ON DETERMINING THE PREDICTOR OF NON-FULL-RANK MULTIVARIATE STATIONARY RAN. (U) NORTH CAROLINA UNIV AT CHAPEL HILL CENTER FOR STOCHASTIC PROC. A G MIAMEE

UNCLASSIFIED MAR 85 TR-96 AFOSR-TR-85-0682 F/G 12/1 NL
ON DETERMINING THE PREDICTOR OF NON-FULL-RANK
MULTIVARIATE STATIONARY RANDOM PROCESSES

by

A.G. Miamee

Technical Report No. 96

March 1985
Algorithms for determining the generating function and the predictor for some non-full-rank multivariate stationary stochastic processes are obtained. In fact it is shown that the well known algorithms given by Wiener and Masani (1958) for the full-rank case, are valid in certain non-full-rank cases exactly in the same form.
ON DETERMINING THE PREDICTOR OF NON-FULL-RANK MULTIVARIATE STATIONARY RANDOM PROCESSES

by

A.G. Miamee

Isfahan University of Technology
Isfahan, Iran

and

Center for Stochastic Processes
Department of Statistics
University of North Carolina
Chapel Hill, NC 27514

Abstract

Algorithms for determining the generating function and the predictor for some non-full-rank multivariate stationary stochastic processes are obtained. In fact it is shown that the well known algorithms given by Wiener and Masani (1958) for the full-rank case, are valid in certain non-full rank cases exactly in the same form.

AMS Subject Classification: primary 60G10.

Key words and phrases: non-full-rank multivariate stationary processes, generating function, and best linear predictor.

Research partially supported by Air Force Office of Scientific Research grant #F49620 82 C 0009.
1. Introduction.

One of the important problems in the prediction theory of multivariate stationary stochastic processes is to obtain some algorithm for determining the best linear predictor in terms of the past observations. Wiener and Masani [9], [10] solved this problem for the full-rank case, when the spectral density $f$ of the processes is bounded above and away from zero, in the sense that there exist positive numbers $c$ and $d$ such that

$$\tag{1.1} cI \leq f(\theta) \leq dI.\]

Masani [2] improved their work substantially showing that the same algorithm is valid if in lieu of (1.1) one assumes that

$$\tag{1.2} \begin{align*}
(i) & \quad f \in L_\infty, \\
(ii) & \quad f^{-1} \in L_1.
\end{align*}$$

several other authors proved the validity of the same algorithm under more general settings, cf. for example Salehi [6], Pourahmadi [8]. However, all these results are under the severe restriction of full-rank and there has been no extension of Wiener and Masani's algorithm beyond the full-rank case.

The purpose of this note is to show that the algorithm remains valid exactly in the same manner for the non-full-rank processes which satisfy the following conditions

(i) The range of $f(\theta)$ is constant a.e. $(d\theta)$,

$$\tag{1.3} \begin{align*}
(ii) & \quad f \in L_\infty, \\
(iii) & \quad f^* \in L_1,
\end{align*}$$

where $A^*$ stands for the generalized inverse (to be defined later) of the matrix $A$. In the full-rank case these conditions clearly reduce to the conditions (1.2), and
hence our result generalizes Masani's algorithm in [2].

Masani's assumption and approach rests on a characterization (Theorem 2.4, [2]) for full-rank minimal multivariate stationary stochastic processes. Our motivation and assumptions are based on a characterization of $J_0$-regularity due to Makagon and Weron [1]. We will employ Wiener and Masani's algorithm to find the predictor of an associated full-rank process (to be clarified later), which is produced using the technique of Salehi and Miamee [5], and using this we will obtain our algorithm for the non-full-rank process.

