

FULL RANK EXTRAPOLATION THEORY

by

Mark Hirsh Isaacs

B.Sc., Simon Fraser University, 1969

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
in the Department
of
Mathematics

© Mark Hirsh Isaacs 1971

SIMON FRASER UNIVERSITY

September, 1971

APPROVAL

Name: Mark Hirsh Isaacs

Degree: Master of Science

Title of Thesis: Full Rank Extrapolation Theory

Examining Committee:

(N. R. Reilly)

Chairman

(D. M. Eaves)

Senior Supervisor

(D. J. Mallory)

Examining Committee

(A. L. Stone)

Examining Committee

(R. R. Rennie)

Examining Committee

Date approved: September 10, 1971

ABSTRACT

In this paper the linear extrapolation theory of full rank multivariate discrete parameter weakly stationary stochastic processes is discussed. The approach used is that which was originally proposed by Zasuhrin and subsequently developed by Wiener and Masani. The fundamental theory required for the study of multivariate processes is first established, culminating in a characterization of full rank processes in terms of their spectral measures. The linear extrapolation problem for full rank processes is then set forth. This is followed by a discussion of the difficulties encountered in seeking an autoregressive representation for the predictor and the partial solutions of Wiener and Masani. The paper concludes with a discussion of the special problems encountered in the theory of degenerate rank processes.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my supervisor, Dr. D.M. Eaves, Mathematics Department, Simon Fraser University, for his continuous assistance and encouragement during the preparation of this thesis.

Thanks are also due to the National Research Council of Canada and to the Simon Fraser University President's Research Grants Committee for their financial support during this period.

Finally, I wish to express my appreciation to Miss Sue Smithe for typing the entire work.

TABLE OF CONTENTS

	<u>Page</u>
Approval	ii
Abstract	iii
Acknowledgements	iv
Introduction	1
Chapter I MATHEMATICAL PREREQUISITES	3
Univariate Concepts	3
Matrix Integration Theory	5
Chapter II MULTIVARIATE STOCHASTIC PROCESSES	12
Complex Valued Random Variables	12
Multivariate Random Variables	13
Time Domain Analysis	17
Chapter III SPECTRAL ANALYSIS	27
Spectral Theory	27
Processes of Full Rank	32
Chapter IV EXTRAPOLATION THEORY	41
The Extrapolation Problem	41
Solutions of the Extrapolation Problem	43
Minimal Processes	50
Determination of the Generating Function	52
Chapter V DEGENERATE RANK PROCESSES	60
Bibliography	63

INTRODUCTION

The study of multivariate prediction theory began in the U.S.S.R. in the early 1940's at approximately the same time as the univariate case was being investigated (for references to the Russian literature see Jang Ze Pei [10]). Problems still exist in the multivariate case, however, whereas the univariate case was essentially solved by the early 1950's (see, for example, Hannan [7]).

The study of univariate processes rests heavily on the theory of complex valued functions on the unit disc, notably the theory of H_p and L_p spaces. The natural multivariate analogue of this is the study of matrix valued functions. Two problems arise with the introduction of matrix methods: the lack of commutativity of multiplication and the existence of singular non-zero matrices. Nonetheless, many results have been obtained, notably in the case of full rank processes (Masani [21], and Rozanov [26]).

There are two closely related approaches to the general problem of multivariate prediction. One technique, adopted by Rozanov [26], Matveev [22], and others is to study a certain Hilbert space H of complex valued random variables. A second, although similar, approach is the one proposed by Zasuhrin and developed by Wiener and Masani [29]. In their approach the vector valued random functions themselves are treated as elements of a Hilbert space ℓ_2 . However, instead of using the normal inner product on

this space a matrix valued inner product called the Gramian is introduced. The results obtained by each approach are almost the same, since ℓ_2 is essentially just H^q . It is this latter approach of Wiener and Masani's that is reviewed in this paper.

Chapter I sets forth the basic definitions of the univariate theory and presents the necessary material on matrix integration theory that is required throughout the paper.

In Chapter II the fundamental Hilbert space ℓ_2 is introduced and its properties developed. The time domain analysis of multivariate discrete parameter weakly stationary (henceforth referred to as q-variate) processes is presented, including the multivariate extension of the classical Wold decomposition theorem.

The spectral theory of q-variate processes is developed in Chapter III, together with a characterization of full rank processes in terms of their spectral measures.

Chapter IV contains an outline of the extrapolation theory of full rank processes as developed by Wiener and Masani [30] and Masani's [19] subsequent generalization of these results.

The first half of the chapter contains various solutions of the extrapolation problem together with the required material on minimal q-variate processes. In the latter half, methods of finding the generating function of a q-variate process are presented.

Chapter V concludes the paper with a discussion of the unsolved problems encountered in the study of degenerate rank processes.

CHAPTER I

MATHEMATICAL PREREQUISITES

UNIVARIATE CONCEPTS

The following univariate definitions are fundamental to the entire paper.

Definition 1 A probability space is an ordered triple (X, \mathcal{B}, P) where X is a set, \mathcal{B} is a Borel field of subsets of X , and P is a positive, real valued, countably additive set function defined on \mathcal{B} with $P(X) = 1$.

Definition 2 A random variable f on X is a complex valued function defined on X which is measurable with respect to \mathcal{B} .

Definition 3 A stochastic process (on X) is a family of random variables $\{f_t: t \in T\}$. If $T = \mathbb{R}$, the reals, the process is called a continuous parameter process; if $T = \mathbb{N}$, the integers, then the process is called a discrete parameter process.

Definition 4 A weakly stationary discrete parameter stochastic process $\{f_n: -\infty < n < \infty\}$ is a process satisfying the following two conditions:

$$1) \quad E f_n \equiv \int_X f_n(x) P(dx) = 0 \quad \forall n$$

$$2) \quad E|f_n|^2 < \infty \quad \forall n, \quad \text{and} \quad E f_m \bar{f}_n = E f_{m+h} \bar{f}_{n+h} \quad \forall n, m, h \in \mathbb{N}.$$

The function $c(m, n) = E f_m \bar{f}_n$ is called the covariance function of the process $\{f_n: -\infty < n < \infty\}$. If the process is weakly stationary, then $c(m, n) = c(m-n, 0) = \Gamma_{m-n}$.

Definition 5 Let the process $\{f_n: -\infty < n < \infty\}$ be weakly stationary. The associated sequence $\{\Gamma_n: -\infty < n < \infty\}$ is called the covariance sequence of the process $\{f_n: -\infty < n < \infty\}$.

Lemma 1 Let $\{\Gamma_n: -\infty < n < \infty\}$ be as above. Then

- i) $\Gamma_0 \geq 0$,
- ii) $\Gamma_{-t} = \bar{\Gamma}_t \quad \forall t \in \mathbb{N}$,
- iii) $|\Gamma_t| \leq \Gamma_0 \quad \forall t \in \mathbb{N}$,
- iv) $\{\Gamma_n: -\infty < n < \infty\}$ is a positive definite sequence; i.e. $\forall t_1, t_2, \dots, t_n$ in \mathbb{N} and any complex numbers c_1, c_2, \dots, c_n , $\sum_{j=1}^n \sum_{k=1}^n c_j \bar{c}_k \Gamma_{t_j - t_k} \geq 0$.

Proof Parts i) and ii) are obvious from the definition. Part iii) is a particular instance of the generalized Cauchy Schwartz inequality in the Hilbert space $L_2(X)$ discussed at the beginning of Chapter 2.

Part iv) follows from the relationship

$$\sum_{j=1}^n \sum_{k=1}^n c_j \bar{c}_k \Gamma_{t_j - t_k} = \sum_{j=1}^n \sum_{k=1}^n c_j \bar{c}_k E f_{t_j} \bar{f}_{t_k} = E \left| \sum_{j=1}^n c_j f_{t_j} \right|^2 \geq 0.$$

Under the usual interpretation of the indexing set as time, saying that a process $\{f_n: -\infty < n < \infty\}$ is weakly stationary is equivalent to

saying that the first and second moments of the random variables are independent of time.

Examples

An uncorrelated process. Let $\{f_n: -\infty < n < \infty\}$ be a sequence of independent identically distributed random variables on some probability space, each having zero mean and unit variance. Then $c(j,k) = E f_j \bar{f}_k = \delta_{jk}$, and hence $\{f_n: -\infty < n < \infty\}$ is a weakly stationary process. Any weakly stationary process having a covariance sequence of the form $\Gamma_t = \delta_{t0} k$, where k is a constant, is called an uncorrelated process.

A moving average process. Let $\{f_n: -\infty < n < \infty\}$ be an uncorrelated process, and let $\{A_k\}_{k=-\infty}^{\infty}$ be a sequence of complex numbers satisfying $\sum_{k=-\infty}^{\infty} |A_k|^2 < \infty$. Let the process $\{g_n: -\infty < n < \infty\}$ be defined by $g_n = \sum_{k=-\infty}^{\infty} A_k f_{n-k}$. The existence of g_n follows from the square summability of the sequence $\{A_k\}_{k=-\infty}^{\infty}$. Then $c(j,k) = \sum_{l=-\infty}^{\infty} A_l \bar{A}_{l+k-j}$, which is a function of only $j-k$. Thus $\{g_n: -\infty < n < \infty\}$ is a weakly stationary process. Such a process is called a moving average process.

MATRIX INTEGRATION THEORY

The material in this section is necessary for the study of multivariate processes. Only the statements of the following theorems and lemmas are included here, since their proofs are not essential to the multivariate theory. For their proofs and a more detailed discussion of this material, the reader is referred to Chapter 3 of Wiener and Masani [29].

The following notation will be used. If A is a square matrix, the symbols ΔA , τA , and A^* will denote the determinant of A , the trace of A , and the conjugate transpose of A respectively. The symbols C , D_+ , and D_- will denote the sets $|z| = 1$, $|z| < 1$, and $1 < |z| \leq \infty$ in the complex plane. The definition and important properties of the Hardy class H_δ can be found in Wiener and Masani [29].

Theorem 1 i) The space of $q \times q$ matrices with complex entries is a Banach algebra under the usual algebraic operations and either of the norms

$$|A|_B = \frac{\text{L.U.B. } \|\text{Ax}\|}{\|\text{x}\| \neq 0} \quad - \text{ Banach norm}$$

$$|A|_E = [\tau A A^*]^{1/2} = \left[\sum_{i=1}^q \sum_{j=1}^q A_{ij}^2 \right]^{1/2} \quad - \text{ Euclidean norm.}$$

ii) This space is a Hilbert space under the same operations and the inner product $(A, B) = \tau A B^* = \sum_{i=1}^q \sum_{j=1}^q A_{ij} \overline{B_{ji}}$.

However, both the Banach norm and the Euclidean norm generate equivalent topologies on this space, since $|A|_B \leq |A|_E \leq \sqrt{q} |A|_B$. In this topology, $A_n \rightarrow A$ as $n \rightarrow \infty$ if and only if each entry of A_n tends to the corresponding entry of A as $n \rightarrow \infty$.

Definition 6 i) $L_\delta, \delta > 0$, is the set of all $q \times q$ matrix valued functions $F = [F_{ij}]$ on the unit circle with complex valued entries

F_{ij} in L_δ .

ii) L_∞ is the set of all $q \times q$ matrix valued functions F on the unit circle with complex valued entries F_{ij} in L_∞ , that is, each F_{ij} is essentially bounded.

Theorem 2 i) $F \in L_\delta, \delta > 0$, if and only if F has measurable entries and $|F|_E \in L_\delta$. $L_\delta, \delta \geq 1$, is a Banach space under the usual algebraic operations and the norm $|F|_\delta = [\frac{1}{2\pi} \int_0^{2\pi} |F(\theta)|_E^\delta d\theta]^{1/\delta}$.

ii) L_2 is a Hilbert space under the usual operations and the inner product $(F, G) = \frac{1}{2\pi} \int_0^{2\pi} \tau\{F(\theta)G^*(\theta)\}d\theta$, with the corresponding norm being $\|F\| = \sqrt{(F, F)} = |F|_2$.

iii) $F \in L_\infty$ if and only if F has measurable entries and $|F|_E$ is essentially bounded. L_∞ is a Banach algebra under the usual algebraic operations and the norm $|F|_\infty = \text{ESS. L.U.B. } |F(\theta)|_E, 0 \leq \theta \leq 2\pi$.

Definition 7 The Lebesgue integral of a function $F \in L_\delta, \delta \geq 1$, is defined by $\int_0^{2\pi} F(\theta)d\theta = [\int_0^{2\pi} F_{ij}(\theta)d\theta]_{ij}$.

