutions to certain Stein equations are given by blocks of the Fisher matrix of an ARMA process.

2 Notation and conventions

Consider the matrix $A \in \mathbb{R}^{n\times n}$ in the following 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 (Hörner recursion) recursively the polynomials $a_k(z)$ 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 $\pi(z)$, 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 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.
We denote the k-th basis vector in $\mathbb{R}^n$ by e_k. Then we can write $A = -e_n a^\top P + S^\top$ and $a(z) = T_a u(z)$.

If $q(.)$ be 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 following identity: $q_k(\alpha;\alpha) = k! q_k^{(k)}(\alpha)$. In the sequel 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 $Dq_k(\alpha;\alpha)$.

3 Companion and confluent Vandermonde matrices

3.1 Formulas related to companion matrices

Lemma 3.1 *Let A be a companian matrix as in equation (1) and let $\pi(\cdot)$ be its characteristic polynomial. We have from the Hörner recursion the following consequences.*

$$a(z)^\top (z - S) = \pi(z) e_n^\top - a^\top \tag{2}$$

and

$$a(z)^\top P(z - A) = \pi(z) e_1^\top. \tag{3}$$

Furthermore

$$\pi(z) - \pi(\alpha) = (z-\alpha) u^*(z)^\top a(\alpha). \tag{4}$$

Proof. Equation (2) is a straightforward consequence of the Hörner recursion and the fact that $a_n(z) = \pi(z)$.
To prove (3) we first notice that $PSP = S^\top$, because S is Toeplitz. Then, using the representation of A, we have

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

where the last equality follows from $a(z)^\top P e_n = 1$.
Finally, equation (4) is well known from the theory of divided differences. $\square$

Next we give some formulas for $\mathrm{adj}(z-A)$. The next proposition is an alternative to formula (31) on [8, page 84].

Proposition 3.2 *The following equations hold true.*

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

and

$$\mathrm{adj}(z-A) = u(z)a(z)^\top P - \pi(z) S \mathcal{R}(S, u(z)), \tag{6}$$

where in general $\mathcal{R}(A, b)$ denotes the reachability matrix of the matrix A and the vector or matrix b.

Proof. We multiply the right hand side of (5) by $(z-A)$. First we consider $a(z)^\top P(z-A)$. In view of (3), 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 $\sum_{j=0}^{n-1} z^{j+1}S^{j+1} - \sum_{j=0}^{n-1} z^j S^{j+1} S^\top$. Use now the equality $SS^\top = I - e_1 e_1^\top$ to rewrite this as $\sum_{j=0}^{n-1} z^j S^j (zS - I + e_1 e_1^\top)$, which equals $\sum_{j=0}^{n-1} z^j S^j (zS - I) + \sum_{j=0}^{n-1} z^j S^j e_{j+1} e_1^\top$. But 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 equations (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. Equation (6) is just another way of writing (5). $\square$

Proposition 3.3 *The following relations hold.*

$$D^k a(z) = k!(I - zS)^{-1} S D^{k-1} a(z) = k!(I - zS)^{-k} S^k a(z). \tag{9}$$

Proof. First we consider the special case $a(z) = u(z)$ and we proceed by induction. Let $k = 1$. One readily verifies that $(I - zS)u(z) = e_1$, hence

$D((I-zS)u(z))=0$. From the product rule for differentiation it then follows that $0=-Su(z)+(I-zS)Du(z)$, which is equivalent to $Du(z)=(I-zS)^{-1}Su(z)$.
Using the induction step and the above result for $k=1$, we get $D^ku(z)=D((k-1)!(I-zS)^{-k+1}S^{k-1}u(z))=(k-1)!((k-1)(I-zS)^{-k}S+(I-zS)^{-k+1}(I-zS)-1S)S^{k-1}u(z)=k!(I-zS)^{-k}S^ku(z)$, what we had to prove.
Finally, we prove the result for arbitrary $a(z)$. Write it as $a(z)=T_au(z)$ and notice that because of its special structure the matrix $T_a=\sum_{k=0}^{n-1}a_kS^k$ commutes with S. Then $D^ka(z)=T_aD^ku(z)=T_a(k!(I-zS)^-kS^ku(z)=k!(I-zS)^{-k}S^kT_au(z)$. The result follows. $\square$

From Cauchy's integral formula (see [5, page 73]) we then obtain

Corollary 3.4

$$(a(\alpha),\frac{Da(\alpha)}{1!},\dots,\frac{D^{n-1}a(\alpha)}{(n-1)!})=\frac{1}{2\pi i}\oint_{\Gamma_\alpha}\frac{1}{z-\alpha}\mathcal{R}(S(I-zS)^{-1},a(z))dz, \quad (10)$$

where Γ_α is a sufficiently small closed contour around α.