In section 2 we set down the necessary preliminaries. Section 3 is devoted to establishing our algorithm for determining the generating function and in section 4 we will show the validity of Wiener and Masani's algorithm for the best linear predictor.
2. Preliminaries

In this section we set down notations and preliminaries. Most of these are standard and can be found in [4], [9] and [10]. Let \( H \) be a complex Hilbert space and \( q \) a positive integer. \( H^q \) denotes the Cartesian product of \( q \)-copies of \( H \), endowed with a Gramian structure as follows: For any two vectors \( x = (x^1, \ldots, x^q)^T \) and \( y = (y^1, \ldots, y^q)^T \) in \( H^q \) their Gramian matrix \( (x,y) \) is defined by

\[
(x,y) = [(x^i, x^j)]_{i,j=1}^q.
\]

It is easy to verify that it has the following properties:

\[
(x,y) \geq 0 ;
\]

\[
(x,x) = 0 \iff x = 0 ;
\]

\[
\left( \sum_{i=1}^m A_i X_i, \sum_{j=1}^n B_j X_j \right) = \sum_{i=1}^m \sum_{j=1}^n A_i (X_i, Y_j) B^* j,
\]

where \( X, Y, X_i, Y_j \) are in \( H^q \), \( A_i, B_j \) are constant \( q \times q \) matrices, and \( A \geq 0 \) means \( A \) is a non-negative definite matrix. We say that \( X \) is orthogonal to \( Y \) if \( (X,Y) = 0 \). It is well known that \( H^q \) is a Hilbert space with the inner product

\[
((X,Y)) = \text{trace} \ (X,Y).
\]

A closed subset \( M \) of \( H^q \) is called a subspace if \( AX + BY \in M \), whenever \( X \) and \( Y \) are in \( M \), \( A \) and \( B \) are \( q \times q \) constant matrices. It is easy to see that \( M \) is a subspace if and only if \( M = M^q \) for some subspace \( M \) of \( H \). For any \( X \) in \( H^q \), \( (X|M) \) denotes the projection of \( X \) onto \( M \), and that is the vector whose \( k \)-th coordinate is \( (X^k|M) \), which is the usual projection of \( X^k \) onto the subspace \( M \).
A bisequence $X_n$, $n \in \mathbb{Z}$, in $H^Q$ is called a \textit{q-variate stationary stochastic process} if the Gramian $(X_m, X_n)$ depends only on $m - n$.

It is well known that every q-variate stationary stochastic process $X_n$ has a non-negative matrix valued measure $F$ on $[0,2\pi]$, called its \textit{spectral measure} such that

$$
(X_m, X_n) = \frac{1}{2\pi} \int_0^{2\pi} e^{-i(m - n)\theta} dF(\theta).
$$

$f$ stands for the Radon-Nikodym derivative of the absolutely continuous (a.c.) part of $F$ with respect to the normalized Lebesgue measure $d\theta$, and it is called the \textit{spectral density} of the process.

To every stationary stochastic process $X_n$, $n \in \mathbb{Z}$ the following subspaces are attached:

$$
M(\pm \infty) = \text{sp} \{X_n, -\infty < n < \infty\}, \text{ i.e. the subspace of } H^Q
$$
generated by all $X_n$, $n \in \mathbb{Z}$,

$$
M(n) = \text{sp} \{X_k, -\infty < k < n\},
$$

$$
M(-\infty) = \bigcap_{n} M(n),
$$

$$
M^-(n) = \overline{\text{sp}} \{X_k, k \neq n\}.
$$

A q-variate stationary stochastic process is called

(a) \textit{non-deterministic} if $M(\pm \infty) \neq M(n)$ for some and hence all $n$ in $\mathbb{Z}$,

(b) \textit{regular} if $M(-\infty) = 0$

(c) \textit{minimal} if $M^-(n) \neq M(\pm \infty)$ for some and hence all $n \in \mathbb{Z}$,