Lemma 2 i) If $F \in L_\delta$ and $G \in L_{\delta'}$, where $\frac{1}{\delta} + \frac{1}{\delta'} = 1$, then $FG \in L_1$.

ii) If $F_n \rightarrow F \in L_\delta$ and $G_n \rightarrow G$ in $L_{\delta'}$ as $n \rightarrow \infty$, where $\frac{1}{\delta} + \frac{1}{\delta'} = 1$, then $\int_0^{2\pi} F_n(\theta)G_n(\theta)d\theta \rightarrow \int_0^{2\pi} F(\theta)G(\theta)d\theta$ as $n \rightarrow \infty$.

iii) If $F \in L_\delta, \delta > 0$, and $G \in L_\infty$, then $FG \in L_\delta$.

iv) If $\delta' > \delta > 0$, then $L_\infty \subset L_{\delta'} \subset L_\delta$ and $|F|_\infty \geq |F|_{\delta'} \geq |F|_\delta$.

v) If $F \in L_\delta, \delta > 0$, then $\Delta F \in L_{\delta/q}$.

It follows from iii), with $G(\theta) = e^{-ni\theta} I$, that every function F in $L_\delta, \delta > 0$, has an n 'th Fourier coefficient

$$A_n = \int_0^{2\pi} e^{-ni\theta} F(\theta)d\theta.$$

If $A_n = [A_{ij}^n]$ and $F(\theta) = [F_{ij}(\theta)]$, then A_{ij}^n is the n 'th Fourier coefficient of the function $F_{ij}(\theta)$.

The following theorem gives the matricial extensions of some well known results of Fourier analysis.

Theorem 3 i) Riemann Lebesgue lemma. If A_n is the n'th Fourier coefficient of $F \in L_\delta$, $\delta \geq 1$, then $A_n \rightarrow 0$ as $n \rightarrow \infty$.

ii) If A_n is the n'th Fourier coefficient of $F \in L_2$, then $\sum_{n=-\infty}^{\infty} |A_n|_E^2 < \infty$. Conversely, if the A_n are such that $\sum_{n=-\infty}^{\infty} |A_n|_E^2 < \infty$, then there is a function $F \in L_2$ such that A_n is its n'th Fourier coefficient.

iii) Parseval's identity. If $F, G \in L_2$ and have n'th Fourier coefficients A_n and B_n respectively, then $\frac{1}{2\pi} \int_0^{2\pi} F(\theta) G^*(\theta) d\theta = \sum_{n=-\infty}^{\infty} A_n B_n^*$.

iv) Convolution rule. If F and G are as in iii) then the n'th Fourier coefficient of FG is $\sum_{k=-\infty}^{\infty} A_k B_{n-k}^*$.

Definition 8 If $F \in L_1$ and A_n is its n'th Fourier coefficient, then the functions defined by

$$F_+(z) = \sum_{n=0}^{\infty} A_n z^n, \quad z \in D_+, \quad \text{and} \quad F_-(z) = \sum_{n=1}^{\infty} A_{-n} z^{-n}, \quad z \in D_-$$

are called respectively the inner and outer functions determined by F .

Theorem 4 If $F \in L_1$ and the values of F are non negative hermitian, then

- 1) $\Delta F_+(z) \in H_{\delta/q}$ on D_+ , and $\Delta F(e^{i\theta}) \in L_{\delta/q}$ on C .
- 2) Either $\Delta F_+(z)$ vanishes identically, or $\log \Delta F \in L_1$ on C and

$$\log |\Delta F_+(0)| \leq \frac{1}{2\pi} \int_0^{2\pi} \log |\Delta F(e^{i\theta})| d\theta.$$

Definition 9 Let F be a $q \times q$ matrix valued function on $[a,b]$.

Then F is of bounded variation if and only if the set of variations

$\left\{ \sum_{k=1}^n |F(x_k) - F(x_{k-1})|_E \right\}$ of F over different finite partitions

$\{x_0, x_1, \dots, x_n\}$ of $[a,b]$ is bounded above F is non decreasing if and

only if its values are Hermitian and $x' > x$ implies that $F(x') - F(x)$ is non negative Hermitian.

Lemma 3 i) $F = [F_{ij}]$ is of bounded variation on $[a,b]$ if and only if each F_{ij} is of bounded variation.

ii) If $F = [F_{ij}]$ is non decreasing and bounded on $[a,b]$, then the F_{ii} , $i = 1, 2, \dots, q$ are real valued, non decreasing and bounded on $[a,b]$. The F_{ij} , $i \neq j$, are functions of bounded variation, in general complex valued.

It follows from the above lemma and known properties of complex valued functions of bounded variation (Hewitt and Stromberg [8]) that if F is of bounded variation on $[a,b]$ it has at most countably many points of discontinuity, all of them are simple, and that F' exists a.e. and is in L_1 on $[a,b]$.

Definition 10 Let F be of bounded variation on $[a,b]$. The functions F^a , F^d , and F^s are defined as follows:

$$F^a(x) = F(a) + \int_a^x F'(t) dt ,$$

$$F^d(x) = F(a+0) - F(a) + \sum_{a < t < x} F(t+0) - F(t-0) + F(x) - F(x-0) ,$$

$$F^s(x) = F(x) - F^a(x) - F^d(x) .$$

These functions are called the absolutely continuous, discontinuous, and singular parts of F respectively.

An immediate consequence of this definition is the following lemma.

Lemma 4 i) $(F^a)' = F'$ a.e.

ii) $(F^d)' = 0$ except at the points of discontinuity of F .

iii) $(F^s)' = 0$ a.e.

iv) If the same superscript notation is adopted for the components of F , then $F^a = [F_{ij}^a]$, $F^d = [F_{ij}^d]$, and $F^s = [F_{ij}^s]$.

Theorem 6 If F is non decreasing and bounded on $[a,b]$, then so are F^a , F^d , and F^s .

Definition 11 Let F and G be $q \times q$ matrix valued functions on

$[a,b]$. Let $\pi = \{x_0, x_1, \dots, x_n\}$ be a partition of $[a,b]$, $|\pi| =$

$\max \{x_k - x_{k-1} : k=1, 2, \dots, n\}$, and $\pi^* = \{t_1, t_2, \dots, t_n\}$, where

$x_{k-1} \leq t_k \leq x_k$ for each $k = 1, 2, \dots, n$.

i) If, as $|\pi| \rightarrow 0$, $S(F, G, \pi, \pi^*) = \sum_{k=1}^n F(t_k) \{G(x_k) - G(x_{k-1})\}$ tends to a limit L , then L is called the left Riemann-Stieltjes integral of F with respect to G from a to b , and it is denoted by $\int_a^b F(x) dG(x)$.

ii) The right Riemann-Stieltjes integral $\int_a^b dG(x) F(x)$ is similarly defined.

iii) If, as $|\pi| \rightarrow 0$, $S'(F, G, \pi, \pi^*) = \sum_{k=1}^n F(t_k) \{G(x_k) - G(x_{k-1})\} F^*(t_k)$ tends to a limit L , then L is called the bilateral Riemann-Stieltjes integral of F and F^* with respect to G from a to b , and it is

denoted by $\int_a^b F(x) dG(x) F^*(x)$.

Theorem 7 i) If F is continuous and G is of bounded variation then the integrals $\int_a^b F(x) dG(x)$, $\int_a^b dG(x) F(x)$, and $\int_a^b F(x) dG(x) F^*(x)$ exist.

ii) $\int_a^b F(x) dG_1(x) \pm \int_a^b F(x) dG_2(x) = \int_a^b F(x) d(G_1(x) \pm G_2(x))$.

iii) $\int_a^b F(x) dG(x) + \int_a^b dF(x) G(x) = F(b)G(b) - F(a)G(a)$.

iv) If $G(x)$ is absolutely continuous on $[a, b]$ (that is, $G(x) = G^a(x) \forall x \in [a, b]$) then $\int_a^b F(x) dG(x) = \int_a^b F(x) G'(x) dx$.

Definition 12 Let G be of bounded variation on $[0, 2\pi]$. Then the n 'th Fourier Stieltjes coefficient of G is defined to be

$$A_n = \frac{1}{2\pi} \int_0^{2\pi} e^{-ni\theta} dG(\theta) = \left[\frac{1}{2\pi} \int_0^{2\pi} e^{-ni\theta} dG_{ij}(\theta) \right].$$

Note that the existence of A_n follows from the part i) of the previous theorem with $F(x) = e^{-ni\theta} I$.

Lemma 5 If F is of bounded variation and for all n $\int_0^{2\pi} e^{-ni\theta} dF(\theta) = 0$, then F is constant valued.

In particular, the above lemma implies that if $\int_0^{2\pi} e^{-ni\theta} dF(\theta) = \int_0^{2\pi} e^{-ni\theta} dG(\theta)$ for all n , then $F(\theta)$ and $G(\theta)$ differ by a constant matrix.

CHAPTER II

MULTIVARIATE STOCHASTIC PROCESSES

In this chapter the structural framework for the study of univariate processes is first presented, and then the multivariate generalizations are developed.

The approach taken to the multivariate theory is that which was originally proposed by Zasuhi in 1941 (see Jang Ze Pei [10]) and subsequently developed by Wiener and Masani [29] in 1958. A similar approach which doesn't rely as much on square matrix analysis can be found in Rozanov [26].

COMPLEX VALUED RANDOM VARIABLES

Let (X, B, P) be a probability space and let $L_2(X)$ denote the set of all complex valued B measurable functions f on X for which $\int_X |f(t)|^2 P(dt) < \infty$, i.e., the set of all random variables on X having finite variance. It is a well known fact (Rozanov[26]) that $L_2(X)$ is a Hilbert space under the usual operations and the inner product $(f, g) = \int_X f(t) \overline{g(t)} P(dt)$, with the corresponding norm being $\|f\| = (f, f)^{\frac{1}{2}}$.

The following notation will be used. If A and B are subsets of $L_2(X)$, then $A+B$ will denote the set of all functions $f+g$ where $f \in A$ and $g \in B$. $\sigma(f_j)_{j \in J}$ will denote the subspace (that is, closed linear manifold) spanned by the functions f_j for $j \in J$. The orthogonal

projection of a function f onto the subspace M of $L_2(X)$ will be written as $(f|M)$.

A weakly stationary discrete parameter stochastic process is then a sequence $\{f_n: -\infty < n < \infty\}$ in $L_2(X)$ which, from a geometric point of view, can be thought of as a single parameter curve in $L_2(X)$.

The basic space studied in the analysis of multivariate processes is similar to the space $L_2(X)$: however, the elements of $L_2(X)$ are replaced by vectors of random variables, $q \times q$ complex matrices are used instead of complex coefficients, and the inner product is replaced by the matrix valued Gramian inner product.

MULTIVARIATE RANDOM VARIABLES

Definition 13 $\ell_2 = \ell_2(X)$ is the set of all q dimensional column vector valued functions F on X whose components F^i , $i = 1, 2, \dots, q$ are elements of $L_2(X)$.

It follows easily from the properties of $L_2(X)$ that ℓ_2 is a Hilbert space under the usual operations and the inner product

$$((F, G)) = \int_X \sum_{i=1}^q \overline{F^i(t)} G^i(t) P(dt) .$$

The corresponding norm is $\|F\| = ((F, F))^{1/2}$. A sequence $\{F_n: -\infty < n < \infty\}$ is said to converge to F in ℓ_2 if and only if $\|F_n - F\| \rightarrow 0$ as $n \rightarrow \infty$. This is equivalent to $F_n^j \rightarrow F^j$ in $L_2(X)$ as $n \rightarrow \infty$ for each $j = 1, 2, \dots, q$.

Definition 14 If F and G are in ℓ_2 , then the matrix

$$(F, G) = [(F^i, G^j)] = \left[\int_X \overline{F^i(t)} G^j(t) P(dt) \right]$$

is called the Gramian of the pair F and G .

Note that $((F,G)) = \tau(F,G)$ and $\|F\| = [\tau(F,F)]^{1/2}$. Although the concept of orthogonality already exists in ℓ_2 , the following definitions which are based on the Gramian rather than on the inner product are used.

Definition 15 Let F_n, F , and G be in ℓ_2 . Then

- i) $F \perp G$ if and only if $(F,G) = 0$.
- ii) F is a normal vector if and only if $(F,F) = I$.
- iii) $\{F_j: j \in J\}$ is an orthonormal set if and only if $(F_m, F_n) = \delta_{mn} I$.

Definition 16 i) A linear manifold in ℓ_2 is a non empty subset M of ℓ_2 such that if $F, G \in M$ then $AF + BG \in M$ for any $q \times q$ complex matrices A and B .

ii) A subspace of ℓ_2 is a linear manifold which is closed in the topology generated by the norm.

iii) The subspace (respectively linear manifold) spanned by a subset M of ℓ_2 is the intersection of all subspaces (respectively linear manifolds) containing M . As before, the subspace spanned by the set $\{F_j: j \in J\}$ will be denoted by $\sigma(F_j)_{j \in J}$.

iv) If $M_j \subset \ell_2$ for $j \in J$, then $\sum_{j \in J} M_j$ will denote the set of all sums $\sum_{j \in J} F_j$, with $F_j \in M_j$, which converge in the topology of the norm.

