On Zero Locations of Predictor Polynomials

Fermin S. V. Bazán*
Departamento de Matemática,
Universidade Federal de Santa Catarina,
Florianópolis, SC 88040-900, Brasil
fermin@mtm.ufsc.br

Licio H. Bezerra†
Departamento de Matemática,
Universidade Federal de Santa Catarina,
Florianópolis, SC 88040-900, Brasil
licio@mtm.ufsc.br

Predictor polynomials are often used in linear prediction methods mainly for extracting properties of physical systems which are described by time series. The aforementioned properties are associated with a few zeros of large polynomials and for this reason the zero locations of those polynomials must be analyzed. We present a linear algebra approach for determining the zero locations of predictor polynomials, which enables us to generalize some early results obtained by Kumaresan in the signal analysis field. We also present an analysis of zero locations for time series having multiple zeros.

KEY WORDS: predictor polynomials, eigenvalues, companion matrices, time series

*Supported by CNPq, Brasil, grant 300487/94-0 (NV).

†Part of this research was performed while this author was at CERFACS, Toulouse, France, and was supported by Fundação CAPES, Brasil, grant BEX 3119/95-5.

1 Introduction

Linear prediction is a technique widely used in developing parametric models that represent dynamic systems from discrete time signals, also known as time series. Once the model is successful, then that model can be very useful among other applications for predicting and forecasting, for example. An excellent survey about the relevance and applications of linear prediction can be found in Makhoul [12] (see also [6] and [14]). The model most currently used for representing a large class of dynamic systems (and which will also be used here) is that described by the impulse response of the system, which is either a real or complex-valued function composed of a weighted sum of n damped/undamped complex exponentials. Thus, the time series under analysis is

$$h_k = \sum_{j=1}^{n} r_j e^{s_j k \Delta t}, \quad k = 0, 1, \ldots \tag{1}$$

where r_j and s_j, $s_i \neq s_j$ for $i \neq j$, are the parameters that hypothetically govern the system, Δt is the sampling interval and n is the *order* of the system. The s_j are called *poles* and are constants which characterize the normal modes of the system while the r_j, called *residues*, describe how much each mode participates in the time series. A predictor polynomial then predicts future samples of the signal by using past samples values. More precisely, a predictor polynomial of degree N

$$P(t) = c_0 + c_1 t + \cdots + c_{N-1} t^{N-1} - t^N, \; N \geq n, \tag{2}$$

predicts h_{j+N} from knowledge of the N preceding $\{h_j, h_{j+1}, \ldots, h_{j+N-1}\}$, according to the rule

$$c_0 h_j + c_1 h_{j+1} + \cdots + c_{N-1} h_{j+N-1} = h_{j+N} \quad j \geq 0. \tag{3}$$

An interesting fact about predictor polynomials is that the system poles s_j, can be extracted from their zeros. Once the nonlinear parameters are found, the task of determining the r_j is simpler. In the linear prediction approach then the problem of constructing a parametric model for a given signal is reduced to that of estimating the coefficients c_i. In practical experimental situations however, although the s_j's are related to n zeros of the polynomial (the signal zeros), there are $N - n$ zeros (the extraneous zeros) without physical meaning that arise as a consequence of using a polynomial of order larger than necessary because the system order, n, is not known in advance. For this reason and mainly to establish criteria that enables us discriminate the signal zeros from the extraneous ones, an analysis on zero locations of predictor polynomials seems to be indispensable. With respect to this, two approaches are known. The first, introduced by Kumaresan [11], uses properties of predictor error filters and is well understood in the signal analysis field; while the other, introduced by Cybenko [4], employes both a classical theorem of Fejer and some properties of orthogonal polynomials in the unit circle [5].

The goal of this work is to present a simplifying approach for the problem based mainly on linear algebra concepts. For this purpose, the concept of predictor matrix is introduced and then the zero locations of predictor polynomials are analyzed from the eigenvalues of their associated companion predictor matrices. Besides this analysis, an extension for the analysis of zero locations of predictor polynomial corresponding to time series containing multiple poles is addressed as well.

The paper is organized as follows: in Section 2 some basic results are presented and the notation is introduced. Section 3 extends the concept of predictor polynomial to the concept of predictor matrix; some theorems which explain much on eigenvalue locations of companion predictor matrices are presented. The paper ends with an extension of results for time series which contain multiples poles.