(d) $J_0$-\textit{regular} if $\bigcap_{n} M^-(n) = 0$. 
If $X_n$ is non-deterministic then $X_n \not\in M(n - 1)$ for all $n$, and hence it has a non-zero \textit{one-sided innovation process}

$$g_n = X_n - (X_n | M(n - 1)).$$

If $X_n$ is minimal then $X_n \not\in M'(n)$ for all $n$, and hence it has a non-zero \textit{two-sided innovation process}

$$\phi_n = X_n - (X_n | M'(n)).$$

The corresponding \textit{one-sided and two-sided predictor error matrices} are defined by

$$G = (g_0, g_0) \text{ and } \Sigma = (\phi_0, \phi_0)$$

respectively. \(\hat{X}_v = (X_v | M(0))\) is called the \textit{best linear predictor of log $v$}. Clearly $X_n$ is non-deterministic if and only if $G \neq 0$ and minimal if and only if $\Sigma \neq 0$. A non-deterministic (regular) process $X_n$ is said to be non-deterministic (regular) of full-rank if $G$ is invertible. The process is called \textit{full-rank minimal} if it is minimal and its two-sided predictor error matrix $\Sigma$ is invertible.

It is useful to note that we have the following inclusions between these various classes of processes

- non-deterministic $\not\Rightarrow$ regular $\not\Rightarrow$ minimal $\not\Rightarrow J_0$-regular $\not\Rightarrow$ full-rank minimal.

The last inclusion is a consequence of Theorems 1 and 2 below, and the others can be easily verified.
It is known that

\[ M(n) = \sum_{k=0}^{\infty} \sp (q_{n-k}) + M(-\infty). \]

Consider G as a linear operator on \( C^q \) to \( C^q \), \( C \) being the complex plane. Let \( J \) be the matrix of the projection on \( C^q \) onto the range of \( G \), and we put \( (\sqrt{G} + J^1)^{-1} = H \). The normalized one-sided innovations are defined by \( h_n = Hq_n \).

One can show that [4]

\[ X_n = \sum_{k=0}^{\infty} A_k \sqrt{G} h_{n-k} + (x_n | M(-\infty)). \]

although \( A_k \)'s in this decomposition are not unique, the coefficients \( A_k \sqrt{G} \) are in fact unique and this enables us to associate the following function to our process

\[ \phi(e^{i\theta}) = \sum_{k=0}^{\infty} A_k e^{ik\theta}, \]

this is called the generating function of the process.

We shall be concerned with the class \( L_p (1 \leq p \leq \infty) \) of all \( q \times q \) matrix valued functions \( g \) on \([0,2\pi]\) whose entries are in the usual Lebesgue space \( L_p \). \( L_0^+ \) will denote the subspace of \( L_2 \) consisting of those matrix valued functions whose \( n \)-th Fourier coefficient vanishes for \( n < 0 \), i.e.

\[ \int e^{-in\theta} g(\theta) d\theta = 0, \quad \text{for all } n < 0. \]

For any \( q \times q \) matrix \( A \) there exists a unique \( q \times q \) matrix \( A^\# \) such that [7]

\[ AA^\# A = A, \quad A^\# A A^\# = A^\# \]

\[ (A^\# A)^* = (A^\# A), \quad (AA^\#)^* = AA^\#. \]
This matrix $A^d$ is called the generalized inverse of $A$ and has the following further properties

$$N^\perp(A) = R(A^d), \quad R^\perp(A) = N(A^d),$$

where $R(B)$ and $N(B)$ denote the range and null space of the matrix $B$, respectively.

For the ease of reference we state the following two theorems which are due to Masani [2], and to Makagon and Weron [1], respectively.

**Theorem 1.** Let $X_n$, $n \in \mathbb{Z}$, be a $q$-variate stationary stochastic process with spectral distribution $F$. $X_n$ is full-rank minimal if and only if $F$ is a.c. and its spectral density $f$ is invertible with $f^{-1} \in L_1$.

**Theorem 2.** Let $X_n$, $n \in \mathbb{Z}$ be a $q$-variate stationary stochastic process with spectral measure $F$. The process $X_n$ is $J_0$-regular if and only if

(i) $F$ is a.c. with respect to $d\theta$, with spectral density $f$,
(ii) $R(f(\theta))$ is constant a.e. $(d\theta)$,
(iii) $f^\# \in L_1$. 