The following lemma is a trivial consequence of the above definitions.

Lemma 6 i) $(G,F) = (F,G)^*$, and (F,F) is non negative definite.

- ii) If $F_n \rightarrow F$ and $G_n \rightarrow G$ as $n \rightarrow \infty$, then $(F_n, G_n) \rightarrow (F, G)$.
- iii) $(\sum_{i=1}^n A_i F_i, \sum_{j=1}^m B_j G_j) = \sum_{i=1}^n \sum_{j=1}^m A_i (F_i, G_j) B_j^*$.
- iv) $F \perp G$ if and only if $F^i \perp G^j$ in $L_2(X)$, for $i, j = 1, 2, \dots, q$.
- v) F is a normal vector if and only if its components form an orthonormal set in $L_2(X)$.
- vi) The set $\{F_j: j \in J\}$ is orthonormal if and only if the set of all components of all the $F_j, j \in J$ forms an orthonormal set in $L_2(X)$.
- vii) Let $\{F_n: -\infty < n < \infty\}$ be a sequence in ℓ_2 satisfying $(F_n, F_m) = \delta_{nm} K$. If the matrix K is invertible and one defines G_n by $G_n = \sqrt{K^{-1}} F_n$, then the sequence $\{G_n: -\infty < n < \infty\}$ is orthonormal.

The proof of the following lemma, although lengthy, is of a routine nature.

- Lemma 7 i) M is a subspace of ℓ_2 if and only if there is a subspace M of $L_2(X)$ such that $M = M^q$. The subspace M is the set of all components of all elements of M .
- ii) Let M be a subspace of ℓ_2 and let $F \in \ell_2$. Then there exists a unique $G \in M$ such that $\|F-G\| \leq \|F-H\|$ for all $H \in M$. For this G , $G^i = (F^i | M)$, where M is as in i) above. A function $G \in M$ satisfies this inequality if and only if $(F-G) \perp M$.
- iii) If M and N are subspaces of ℓ_2 and $M \subset N$, then there exists a unique subspace M' of ℓ_2 such that $N = M + M'$ and $M \perp M'$.
- iv) Let $\{G_j: j \in J\}$ be a subset of ℓ_2 . Then $F \in \sigma(G_j)_{j \in J}$ if and only if $F = \lim_{n \rightarrow \infty} F_n$, where each F_n is a finite linear combination, with matrix coefficients, of the $G_j, j \in J$.

v) Let $M = \sigma(F_j)_{j \in J} \subset \ell_2$ and $M^i = \sigma(F_j^i)_{j \in J} \subset L_2(X)$. If M is as in i) above, then $M = \text{closure} \sum_{j=1}^{\infty} M^j$.

Definition 17 The function G of part ii) above is called the orthogonal projection of F onto M and is denoted by $(F|M)$.

The last part of lemma 6 illustrates one way in which the multivariate theory differs from the univariate case. Whereas in the univariate theory any orthogonal sequence in $L_2(X)$ can be normalized, this is not in general possible in the space ℓ_2 . If the matrix K of lemma 6, part vii) is not invertible, very little can be done (see however Masani [21]). Therefore it becomes necessary to take Fourier expansions of functions with respect to sequences that are orthogonal but not necessarily orthonormal.

Theorem 8 Let $\{\varphi_n : -\infty < n < \infty\}$ be a sequence in ℓ_2 with $(\varphi_m, \varphi_n) = \delta_{mn} K$, $K \neq 0$. If $F = \sum_{j=-\infty}^{\infty} A_j \varphi_j$ and $G = \sum_{j=-\infty}^{\infty} B_j \varphi_j$, then

i) $(F, G) = \sum_{j=-\infty}^{\infty} A_j K B_j^*$, $\|F\|^2 = \sum_{j=-\infty}^{\infty} |A_j K^{\frac{1}{2}}|_E^2 < \infty$, and $(F, \varphi_n) = A_n K$.

ii) The linear manifold $\sum_{j=-\infty}^{\infty} \sigma(\varphi_j)$ is closed and identical to $\sigma(\varphi_j)_{j=-\infty}^{\infty}$.

iii) For any $H \in \ell_2$ there exist matrices C_j such that $(H|\sigma(\varphi_j)_{j=-\infty}^{\infty}) = \sum_{j=-\infty}^{\infty} C_j \varphi_j$, and $(H, \varphi_n) = C_n K$.

Proof i) This follows from the linearity and convergence properties of the Gramian as given in lemma 6; noting that $\|F\|^2 = \tau(F, F)$ gives the equation for $\|F\|^2$. If one takes $G = \varphi_n$, then $B_j = \delta_{jn} I$ and hence $(F, \varphi_n) = (F, G) = \sum_{j=-\infty}^{\infty} A_j K B_j^* = A_n K$.

ii) It is easily seen that $\sum_{j=-\infty}^{\infty} \sigma(\varphi_j) \subset \sigma(\varphi_j)_{j=-\infty}^{\infty}$. The proof of the reverse inclusion, which requires a lemma on Hermitian matrices, can be found in Wiener and Masani [29].

iii) Let $\hat{H} = (H | \sigma(\varphi_j)_{j=-\infty}^{\infty})$. Since $\sigma(\varphi_j)_{j=-\infty}^{\infty} = \sum_{j=-\infty}^{\infty} \sigma(\varphi_j)$ by part ii), there exist matrices C_j such that $\hat{H} = \sum_{j=-\infty}^{\infty} C_j \varphi_j$ and $(\hat{H}, \varphi_n) = C_n K$. By part ii) of lemma 7 $(H - \hat{H}) \perp \sigma(\varphi_j)_{j=-\infty}^{\infty}$; hence $(H - \hat{H}, \varphi_n) = (H, \varphi_n) - (\hat{H}, \varphi_n) = 0$. Thus $(H, \varphi_n) = C_n K$.

TIME DOMAIN ANALYSIS

The following is the multivariate generalization of definitions 3 and 4.

Definition 18 A q dimensional discrete parameter weakly stationary stochastic process (or equivalently, a q -variate process) is a sequence $\{F_n : -\infty < n < \infty\}$ in ℓ_2 having the property that the Gramian $(F_m, F_n) = \Gamma_{m-n} = [\gamma_{ij}^{(m-n)}]$ depends only on the difference $m-n$ and not on m or n separately. The sequence $\{\Gamma_n : -\infty < n < \infty\}$ is called the covariance sequence of the process, and Γ_n is called the covariance matrix for lead n .

The study of multivariate processes as sequences in ℓ_2 is referred to as time domain analysis. This is as opposed to spectral analysis, which is examined in the next chapter.

Every q -variate process $\{F_n : -\infty < n < \infty\}$ has associated with it q univariate weakly stationary processes, since $(F_m^j, F_n^j) = \gamma_{jj}^{(m-n)}$. The converse however is obviously false.

The next lemma is the multivariate extension of lemma 1.

Lemma 8 Let $\{\Gamma_n : -\infty < n < \infty\}$ be the covariance sequence of the q -

variate process $\{\Gamma_n: -\infty < n < \infty\}$. Then

i) $\Gamma_0 \geq 0$, i.e. the matrix Γ_0 is non negative definite,

ii) $\Gamma_{-n} = \Gamma_n^*$,

iii) The sequence $\{\Gamma_n: -\infty < n < \infty\}$ is non negative definite: if

C_1, C_2, \dots, C_n are any $n \times q \times q$ complex matrices and t_1, t_2, \dots, t_n are any integers, then $\sum_{i=1}^n \sum_{j=1}^n C_i \Gamma_{t_i - t_j} C_j^* \geq 0$.

Proof Part i) follows from lemma 6, and ii) is obvious. For iii), let x be any row vector in C^q . Then

$$x \left(\sum_{i=1}^n \sum_{j=1}^n C_i \Gamma_{t_i - t_j} C_j^* \right) x^* = \sum_{i=1}^n \sum_{j=1}^n x C_i (F_{t_i}, F_{t_j}) C_j^* x^* = E \left| \sum_{i=1}^n x C_i F_{t_i} \right|^2 \geq 0.$$

Let $\{F_n: -\infty < n < \infty\}$ be a q -variate process. It is well known (see Rozanov [26]) that there exists a unitary operator U on $L_2(X)$ satisfying $UF_n^j = F_{n+1}^j$ for $j = 1, 2, \dots, q$ and any n . The operator U may not be unique: however, if V is another unitary operator satisfying the above conditions then U and V will agree on the subspace of $L_2(X)$ spanned by $\{F_n^j: -\infty < n < \infty, j = 1, 2, \dots, q\}$. Since this is the largest subspace of $L_2(X)$ that is ever considered, U can then be considered unique.

Definition 19 The operator U is called the shift operator of the process $\{F_n: -\infty < n < \infty\}$, and one writes $UF_n = F_{n+1}$.

Definition 20 Let $\{F_n: -\infty < n < \infty\}$ be a q -variate process. The subspace of ℓ_2 spanned by $\{F_k: -\infty < k \leq n\}$ is called the present and past of F_n , and is denoted by M_n . The subspace $M_{-\infty} = \bigcap_{n=-\infty}^{\infty} M_n$ is called

the remote past of the process, while $M_\infty = \text{closure } \bigcup_{n=-\infty}^{\infty} M_n$ is called the space spanned by the process.

The present and past of F_n^j of the component processes $\{F_n^j: -\infty < n < \infty\}$ will be denoted by M_n^j , which is a subspace of $L_2(X)$. The spaces $M_{-\infty}^j$ and M_∞^j are defined analogously to $M_{-\infty}$ and M_∞ .

The proof of the next lemma is clear from the above definitions and Lemma 7.

Lemma 9 Let $\{F_n: -\infty < n < \infty\}$ be a q -variate process. Then

- i) $M_{-\infty} \subset M_n \subset M_{n+1} \subset M_\infty$,
- ii) $M_{-\infty}^j \subset M_n^j \subset M_{n+1}^j \subset M_\infty^j$ for $j = 1, 2, \dots, q$,
- iii) $U^n M_k = M_{k+n}$,
- iv) $U^n (F_j | M_k) = (F_{j+n} | M_{k+n})$,
- v) $G \in M_n$ if and only if $G^j \in \text{closure } \sum_{i=1}^q M_n^i$ for $j = 1, 2, \dots, q$.

Definition 21 The q -variate process $\{F_n: -\infty < n < \infty\}$ is non deterministic if for some n , $F_n \notin M_{n-1}$. The process is minimal if for some n , $F_n \notin \sigma(F_k)_{k \neq n}$.

Since the process $\{F_n: -\infty < n < \infty\}$ is weakly stationary, it follows that $F_n \notin M_{n-1}$ for some particular n if and only if $F_n \notin M_{n-1}$ for all n . Similarly $F_n \notin \sigma(F_k)_{k \neq n}$ holds for some particular n if and only if it holds for all n .

Let $\{F_n: -\infty < n < \infty\}$ be a non deterministic process. Then $G_n = F_n - (F_n | M_{n-1})$ is different from 0 for every n . Intuitively, this

means that no matter how much information about the past behaviour of the process one is given, it is impossible to linearly predict exactly the behaviour of the process at the next moment in time.

Lemma 10 Let $\{F_n: -\infty < n < \infty\}$ and $\{G_n: -\infty < n < \infty\}$ be as above.

Then

i) $\{G_n: -\infty < n < \infty\}$ is a q -variate process with the same shift operator as $\{F_n: -\infty < n < \infty\}$, i.e., $UG_n = G_{n+1}$.

ii) The process $\{G_n: -\infty < n < \infty\}$ is uncorrelated: $(G_m, G_n) = \delta_{mn} G$, where $G = (G_0, G_0) = (G_n, G_n)$.

Proof i) $(G_m, G_n) = (F_m - (F_m | M_{m-1}), F_n - (F_n | M_{n-1})) = (U^h[F_m - (F_m | M_{m-1})], U^h[F_n - (F_n | M_{n-1})]) = (F_{m+h} - (F_{m+h} | M_{m+h-1}), F_{n+h} - (F_{n+h} | M_{n+h-1})) = (G_{m+h}, G_{n+h})$. $UG_n = UF_n - U(F_n | M_{n-1}) = F_{n+1} - (F_{n+1} | M_n) = G_{n+1}$.

ii) If $m = n$, then $(G_m, G_n) = (G_n, G_n) = (U^n G_0, U^n G_0) = (G_0, G_0) = \delta_{nn} G$.