3.2 Lagrange interpolation

Assume to be given n distinct complex numbers α_i. The problem is then to find n polynomials p_i of degree at most $n-1$ such that $p_i(\alpha_j)=\delta_{ij}$. Let $p(z)=(p_1(z),\dots,p_n(z))^\top$. Notice that we have n^2 conditions and equally many unknowns (n parameters for each of the n polynomials p_i). It is well known (see e.g.[7, section 2.10] or [16, section 9.2] for an equivalent form) that $p(z)$ is given by

$$p(z)=V(\alpha)^{-1}u(z), \quad (11)$$

where $V(\alpha)=(u(\alpha_1),\dots,u(\alpha_n))$, the Vandermonde matrix associated with the α_i.
The situation becomes slightly more complicated if we look at the generalized interpolation problem below. Now we are given s different complex numbers $\alpha_1,\dots,\alpha_s$ and we want to find n polynomials $p_1,\dots,p_n$ of degree at most $n-1$ such that the $p_i(\alpha_j)$ take on certain given values. Notice that we still have n^2 unknown parameters to determine, but only ns conditions. Therefore we add constraints by prescribing certain values of $p_i^{(k)}(\alpha_j)$ for $k=1,\dots,m_j-1$, where the m_j are such that $\sum_{j=1}^s m_j=n$. Notice that in this way we again have n^2 constraints.
In order to give the solution to this interpolation problem an elegant form we present the conditions as described below. We need some notation. For each i we denote by $\Pi(i)$ the $n\times m_i$ matrix with columns $\Pi(i)_k=\frac{p^{(k-1)}(\alpha_k)}{(k-1)!}$, with $k=0,\dots,m_j-1$. The contraints are now given by the equality $(\Pi(1),\dots,\Pi(s))=I$. Notice that if all the m_i are equal to 1, we are back in the first situation.
Write $\pi(z)=\prod_{i=1}^s(z-\alpha_i)^{m_i}=\sum_{j=0}^n a_jz^{n-j}$ and let A be the associated

companion matrix of equation (1), so that π is its characteristic polynomial. Let $U_i(z)$ be the $n \times m_i$ matrix with $k-th$ column equal to $\frac{1}{(k-1)!}u^{(k-1)}(z)$ and write $U_i = U_i(\alpha_i)$ and let V be the $n \times n$ matrix (often called the *confluent* Vandermonde matrix associated with the eigenvalues of A) defined by $V = (U_1, \dots, U_s)$. An integral representation of V can be given using equation (10) for each of the U_i.

Like in the case were all the α_i $(i = 1, \dots, n)$ were different we have a similar expression for the solution of the interpolation problem.

Proposition 3.5 *The unique solution is* $p(z) = V^{-1}u(z)$.

The result is known and almost trivial now. For completeness, we give the simple proof. Enumerate the standard basis of $\mathbb{R}^n$ as $e_{11}, \dots, e_{1m_1}, \dots, e_{s1}, \dots, e_{sm_s}$. By definition of V we have $Ve_{kj} = \frac{1}{(j-1)!}D^{j-1}u(\alpha_k)$. On the other hand for $p(z)$ defined by $p(z) = V^{-1}u(z)$ we have $V\pi^{(j-1)}(\alpha_k) = D^{j-1}u(\alpha_k)$, hence $\pi^{(j-1)}(\alpha_k) = (j-1)!e_{kj}$, since V is non-singular. Hence this $p(z)$ satisfies the conditions.

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 3.6 *The polynomials p and p^* are related by*

$$p^*(z) = V^{-1}PVp(z) = \Pi^* p(z). \tag{12}$$

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

Proof. This follows from

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

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

3.3 Inverse of the Vandermonde matrix

For the application to time series that we have in mind, we need the inverse of a (confluent) Vandermonde matrix. Especially in the 1970's this was a 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 to a related subject, the inversion of Cauchy-Vandermonde matrices. These appear in rational interpolation problems and are beyond the scope of the present paper.
In this subsection we provide inversion formulae 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], [6] that give

results in the spirit of our corollary 3.9 and proposition 3.11 (but obtained by different methods) and [19] who has elementwise expressions. Related results of a different nature can be found in e.g. [10], [3] or [17].
We need the Jordan decomposition of A. We use the notation S_{m_i} to denote the shift matrix of size $m_i \times m_i$.