3. Determination of the generating function.

In this section we give an algorithm for determining the generating function of a (not necessarily full-rank) stationary stochastic process. The result of this section extends Masani's algorithm developed in [2] to the non-full-rank case. Our technique is essentially that used by Salehi and Miamee in [5] where the following formula for the two-sided prediction error matrix $\Sigma$ of a $J_0$-regular process was obtained

$$
\Sigma = \left[ \begin{array}{c}
\int_{2\pi}^{2\pi} i\#(\theta)d\theta \\
0
\end{array} \right].
$$

We will continue this work under the assumption that our process is $J_0$-regular or equivalently assuming that conditions (i), (ii), and (iii) of Theorem 2 are valid. Let $h_1, h_2, ..., h_p, h_{p+1}, ..., h_q$ be an orthonormal basis for the $q$-dimensional complex Euclidean space $\mathbb{C}^q$ such that

$$
R = R(f(\theta)) = \overline{sp}(h_i, 1 \leq i \leq p) \ a.e. \ (d\theta),
$$

and

$$
N = R^\perp = N(f(\theta)) = \overline{sp}(h_i, p+1 \leq i \leq q).
$$

Let $e_1, e_2, ..., e_q$ be the standard basis of $\mathbb{C}^q$. Define the unitary operator $U$ on $\mathbb{C}^q$ by $U e_i = e_i, 1 \leq i \leq q$. Letting $R_1 = \overline{sp}(e_i, 1 \leq i \leq p)$ then $R_1^\perp = \overline{sp}(e_i, p+1 \leq i \leq q)$. Clearly $U$ maps $R$ onto $R_1$ and $R_1^\perp$ onto $R_1^\perp$ and $U^*$ maps $R_1$ onto $R$ and $R_1^\perp$ onto $R_1^\perp$. As usual we will identify any linear operator on $\mathbb{C}^q$ with its matrix with respect to the standard basis of $\mathbb{C}^q$. By our choice of $U$ we have

$$
(3.1) \quad Uf(\theta)U^* = \begin{bmatrix}
q(\theta) & 0 \\
0 & 0
\end{bmatrix}
$$
where \( \varphi(\theta) \) is a \( p \times p \) non-negative matrix valued function whose rank is a.e. equal to \( p \). Let
\[
Y_n = UX_n, \quad n \in \mathcal{Z}
\]
be a new stationary stochastic process, then we have
\[
(Y_m^n, Y_n) = (UX_m^n, UX_n) = U \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i(m-n) \theta} f(\theta) d\theta U^* \]
\[
= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i(m-n) \theta} U f(\theta) U^* d\theta \]
\[
= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i(m-n) \theta} \begin{bmatrix} \varphi(\theta) & 0 \\ 0 & 0 \end{bmatrix} d\theta.
\]

(3.2)

This shows that, for \( p+1 \leq k \leq q \), the \( k \)-th component \( Y_n^k \) of \( Y_n \) is zero for all \( n \in \mathcal{Z} \). The \( p \)-variate stationary stochastic process \( Z_n = (Y_n^1, \ldots, Y_n^n)^T \) has spectral density \( q \). Since \( U \) takes \( \mathbb{R} \) onto \( \mathbb{R}^p \) and \( \mathbb{R}^i \) onto \( \mathbb{R}^i \), one can see that
\[
\begin{bmatrix}
\varphi^{-1} & 0 \\
0 & 0
\end{bmatrix}
= \begin{bmatrix}
\varphi & 0 \\
0 & 0
\end{bmatrix} = (U f(\theta) U^*)^* = U f(\theta) U^*.
\]

(3.3)

Now since \( X_n \) is assumed to be \( \mathcal{J}_0 \)-regular, Theorem 2 implies that \( f(\theta) \) is integrable. Thus (3.2) implies that \( \varphi^{-1} \) is integrable and hence by Theorem 1, \( Z_n \) is full-rank minimal.