One may assume without loss of generality that $m < n$. Then $(G_m, G_n) = (F_m - (F_m | M_{m-1}), F_n - (F_n | M_{n-1}))$. Since $(F_m | M_{m-1}) \in M_{m-1} \subset M_m$, it follows that $F_m - (F_m | M_{m-1}) \in M_m \subset M_n$, since $m < n$. From lemma 7, $H - (H | S) \perp S$ for any $H \in \ell_2$ and any subspace S of ℓ_2 . Hence

$(F_m - (F_m | M_{m-1}), F_n - (F_n | M_{n-1})) = 0$. Thus $(G_m, G_n) = 0 = \delta_{mn} G$.

Definition 22 Let the q -variate processes $\{F_n: -\infty < n < \infty\}$ and

$\{G_n: -\infty < n < \infty\}$ be as above. The process $\{G_n: -\infty < n < \infty\}$ is

called the innovation process of the process $\{F_n: -\infty < n < \infty\}$. The

Gramian matrix $G = (G_0, G_0)$ is called the prediction error matrix for lag 1.

Definition 23 The rank of the q -variate process $\{F_n: -\infty < n < \infty\}$ is the rank of the prediction error matrix for lag 1, and is denoted by ρ . If $\rho = q$ then the process is said to be of full rank: otherwise, it is of degenerate rank.

It follows that a q -variate process is non deterministic if and only if $\rho \geq 1$. If $\rho \neq q$, however, then the prediction error matrix for lag 1 is not invertible, and hence, by previous remarks, the innovation process cannot be normalized. Thus problems arise in the multivariate theory of degenerate rank processes that do not arise in the univariate theory.

The following definition and theorem delineate those processes which are, in a sense, the most non deterministic processes of all.

Definition 24 A process $\{F_n: -\infty < n < \infty\}$ is regular if and only if $(F_0 | M_{-n}) \rightarrow 0$ as $n \rightarrow \infty$.

Note In the literature there are cases of conflicting use of terminology, and also cases of different terms used for the same concept. For example, Doob [4] uses the term regular in a sense different from the above. In other papers, such as Masani [21] a process that is regular by the above definition is called a purely non deterministic process. Conflicting definitions of rank also exist: see chapter 4.

The proof of the following theorem can be found in Wiener and Masani [29].

Theorem 9 The following conditions on a q -variate process are equivalent:

- i) $\{F_n: -\infty < n < \infty\}$ is a regular process,

ii) $\{F_n: -\infty < n < \infty\}$ is a one sided moving average process, i.e.,

$$F_n = \sum_{k=0}^{\infty} A_k \varphi_{n-k}, \text{ where } (\varphi_m, \varphi_n) = \delta_{mn} K,$$

iii) $M_{-\infty} = \{0\}$.

The next lemma is the justification of the name innovation process for the process $\{G_n: -\infty < n < \infty\}$. It is the basis for the Wold decomposition theorem.

Lemma 11 Let $\{F_n: -\infty < n < \infty\}$ be a q -variate process, and let

$\{G_n: -\infty < n < \infty\}$ be its innovation process. Let M_n be the present and past of F_n , and let N_n be the present and past of G_n . Then

i) if $m < n$, $M_n = M_m + \sigma(G_k)_{k=m+1}^n$, and $M_m \perp \sigma(G_k)_{k=m+1}^n$.

ii) $M_n = M_{-\infty} + N_n$, $M_{-\infty} \perp N_{\infty}$.

Proof i) Since $G_k = F_k - (F_k | M_{k-1})$ and $M_{k-1} \subset M_k$, $G_k \in M_k$ for any k .

Thus $M_m + \sigma(G_k)_{k=m+1}^n \subset M_n$. On the other hand, $F_n = F_n - (F_n | M_{n-1}) +$

$(F_n | M_{n-1}) = G_n + (F_n | M_{n-1}) \in M_{n-1} + \sigma(G_k)_{k=m+1}^n$. Thus $M_n = M_{n-1} +$

$\sigma(G_k)_{k=n}^n$. Continuing this decomposition, $M_n = M_{n-2} + \sigma(G_k)_{k=n-1}^n = \dots$

$= M_m + \sigma(G_k)_{k=m+1}^n$. Since $G_k = F_k - (F_k | M_{k-1})$, $G_k \perp M_{k-1}$ for any k .

In particular, $G_{m+1} \perp M_m$, $G_{m+2} \perp M_{m+1} \supset M_m$, ..., and $G_n \perp M_{n-1} \supset M_m$.

Since each G_k , $k = m+1, \dots, n$ is orthogonal to M_m , it follows that

$$M_m \perp \sigma(G_k)_{k=m+1}^n.$$

ii) Let $H \in M_{-\infty}$. Then $H \in M_n$ for each n . Since $M_n \perp G_{n+1}$, it

follows that $H \perp G_n$ for each n . Hence $H \perp \sigma(G_k)_{k=-\infty}^{\infty} = N_{\infty}$. Because

this is true for every $H \in M_{-\infty}$, $M_{-\infty} \perp N_{\infty}$. Since $G_n \in M_n$ and $M_{-\infty} \subset$

M_n it follows that $M_{-\infty} + N_n \subset M_n$. To prove the reverse inclusion, it will be shown that if $H \in M_n$, then $(H | M_{-\infty} + N_n)$ is just H . Since $M_{-\infty} \perp N_n$, by an earlier lemma $(H | M_{-\infty} + N_n) = (H | M_{-\infty}) + (H | N_n)$.

Hence $(H | M_{-\infty}) = \lim_{m \rightarrow -\infty} (H | M_m) + \lim_{m \rightarrow -\infty} (H | \sigma(G_k)_{k=m+1}^n) = \lim_{m \rightarrow -\infty} (H | M_m + \sigma(G_k)_{k=m+1}^n)$
 $= \lim_{m \rightarrow -\infty} (H | M_m) = (H | M_n) = H$, the second and 3rd equalities following from part i) above. Thus $M_n = M_{-\infty} + N_n$.

The next theorem is the multivariate extension of the classical Wold decomposition theorem. Only an outline of the proof given by Wiener and Masani is given here. For details, see Wiener and Masani [29].

Theorem 10 (Wold Decomposition) Let $\{F_n : -\infty < n < \infty\}$ be a q -variate

process and let $\{G_n : -\infty < n < \infty\}$ be its innovation process. If M_n and N_n are the present and past of F_n and G_n respectively, then

i) $F_n = U_n + V_n$, where $U_n = (F_n | N_n) \perp V_n = (F_n | M_{-\infty})$.

ii) $\{U_n : -\infty < n < \infty\}$ and $\{V_n : -\infty < n < \infty\}$ are q -variate processes.

iii) The process $\{V_n : -\infty < n < \infty\}$ is deterministic, and for each n , $\sigma(V_k)_{k=-\infty}^n = M_{-\infty}$.

iv) The process $\{U_n : -\infty < n < \infty\}$ is regular, and can be written as the following moving average:

$$U_n = \sum_{k=0}^{\infty} A_k G_{n-k}, \quad \|U_n\|^2 = \sum_{k=0}^{\infty} |A_k G^{\frac{1}{2}}|^2_E < \infty,$$

where as before $G = (G_0, G_0)$ and the A_k are matrices satisfying

$$A_k G = (U_0, G_{-k}) = (F_0, G_{-k}), \quad A_0 G = G = G A_0^*.$$

Proof Part i) is obvious from the previous lemma, while the verification

of ii) is a simple consequence of the linearity of the Gramian inner product. Part iii) follows essentially from part i) of lemma 7. Since $U_n \in \sigma(G_k)_n$, it follows that $U_n = \sum_{k=0}^{\infty} A_{nk} G_{n-k}$. By use of the shift operator U of the processes $\{F_n: -\infty < n < \infty\}$ and $\{G_n: -\infty < n < \infty\}$, it can be shown that the coefficient A_{nk} is independent of n . That the process $\{U_n: -\infty < n < \infty\}$ is regular follows from theorem 9.

Corollary If the process $\{F_n: -\infty < n < \infty\}$ is of full rank, then the process $\{U_n: -\infty < n < \infty\}$ can be represented as

$$U_n = \sum_{k=0}^{\infty} C_k H_{n-k}, \quad \|U_n\|^2 = \sum_{k=0}^{\infty} |C_k|_E^2 < \infty$$

where $\{H_n: -\infty < n < \infty\}$ is the normalized innovation process:

$$H_n = \sqrt{G^{-1}} G_n, \quad C_k = (U_0, H_{-k}) = A_k G^{\frac{1}{2}} \text{ where } A_k G \text{ is as above, and } A_0 \sqrt{G} = \sqrt{G}.$$

Definition 25 The function $\phi(e^{i\theta}) = \sum_{k=0}^{\infty} A_k G^{\frac{1}{2}} e^{ki\theta}$ is called the generating function of the process $\{F_n: -\infty < n < \infty\}$.

The converse to the above theorem is not true. Robertson [24] has an example of a decomposition of a process into orthogonal regular and deterministic processes which are not the component processes in the Wold decomposition. He has also given necessary and sufficient conditions for such a decomposition to in fact be Wold's decomposition (see also Jang Ze Pei [10]).

The importance of the Wold decomposition theorem cannot be overestimated. In the study of prediction theory it says, in effect, that the only kind of processes which need really be investigated are regular processes.

Examples

Uncorrelated q-variate process. Let K be a non negative Hermitian matrix, let $\{f_n^i: -\infty < n < \infty\}$ be an uncorrelated univariate process in $L_2(X)$ for each $i = 1, 2, \dots, q$, such that any two of these processes are uncorrelated with each other except at the same time, i.e. $E f_n^{(i)} \overline{f_n^{(j)}} = \delta_{mn} K_{ij}$. If $F_n = [f_n^1, f_n^2, \dots, f_n^q]^T$, then $\{F_n: -\infty < n < \infty\}$ is an uncorrelated q-variate process with $(F_m, F_n) = \delta_{mn} K$.

Moving average process. The simplest example of such a q-variate process

is that which is constructed from q-variate moving averages and placing them side by side as in the above example. Let $\{f_n^i: -\infty < n < \infty\}$ be a

univariate moving average process for each $i = 1, 2, \dots, q$: $f_n^i = \sum_{k=-\infty}^{\infty} c_k^{(i)} g_{n-k}^{(i)}$, $\sum_{k=-\infty}^{\infty} |c_k^{(i)}|^2 < \infty$, and $(g_m^{(i)}, g_n^{(i)}) = \delta_{mn} K_{ii}$, $K_{ii} > 0$.

If the uncorrelated univariate processes $\{g_n^i: -\infty < n < \infty\}$ also satisfy

$E g_m^{(i)} \overline{g_n^{(j)}} = \delta_{mn} K_{ij}$ then $\{F_n: -\infty < n < \infty\}$ is a q-variate moving

average process: $F_n = \sum_{k=-\infty}^{\infty} C_k G_{n-k}$, where $F_n = [f_n^{(1)}, f_n^{(2)}, \dots, f_n^{(q)}]^T$, $G_n = [g_n^{(1)}, g_n^{(2)}, \dots, g_n^{(q)}]^T$, $C_k = \text{diag}(c_k^{(1)}, c_k^{(2)}, \dots, c_k^{(q)})$, $\sum_{k=-\infty}^{\infty} |C_k|^2 < \infty$.

It follows then that $(F_m, F_n) = \sum_{j=-\infty}^{\infty} C_j K C_{n-m+j}$, since $(G_m, G_n) = \delta_{mn} K$.

A more interesting example of a q-variate moving average process is the following. Let the process $\{g_n^{(i)}: -\infty < n < \infty\}$ be as in the above example, and define the process $\{f_n^{(i)}: -\infty < n < \infty\}$ as follows:

$f_n^{(i)} = \sum_{k=-\infty}^{\infty} c_k^{(i)} g_{n-k}^{(i)} + \sum_{k=-\infty}^{\infty} c_k^{(i+1)} g_{n-k}^{(i+1)}$, for $i = 1, 2, \dots, q-1$ and $f_n^{(q)} = \sum_{k=-\infty}^{\infty} c_k^{(q)} g_{n-k}^{(q)}$. Then $F_n = \sum_{k=-\infty}^{\infty} D_k G_{n-k}$ where F_n and G_n are as above and $D_k =$

$$\begin{bmatrix} c^{(1)} & c^{(2)} & 0 & \dots & 0 \\ 0 & c^{(2)} & c^{(3)} & \dots & 0 \\ \vdots & & & & \vdots \\ \vdots & & & c^{(q-1)} & c^{(q)} \\ 0 & \dots & \dots & 0 & c^{(q)} \end{bmatrix}.$$

Thus $\{F_n: -\infty < n < \infty\}$ is a q -variate moving average process.