2 Preliminary Results and Notation

Let $\mathsf{H}(l)$ be the $M \times N$ Hankel matrix indexed by an integer $l \geq 0$, whose ith column vector is $\hat{\mathbf{h}}_i = [h_{l+i-1}\, h_{l+i} \cdots h_{l+M+i-2}]^T$.

$$\mathsf{H}(l) = [\hat{\mathbf{h}}_l \;\; \hat{\mathbf{h}}_{l+1} \;\; \cdots \; \hat{\mathbf{h}}_{l+N-1}] = \begin{bmatrix} h_l & h_{l+1} & \cdots & h_{l+N-1} \\ h_{l+1} & h_{l+2} & \cdots & h_{l+N} \\ \vdots & \vdots & \cdots & \vdots \\ h_{l+M-1} & h_{l+M} & \cdots & h_{l+M+N-2} \end{bmatrix}_{M\times N} . \quad (4)$$

In relation to $\mathsf{H}(l)$, it is easy to see that

$$\mathsf{H}(l) = \mathsf{V}\mathsf{Z}^l\mathsf{R}\mathsf{W}, \quad (5)$$

where $\mathsf{Z} = diag(z_1, \cdots, z_n)$, in which $z_j = e^{s_j \Delta t}$, $j = 1, 2, \cdots, n$, $\mathsf{R} = diag(r_1, \cdots, r_n)$, $\mathsf{V} = \mathsf{V}(z_1, z_2, \cdots, z_n)$ is the $M \times n$ Vandermonde matrix described below

$$\mathsf{V} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ z_1 & z_2 & \cdots & z_n \\ \vdots & \vdots & \vdots & \vdots \\ z_1^{M-1} & z_2^{M-1} & \cdots & z_n^{M-1} \end{bmatrix}_{M \times n}, \tag{6}$$

and W is the transpose of the submatrix of V formed by taking its N first rows. A consequence of decomposition (5) is that for all $l \geq 0$, $rank(\mathsf{H}(l)) = n$ whenever $M \geq N \geq n$ and $s_i \neq s_j$, $i \neq j$, [1]. It also follows that the column space of $\mathsf{H}(l)$, denoted by $\mathcal{R}(\mathsf{H}(l))$, is spanned by the column vectors of matrix V and that the null space of $\mathsf{H}(l)$, denoted by $\mathcal{N}(\mathsf{H}(l))$, is the same as that of W. The pseudo- inverse of a matrix A of order $M \times N$ is defined as the unique matrix $\mathsf{A}^\dagger$ of order $N \times M$ satisfying the conditions: $i)$ $\mathsf{A}\mathsf{A}^\dagger\mathsf{A} = \mathsf{A}$, $ii)$ $\mathsf{A}^\dagger\mathsf{A}\mathsf{A}^\dagger = \mathsf{A}^\dagger$, $iii)$ $(\mathsf{A}^\dagger\mathsf{A})^H = \mathsf{A}^\dagger\mathsf{A}$; and $iv)$ $(\mathsf{A}\mathsf{A}^\dagger)^H = \mathsf{A}^\dagger\mathsf{A}$; where the supercript ^{H} is used to denote conjugate transposition. If A has the full rank factorization $\mathsf{A} = \mathsf{B}\mathsf{C}$, where rank($\mathsf{A}$)=rank($\mathsf{B}$)= rank($\mathsf{C}$), then

$$\mathsf{A}^\dagger = \mathsf{C}^\dagger\mathsf{B}^\dagger, \text{ with } \mathsf{B}^\dagger = (\mathsf{B}^H\mathsf{B})^{-1}\mathsf{B}^H, \ \mathsf{C}^\dagger = \mathsf{C}^H(\mathsf{C}\mathsf{C}^H)^{-1}. \tag{7}$$

Further properties about pseudo-inverses can be found in Bjork [3] and Stewart [13]. The spectrum of a matrix will be denoted by $\lambda(\mathsf{A})$. In the next section we present a theorem for which the following two lemmas will be required.

Lemma 2.1 Let $\mathsf{A} \in \mathbb{C}^{M \times N}$ and $\mathsf{B} \in \mathbb{C}^{N \times M}$, $M \geq N$, where $\mathbb{C}^{M \times N}$ denotes the set of all complex $M \times N$ matrices. Then $\lambda(\mathsf{A}\mathsf{B}) = \lambda(\mathsf{B}\mathsf{A}) \cup \{0\}$.

A proof for this lemma can be found in Horn [9].