We are going to utilize Masani's algorithm to obtain the generating function \( \psi \) and predictor \( \hat{Z}_n \) of this full-rank minimal process \( Z_n \), and then use this to get the generating function \( \phi \) and predictor \( \hat{X}_n \) of our process \( X_n \). The following lemma, which reveals the close tie between \( \psi \) and \( \phi \), is crucial in the development of our algorithm.
Lemma. Let $X_n, n \in \mathbb{Z}$ be a $\mathcal{L}_0$-regular stationary stochastic process with spectral density $f$. Let $g$ be the spectral density of the corresponding full-rank minimal process $Z_n$ discussed above. If $\phi$ and $\psi$ are the generating functions of $X_n$ and $Z_n$ respectively then

$$\phi = U^* \begin{bmatrix} \psi & 0 \\ 0 & 0 \end{bmatrix} U,$$

where $U$ is the unitary matrix obtained above.

Proof. We first note that, since $\phi$ and $\psi$ as generating functions are optimal (cf. Lemma 3.7 and Definition 4.1 in [3]). Now from (3.1) we get

$$(3.4) \quad f = U^* \begin{bmatrix} g & 0 \\ 0 & 0 \end{bmatrix} U = (U^* \begin{bmatrix} \psi & 0 \\ 0 & 0 \end{bmatrix} U) (U^* \begin{bmatrix} \psi & 0 \\ 0 & 0 \end{bmatrix} U)^*$$

on the other hand

$$f = \phi \phi^*.$$

Since $f$ has two factors $\phi$ and

$$\delta = U^* \begin{bmatrix} \psi & 0 \\ 0 & 0 \end{bmatrix} U,$$

belonging to $L^2_0$, to complete the proof it suffices to show that the latter one is also optimal (cf. uniqueness Theorem 4.4 of [3]). To prove this we first note that since the 0-th coefficient $\psi_+^*(0)$ of $\psi$ is nonnegative definite and

$$\delta_+(0) = U^* \begin{bmatrix} \psi_+^*(0) & 0 \\ 0 & 0 \end{bmatrix}$$

we have
On the other hand if
\begin{equation}
\delta_+(0) \geq 0.
\end{equation}

(3.6) \hspace{1cm} f = \gamma \gamma^*, \quad \gamma \in L^2_2

is another factorization of \( f \), then
\begin{equation}
\begin{bmatrix}
g & 0 \\ 0 & 0
\end{bmatrix} = UfU^* = (U\gamma U^*)(U\gamma U^*)^*
\end{equation}

but \( g = \psi \psi^* \) implies that
\begin{equation}
\begin{bmatrix}
g & 0 \\ 0 & 0
\end{bmatrix} = \begin{bmatrix}
\psi & 0 \\ 0 & 0
\end{bmatrix} \begin{bmatrix}
\psi & 0 \\ 0 & 0
\end{bmatrix}^*
\end{equation}

Since \( \psi \) is the generating function of \( Z_n \), one can prove that the function
\begin{equation}
\begin{bmatrix}
\psi & 0 \\ 0 & 0
\end{bmatrix}
\end{equation}

is the generating function of \( Y_n \). In fact we know that the generating function \( \phi \) of a \( q \)-variate stationary stochastic process \( X_n \) is given by
\begin{equation}
\phi = \sum_{n=0}^{\infty} A_n \sqrt{\mathcal{G}} e^{in\theta},
\end{equation}

where \( A_n \)'s are the coefficients in the representation
\begin{equation}
X_0 = \sum_{n=0}^{\infty} A_n g_{-n} + (X_0 | M(-\infty))
\end{equation}

of \( X_n \) in terms of its innovation process
\begin{equation}
g_n = X_n - (X_n | M(n-1))
\end{equation}

and \( G = (g_0, g_0) \) is the predictor error matrix. Comparing \( Z_n \) with \( Y_n = [Z_n | 0]^T \) we note that
\[ q_Y = \begin{bmatrix} g & 0 \\ 0 & 0 \end{bmatrix}, \quad G_Y = \begin{bmatrix} G & 0 \\ 0 & 0 \end{bmatrix}, \quad \text{and} \quad \sqrt{G} = \begin{bmatrix} \sqrt{G} & 0 \\ 0 & 0 \end{bmatrix} \]

\[ Y_0 = \begin{bmatrix} Z_0 \\ 0 \end{bmatrix} = \begin{bmatrix} \sum_{n=0}^{\infty} A_n Z_n + (Z_0 | M^Z(-\infty)) \\ 0 \end{bmatrix} = \Sigma \]

\[ = \sum_{n=0}^{\infty} \begin{bmatrix} A_n & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} g \\ 0 \end{bmatrix} + (Y_0 | M^Y(-\infty)). \]