Processes of arbitrary rank. Consider the example of an uncorrelated

process as given above. It follows from lemma 7 and the condition

$E f_m^{(i)} \overline{f_n^{(j)}} = \delta_{mn} K_{ij}$ that $(F_v | M_0) = 0$ if $v > 0$. Thus if

$\{G_n: -\infty < n < \infty\}$ is the innovation process of the process

$\{F_n: -\infty < n < \infty\}$, then $G_n = F_{n+1}$. Thus the prediction error matrix

for lag 1 of the process $\{F_n: -\infty < n < \infty\}$ is just $(G_0, G_0) = (F_1, F_1) = K$.

Thus the rank of the process $\{F_n: -\infty < n < \infty\}$ is just the rank of the

matrix K , and hence through suitable choices of K processes of any given

rank can be constructed.

CHAPTER III

SPECTRAL ANALYSIS

SPECTRAL THEORY

In this section a more detailed investigation will be made of the shift operator U of a q -variate process. This will lead to the study of the space $L_{2,m}$ of functions on C which are square integrable with respect to a particular matrix valued measure M closely related to U . It will then be shown that this space $L_{2,m}$ is in fact isomorphic to M_∞ .

Let U be the shift operator of the q -variate process $\{F_n : -\infty < n < \infty\}$. It is known (see Wiener and Masani [29]) that because U is unitary it can be written as

$$U = \int_0^{2\pi} e^{-i\theta} E(d\theta) ,$$

where E is a projection valued measure over $([0, 2\pi], \mathcal{B})$, \mathcal{B} being the class of Borel subsets of $[0, 2\pi]$. Two measures can then be associated with the process $\{F_n : -\infty < n < \infty\}$.

i) An ℓ_2 valued, countably additive, orthogonally scattered (c.a.o.s.) measure ξ defined by

$$\xi(B) = E(B)F_0 \quad B \in \mathcal{B} .$$

The measure has the property that if $B \cap C = \emptyset$, then $(\xi(B), \xi(C)) = 0$.

ii) A $q \times q$ non negative Hermitian matrix valued measure M defined by

$$M(B) = (\xi(B), \xi(B)) \quad B \in \mathcal{B}.$$

Definition 26 The matrix valued measure M described above is called the spectral measure of the process $\{F_n : -\infty < n < \infty\}$. The matrix valued function $F(\theta), \theta \in [0, 2\pi]$, defined by $F(\theta) = 2\pi M([0, \theta])$, is called the spectral distribution function of the process.

It is possible (see Doob [4]) to define integrals of a complex valued function ϕ on $[0, 2\pi]$ with respect to the components of the measures and M , since they have properties akin to those of a process with orthogonal increments. Hence the following makes sense.

Definition 27 Let ξ , M , and F be as above, and let $\phi: [0, 2\pi] \rightarrow \mathbb{C}$.

The integrals

$$\int_0^{2\pi} \phi(\theta) \xi(d\theta), \quad \int_0^{2\pi} \phi(\theta) M(d\theta), \quad \text{and} \quad \int_0^{2\pi} \phi(\theta) dF(\theta)$$

are defined to be, respectively,

$$\left[\int_0^{2\pi} \phi(\theta) \xi^j(d\theta) \right], \quad \left[\int_0^{2\pi} \phi(\theta) M^{ij}(d\theta) \right], \quad \text{and} \quad \left[\int_0^{2\pi} \phi(\theta) dF^{ij}(\theta) \right].$$

The following theorem relates these measures to the process $\{F_n : -\infty < n < \infty\}$.

Theorem 11 i) $F_n = \int_0^{2\pi} e^{-ni\theta} E(d\theta) F_0 = \int_0^{2\pi} e^{-ni\theta} \xi(d\theta).$

ii) $\Gamma_n = \int_0^{2\pi} e^{-ni\theta} M(d\theta) = \frac{1}{2\pi} \int_0^{2\pi} e^{-ni\theta} dF(\theta).$

Proof i) It is easily seen that $U^n = \int_0^{2\pi} e^{-ni\theta} E(d\theta)$; this follows

from the orthogonality of the projection valued measure E . Since

$$F_n = U^n F_0, \text{ it follows that } F_n = \int_0^{2\pi} e^{-ni\theta} E(d\theta) F_0.$$

ii) $\Gamma_n = (F_n, F_0) = (\int_0^{2\pi} e^{-ni\theta} E(d\theta) F_0, \int_0^{2\pi} E(d\theta) F_0)$. By representing these integrals as limits of approximating sums, the result follows from the linearity and limiting properties of the Gramian in lemma 6.

Theorem 12 Let $\{F_n: -\infty < n < \infty\}$ be a q -variate process and let F be its spectral distribution function. Then F is bounded, non decreasing, and right continuous on $[0, 2\pi]$ with $F(0) = 0$.

Proof It is clear from the properties of the measure E that F is non decreasing and $F(0) = 0$. F is bounded and right continuous because $E([0, 2\pi]) = I$ and $\lim_{x \rightarrow x_0} E([0, x]) = E([0, x_0])$ (see [29]).

It is extremely useful to know when a given matrix valued function F is in fact the spectral distribution function of some q -variate process. In [2] Cramer proved that the above conditions are in fact sufficient. This result is extremely useful in the construction of examples.

Theorem 13 (Cramer) Let F be a $q \times q$ complex matrix valued function on $[0, 2\pi]$ which satisfies the following: F is bounded, non decreasing, right continuous, and $F(0) = 0$. Then there exists a q -variate process $\{F_n: -\infty < n < \infty\}$ such that $F(\theta)$ is its spectral distribution function.

From the above theorems and the theory of complex valued functions (see Hewitt and Stromberg [8]) the following theorem can be proved.

Theorem 14 Let $F(\theta)$ be the spectral distribution function of the process $\{F_n: -\infty < n < \infty\}$. Then F has a derivative a.e. (Leb.) which has non negative Hermitian values and is in L_1 .

Definition 28 If the spectral distribution function F is in fact absolutely continuous (with respect to Lebesgue measure) on $[0, 2\pi]$ then the matrix valued function $F'(\theta)$ is called the spectral density function of the process.

Let $\{f_n: -\infty < n < \infty\}$ be a 1-variate process. It is a well known fact that the space M_∞ of this process is isomorphic to the space of complex valued functions on $[0, 2\pi]$ which are square integrable with respect to the spectral measure of $\{f_n: -\infty < n < \infty\}$. Under this isomorphism the function f_n is mapped into $e^{-ni\theta}$.

The generalization of this isomorphism for a multivariate process was first accomplished by Wiener and Masani [30] under the assumption that the process is of full rank. The work of Rosenberg [27] and Rozanov [26] extended this result to all q -variate processes.

Let $\{F_n: -\infty < n < \infty\}$ be a q -variate process with spectral measure M . Let $U(\theta)$ and $V(\theta)$, $\theta \in [0, 2\pi]$, be $q \times q$ matrix valued functions. Since M is non negative Hermitian valued, it follows that for any Borel set Δ , $\tau M(\Delta) = 0$ implies $M(\Delta) = 0$. Thus the entries of M are all absolutely continuous with respect to the measure τM . Let $\frac{dM(\theta)}{d\tau M(\theta)}$ be the matrix of Radon-Nikodym derivatives of the entries of M with respect to τM .

Definition 29 The integral $\int_0^{2\pi} U(\theta)M(d\theta)V(\theta)$ is defined to be $\int_0^{2\pi} U(\theta) \frac{dM(\theta)}{dM(\theta)} V(\theta) \tau M(d\theta)$.

It can be shown (Rozanov [26]) that the value of the above integral does not in fact depend on the measure M . Any other measure P satisfying $M_{ij} \ll P$, $i, j = 1, 2, \dots, q$ could have been used instead of τM and would yield the same value for the above integral.

Definition 30 The class $L_{2,m}$ is defined as the set of all $q \times q$ matrix valued functions $V(\theta)$ on $[0, 2\pi]$ such that $\int_0^{2\pi} V(\theta)M(d\theta)V^*(\theta)$ exists.

Note that if U and V are in $L_{2,m}$ and A and B are $q \times q$ matrices, then $AU + BV$ is also in $L_{2,m}$. It is also clear that if U and V are in $L_{2,m}$ then $\int_0^{2\pi} U(\theta)M(d\theta)V^*(\theta)$ exists.

Definition 31 Let U and V be in $L_{2,m}$. The $L_{2,m}$ matrix and complex valued inner products are defined as follows:

$$(U, V)_m = \int_0^{2\pi} U(\theta)M(d\theta)V^*(\theta),$$

$$((U, V))_m = \tau(U, V)_m.$$

The following important theorem is due independently to both Rosenberg and Rosanov. For its proof see Rosenberg [27].

Theorem 15 The space $L_{2,m}$ is complete under the norm $|V|_m = ((V, V))_m^{\frac{1}{2}}$.

Recall that $M(B) = (\xi(B), \xi(B))$ for $B \in \mathcal{B}$. As a consequence of the above theorem it is possible to define integrals of the form

$\int_0^{2\pi} v(\theta) \xi(d\theta)$ for any v in $L_{2,m}$. It is then possible to prove the following theorem. For the details, see Rosenberg [27].

Theorem 16 Let ξ be the c.a.o.s. measure and M the spectral measure associated with the process $\{F_n : -\infty < n < \infty\}$. Let $S = \sigma(\xi(B)) \subset \mathcal{L}_2$, $B \in \mathcal{B}$.

Then

- i) $G \in S$ if and only if there exists $v \in L_{2,m}$ such that $G = \int_0^{2\pi} v(\theta) \xi(d\theta)$,
- ii) The function v above is uniquely defined up to a set of zero M measure,
- iii) The correspondence $v \rightarrow \int_0^{2\pi} v(\theta) \xi(d\theta)$ is an isomorphism of $L_{2,m}$ onto S : it is 1:1, onto, linear, and $(U, v)_m = (\int_0^{2\pi} U(\theta) \xi(d\theta), \int_0^{2\pi} v(\theta) \xi(d\theta))$ for all U and v in $L_{2,m}$.

Since $S = \sigma(\xi(B)) = \sigma(U^n F_0)_{n=-\infty}^{\infty} = M_{\infty}$, it follows that $L_{2,m}$ and M_{∞} are isomorphic Hilbert spaces. From theorem 11 it is known that

$F_n = \int_0^{2\pi} e^{-ni\theta} \xi(d\theta)$; hence the $L_{2,m}$ image of F_n under this isomorphism is $e^{-ni\theta} I$. The image of the shift operator U in M_{∞} is the operator $e^{-i\theta}$ (multiplication by $e^{-i\theta}$), the values of the measure $\xi(B)$ correspond to the functions χ_B , and the projection operators $E(B)$ correspond to multiplication by χ_B .

PROCESSES OF FULL RANK

In this section a characterization of full rank processes in terms of their spectral measures will be found. In the univariate theory a process

is of full rank if and only if the process is non deterministic. Since the spectral criteria for non determinism is known in the univariate case, a characterization of such full rank processes is available (see, for example, Rozanov [26]).

The solution in the multivariate case was obtained by Wiener and Masani in 1958. Rozanov also has a solution to the question of full rank; however, the reader should be warned that his definition of rank is different from that adopted here. According to Rozanov, a q -variate process

$\{F_n : -\infty < n < \infty\}$ has rank m if and only if

- i) the process has a spectral density $F'(\theta)$,
- ii) The rank of $F'(\theta)$ is m for almost all θ .

This definition has an obvious drawback in that not all processes have a rank. It can be shown that a process that is of full rank by the Wiener and Masani definition is also of full rank according to Rozanov's definition. The converse, however, is not true: see Jang Ze Pei [10].

The following theorem shows that a large class of q -variate processes do in fact have spectral density functions.

Theorem 17 Let $\{F_n : -\infty < n < \infty\}$ be a moving average process:

$$F_n = \sum_{k=-\infty}^{\infty} A_k G_{n-k}, \quad (G_m, G_n) = \delta_{mn} G, \quad \sum_{k=-\infty}^{\infty} |A_k G^{\frac{1}{2}}|^2 < \infty.$$

Then its spectral distribution function $F(\theta)$ is absolutely continuous with respect to Lebesgue measure, and $F'(e^{i\theta})$ admits the factorization

$$F'(e^{i\theta}) = \Phi(e^{i\theta})\Phi^*(e^{i\theta}), \quad \text{where } \Phi(e^{i\theta}) = \sum_{k=-\infty}^{\infty} A_k G^{\frac{1}{2}} e^{ki\theta}.$$

Proof It follows from theorem 3 that the functions Φ and Φ^* are in L_2 , which implies $\Phi\Phi^* \in L_1$. The k 'th Fourier coefficients of Φ and Φ^* are $A_k G^{\frac{1}{2}}$ and $G^{\frac{1}{2}} A_{-k}^*$ respectively. Hence, from theorem 3, the N 'th Fourier coefficient of $\Phi\Phi^*$ is $\sum_{j=-\infty}^{\infty} A_j G^{\frac{1}{2}} G^{\frac{1}{2}} A_{j-N}^* = \sum_{j=-\infty}^{\infty} A_j G A_{j-N}^*$. However, this is just Γ_n , since $\Gamma_n = (F_n, F_0) = (\sum_{k=-\infty}^{\infty} A_k G_{n-k}, \sum_{k=-\infty}^{\infty} A_k G_{-k}) = \sum_{k=-\infty}^{\infty} A_k G A_{k-n}^*$. Thus $2\pi\Gamma_n = \int_0^{2\pi} e^{-ni\theta} \Phi(e^{i\theta}) \Phi^*(e^{i\theta}) d\theta = \int_0^{2\pi} e^{-ni\theta} dF(e^{i\theta})$ by theorem 11. It then follows from theorem 7 and lemma 7 that $F(e^{i\theta}) = \int_0^\theta \Phi(e^{it}) \Phi^*(e^{it}) dt$, i.e. $F'(e^{i\theta}) = \Phi(e^{i\theta}) \Phi^*(e^{i\theta})$.