Lemma 2.2 Let $\mathsf{A} \in \mathbb{C}^{M \times M}$ be a hermitian matrix whose ith column vector is $\mathbf{a}_i$; that is, $\mathsf{A} = [\mathbf{a}_1 \ \mathbf{a}_2 \ \cdots \ \mathbf{a}_M]$. Let $\mathsf{A}^\uparrow = [\mathbf{a}_2 \ \mathbf{a}_3 \ \cdots \ \mathbf{a}_M \ \mathbf{0}]$ and $\mathsf{A}^\downarrow = [\mathbf{0} \ \mathbf{a}_1 \ \mathbf{a}_2 \ \cdots \ \mathbf{a}_{M-1}]$. Then $\lambda(\mathsf{A}^\uparrow) = \overline{\lambda(\mathsf{A}^\downarrow)}$, where the bar denotes complex conjugation.

Proof: First, write $\mathsf{A}^{\uparrow}$ and $\mathsf{A}^{\downarrow}$ in block triangular form

$$\mathsf{A}^{\uparrow} = \left[\begin{array}{c|c} \mathsf{F} & \mathbf{0} \\ \hline f^H & 0 \end{array}\right], \qquad \mathsf{A}^{\downarrow} = \left[\begin{array}{c|c} 0 & b^H \\ \hline \mathbf{0} & \mathsf{G} \end{array}\right],$$

where F and G are both $(M-1)\times(M-1)$ submatrices of A, and $\mathbf{0}$, f, b are all complex vectors in $\mathbb{C}^{M-1}$. Since A is hermitian, by simple inspection one sees that $\mathsf{F} = \mathsf{G}^H$. The assertion of the lemma follows as a consequence of the structure of those matrices. ■

3 The Zeros of Predictor Polynomials

We start by extending the notion of predictor polynomial to the concept of predictor matrix. Let $\mathsf{H}(l)$, $l \geq 0$, be a Hankel matrix of order $M \times N$, we say that a matrix $\mathcal{C}$ of order $N \times N$ is a *predictor matrix* if and only if for all integers $l \geq 0$ $\mathsf{H}(l+1) = \mathsf{H}(l)\mathcal{C}$. The matrix $\mathcal{C}$ has an analogous role as that developed of predictor polynomials: it predicts the new data sample h_{l+N} from knowledge of the preceding N samples $\{h_l,\ h_{l+1}, \ldots,\ h_{l+N-1}\}$. We observe that there are an infinite collection of matrices $\mathcal{C}$ satisfying this definition and that the signal zeros can always be extracted from the eigenvalues of any predictor matrix. Of course, from the equality above, using equation (5) and the pseudoinverse properties expressed in (7), one sees that

$$\mathsf{H}(l+1) = \mathsf{H}(l)\mathcal{C} \Leftrightarrow \mathsf{W}\mathcal{C} = \mathsf{Z}\mathsf{W}. \tag{8}$$

The assertion follows after observing that the rows of W are left eigenvectors of $\mathcal{C}$ corresponding to the signals. We shall analyze the locations of the eigenvalues of the predictor polynomial $P(t) = t^N - (c_0 + c_1 t + \cdots + c_{N-1}t^{N-1})$, $N \geq n$, by analyzing the locations of the eigenvalues of a predictor companion matrix

$$\mathcal{C} = [e_2\ e_3\ \cdots\ e_N\ \mathbf{c}] = \begin{bmatrix} 0 & 0 & \cdots & 0 & c_0 \\ 1 & 0 & \cdots & 0 & c_1 \\ 0 & 1 & \cdots & 0 & c_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & c_{N-1} \end{bmatrix}_{N\times N}, \tag{9}$$

whose Nth column vector, $\mathbf{c}$, is solution of the system of linear equations:

$$\mathsf{H}(l)\mathbf{c} = \hat{\mathbf{h}}_{l+N} \;\Leftrightarrow\; \mathsf{W}\mathbf{c} = \mathsf{Z}^N \lfloor\mathbf{1}, \tag{10}$$

where $\lfloor\mathbf{1}$ denotes a vector with components equal to unity. It is obvious that the above system is consistent and that it has an infinite number of solutions, since H is rank defficient and because $\hat{\mathbf{h}}_{l+N} \in \mathcal{R}(\mathsf{H}(l))$. Therefore, the zeros of $P(t)$, or equivalently, the eigenvalues of the companion matrix $\mathcal{C}$ depend on how one chooses the vector $\mathbf{c}$. Interesting work regarding zeros of polynomials can be found in [7], [9] and [10]. Generalizations on predictor matrices may be found in [2].