Although the coefficients arising in this sum are not unique they will give us the generating function uniquely, and we have

\[ \phi_Y = \sum_{n=0}^{\infty} \begin{bmatrix} A_n Z_n & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \sqrt{G} & 0 \\ 0 & 0 \end{bmatrix} e^{-in\theta} \]

\[ = \sum_{n=0}^{\infty} \begin{bmatrix} A_n & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \sqrt{G} & 0 \\ 0 & 0 \end{bmatrix} e^{-in\theta} = \sum_{n=0}^{\infty} \begin{bmatrix} A_n & 0 \\ 0 & 0 \end{bmatrix} e^{-in\theta} \cdot \begin{bmatrix} e^Z & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} A_n e^Z & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \psi & 0 \\ 0 & 0 \end{bmatrix}. \]

Thus \( \begin{bmatrix} \psi & 0 \\ 0 & 0 \end{bmatrix} \) is the optimal factor of \( \begin{bmatrix} g & 0 \\ 0 & 0 \end{bmatrix} \). (3.7) and (3.8) together with the optimality of \( \begin{bmatrix} \psi & 0 \\ 0 & 0 \end{bmatrix} \).
imply that
\[
\begin{bmatrix}
\psi_+(0) & 0 \\
0 & 0
\end{bmatrix} \begin{bmatrix}
\psi_+(0) & 0 \\
0 & 0
\end{bmatrix} \geq (U\psi_+(0)U^*)(U\psi_+(0)U^*)\ast.
\]

This in turn implies that
\[
(\delta_+(0))^2 = (U\begin{bmatrix}
\psi_+(0) & 0 \\
0 & 0
\end{bmatrix} U)(U\begin{bmatrix}
\psi_+(0) & 0 \\
0 & 0
\end{bmatrix} U) \geq (U^*\psi_+(0)U)(U^*\psi_+(0)U)\ast.
\]

This together with (3.5) shows that \(\delta_0\) is the optimal factor of \(f\). Thus by
the uniqueness theorem mentioned above

\[
\phi = \delta = U^* \begin{bmatrix}
\psi & 0 \\
0 & 0
\end{bmatrix} U. \quad \text{Q.E.D.}
\]

Now we are ready to give the algorithm determining the generating function
of our \(J_0\)-regular \(q\)-variate stationary stochastic process \(X_n\). Since \(f\) satisfies
the conditions (i), (ii), and (iii) of (1.3) one can see that these imply that
\(g\) satisfies the corresponding conditions (i) and (ii) of (1.2).

Thus we can use Masani's algorithm developed in section 4 in [2]
to compute the generating function \(\phi\) of the desired process \(X_n\) via the formula

\[
\phi = U^* \begin{bmatrix}
\psi & 0 \\
0 & 0
\end{bmatrix} U.
\]

Remark. One can similarly extend the other available algorithms (such
as that in [8]) to obtain corresponding algorithms for the non-full-rank case.
4. Determination of the Predictor.

In this section we show that the unique autoregressive series, of [2], giving the linear predictor in the full-rank case, can be used to obtain the predictor in our non-full-rank case. In fact as we will see, exactly the same formula works in this case as well. We continue to assume that the density $f$ of our stationary stochastic process $X_n$ satisfies conditions (1.3). Using the notations and results of section 3 we know that

$$f = U \begin{bmatrix} g & 0 \\ 0 & 0 \end{bmatrix} U,$$

and the density $g$ satisfies conditions (i) and (ii) of (1.2). Thus, using the technique developed in [2] one can show that

$$\hat{Z}_\nu = \sum_{k=0}^{\infty} E_{\nu k} Z_{-k}, \quad \text{in } H^p,$$

where

$$E_{\nu k} = \sum_{n=0}^{k} C_{\nu+n} D_{k-n}$$

with $C_k$ and $D_k$ being the $k$-th Fourier coefficients of $\psi$ and $\psi^{-1}$ respectively.