Proposition 3.7 *The following relations hold true.*

1. $(A-\alpha_i)u(\alpha_i) = \alpha_i u(\alpha_i)$ *and for* $j = 1, \dots, m_i - 1$ *one has*
$$(A-\alpha_i)D^j U(\alpha_i) = jD^{j-1}U(\alpha_i).$$
2. $(A-\alpha_i)U_i = U_i S_{m_i}^\top$.
3. *The Jordan form of* A *is given by the block diagonal matrix* J_A *with* i*-th block given by* $\alpha_i I_{m_i} + S_{m_i}^\top$, *so* $V^{-1}AV = J_A$.

Proof. The first assertion implies the other two. The expression for the Jordan form is essentially exercise 22 on page 69 of [16]. □

As a first step towards expressions for the inverse of a Vandermonde matrix we will use

Proposition 3.8 *Let* J_A *be the Jordan form of the companion matrix* A. *Then*

$$\mathrm{adj}(z - J_A) = p(z)a(z)^\top PV - \pi(z)V^{-1}S\mathcal{R}(S, u(z))V. \tag{13}$$

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

Proof. This follows from propositions 3.2 and 3.5. □

As an easy corollary (see also [18],[11],[20] or [4, page 64]) we obtain the inverse of a genuine Vandermonde matrix.

Corollary 3.9 *If all the* α_k *are different, then*

$$V^{-1} = \mathrm{diag}\,(\pi(\alpha_k;\alpha_k))^{-1} \begin{pmatrix} a(\alpha_1)^\top \\ \vdots \\ a(\alpha_n)^\top \end{pmatrix} P. \tag{14}$$

Proof. Observe first that in this case J_A is diagonal and so is $\mathrm{adj}(z - J_A)$, with elements $\pi(z;\alpha_k)$ and notice that $\pi(\alpha_j;\alpha_k)$ for $j \neq k$. Then we obtain from the above proposition and the defining properties of the interpolation polynomials that $\pi(\alpha_k;\alpha_k)e_k e_k^\top = e_k a(\alpha_k)^\top PV$. Summing these equations over k and taking inverses on both sides yields the assertion of the corollary. □

The other extreme case of a confluent Vandermonde matrix is the one where all the zeros of π coincide. Also this case is relatively simple to analyze and we have the following result.

Proposition 3.10 *Assume that π has one zero α (of multiplicity n). Then the inverse of V is given by*

$$V^{-1} = \begin{pmatrix} \frac{1}{(n-1)!} a^{(n-1)}(\alpha)^\top \\ \vdots \\ a(\alpha)^\top \end{pmatrix} P. \tag{15}$$

Proof. In this case one easily verifies that $\mathrm{adj}(z - J_A) = \sum_{l=0}^{n-1} (z-\alpha)^l (S^\top)^{n-1-l}$ and so

$$D^k \mathrm{adj}(\alpha - J_A) = k!(S^\top)^{n-1-l}. \tag{16}$$

On the other hand we obtain from equation (13) that

$$D^k \mathrm{adj}(\alpha - J_A) = \sum_{l=0}^{k} \binom{k}{l} D^l p(\alpha) D^{k-l} a(\alpha)^\top PV,$$

which by definition of the interpolation polynomial p results in

$$D^k \mathrm{adj}(\alpha - J_A) = \sum_{l=0}^{k} l! \binom{k}{l} e_{l+1} D^{k-l} a(\alpha)^\top PV. \tag{17}$$

Comparison of equations (16) and (17) yields

$$(S^\top)^{n-1-k} = \sum_{l=0}^{k} \frac{1}{(k-l)!} e_{l+1} D^{k-l} a(\alpha)^\top PV.$$

Since $S^\top e_{l+1} = e_l$ (with the convention $e_0 = 0$) we get by premultiplication with $S^\top$ the equality

$$(S^\top)^{n-k} = \sum_{l=0}^{k} \frac{1}{(k-l)!} e_l D^{k-l} a(\alpha)^\top PV.$$

Hence

$$(S^\top)^{n-1-k} - (S^\top)^{n-k} = \sum_{l=0}^{k} \frac{1}{(k-l)!} (e_{l+1} - e_l) D^{k-l} a(\alpha)^\top PV.$$

Summation of the left hand side of this equation over $k = 0, \ldots, n-1$ yields the identity matrix, whereas summation of the right hand side over the same range of values of k yields $\sum_{k=0}^{n-1} \frac{1}{(n-k-1)!} e_{k+1} D^{n-k-1} a(\alpha)^\top PV$. This proves the assertion. □