Theorem 3.1 Let $\mathcal{C}^{\dagger}$ be a companion matrix whose Nth column vector is the minimum norm solution of (10). Then $\mathcal{C}^{\dagger}$ has n of its eigenvalues located at $z_i = e^{s_i \Delta t}$, $i = 1, \ldots, n$; the others, which are called the extraneous ones, fall inside the unit circle.

Proof: As has already been stated, $e^{s_i \Delta t}$, $i = 1, \ldots, n$ forms part of the spectrum of any predictor matrix. Therefore we need to prove the second part of the theorem. We start by observing that the right eigenvectors associated with extraneous eigenvalues belong to the null space of W. Now, the Nth component of these eigenvectors cannot be zero. To see this, let ϕ be a unit lenght eigenvector associated to the extraneous eigenvalue γ.

$$\mathcal{C}^{\dagger}\phi = \gamma\,\phi \Leftrightarrow \begin{bmatrix} 0 \\ \phi_1 \\ \phi_2 \\ \cdots \\ \phi_{N-1} \end{bmatrix} + \phi_N \mathbf{c} = \gamma \begin{bmatrix} \phi_1 \\ \phi_2 \\ \cdots \\ \phi_{N-1} \\ \phi_N \end{bmatrix}.$$

If $\gamma = 0$ there is nothing to prove; assume then $\gamma \neq 0$. By observing the above relation, one sees that $\phi_N \neq 0$, otherwise ϕ would be the zero vector, which is a contradiction. Now, let $\phi^{\downarrow H} = [0\,\bar{\phi}_1, \cdots, \bar{\phi}_{N-1}]$. Left multiplying in the above relation by ϕ^H yields

$$\gamma = \phi^H \mathcal{C} \phi = \phi^H \phi^{\downarrow} + \phi_N \phi^H \mathbf{c}$$

But, since $\mathbf{c}$ is the minimum norm solution of (10), $\mathbf{c} \in [\mathcal{N}(\mathsf{W})]^{\perp}$, and so $\phi^H \mathbf{c} = 0$. Hence,

$$|\gamma| = |\phi^H \phi^{\downarrow}| \leq \|\phi^{\downarrow}\| < 1.$$

∎

Linear prediction may also be carried out in the reverse direction. For this, it is sufficient to seek a $N \times N$ matrix which enables us to come back from the state $l+1$ of the system to the state l. Thus, one may define a backward predictor matrix in the following way: we say that a $N \times N$ matrix, $\mathcal{D}$ is a backward predictor matrix if and only if for all $l \geq 0$, $\mathsf{H}(l) = \mathsf{H}(l+1)\mathcal{D}$. As before, one has the following equivalence

$$\mathsf{H}(l) = \mathsf{H}(l+1)\mathcal{D} \Leftrightarrow \mathsf{W}\mathcal{D} = \mathsf{Z}^{-1}\mathsf{W}, \tag{11}$$

from which, as before, $\{z_1^{-1},\ z_2^{-1}, \ldots, z_n^{-1}\}$ are eigenvalues of $\mathcal{D}$ with the rows of W as associated left eigenvectors. This is an important fact that may be useful for discriminating signal zeros corresponding to exponentially damped signals. We will see later that provided the signal zeros are extracted from the eigenvalues of a suitable companion matrix, the extraneous zeros separate from the signal zeros in a natural way: the signal zeros fall outside the unit circle while the extraneous ones lie inside. To see this, we will consider backward predictor companion matrices of the form

$$\mathcal{C}^{\downarrow} = [\mathbf{d}\, e_1\ e_2\ \cdots\ e_{N-1}] = \begin{bmatrix} d_0 & 1 & 0 & \cdots & 0 \\ d_1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ d_{N-2} & 0 & 0 & \cdots & 1 \\ d_{N-1} & 0 & 0 & \cdots & 0 \end{bmatrix}_{N \times N}, \tag{12}$$

where the first column vector, $\mathbf{d}$, is a solution of the system of linear equations:

$$\mathsf{H}(l+1)\mathbf{d} = \hat{\mathbf{h}}_l \Leftrightarrow \mathsf{W}\mathbf{d} = \mathsf{Z}^{-1}\lfloor \mathbf{1}, \tag{13}$$

with associated backward predictor polynomial $Q(t) = d_{N-1} + d_{N-2}t + \cdots + d_0 t^{N-1} - t^N$. We will show that, provided the vector $\mathbf{d}$ is selected via pseudoinversion, then the extraneous eigenvalues of $\mathcal{C}^{\downarrow}$, or equivalently the extraneous zeros of $Q(t)$, are the complex conjugates of those of $\mathcal{C}^{\uparrow}$. We first prove an auxiliary result in the following lemma.