Theorem 18 Let $\{F_n: -\infty < n < \infty\}$ be a regular process:

$$F_n = \sum_{k=0}^{\infty} A_k G_{n-k}, \quad (G_m, G_n) = \delta_{mn} G, \quad \sum_{k=0}^{\infty} |A_k G^{\frac{1}{2}}|_E^2 < \infty.$$

Then $F'(e^{i\theta}) = \Phi(e^{i\theta}) \Phi^*(e^{i\theta})$, where $\Phi(e^{i\theta}) = \sum_{k=0}^{\infty} A_k G^{\frac{1}{2}} e^{ki\theta}$ and either

i) $\Delta\Phi_+(z) = 0$ for all $z \in D_+$,

ii) $\log \Delta F'(e^{i\theta}) \in L_1$ on C and $\log(\Delta A_0 G A_0^*) \leq \frac{1}{2\pi} \int_0^{2\pi} \log \Delta F'(e^{i\theta}) d\theta$.

Proof The factorization of $F'(e^{i\theta})$ is obvious from the last theorem.

Since $\Phi \in L_2$ and its n 'th Fourier coefficient is zero if $n < 0$, by theorem 4 either the first alternative above holds or else $\log \Delta\Phi \in L_1$ on C and $\log |\Delta(A_0 G^{\frac{1}{2}})| \leq \frac{1}{2\pi} \int_0^{2\pi} \log |\Delta\Phi(e^{i\theta})| d\theta$. However, this implies $\log(A_0 G A_0^*) \leq \frac{1}{2\pi} \int_0^{2\pi} \log \Delta F'(e^{i\theta}) d\theta$ since $\Delta A_0 G A_0^* = |\Delta A_0 G^{\frac{1}{2}}|^2$ and $\Delta F' = |\Delta\Phi|^2$.

Rozanov has shown that the above factorization is in fact sufficient for the process $\{F_n: -\infty < n < \infty\}$ to be regular. For details of the proof see [26].

Theorem 19 (Rozanov) The process $\{F_n: -\infty < n < \infty\}$ is regular if and only if its spectral distribution $F(\theta)$ is absolutely continuous with respect to Lebesgue measure, $R(F') = \rho$ a.e., and $F'(e^{i\theta})$ admits the factorization $F'(e^{i\theta}) = \Phi(e^{i\theta})\Phi^*(e^{i\theta})$ with $\Phi(e^{i\theta}) \in L_2^{0+}$ on C , i.e., in the set of all functions in L_2 whose n 'th Fourier coefficient vanishes for $n < 0$.

Upon combining the above results with the Wold decomposition theorem one obtains the following theorem.

Theorem 20 Let $\{F_n: -\infty < n < \infty\}$ be a non deterministic q -variate process. Let $\{U_n: -\infty < n < \infty\}$, $\{V_n: -\infty < n < \infty\}$, $\{A_k\}_{k=-\infty}^{\infty}$ and G be as in the Wold decomposition theorem, and let F , F_u , and F_v be the spectral distribution functions of the processes $\{F_n: -\infty < n < \infty\}$, $\{U_n: -\infty < n < \infty\}$, and $\{V_n: -\infty < n < \infty\}$ respectively. Then

i) $F = F_u + F_v$,

ii) F_u is absolutely continuous and $F_u' = \Phi\Phi^*$, where $\Phi(e^{i\theta}) = \sum_{k=0}^{\infty} A_k G_k^{1/2} e^{ki\theta}$,

iii) If $\{F_n: -\infty < n < \infty\}$ is of full rank then $\log \Delta F_u' \in L_1$ on C and $\log \Delta G = \frac{1}{2\pi} \int_0^{2\pi} \log \Delta F_u'(e^{i\theta}) d\theta$.

Proof i) Since the processes $\{U_n: -\infty < n < \infty\}$ and $\{V_n: -\infty < n < \infty\}$ are orthogonal, $(F_n, F_0) = (U_n + V_n, U_0 + V_0) = (U_n, U_0) + (U_n, V_0) + (V_n, U_0) + (V_n, V_0) = (U_n, U_0) + (V_n, V_0)$. Thus

$$\frac{1}{2\pi} \int_0^{2\pi} e^{-ni\theta} dF(\theta) = \frac{1}{2\pi} \int_0^{2\pi} e^{-ni\theta} dF_u(\theta) + \frac{1}{2\pi} \int_0^{2\pi} e^{-ni\theta} dF_v(\theta) = \frac{1}{2\pi} \int_0^{2\pi} e^{-ni\theta} d(F_u(\theta) + F_v(\theta)).$$

Since this is true for all n , by lemma 7 $F(\theta) = F_u(\theta) + F_v(\theta) + K$,

where K is a constant matrix. However $F(0) = F_u(0) = F_v(0) = 0$: hence

$K = 0$ and $F(\theta) = F_u(\theta) + F_v(\theta)$.

ii) This is obvious from the theorems 10 and 18.

iii) Recall that $A_0 G = G$. If the process $\{F_n: -\infty < n < \infty\}$ is of full rank G is invertible and hence $A_0 = A_0^* = I$. Thus $\Delta(\Phi_+(0)) = \Delta(G^{\frac{1}{2}}) \neq 0$ and hence the second alternate conclusion of theorem 18 must hold. The required result follows since $A_0 G A_0^* = G$.

The following lemma is required for the proof of the main theorem of this section. Its proof, while not difficult, requires some extraneous material on matrix integration theory and hence is not presented. For details of the proof the reader should consult Wiener and Masani [29].

Lemma 12 Let F be the spectral distribution function of the q -variate process $\{F_n: -\infty < n < \infty\}$. Let $P(z) = \sum_{n=0}^{\infty} A_n z^n$ and $P(F) = \sum_{n=0}^{\infty} A_n F_{-n}$.

Then

$$i) \quad (P(F), P(F)) = \frac{1}{2\pi} \int_0^{2\pi} P(e^{i\theta}) dF(e^{i\theta}) P^*(e^{i\theta}),$$

ii) $\log \Delta(P(F), P(F)) \geq \frac{1}{2\pi} \int_0^{2\pi} \log \Delta F'(e^{i\theta}) d\theta + \log |\Delta A_0|^2$; the integral on the right may be equal to $-\infty$.

Theorem 21 Let $\{F_n: -\infty < n < \infty\}$ be a q -variate process. Then $\rho = q$, i.e., the process is of full rank, if and only if $\log \Delta F' \in L_1$ on C .

When this occurs

$$\Delta G = \exp\left[\frac{1}{2\pi} \int_0^{2\pi} \log \Delta F'(e^{i\theta}) d\theta\right],$$

where G is the prediction error matrix for lag 1.

Proof \Rightarrow Let the process $\{F_n: -\infty < n < \infty\}$ be of full rank, and let $F_n = U_n + V_n$ be its Wold decomposition. From theorem 20 $\log \Delta F'_u \in L_1$ on C and $\log \Delta G \leq \frac{1}{2\pi} \int_0^{2\pi} \log \Delta F'_u(e^{i\theta}) d\theta$. Since $F' = F'_u + F'_v$ and the values of F'_v are non negative Hermitian, it is easily shown that $\Delta F'_u(e^{i\theta}) \leq \Delta F'(e^{i\theta})$. This in turn implies $\frac{1}{2\pi} \int_0^{2\pi} \log \Delta F'_u(e^{i\theta}) d\theta \leq \frac{1}{2\pi} \int_0^{2\pi} \log \Delta F'(e^{i\theta}) d\theta$. Thus $\log \Delta G \leq \frac{1}{2\pi} \int_0^{2\pi} \log \Delta F'(e^{i\theta}) d\theta$. Since the process is of full rank $\Delta G \neq 0$, which implies $\int_0^{2\pi} \log \Delta F'(e^{i\theta}) d\theta$ is greater than $-\infty$. From theorem 4 $\int_0^{2\pi} \log \Delta F'(e^{i\theta}) d\theta \leq \log \Delta \left[\frac{1}{2\pi} \int_0^{2\pi} F'(e^{i\theta}) d\theta \right] < \infty$ since $F' \in L_1$. Thus when the process is of full rank $\log \Delta F' \in L_1$ on C .

\Leftarrow Let $\log \Delta F' \in L_1$ on C . The innovation vector G_0 is, by definition, equal to $F_0 - (F_0 | M_{-1})$. Since $(F_0 | M_{-1}) \in M_{-1}$, $G_0 = \lim_{n \rightarrow \infty} G^{(n)}$, where $G^{(n)} = F_0 - \sum_{j=1}^n B_j^{(n)} F_{-j}$ by lemma 7. By part ii) of the above lemma $\log \Delta(G^{(n)}, G^{(n)}) \geq \frac{1}{2\pi} \int_0^{2\pi} \log \Delta F'(e^{i\theta}) d\theta + \log |\Delta A_0|^2$. But A_0 , the coefficient of F_0 , is just I : thus $\log |\Delta A_0|^2 = 0$. Therefore $\log \Delta(G^{(n)}, G^{(n)}) \geq \frac{1}{2\pi} \int_0^{2\pi} \log \Delta F'(e^{i\theta}) d\theta$. Upon taking the limit as $n \rightarrow \infty$, one obtains $\log \Delta G \geq \frac{1}{2\pi} \int_0^{2\pi} \log \Delta F'(e^{i\theta}) d\theta$. By assumption the last integral is finite, so $\Delta G > 0$. Thus the process $\{F_n: -\infty < n < \infty\}$ is of full rank.

When the process is in fact of full rank all of the above inequalities hold. From the first part of the above proof $\log \Delta G \leq \frac{1}{2\pi} \int_0^{2\pi} \log \Delta F'(e^{i\theta}) d\theta$, and from the second part of the proof the reverse inequality holds. Thus when the process is of full rank $\log \Delta G = \frac{1}{2\pi} \int_0^{2\pi} \log \Delta F'(e^{i\theta}) d\theta$.

Theorem 22 Let $\{F_n: -\infty < n < \infty\}$ be a full rank q -variate process, $\{U_n: -\infty < n < \infty\}$ and $\{V_n: -\infty < n < \infty\}$ the component processes in its Wold decomposition, and F, F_u , and F_v the spectral distribution functions

of these processes respectively. If $F = F^{(a)} + F^{(d)} + F^{(s)}$ is the decomposition of F into its absolutely continuous, discontinuous, and singular parts, then

$$F_u = F^{(a)}, \text{ and } F_v = F^{(d)} + F^{(s)}.$$

Proof From theorem 20 it is known that $F = F_u + F_v$ and that F_u is absolutely continuous. Hence one need only show that $F_v' = 0$ a.e. on C . Since the process is of full rank, by the previous theorem $\log \Delta F' \in L_1$ on C . Hence $\Delta F'$ is different from zero a.e., or equivalently, $R(F') = q$ a.e. The required result is then a consequence of the following two lemmas. The proof of lemma 13 can be found in Jang Ze Pei [10].

Lemma 13 $R(F') = R(F_u') + R(F_v')$ a.e. Leb.

Lemma 14 Let $\{F_n: -\infty < n < \infty\}$ be any non deterministic q -variate process and let $F_n = U_n + V_n$ be its Wold decomposition. Let F_u' be the spectral density of the process $\{U_n: -\infty < n < \infty\}$. Then $R(F_u') = \rho$ a.e. Leb., where ρ is the rank of the process $\{F_n: -\infty < n < \infty\}$.

Proof Masani [20] has shown that the generating function Φ of the process $\{F_n: -\infty < n < \infty\}$ can be written in the form $\Phi(e^{i\theta}) = \Omega(e^{i\theta})G^{\frac{1}{2}}$ where the matrix valued function Ω is invertible a.e. Leb. Since $F_u' = \Phi\Phi^*$ by theorem 20, it follows that $R(F_u') = R(G) = \rho$ a.e. Leb.