When the α_i are zeros of π with arbitrary multiplicity the situation is more complicated. First we proceed with some results of a more general nature. Let M be the block diagonal matrix with s blocks $M(i)$ of size $m_i \times m_i$ specified by

$$M(i)_{kl} = \begin{cases} \frac{1}{(k+l-m_i-1)!} D^{k+l-m_i-1} \pi_{m_i}(\alpha_i; \alpha_i) & \text{if } k+l-m_i-1 \geq 0 \\ 0 & \text{else.} \end{cases} \tag{18}$$

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 following alternative expression:

$$M(i) = \sum_{l=0}^{m_i-1} \delta_l S^l P, \text{ where } \delta_l = \frac{1}{l!} D^l \pi_{m_i}(\alpha_i; \alpha_i). \tag{19}$$

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 $r_1 S^{j-1}$, where S stands for the shift matrix of size $m_i \times m_i$. This can be seen as follows.
Denote now by e_j the j-th standard basis (column) vector in $\mathbb{R}^{m_i}$. Below we will use the following facts: $Se_j = e_{j+1}$ and $SP = PS^\top$, since S is Toeplitz.
The first row r_1 of $M(i)^{-1}$ satisfies $r_1 M(i) = e_1^\top$, so $\sum_{l=0}^{m_i-1} \delta_i r_1 S^l P = e_1^\top$. Then we get $r_1 S^{j-1} M(i) = \sum_{l=0}^{m_i-1} \delta_i r_1 S^{l+j-1} P = \sum_{l=0}^{m_i-1} \delta_i r_1 S^l P (S^\top)^{j-1} = e_1^\top (S^\top)^{j-1} = e_j^\top$, which implies that $r_1 S^{j-1}$ is indeed the j-th row if $M(i)^{-1}$.

As a matter of fact the inverses of the matrices $M(i)$ have a particular simple structure. To enlighten this we make the notation a little easier. Let f be an $m-1$ times continuously differentiable function and let the matrix valued function $M(z)$ of size $m \times m$ be defined by

$$M_{kl}(z) = \begin{cases} \frac{1}{(k+l-m-1)!} D^{k+l-m-1} f(z) & \text{if } k+l-m-1 \geq 0 \\ 0 & \text{else.} \end{cases} \tag{20}$$

Alternatively, we can write

$$M(z) = \sum_{l=0}^{m-1} \gamma_l S^l P \text{ with } \gamma_l = \frac{1}{l!} D^l f(z), \tag{21}$$

where now S stands for the $m \times m$ shift matrix and P for the $m \times m$ backward identity matrix. We claim that $(M(z))^{-1}$ has the following elements

$$(M(z))^{-1}_{kl} = \begin{cases} \frac{1}{(m+1-k-l)!} D^{m+1-k-l} (\frac{1}{f(z)}) & \text{if } m+1-k-l \geq 0 \\ 0 & \text{else.} \end{cases} \tag{22}$$

Alternatively, we have that

$$M(z)^{-1} = \sum_{l=0}^{m-1} \eta_l P S^l \text{ with } \eta_l = \frac{1}{l!} D^l (\frac{1}{f(z)}). \tag{23}$$

The claim is verified as follows. Application of the Leibniz rule to k-fold differentiation of the product $f\frac{1}{f}$ gives $\sum_{j=0}^{k}\gamma_j\eta_{k-j}=\delta_{0k}$. Next we compute

$$\begin{aligned}\sum_{j=0}^{m-1}\gamma_j S^j P\sum_{l=0}^{m-1}\eta_l P S^l &= \sum_{j,l=0}^{m-1}\gamma_j\eta_l S^{j+l}\\ &= \sum_{j=0}^{m-1}\gamma_j\sum_{k=j}^{m-1}\eta_{k-j}S^k\\ &= \sum_{k=0}^{m-1}(\sum_{j=0}^{k}\gamma_j\eta_{k-j})S^k = I,\end{aligned}$$

which shows that the claim holds true.
In particular, the first row of $(M(z))^{-1}$ is given by

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

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 $M(i)$. Its first row is then

$$(\frac{1}{\pi_{m_i}(\alpha_i;\alpha_i)},\dots,\frac{1}{(m_i-1)!}D^{m_i-1}(\frac{1}{\pi_{m_i}(\alpha_i;\alpha_i)})). \tag{24}$$

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 $\pi_k^*(z;\alpha)=z^{n-1}\pi_k(\alpha)$. 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 3.11 *The following equalities hold.*

$$\begin{aligned}u^*(z)^\top T_a &= a(z)^\top P &(25)\\ u^*(z)^\top T_a V &= (\pi_1(z;\alpha_1),\dots,\pi_{m_1}(z;\alpha_1),\dots,\pi_1(z;\alpha_s),\dots,\pi_{m_s}(z;\alpha_s)) &(26)\\ V^\top P T_a V &= M &(27)\\ u(z)^\top T_a V &= (\pi_1^*(z;\alpha_1),\dots,\pi_{m_1}^*(z;\alpha_1),\dots,\pi_1^*(z;\alpha_s),\dots,\pi_{m_s}^*(z;\alpha_s)) &(28)\\ V^\top T_a V &= N &(29)\\ V^{-1} &= M^{-1}V^\top P T_a = M^{-1}(T_a V)^\top P &(30)\end{aligned}$$

Proof. The equality (25) 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 (26). Consider (4) and differentiate k times w.r.t. α.