Lemma 3.1 Let $\mathcal{P}$ be the orthogonal projection onto the subspace $[N(\mathsf{W})]^{\perp}$, and let $\mathcal{Q} = I - \mathcal{P}$, where I is the $N \times N$ identity matrix. Since $\mathcal{C}^{\downarrow} = \mathcal{S}^{\downarrow} + \mathcal{R}^{\downarrow}$, where $\mathcal{S}^{\downarrow} = \mathcal{P}\mathcal{C}^{\downarrow}$, $\mathcal{R}^{\downarrow} = \mathcal{Q}\mathcal{C}^{\downarrow}$, the signal eigenvalues of $\mathcal{C}^{\downarrow}$ are the non-zero eigenvalues of $\mathcal{S}^{\downarrow}$ while the non-zero extraneous ones are the non-zero eigenvalues of $\mathcal{R}^{\downarrow}$.

Proof: Observe that $\mathcal{P} = \mathsf{W}^{\dagger}\mathsf{W}$. Now, by using (11), $\mathcal{S}^{\downarrow} = \mathcal{P}\mathcal{C}^{\downarrow} = \mathsf{W}^{\dagger}\mathsf{Z}^{-1}\mathsf{W}$. Hence, by applying Lemma 2.1, one sees that $\lambda(\mathcal{S}^{\downarrow}) = \lambda(\mathsf{W}^{\dagger}\mathsf{Z}^{-1}\mathsf{W}) = \lambda(\mathsf{Z}^{-1})\mathsf{W}\mathsf{W}^{\dagger} \cup \{0\} = \lambda(\mathsf{Z}^{-1}) \cup \{0\}$ and so the assertion holds. On the other hand, let γ be an extraneous eigenvalue of $\mathcal{C}^{\downarrow}$, and let ϕ be an associated right eigenvector. We prove that ϕ and γ are an eigenpair of $\mathsf{R}^{\downarrow}$. In fact,

$$\gamma\phi = \mathcal{C}^{\downarrow}\phi = \mathcal{S}^{\downarrow}\phi + \mathcal{R}^{\downarrow}\phi = \mathsf{W}^{\dagger}\mathsf{Z}^{-1}\mathsf{W}\phi + \mathcal{R}^{\downarrow}\phi = \mathcal{R}^{\downarrow}\phi,$$

because $\phi \in \mathcal{N}(\mathsf{W})$. ■

Remark: A similar result can be established for $\mathcal{C}^{\uparrow}$: its signal eigenvalues are the non-zero eigenvalues of $\mathcal{S}^{\uparrow}$ while its non-zero extraneous eigenvalues are the non-zero eigenvalues of $\mathcal{R}^{\uparrow}$, where $\mathcal{S}^{\uparrow} = \mathcal{P}\mathcal{C}^{\uparrow}$ and $\mathcal{R}^{\uparrow} = \mathcal{Q}\mathcal{C}^{\uparrow}$.

Theorem 3.2 Let $\mathcal{C}^{\downarrow}$ be a companion matrix whose column vector $\mathbf{d}$ is the minimum norm solution of the system of equations (13). Then $\mathcal{C}^{\downarrow}$ has n eigenvalues located at $z_i = e^{-s_i\Delta t}$, $i = 1, \ldots, n$ and the extraneous ones fall inside the unit circle and are the complex conjugates of those of $\mathcal{C}^{\uparrow}$.

Proof: By Lemma 3.1, the system eigenvalues are the non-zero eigenvalues of $\mathcal{S}^{\downarrow}$, that is, $z_i = e^{-s_i\Delta t}$, $i = 1, \ldots, n$; and the non-zero extraneous eigenvalues are non-zero eigenvalues of $\mathsf{R}^{\downarrow}$. Next, let q_i be the ith column vector of $\mathcal{Q}$, ie, $\mathcal{Q} = [q_1\ q_2\ \cdots\ q_N]$. From this, we have

$$\mathcal{R}^{\uparrow} = [q_2\ q_3\ \cdots q_{N-1}0] \quad \text{and} \quad \mathcal{R}^{\downarrow} = [0\ q_1\ q_2\ \cdots q_N]$$

because both column vectors $\mathbf{c}$ and $\mathbf{d}$ belong to $[N(\mathsf{W})]^{\perp}$. But $\mathcal{Q}$ is hermitian; therefore, the final assertion of the theorem follows as consequence of Lemma 2.2 and Theorem 3.1. ■

When modeling is performed with time domain data, as is the case when one analyzes vibrating systems in modal analysis, for example, one deals with polynomials whose coefficients are real. In this case, as consequence of Theorem 3.2, we have

Corollary 3.1 Let $\mathcal{C}^{\uparrow}$ and $\mathcal{D}^{\downarrow}$ be the matrices introduced in the above theorems. Then, provided that real-valued time series are analyzed, the extraneous eigenvalues of $\mathcal{C}^{\uparrow}$ match exactly with those of $\mathcal{D}^{\downarrow}$.