Thus if $\{F_n: -\infty < n < \infty\}$ is of full rank, $R(F_u') = q$ a.e., $R(F_v') = 0$ a.e., and hence $F_v' = 0$ a.e.

Examples

Spectra of uncorrelated processes. Let the process $\{F_n : -\infty < n < \infty\}$ be an uncorrelated q -variate process with $(F_m, F_n) = \delta_{mn} K$. Such a process has a spectral density $F'(\theta)$ which is just the constant matrix K . This is because $\frac{1}{2\pi} \int_0^{2\pi} e^{-ni\theta} K d\theta = \frac{1}{2\pi} K \int_0^{2\pi} e^{-ni\theta} d\theta = 0$ if $n \neq 0$ and equals K if $n = 0$.

Discrete spectrum. Let the process $\{F_n : -\infty < n < \infty\}$ be defined by $F_n = \sum_{k=1}^p e^{-ni\theta_k} \varphi_k$, where $\varphi_k \in \ell_2$ for each $k = 1, 2, \dots, p$ and satisfy $(\varphi_m, \varphi_n) = \delta_{mn} K$, the θ_k are arbitrary numbers in $[0, 2\pi]$, and p is some fixed finite number. Then $F_n = \int_0^{2\pi} e^{-ni\theta} E(d\theta) F_0$ where $E(\theta_k) F_0 = \varphi_k$ and $E(\Delta) F_0 = 0$ if $\Delta \cap \{\theta_1, \theta_2, \dots, \theta_p\} = \emptyset$. In this case the spectral distribution function $F(\theta)$ is a step function with jumps at the points θ_k , $k = 1, 2, \dots, p$. Thus $F(\theta) = 2\pi j K$ where the integer j satisfies $\theta_j \leq \theta \leq \theta_{j+1}$, and the process $\{F_n : -\infty < n < \infty\}$ is said to have a discrete spectrum.

It is clear from the Wold decomposition theorem that $\{F_n : -\infty < n < \infty\}$ is deterministic: if it were non deterministic it would have a regular component which has an absolutely continuous spectral distribution function, contrary to the nature of $F(\theta)$.

Spectra of moving average processes. Let $\{F_n : -\infty < n < \infty\}$ be a q -variate moving average process: $F_n = \sum_{k=-\infty}^{\infty} C_k G_{n-k}$, $\sum_{k=-\infty}^{\infty} |C_k|_E^2 < \infty$, $(G_m, G_n) = \delta_{mn} K$. From the properties of the Gramian in lemma 6 it follows that $(F_n, F_0) =$

$$\left(\sum_{k=-\infty}^{\infty} C_k G_{n-k}, \sum_{j=-\infty}^{\infty} C_j G_{-j} \right) = \sum_{k=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} C_k (G_{n-k}, G_{-j}) C_j^* =$$

$$\frac{1}{2\pi} \int_0^{2\pi} \left(\sum_{k=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} C_k e^{-(n-k+j)i\theta} C_j^* \right) d\theta = \frac{1}{2\pi} \int_0^{2\pi} e^{-ni\theta} \left(\sum_{k=-\infty}^{\infty} C_k e^{ki\theta} \right) \left(\sum_{j=-\infty}^{\infty} C_j e^{ji\theta} \right)^* d\theta .$$

Since this is true for all n , it follows that the process $\{F_n : -\infty < n < \infty\}$

has an absolutely continuous spectral distribution function F , and

$$F'(e^{i\theta}) = \left(\sum_{k=-\infty}^{\infty} C_k e^{ki\theta} \right) \left(\sum_{j=-\infty}^{\infty} C_j e^{ji\theta} \right)^* .$$

In [22] Matveev has a rather complicated example of a spectral density function of a deterministic bivariate process which has regular component processes. This example shows the great extent to which component processes can interact with each other. It also illustrates the general lack of concordance between the Wold and Lebesgue-Cramér decompositions of a q -variate process.

CHAPTER IV

EXTRAPOLATION THEORY

THE EXTRAPOLATION PROBLEM

The prediction problems associated with a q -variate process $\{F_n: -\infty < n < \infty\}$ are generally divided into three closely related categories.

Filtering theory is concerned with trying to determine F_k when all that is known are the random variables $\{F_j + N_j: j \in T\}$, where $\{N_n: -\infty < n < \infty\}$ is a q -variate "noise" process and $T \subset N$. Such problems arise frequently in communications theory, where the process $\{F_n: -\infty < n < \infty\}$ is a signal and the process $\{N_n: -\infty < n < \infty\}$ is the unwanted static that is mixed in with the signal. For further information on filtering theory see Rozanov [26] and Hannan [7].

Interpolation theory deals with the problem of determining F_k when only the part $\{F_n: n \in T\}$ is known, where $T \subset N$. Such problems are encountered in control theory, where, for example, the process $\{F_n: -\infty < n < \infty\}$ may be sampled at regular time intervals and inferences are to be drawn about the unobserved random variables.

The special case of interpolation theory that arises when T is of the form $\{k \in N: k \leq k'\}$ is called extrapolation theory.

In this paper only linear least squares extrapolation will be considered.

This is equivalent to saying that the best predictor \hat{F}_v for F_v , given the random variables $\{F_j: j \leq k'\}$, is simply $(F_v | M_{k'})$. Because of the stationarity of the process $\{F_n: -\infty < n < \infty\}$, one need only consider the case $k' = 0$.

Definition 32 $\hat{F}_v = (F_v | M_0)$, $v > 0$, is called the predictor for F_v .

There are three basic questions that can be asked about \hat{F}_v :

- i) Since $\hat{F}_v \in M_0 \subset M_\infty$, it has an isomorphic image $Y_v(e^{i\theta})$ in $L_{2,m}$. When can $Y_v(e^{i\theta})$ be found?
- ii) When does there exist an autoregressive representation of \hat{F}_v : i.e., when does there exist a series representation $\hat{F}_v = \sum_{k=0}^{\infty} A_k F_{-k}$?
- iii) When such a series exists, how can the coefficients A_k , $k = 0, 1, \dots$ be found?

Question i) was first answered by Wiener and Masani in [30] under the assumption that the process $\{F_n: -\infty < n < \infty\}$ is of full rank. The generalization of this result to processes of arbitrary rank can be found in Masani [21].

Only sufficient conditions for the existence of an autoregressive representation of \hat{F}_v have been found, and in these cases methods of determining the coefficients A_k , $k = 0, 1, \dots$ are known.

The remainder of this chapter is devoted to the papers [30] by Wiener and Masani and [19] by Masani. Only full rank processes will be considered, although a few minor generalizations of some of this material are known (Masani [21]).

SOLUTIONS OF THE EXTRAPOLATION PROBLEM

Let $\{F_n: -\infty < n < \infty\}$ be a q -variate full rank process, and let $F_n = U_n + V_n$ be its Wold decomposition. Then $\hat{F}_v = (F_v | M_0) = (U_v + V_v | M_0) = (U_v | M_0) + (V_v | M_0) = (U_v | M_0) + V_v$, since $V_v \in M_{-\infty} \subset M_0$. Thus the only process that is interesting from a probabilistic viewpoint is the regular process $\{U_n: -\infty < n < \infty\}$, since the prediction error is independent of the process $\{V_n: -\infty < n < \infty\}$: $F_v - \hat{F}_v = U_v + V_v - (U_v | M_0) - V_v = U_v - (U_v | M_0)$. It is therefore assumed throughout the remainder of this chapter that the full rank q -variate process is regular. It follows from theorem 22 that this is equivalent to assuming that the spectral density F' of the process exists.

The next two lemmas and theorem provide the solution to question i). Since the spectral density F' of the process $\{F_n: -\infty < n < \infty\}$ exists, Lebesgue measure can be used in the definition of $L_{2,m}$ instead of the measure τM . It then follows that $L_{2,m}$ is the class of all $q \times q$ matrix valued functions ϕ on C for which $\int_0^{2\pi} \phi(e^{i\theta}) F'(e^{i\theta}) \phi^*(e^{i\theta}) d\theta$ exists. With this characterization of $L_{2,m}$ the following lemma is easily proved.

Lemma 15 i) $\phi \in L_{2,m}$ if and only if $\phi \sqrt{F'} \in L_2$.

ii) If $\phi, \psi \in L_{2,m}$ then $((\phi, \psi))_m = ((\phi \sqrt{F'}, \psi \sqrt{F'}))$.

iii) If $\phi \in L_\infty$ and $\psi \in L_{2,m}$, then $\phi \psi \in L_{2,m}$.

Lemma 16 Let $\{H_n: -\infty < n < \infty\}$ be the normalized innovation process of the process $\{F_n: -\infty < n < \infty\}$, and let Φ be its generating function.

Then

- i) $e^{-ni\theta}\phi^{-1} \in L_{2,m}$ and is the isomorphic image of $H_n \in M_\infty$.
- ii) If $\psi \in L_{2,m}$ then $\psi\phi \in L_2$.
- iii) Let $\psi \in L_{2,m}$ and let A_k be the k 'th Fourier coefficient of $\psi\phi$. Then as $n \rightarrow \infty$ $(\sum_{k=-n}^n A_k e^{ki\theta})\phi^{-1} \rightarrow \psi$ in $L_{2,m}$.

Proof i) That ϕ^{-1} actually exists a.e. follows from the invertibility of F' a.e. and the fact that $F' = \phi\phi^*$ a.c.. Let ψ be the isomorphic image in $L_{2,m}$ of H_n . By using the isomorphism of $L_{2,m}$ and M_∞ and comparing Fourier coefficients one can show that $F'\psi^* = e^{ni\theta}\phi$. Since $F' = \phi\phi^*$ and ϕ' exists a.e., it follows that $\psi = e^{-ni\theta}\phi^{-1}$.

ii) Let $\psi \in L_{2,m}$. Then by the above lemma $\psi\sqrt{F'} \in L_2$, and hence $|\psi\sqrt{F'}|_E^2 \in L_1$. Since $F' = \phi\phi^*$ a.e., $|\psi\sqrt{F'}|_E^2 = \tau|\psi F'\psi^*| = \tau|\psi\phi\phi^*\psi^*| = |\psi\phi|_E^2$. Thus $|\psi\phi|_E^2 \in L_1$ and hence $\psi\phi \in L_2$.

iii) For each n $(\sum_{k=-n}^n A_k e^{ki\theta})\phi^{-1}$ is actually in $L_{2,m}$ since $\phi^{-1} \in L_{2,m}$ and $\sum_{k=-n}^n A_k e^{ki\theta} \in L_\infty$. Then $\|(\sum_{k=-n}^n A_k e^{ki\theta})\phi^{-1} - \psi\|_m = \|(\sum_{k=-n}^n A_k e^{ki\theta} - \psi\phi)\phi^{-1}\sqrt{F'}\| = \|\sum_{k=-n}^n A_k e^{ki\theta} - \psi\phi\|$ since $\sqrt{F'} = \phi$. By definition of the coefficients A_k the right inner product tends to 0, and hence $(\sum_{k=-n}^n A_k e^{ki\theta})\phi^{-1} \rightarrow \psi$ in $L_{2,m}$ as $n \rightarrow \infty$.

Theorem 23 Let ϕ be as in the above lemma. Then

$$Y_V(e^{i\theta}) = [e^{-vi\theta}\phi(e^{i\theta})]_{0^+}\phi^{-1}(e^{i\theta}),$$

where $[F]_{0^+}$ denotes the function whose Fourier series is $\sum_{k=0}^{\infty} A_k e^{ki\theta}$, A_k being the k 'th Fourier coefficient of F .

Proof It must first be shown that the above expression for $Y_V(e^{i\theta})$ is in fact in $L_{2,m}$. Since $F' = \Phi\Phi^*$ a.e., $Y_V(e^{i\theta})_{F'}(e^{i\theta})Y_V^*(e^{i\theta}) = [e^{-vi\theta}\Phi(e^{i\theta})]_{0^+}[e^{-vi\theta}\Phi(e^{i\theta})]_{0^+}^*$. This is in L_1 since each factor is in L_2 . Thus, by lemma 15, $[e^{-vi\theta}\Phi(e^{i\theta})]_{0^+}\Phi(e^{i\theta})^{-1}$ is in $L_{2,m}$.