We obtain $-D^k\pi(\alpha) = u^*(z)^\top((z-\alpha)D^k a(\alpha) - kD^{k-1}a(\alpha))$.
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 $u^*(z)^\top((z-\alpha)D^k a(\alpha) - kD^{k-1}a(\alpha)$ for $1 \leq k \leq m-1$ and $\pi(z) = (z-\alpha)u^*(z)^\top a(\alpha)$. Write now $q_k(z) = u^*(z)^\top D^k a(\alpha)$, then $q_0(z) = \frac{\pi(z)}{z-\alpha}$ and we have the recursive system of equations $0 = (z-\alpha)q_k(z) - kq_{k-1}(z)$ for $k = 1, \ldots, m-1$. 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

$$u^*(z)^\top D^k a(\alpha) = k!\pi_{k+1}(z;\alpha). \tag{31}$$

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 the α_i for w and use the definition of the interpolation polynomial to get

$$u^*(z)^\top T_a V = (a(\alpha_1), \ldots, \frac{D^{m_1-1}a(\alpha_1)}{(m_1-1)!}, \ldots, a(\alpha_s), \ldots, \frac{D^{m_s-1}a(\alpha_s)}{(m_s-1)!}). \tag{32}$$

Combination of equations (31) and (32) yields (26).
To prove (27) we start from equation (26). Take the appropriate j-th order derivatives, divide by $j!$ and substitute the α_i in the resulting expression. Doing so results in a block-diagonal matrix, with on the diagonal exactly the $M(i)$ as above.
To prove (28) we first notice that from (25) we get $u(z)^\top T_a V = z^{n-1}u^*(z)^\top T_a V$ and then that it follows from (26) that a typical element of this row vector is $z^{n-1}\pi_k(\frac{1}{z};\alpha_i) = \pi_k^*(z;\alpha_i)$ by definition of the polynomials $\pi_k^*(z;\alpha)$.
The proof of (29) completely parallels that of (27) and is therefore omitted. Now we turn to (30). First we observe that all the matrices $M(i)$ are invertible, because of their triangular structure and the non-zero elements $\pi_{m_i}(\alpha_i;\alpha_i)$ (α_i had multiplicity m_i) on the antidiagonal. Therefore also M is invertible and taking inverses in equation (27) yields the first equality of (30). The second one then follows from $PT_a = T_a^\top P$. □

From formula (30) we obtain the equations (14) and (15) as special cases. Indeed if all the zeros of $\pi(z)$ are different, the matrix M is diagonal with elements $\pi_1(\alpha_i;\alpha_i)$. We have also seen in the proof of (26) that $T_a V$ (in this case) equals $(a(\alpha_1), \ldots, a(\alpha_n))$. So the expression (14) follows from the second formula for V^{-1} in (30).
If $\pi(z)$ has only one zero α, so $\pi(z) = (z-\alpha)^n$, the story is equally simple. In this case the matrix M is nothing else but the matrix P and $T_a V$ is now $(a(\alpha), \ldots, \frac{D^{n-1}a(\alpha)}{(n-1)!})$. Hence equation (15) follows.

Corollary 3.12 *The matrices M and N are related through the following two identities*

$$M = N^\top \Pi^* \tag{33}$$
$$N = (\Pi^*)^\top M \tag{34}$$

Moreover $NM^{-1} = MN^{-1}$.

Proof. From equation (29) 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 equation (27) 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 (33). But then also $N = (\Pi^*)^{-\top} M = (\Pi^*)^\top M$, since Π^* is involutive. For the same reason the final assertion of the corollary follows. □

As a side remark we mention that proposition 3.11 has as a corollary an expression for the partial fraction expansion of $\frac{1}{\pi(z)}$ (see also [8, section V.2]):

$$\frac{1}{\pi(z)} = (\frac{1}{z-\alpha_1}, \ldots, \frac{1}{(z-\alpha_1)^{m_1}}, \ldots, \frac{1}{z-\alpha_s}, \ldots, \frac{1}{(z-\alpha_s)^{m_s}}) V^{-1} e_n. \tag{35}$$