We would like to mention that the zero locations for predictor polynomials was earlier studied by Kumaresan [11], in the context of linear prediction-error filter polynomials for deterministic signals. He stated the results of Corollary 3.1 regardless of whether the time series is real or complex valued. However, as we saw, this result holds only for real-valued time series.

4 Extension for Time Series with Multiple Poles

Time series containing multiple poles arise very often in mechanical and electromagnetic systems, for instance after sampling response funtions [8], [15] of the type

$$h(t) = \sum_{k=1}^{K} \sum_{q=1}^{n_k} r_{k,q} t^{q-1} e^{s_k t}. \qquad (14)$$

We will carry out an analysis regarding zero locations of polynomials which come in connection with time series related to the above function. For this, we assume $s_i \neq s_j$ for $i \neq j$, $n_k \geq 1$, $r_{k,n_k} \neq 0$, for $k = 1, 2, \ldots, K$, and the multiplicities satisfy $n_1 + \cdots + n_K = n$. With the above assumptions we shall show that the aforementioned polynomials have zeros which behave similarly as those in the simple pole case. For simplicity we assume the signal has a single pole of multiplicity n; that is, we assume time series of the form

$$h_k = [(r_1 + (k\Delta t) r_2 + (k\Delta t)^2 r_3 + \cdots + (k\Delta t)^{n-1} r_n] e^{s\Delta t}, \; k = 0, 1, \ldots \qquad (15)$$

The basic idea is to follow a procedure like that one developed for the single zero case. So, we start by seeking a factorization for the Hankel matrix $\mathsf{H}(l)$. For this, let $z = e^{s\Delta t}$, and let $\hat{\mathbf{h}}_j$ be the jth column vector of $\mathsf{H}(l)$. Using

(15), we write

$$\hat{\mathbf{h}}_j = \begin{bmatrix} h_{l+j-1} \\ h_{l+j} \\ \vdots \\ h_{l+j+M-2} \end{bmatrix} = \mathsf{V}\mathcal{P}^j \mathsf{r}, \tag{16}$$

where V is an $M \times n$ matrix defined by

$$\mathsf{V} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ z & (\Delta t)z & \cdots & (\Delta t)^{n-1}z \\ z^2 & (2\Delta t)z^2 & \cdots & (2\Delta t)^{n-1}z^2 \\ \vdots & \vdots & \cdots & \vdots \\ z^{M-1} & ((M-1)\Delta t)z^{M-1} & \cdots & ((M-1)\Delta t)^{n-1}z^{M-1} \end{bmatrix}, \tag{17}$$

$$\mathsf{P} = z\Delta^{-1}\mathsf{G}\Delta, \tag{18}$$

with $\Delta = diag(1, \Delta t, (\Delta t)^2, \ldots, (\Delta t)^{n-1})$, and G the $n \times n$ upper triangular matrix defined so that its i, j entry is $\mathsf{G}_{i,j} = \dfrac{j}{(i-1)!(j-i+1)!}$, for $j \geq i$. That is, the non-null elements of the j-th column of G are the binomial coefficients of an expansion of the type $(a+b)^{j-1}$. On the other hand, r is a vector containing the residues r's. Next, let R be a symmetric $n \times n$ matrix defined so that its i, j entry is

$$\begin{cases} \mathsf{R}_{i,j} = \dfrac{(i+j-2)!}{(i-1)!(j-1)!} r_{i+j-1}, & \text{for } i \leq n-j+1 \\ \mathsf{R}_{i,j} = 0\ , & \text{otherwise.} \end{cases} \tag{19}$$

In other words, the elements of the jth cross diagonal of R are obtained by the non-vanishing entries of the jth column of G times the jth element of vector r. If $n = 4$ for example, G and R are

$$\mathsf{G} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathsf{R} = \begin{bmatrix} r_1 & r_2 & r_3 & r_4 \\ r_2 & 2r_3 & 3r_4 & 0 \\ r_3 & 3r_4 & 0 & 0 \\ r_4 & 0 & 0 & 0 \end{bmatrix}. \tag{20}$$