Now one can easily verify that

$$\hat{\gamma}_{\nu} = \left[ \hat{Z}_\nu \right] = \sum_{k=0}^{\infty} \left[ E_{\nu k} 0 \right] Y_{-k}, \quad \text{in } H^q,$$

and

$$\begin{bmatrix} E_{\nu k} & 0 \\ 0 & 0 \end{bmatrix} = \sum_{n=0}^{k} \begin{bmatrix} C_{\nu+n} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} D_{k-n} & 0 \\ 0 & 0 \end{bmatrix}$$

(4.1)

Since $Y_n = UX_n$, one can also verify that

$$\hat{X}_n = U*\hat{Y}_n = U*\hat{Y}_n.$$
Hence we have
\[ \hat{X}_n = U^* \left( \sum_{k=0}^{\infty} \begin{bmatrix} E_{vk} & 0 \\ 0 & 0 \end{bmatrix} Y_{-k} \right) = \]
\[ \sum_{k=0}^{\infty} \begin{bmatrix} E_{vk} \\ 0 \end{bmatrix} U^* Y_{-k} \]
\[ \in H^q. \]

Letting \( F_{vk} \) to be
\[ F_{vk} = U^* \begin{bmatrix} E_{vk} \\ 0 \end{bmatrix} U \]
we get the following autoregressive series representation for the best linear predictor \( \hat{X}_n \):
\[ \hat{X}_n = \sum_{k=0}^{\infty} F_{vk} X_{-k}. \]

Now let us examine the coefficients \( F_{vk} \) in (4.3) more carefully. Doing this we will be able to write \( F_{vk} \) in terms of the Fourier coefficients of the generating function \( \Phi \) of our original process \( X_n \) rather than that of the auxiliary process \( Z_n \). From (4.2) we can write
\[ F_{vk} = U^* \begin{bmatrix} E_{vk} \\ 0 \end{bmatrix} U. \]

Now using (4.1) we have
\[ F_{vk} = U^* \left( \sum_{n=0}^{k} \begin{bmatrix} C_{v+n} \\ 0 \end{bmatrix} \begin{bmatrix} D_{k-n} \\ 0 \end{bmatrix} \right) U. \]
\[ = \sum_{n=0}^{k} (U^* \begin{bmatrix} C_{v+n} \\ 0 \end{bmatrix} U) (U^* \begin{bmatrix} D_{k-n} \\ 0 \end{bmatrix} U). \]
Thus
\[ F_{vk} = \sum_{n=0}^{k} M_{v+n} N_{k-n}, \]
with
\[
M_n = U^* \begin{bmatrix} C_n & 0 \\ 0 & 0 \end{bmatrix} U \quad \text{and} \quad N_n = U^* \begin{bmatrix} D_n & 0 \\ 0 & 0 \end{bmatrix} U.
\]

But by the Lemma we have
\[
\phi = U^* \begin{bmatrix} \psi & 0 \\ 0 & 0 \end{bmatrix} U \quad \text{and} \quad \phi^# = U^* \begin{bmatrix} \psi^{-1} & 0 \\ 0 & 0 \end{bmatrix} U.
\]

Thus we observe that \( M_n \) and \( N_n \) are exactly the \( n \)-th Fourier coefficients of \( \phi \) and \( \phi^# \) respectively.

Summarizing, we have shown that the best linear predictor \( \hat{X}_\psi \) can be written exactly in the same form obtained in [2] for the full-rank processes. i.e. we have
\[
\hat{X}_\psi = \sum_{k=0}^{\infty} (\sum_{n=0}^{k} M_{n+k} D_{k-n})X_{-k}, \quad \text{in } H^0.
\]

where \( M_n \) and \( N_n \) are the \( n \)-th Fourier coefficients of \( \phi \) and its generalized inverse \( \phi^# \) (instead of \( \phi \) and its inverse \( \phi^{-1} \) in the full-rank case).
REFERENCES