An easy way to see this is as follows. First postmultiply equation (26) by $V^{-1}e_n$. Because $T_a e_n = e_n$ and $u^*(z)^\top e_n = 1$ we obtain

$$1 = (\pi_1(z;\alpha_1), \ldots, \pi_{m_1}(z;\alpha_1), \ldots, \pi_1(z;\alpha_s), \ldots, \pi_{m_s}(z;\alpha_s)) V^{-1} e_n.$$

Now divide by $\pi(z)$. □

Equation (35) can also be used to evaluate the elements of the last column of V^{-1}. Enumerate the standard basis vectors of $\mathbb{R}^n$ as in the proof of proposition 3.5. Multiply equation (35) with $(z-\alpha_i)^k$ for $k = 0, \ldots, m_i - 1$. Integrate over a sufficiently small contour around α_i.
The right hand side of (35) then becomes $e_{i,k+1}^\top V^{-1} e_n$, whereas the left hand side becomes $\frac{1}{(m_i-1)!} D^{m_i-k-1} \frac{1}{\pi_{m_i}(\alpha_i;\alpha_i)}$.
Further down we will compare this with equation (38).

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

Proposition 3.13 *The following integral representations for the matrices $M(i)$ and $M(i)^{-1}$ are valid.*

$$M(i) = \frac{1}{2\pi i} \oint_{\Gamma_{\alpha_i}} u^*_{m_i}(z-\alpha_i) u^*_{m_i}(z-\alpha_i)^\top \frac{\pi(z)}{(z-\alpha_i)^{2m_i}} dz \tag{36}$$

$$M(i)^{-1} = \frac{1}{2\pi i} \oint_{\Gamma_{\alpha_i}} u_{m_i}(z-\alpha_i) u_{m_i}(z-\alpha_i)^\top \frac{1}{\pi(z)} dz. \tag{37}$$

As we have noticed before, $M(i)^{-1}$ is completely determined by its first row (or column). From proposition 3.13 we get, using Cauchy's theorem, that this first row is given by

$$\frac{1}{2\pi i} \oint_{\Gamma_{\alpha_i}} u_{m_i}(z-\alpha_i)^\top \frac{1}{\pi(z)} dz = (\frac{1}{\pi_{m_i}(\alpha_i;\alpha_i)}, \ldots, \frac{1}{(m_i-1)!} D^{m_i-1} \frac{1}{\pi_{m_i}(\alpha_i;\alpha_i)}) P, \tag{38}$$

in agreement with what we already found in (24).
In contrast with this, the matrix $M(i)$ is determined by its last row, which is

given by the vector whose j-th entry is given by $\frac{1}{(j-1)!}D^{j-1}a_{m_i}(\alpha_i, \alpha_i)$ ($j = 1, \ldots, m_i$).
The expressions for $M(i)^{-1}$ enable us also to get a formula for $p^*(0) = V^{-1}e_n$. From equation (30) we obtain $p^*(0) = M^{-1}V^\top e_1$. But the first column of $V^\top$ is given by the vectors $e_{m_i 1}$ stacked underneath each other. Hence $p^*(0)$ is the vector consisting of the subvectors $M(i)^{-1}e_{m_i 1}$, where the unit vectors $e_{m_i 1}$ are those of the proof of proposition 3.5.

4 Application to Stein's equation and time series

The goal of the present section is to obtain a way to compute the stationary Fisher information matrix of an ARMA process. Below, in section 4.2, we use that this matrix can be decomposed into block matrices that satisfy certain Stein equations. Therefore we first focus on the solutions of Stein equations.

4.1 Stein's equation

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 simple known result (see for instance [16, 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*

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

Consider again the polynomial $\pi(x) = \prod_{i=1}^{s}(z-\alpha_i)^{m_i} = \sum_{j=0}^{n} a_j z^j$ as in the previous sections and let A be its companion matrix. The purpose is to give in this case an expression for $\frac{1}{2\pi i}\oint_\Gamma (z-A)^{-1}f(z)dz$, where Γ is a closed contour surrounding all the α_k. From proposition 3.2 we obtain that this integral is given by $\frac{1}{2\pi i}\oint_\Gamma \frac{1}{\pi(z)}u(z)a(z)^\top Pf(z)dz$. Then for sufficiently small contours Γ_{α_k} around α_k this integral can be split into a sum of integrals $\frac{1}{2\pi i}\oint_{\Gamma_{\alpha_k}} \frac{1}{\pi(z)}u(z)a(z)^\top Pf(z)dz$. The value of each of these is equal to $\frac{1}{(m_k-1)!}D^{m_k-1}\left(\frac{u(z)a(z)^\top Pf(z)}{\pi_k(z;\alpha_k)}\right)$.