Using (16), the Hankel matrix can now be written as

$$\mathsf{H}(l) = [\hat{\mathbf{h}}_l \ \ \hat{\mathbf{h}}_{l+1} \ \cdots \ \hat{\mathbf{h}}_{l+N-1}] = \mathsf{V}\mathcal{P}^l[\mathsf{R}e_1, \mathcal{P}\mathsf{R}e_1, \ldots, \mathcal{P}^{N-1}\mathsf{R}e_1]. \tag{21}$$

But, by direct calculation one can verify that $\mathsf{PR} = \mathsf{RP}^T$ and that

$$(\mathsf{P}^i)^T e_1 = \begin{bmatrix} z^{i-1} \\ ((i-1)\Delta t) z^{i-1} \\ \vdots \\ ((i-1)\Delta t)^{n-1} z^{i-1} \end{bmatrix}; \tag{22}$$

thus, a factorization for $\mathsf{H}(l)$ is

$$\mathsf{H}(l) = \mathsf{VR}(\mathsf{P}^l)^T\mathsf{W}; \tag{23}$$

in which W is obtained by transposing the submatrix of V formed by its N first rows. If we assume $M \geq N \geq n$, it can be proved that V is of rank n, see [15], and that R is non-singular since we assume $r_n \neq 0$. Hence, the factorization above is a full rank factorization of $\mathsf{H}(l)$ and thus $rank(\mathsf{H}(l)) = n$. With this decomposition at hand, one can always compute solutions for our familiar linear prediction equation, $\mathsf{H}(l+1) = \mathsf{H}(l)\mathcal{C}$. Furthermore, we should emphasize that if for some l a solution is computed, it serves to predict exactly the whole signal since that solution does not depend on l, but only on the poles:

$$\mathsf{H}(l+1) = \mathsf{H}(l)\mathcal{C} \Leftrightarrow \mathsf{P}^T\mathsf{W} = \mathsf{W}\mathcal{C}. \tag{24}$$

After replacing P for its canonical Jordan decomposition, $\mathsf{P} = Q\,J\,Q^{-1}$, one sees that $z = e^{s\Delta t}$ is an eigenvalue of $\mathcal{C}$ of multiplicity n, with associated left generalized eigenvectors, the rows of $Q^T\mathsf{W}$; that is, any predictor matrix has an eigenvalue of multiplicity n, $z = e^{s\Delta t}$. Hence, if the predictor matrix is a companion one, the associated predictor polynomial has a zero of multiplicity n, which is $z = e^{s\Delta t}$. Here again, the location in the complex plane of the extra $N - n$ zeros depends on the choice of $\mathcal{C}$. But, if one chooses the companion matrix so that its Nth column vector is the minimal 2-norm solution of the system of linear equations $\mathsf{H}(l)\mathbf{c} = \hat{\mathbf{h}}_{l+N}$, the procedure employed in the proof of Theorem 3.1 allow us in the current situation to prove that $\mathcal{C}$ (or equivalently the associated predictor polynomial $P(t)$) has n repeated eigenvalues located at $z = e^{s\Delta t}$ and the remaining ones fall inside the unit circle, analogous as in the single pole case. If one carries out linear prediction in the reverse direction, one is able to obtain analogous results, that is, one may show that if $\mathcal{C}^{\downarrow}$ is a backward predictor companion matrix whose first column vector $\mathbf{d}$ is the minimal 2-norm solution of $\mathsf{H}(l+1)\mathbf{d} = \hat{\mathbf{h}}_l$, then $\mathcal{C}^{\downarrow}$ has an eigenvalue of multiplicity n located at $z^{-1} = e^{-s\Delta t}$ while the extraneous ones correspond to the conjugates of those of $\mathcal{C}^{\uparrow}$.

To analyze time series arising from (14), notice that in this case the column vector $\hat{\mathbf{h}}_j$ can be written as

$$\hat{\mathbf{h}}_j = \sum_{k=1}^{K} \mathsf{V}_k \mathsf{P}_k^j \mathsf{r}_k, \tag{25}$$

where V_k is an $M \times n_k$ matrix, P_k is an $n_k \times n_k$ matrix, both defined by replacing z by $z_k = e^{s_k \Delta t}$, in (17) and (18) respectively, and r_k is the residues vector associated to the pole s_k. Hence, one may prove that the decomposition expressed in (23) actually holds, with the difference that V, P and R are now block matrices:

$$\mathsf{V} = [\mathsf{V}_1 \cdots \mathsf{V}_K], \quad \mathsf{P} = diag(\mathsf{P}_1, \ldots, \mathsf{P}_K), \quad \text{and} \quad \mathsf{R} = diag(\mathsf{R}_1, \ldots, \mathsf{R}_K),$$

where R_k are symmetric $n_k \times n_k$ matrices, defined analogously as in (19) using the components of the residues related to the pole s_k. On its turn, W^T is an $N \times n$ block matrix, obtained by taking the N first rows of V.

5 Concluding Remarks

A decomposition of $\mathsf{H}(l)$ for the particular case $l = 0$ was presented by Wei [15] and was employed to prove that the poles of the time series (14) can be extracted from any polynomial whose coefficients are solutions of a linear system with coefficient matrix $\mathsf{H}(0)$, although nothing was said regarding the extra $N - n$ zeros. Here, using the fact that V is full rank [15], we see from (23) that for all $l \geq 0$, $rank(\mathsf{H}(l)) = n$, which allow us to conclude that the number of exponentials contained in any time series either expressed by (1) or (14) can always be detected from the rank of $\mathsf{H}(l)$. Another consequence of using the factorization (23) is that the involved work developed by Wei in proving that result related to the extraction of poles from polynomials now becomes very easy, since this is a direct consequence of (24).

As final remark, we would like to emphasize that the actual factorization of $\mathsf{H}(l)$ and the results obtained in the previous analysis enable us to prove theorems which in some sense are the companion ones of Theorem 3.1 and

Theorem 3.2. That is, both Theorem 3.1 and Theorem 3.2 hold, regardless of whether the time series has multiple poles or not. As in practical applications the polynomial coefficients are only mere approximations of the true ones, an interesting problem which should be investigated is how the signal zeros change when the polynomial coefficients suffer small perturbations. We hope this work is useful for developing a zero perturbation analysis taking advantage of the rich resources of the linear algebra. This done, the problem of separating the signal zeros from the extraneous ones may be tackled by using backward predictor companion matrices.

References

[1] F. S. V. Bazán and C. Bavastri, An Optimized Pseudo-Inverse Algorithm (OPIA) for Multi-Input Multi-Output Modal Parameter Identification, *Mechanical Systems and Signal Processing*, 10, 365-380, 1996.

[2] L. H. Bezerra and F. S. V. Bazán, Eigenvalue Locations of Generalized Companion Predictor Matrices, Technical Report TR/PA/97/01, CERFACS, Toulouse, 1997.

[3] Å. Björk, *Numerical Methods for Least Squares Problems*, SIAM, Philadelphia, 1996.

[4] G. Cybenko, Locations of Zeros of Predictor Polynomials, *IEEE Trans. on Automat. Control*, AC-27, 235-237, 1982.

[5] P. J. Davies, *Interpolation and Approximation*, Blaisdell, New York, NY, 1963.

[6] B. De Moor, The Singular Value Decomposition and Long and Short Spaces of Noisy Matrices, *IEEE Trans. on Signal Processing*, 41, 2826-2838, 1993.

[7] A. Edelman and H. Murakami, Polynomial Roots from Companion Matrix Eigenvalues, *Math. Comp.*, 64, 763-776, 1995.

[8] T. L. Henderson, Geometric Methods for Determining System Poles from Transient Response, *IEEE Trans. on ASSP*, 29, 982-988, 1981.

[9] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.

[10] F. Kittaneh, Singular Values of Companion Matrices and Bounds on Zeros of Polynomials, *SIAM J. Matrix Anal. Appl.*, 16, 333-340, 1995.

[11] R. Kumaresan, On the Zeros of the Linear Prediction-Error Filters for Deterministic Signals, *IEEE Trans. on ASSP*, ASSP-31, 217-221, 1983.

[12] J. Makhoul, Linear Prediction: A Tutorial Review, *Proc. of the IEEE*, 63, 561-580, 1975.

[13] G. W. Stewart and J.-G. Sun, *Matrix Perturbation Theory*, Academic Press, San Diego, 1990.

[14] A. van der Veen, E. F. Deprettere and A. L. Swindlehurst, Subspace-Based Signal Analysis Using Singular Value Decomposition, *Proc. of the IEEE*, 81, 1277-1309, 1993.

[15] M. Wei and G. Majda, A New Theoretical Approach for Prony's Method, *Linear Algebra Appl.*, 136, 119-132, 1990.