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

$$S - ASC^\top = H. \tag{39}$$

This equation, apart from its general interest in matrix and operator theory (e.g. the operator that takes S to $S - ASC$ is called a displacement operator, see [9]),

also shows up in the context of the analysis of stationary ARMA processes. More precisely, the asymptotic Fisher information matrix for such a process satisfies an equation of type (39), see [13]. In section 4.2 we will apply the results of the present section. In [14] we study this equation further and relate solutions of various Stein equations to a certain Fisher information matrix.
The solution (see [15]) 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 equal to $\sum_{k=0}^{\infty} A^k H(C^\top)^k$.
We continue with presenting an alternative way to obtaining 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 likewise associated with C. Let $\hat{S} = V_A^{-1} S V_C^{-\top}$ and $\hat{H} = V_A^{-1} H V_C^{-\top}$. Notice that the results of section 3.3 indicate an efficient way to compute $\hat{H}$.
Premultiplication of (39) with V_A^{-1} together with postmultiplication with $V_C^{-\top}$ results in

$$\hat{S} - J_A \hat{S} J_C^\top = \hat{H}, \tag{40}$$

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 [15]) 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 the AR and the MA-polynomial and thus lie inside the unit circle. More on this in section 4.2.
The computation of the inverse of the matrix $I - J_C \otimes J_A$ can also be done in an efficient way. Let $J_{A,i}$ be the i-th Jordan block of J_A and $J_{C,j}$ be the j-th Jordan block of J_C. Then $I - J_C \otimes J_A$ is block diagonal with diagonal blocks $I - J_{A,i} \otimes J_{C,j}$. Moreover, these blocks are upper triangular and even almost block diagonal. On the diagonal we find the blocks $I - \alpha_i J_{C,j}$ and on the subdiagonal just above it the blocks $-J_{C,j}$. Therefore, $(I - J_{A,i} \otimes J_{C,j})^{-1}$ is again upper triangular with on the diagonal the blocks $(I - \alpha_i J_{C,j})^{-1}$ and on the k-th subdiagonals above it one finds the blocks $(-1)^k P_k((I - \alpha_i J_{C,j})^{-1}, J_{C,j})$, where for matrices F and G of appropriate dimensions the operator P_k is defined as the alternating product $P_k(F, G) = FGFG \cdots FGF$, with k occurrences of G and $k+1$ occurrences of F.

The generic case is those where all the eigenvalues of A have multiplicity 1 and likewise for the eigenvalues of C. Consequently the matrices J_A and J_C are diagonal. In this case equation (40) is easy to solve: $\hat{S}$ has elements $\hat{S}_{ij} = \frac{1}{1-\alpha_i\gamma_j}\hat{H}_{ij}$.

4.2 Application to ARMA processes

In this subsection we specialize to the situation where we deal with the asymptotic Fisher information matrix for an ARMA(p, q) process. This process is the stationary discrete time stochastic process y that satisfies

$$y_t + a_1 y_{t-1} + \cdots + a_p y_{t-p} = \varepsilon_t + c_1 \varepsilon_{t-1} + \cdots + c_q \varepsilon_{t-q}, \tag{41}$$

where ε is a Gaussian white noise sequence with unit variance.
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. 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 $\theta = (a_1, \ldots, a_p, c_1, \ldots, c_q)^\top$. 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 Gaussianity, 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 [12].
The matrix $F(\theta)$ has a representation in the spectral domain given by the following block decomposition

$$F(\theta) = \begin{pmatrix} F_{aa} & F_{ac} \\ F_{ac}^\top & F_{cc} \end{pmatrix} \tag{42}$$

where the matrices appearing here have the following elements

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

With $k(z) = A(z)A^*(z)C(z)C^*(z)$ we have the following compact expression for the whole Fisher information matrix.

$$F(\theta) = \frac{1}{2\pi i} \oint_{|z|=1} \frac{1}{k(z)} \begin{pmatrix} C^*(z)u_p(z) \\ -A^*(z)u_q(z) \end{pmatrix} \begin{pmatrix} C(z)u_p^*(z)^\top & -A(z)u_q^*(z)^\top \end{pmatrix} dz. \tag{43}$$

As in section 2 we let $A \in \mathbb{R}^{p \times p}$ be the matrix given by (1) for $n = p$ and the matrix $C \in \mathbb{R}^{q \times q}$ of analogous form. We assume that both A and C have all their eigenvalues inside the unit circle.
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 [13] we showed that the Fisher information matrix $F(\theta)$ is the solution of the following Lyapunov equation

$$F(\theta) - \tilde{A}F(\theta)\tilde{A}^\top = ee^\top, \tag{44}$$

where $e^\top = (e_{pp}^\top, e_{qq}^\top)$ with e_{pp} the last standard basis vector in $\mathbb{R}^p$ and e_{qq} the last standard basis vector in $\mathbb{R}^q$. Take for $F(\theta)$ the block decomposition

$$F(\theta) = \begin{pmatrix} F_{aa} & F_{ac} \\ F_{ca} & F_{cc} \end{pmatrix} \tag{45}$$

where $F_{aa} \in \mathbb{R}^{p\times p}$, $F_{ca} \in \mathbb{R}^{q\times p}$ etc. Then each of these blocks solves a corresponding Stein equation. For instance for F_{ac} we have

$$F_{ac} - AF_{ac}C^\top = H_{ac}. \tag{46}$$

In this case the matrix H_{ac} takes the simple form $H_{ac} = e_{pp}e_{qq}^\top$. Next we bring this equation in Jordan form, parallel to equation (40). The matrix $\hat{H}_{ac}$ is now $\hat{H}_{ac} = V_A^{-1}e_{pp}(V_C^{-1}e_{qq})^\top$, both the last column of the inverse of a Vandermonde matrix. We have seen in section 3.3 how to compute such a column and that for instance $V_A^{-1}e_{pp}$ is equal to $p_A^*(0)$, where $p_A^*(z) = z^{p-1}p_A(\frac{1}{z})$ and p_A the interpolation polynomial related to the eigenvalues of A as described in section 3.2. Likewise $V_C^{-1}e_{qq} = p_C^*(0)$.
Letting $\hat{F}_{ac} = V_A^{-1}F_{ac}V_C^{-\top}$, it follows from the results of section 4.1 that, assuming that A and C only have eigenvalues of multiplicity 1, $\hat{F}_{ac}$ has as it's ij-th element

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

Now it is again 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.

References

[1] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel. *Time series analysis.* Prentice Hall Inc., Englewood Cliffs, NJ, third edition, 1994. Forecasting and control.

[2] P. J. Brockwell and R. A. Davis. *Time Series: Theory and Methods.* Springer-Verlag, second edition, 1991.

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

[4] E. Cheney. *Introduction to matrix interpolation theorey, 2nd ed.* Chelsea, 1982.

[5] J. Conway. *Functions of one complex variable.* Springer, 1975.

[6] F. G. Csáki. Some notes on the inversion of confluent Vandermonde matrices. *IEEE Trans. Automatic Control*, AC-20:154–157, 1975.

[7] P. Fuhrmann. *A polynomial approach to linear algebra.* Springer, 1996.

[8] F. Gantmacher. *The theory of matrices vol. 1.* AMS Chelsea Publishing, 1959.

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

[10] I. C. Göknar. Obtaining the inverse of the generalized Vandermonde matrix of the most general type. *IEEE Trans. Automatic Control*, AC-18(5):530–532, 1973.

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

[12] A. Klein and P. Spreij. On Fisher's information matrix of an ARMAX process and Sylvester's resultant matrices. *Linear Algebra Appl.*, 237/238:579–590, 1996. Special issue honoring Calyampudi Radhakrishna Rao.

[13] A. Klein and P. Spreij. On Fisher's information matrix of an ARMA process. In *Stochastic differential and difference equations (Győr, 1996)*, pages 273–284. Birkhäuser Boston, Boston, MA, 1997.

[14] A. Klein and P. Spreij. On Stein's equation, Vandermonde matrices and Fisher's information matrix of time series processes. Part I: The autoregressive moving average process. Universiteit van Amsterdam, AE-Report 7/99, 1999.

[15] P. Lancaster and L. Rodman. *Algebraic Riccati Equations.* Oxford Science Publications, 1995.

[16] P. Lancaster and M. Tismenetsky. *The theory of matrices.* Academic Press Inc., Orlando, Fla., second edition, 1985.

[17] L. Lupaş. On the computation of the generalized Vandermonde matrix inverse. *IEEE Trans. Automatic Control*, AC-20:559–561, 1975.

[18] V. Olshevsky. Eigenvector computation for almost unitary Hessenberg matrices and inversion of Szegő-Vandermonde matrices via discrete transmission lines. *Linear Algebra Appl.*, 285(1-3):37–67, 1998.

[19] R. H. Schappelle. The inverse of the confluent Vandermonde matrix. *IEEE Trans. Automatic Control*, AC-17(5):724–725, 1972.

[20] J. Traub. Associated polynomials and uniform methods for the solution of linear problems. *SIAM Review*, 8:277–301, 1966.