Let $F_n = \sum_{k=0}^{\infty} C_k H_{n-k}$ as in the corollary to the Wold decomposition theorem, where $\{H_n: -\infty < n < \infty\}$ is the normalized innovation process of the regular process $\{F_n: -\infty < n < \infty\}$. Due to the orthogonality of the process $\{H_n: -\infty < n < \infty\}$ $(F_V | M_0) = \sum_{k=V}^{\infty} C_k H_{V-k} = \sum_{k=0}^{\infty} C_{k+V} H_{-k}$. Since C_k is also the k 'th Fourier coefficient of Φ , it follows that $[e^{-vi\theta}\Phi(e^{i\theta})]_{0^+}\Phi^{-1}(e^{i\theta})\Phi(e^{i\theta}) = [e^{-vi\theta}\Phi(e^{i\theta})]_{0^+} = \sum_{k=0}^{\infty} C_{v+k} e^{ki\theta} \in L_2$. By iii) of the previous lemma this implies $\lim_{n \rightarrow \infty} (\sum_{k=0}^n C_{v+k} e^{ki\theta})\Phi^{-1}(e^{i\theta}) = [e^{-vi\theta}\Phi(e^{i\theta})]_{0^+}\Phi^{-1}(e^{i\theta})$. By i) of the last lemma it follows that $[e^{-vi\theta}\Phi(e^{i\theta})]_{0^+}\Phi^{-1}(e^{i\theta})$ is the $L_{2,m}$ image of $\lim_{n \rightarrow \infty} \sum_{k=0}^n C_{v+k} H_{-k} = \hat{F}_V$. Thus $Y_V(e^{i\theta}) = [e^{-vi\theta}\Phi(e^{i\theta})]_{0^+}\Phi^{-1}(e^{i\theta})$.

Although the above result was relatively easy to obtain, the determination of \hat{F}_V in M_0 is a much more difficult problem.

It is possible to obtain approximations of \hat{F}_V in M_0 by solving systems of linear equations. If the $q \times q$ matrices $A_j^{(n)}$ are chosen so that $\sum_{j=0}^n A_j^{(n)} F_{-j} = (F_V | \sigma(F_{-k})_{k=0}^n)$, it can be shown that $\hat{F}_V = (F_V | M_0) = \lim_{n \rightarrow \infty} \sum_{j=0}^n A_j^{(n)} F_{-j}$. If the $A_j^{(n)}$ are so chosen it follows that $(F_V - \sum_{j=0}^n A_j^{(n)} F_{-j}) \perp F_{-k}$ for $k = 0, 1, \dots, n$. Thus $0 = (F_V - \sum_{j=0}^n A_j^{(n)} F_{-j}, F_{-k}) = (F_V, F_{-k}) - \sum_{j=0}^n A_j^{(n)} (F_{-j}, F_{-k})$, and hence $\Gamma_{v+k} = \sum_{j=0}^n A_j^{(n)} \Gamma_{k-j}$ for $k = 0, 1, \dots, n$. This is a system of $n+1$ equations in the $n+1$ unknown matrices $A_0^{(n)}, A_1^{(n)}, \dots, A_n^{(n)}$, and can be

written as the single matrix equation

$$[A_0^{(n)} \quad A_1^{(n)} \quad \dots \quad A_n^{(n)}] \begin{bmatrix} \Gamma_0 & \dots & \Gamma_n \\ \vdots & & \vdots \\ \vdots & & \vdots \\ \Gamma_{-n} & & \Gamma_0 \end{bmatrix} = [\Gamma_v \quad \Gamma_{v+1} \quad \dots \quad \Gamma_{v+n}] .$$

Wiener and Masani [30] have shown that the second matrix on the left is invertible if the process $\{F_n : -\infty < n < \infty\}$ is of full rank. When this occurs the unknown coefficients $A_0^{(n)}, A_1^{(n)}, \dots, A_n^{(n)}$ are given by

$$[A_0^{(n)} \quad A_1^{(n)} \quad \dots \quad A_n^{(n)}] = [\Gamma_v, \Gamma_{v+1} \quad \dots \quad \Gamma_{v+n}] \begin{bmatrix} \Gamma_0 & \dots & \Gamma_n \\ \vdots & & \vdots \\ \vdots & & \vdots \\ \Gamma_{-n} & & \Gamma_0 \end{bmatrix}^{-1}$$

The problem with this solution is that it involves the inversion of an $(n+1)q \times (n+1)q$ matrix. As $n \rightarrow \infty$ it becomes more and more difficult to actually compute this inverse. Thus more subtle techniques are required to solve the time domain extrapolation problem. To see the difficulties involved the following lemma is required. Its proof can be found in Masani [20].

Lemma 17 Let $\{F_n : -\infty < n < \infty\}$ be a q -variate process (of any rank ρ) and let Φ be its generating function. Then

- i) Φ is an optimal function in L_2^{0+} , i.e., $\Phi_+(0) \geq 0$, and if $\psi \in L_2^{0+}$ and $\psi\psi^* = \Phi\Phi^*$ a.e., leb., then $(\psi_+(0)\psi_+^*(c))^{\frac{1}{2}} \leq \Phi_+(0)$,
- ii) if $\rho = q$ then

$$\Delta\Phi_+(z) = \exp \left[\frac{1}{2\pi} \int_0^{2\pi} \frac{e^{i\theta} + z}{e^{i\theta} - z} \log \Delta[F'(\theta)]^{\frac{1}{2}} d\theta \right] .$$

Since the process $\{F_n : -\infty < n < \infty\}$ is assumed to be of full rank, by the above lemma the holomorphic extension $\Phi_+(z)$ of Φ is invertible at each $z \in D_+$. Let $\Phi_+(z) = \sum_{k=0}^{\infty} C_k z^k$, where $C_k = A_k G^k$, and let $[\Phi_+(z)]^{-1} = \sum_{k=0}^{\infty} D_k z^k$. The coefficients D_k satisfy the relations $C_0 D_0 = I$, $C_0 D_1 + C_1 D_0 = 0$, $C_0 D_2 + C_1 D_1 + C_2 D_0 = 0$, etc. It then follows that $Y_V(e^{i\theta})$ has a holomorphic extension to D_+ which is given by

$$(Y_V)_+(z) = \left(\sum_{k=v}^{\infty} C_k z^k \right) \left(\sum_{k=0}^{\infty} D_k z^k \right) = \sum_{k=0}^{\infty} E_{vk} z^k,$$

where $E_{vk} = \sum_{j=0}^k C_{v+j} D_{k-j}$. One would then hope that $Y_V(e^{i\theta})$ is some sort of radial limit of $(Y_V)_+(z)$ and thus be expressible as $Y_V(e^{i\theta}) = \sum_{k=0}^{\infty} E_{vk} e^{ki\theta}$. Problems arise however, since $Y_V(e^{i\theta})$ may not be in L_1 on C and thus would not have the matrices E_{vk} as Fourier coefficients.

The problem of determining under what restrictions $(Y_V)_+(z)$ does converge in $L_{2,m}$ to $Y_V(e^{i\theta})$ has been studied by Rozanov [26], Wiener and Masani [30], and Masani [19]. Rozanov assumes that the given regular q -variate process $\{F_n : -\infty < n < \infty\}$ forms a basis; i.e., any element H of M_{∞} can be represented as $H = \sum_{k=-\infty}^{\infty} A_k F_k$, where the series converges in M_{∞} , the representation is unique, and the sum of the series doesn't change under arbitrary permutations of its terms. He then shows that this property of a process implies the boundedness condition of Wiener and Masani, and his development is almost exactly the same as theirs. Hence the treatment of Wiener and Masani will be outlined in the remainder of this section.

The following restriction is placed on the regular, full rank q -variate process $\{F_n : -\infty < n < \infty\}$.

Boundedness Condition: The spectral density F' satisfies the following:
 $0 < \lambda I \leq F'(e^{i\theta}) \leq \lambda' I < \infty$ for some real numbers λ and λ' .

Wiener and Masani [30] have shown that in practical applications of the theory this condition will usually be met due to the errors inherent in measuring data and estimating the spectral density F' .

Lemma 18 i) $L_{2,m} = L_2$.

ii) For every $\phi \in L_2$,

$$\lambda \int_0^{2\pi} \phi(e^{i\theta}) \phi^*(e^{i\theta}) d\theta \leq 2\pi(\phi, \phi)_m \leq \lambda' \int_0^{2\pi} \phi(e^{i\theta}) \phi^*(e^{i\theta}) d\theta$$

iii) $L_{2,m}$ convergence, and L_2 convergence are equivalent.

Proof From the boundedness condition it follows that both $\sqrt{F'}$ and $(\sqrt{F'})^{-1}$ are in L_∞ . The criteria $\phi \in L_{2,m}$ if and only if $\phi \sqrt{F'} \in L_2$ then establishes i). Part ii) follows from lemma 1.5 of Wiener and Masani [30] on Hermitian matrices. From ii) it is easily seen that for any $\phi \in L_2$ $\lambda \|\phi\| \leq \|\phi\|_m \leq \lambda' \|\phi\|$. This readily establishes iii).

The proof of the following lemma can be found in Wiener and Masani [30].

Lemma 19 i) For any matrices A_0, A_1, \dots, A_n

$$\lambda \sum_{k=0}^n A_k A_k^* \leq \left(\sum_{k=0}^n A_k F^{-k}, \sum_{k=0}^{\infty} A_k F^{-k} \right) \leq \lambda' \sum_{k=0}^n A_k A_k^*.$$

ii) $M_0 = \sum_{k=0}^{\infty} \sigma(F^{-k})$

iii) If $G = \sum_{k=0}^{\infty} B_k F^{-k}$ then

$$\lambda \sum_{k=0}^{\infty} |B_k|_E^2 \leq \|G\|^2 \leq \lambda' \sum_{k=0}^{\infty} |B_k|_E^2 .$$

As an immediate corollary to ii) above one has the series representation $H_n = \sum_{k=0}^{\infty} D_k F_{n-k}$, where $\{H_n: -\infty < n < \infty\}$ is the normalized innovation process of the process $\{F_n: -\infty < n < \infty\}$, and the matrices D_k $k = 0, 1, \dots$ satisfy $\sum_{k=0}^{\infty} |D_k|_E^2 < \infty$.

Theorem 24 Let $\Phi(e^{i\theta})$ be the generating function of the process $\{F_n: -\infty < n < \infty\}$, and let the process $\{H_n: -\infty < n < \infty\}$ and the matrices D_k $k = 0, 1, \dots$ be as above. Then

- i) both Φ and Φ^{-1} are in L_{∞}^{0+} ,
- ii) The matrix D_k is the k 'th Fourier coefficient of Φ^{-1} .

Proof From the definition of Φ and theorem 18 it follows that $\Phi \in L_2^{0+}$. Since $\Phi\Phi^* = F'$ a.e. and the boundedness condition is assumed it follows that $\Phi \in L_{\infty}^{0+}$ and $\Phi^{-1} \in L_{\infty}$. From lemma 16 it is known that H_0 corresponds to Φ^{-1} under the isomorphism between $L_{2,m}$ and M_{∞} . Since $\sum_{k=0}^n D_k F_{-k}$ converges to H_0 by the last lemma, it follows that $\sum_{k=0}^n D_k e^{ki\theta}$ converges to Φ^{-1} in $L_{2,m}$. But $L_{2,m} = L_2$ and they have identical topologies: hence $\sum_{k=0}^n D_k e^{ki\theta}$ converges to Φ^{-1} in L_2 . Thus Φ^{-1} is in fact in L_{∞}^{0+} and D_k is its k 'th Fourier coefficient.

Theorem 25 Let $\{F_n: -\infty < n < \infty\}$ be a regular, full rank q -variate process whose spectral density F' satisfies the boundedness condition. Then

$$\hat{F}_v = \sum_{k=0}^{\infty} E_k F_{-k} ,$$

where $E_{vk} = \sum_{j=0}^{\infty} C_{v+j} D_{k-j}$, C_k and D_k being the k 'th Fourier coefficients of ϕ and ϕ^{-1} respectively.

Proof Let $Y_v(e^{i\theta}) = [e^{-vi\theta} \phi(e^{i\theta})]_{0+} \phi^{-1}(e^{i\theta})$. Since $[e^{-vi\theta} \phi]_{0+} = \sum_{k=0}^{\infty} C_{v+k} e^{ki\theta} = e^{-vi\theta} \{ \phi(e^{i\theta}) - \sum_{k=0}^{v-1} C_k e^{ki\theta} \}$, it follows that $[e^{-vi\theta} \phi]_{0+} \in L_{\infty}^{0+}$. From the previous lemma it is known that $\phi^{-1} \in L_{\infty}^{0+}$; hence $Y_v(e^{i\theta}) \in L_{\infty}^{0+}$.

It follows from the convolution rule in theorem 3 that the k 'th Fourier coefficient of $Y_v(e^{i\theta})$ is E_{vk} when $k \geq 0$, and is clearly zero when $k < 0$. Thus as $n \rightarrow \infty$, $\sum_{k=0}^n E_{vk} e^{ki\theta} \rightarrow Y_v(e^{i\theta})$ in L_2 . However, by lemma 18, this implies $\sum_{k=0}^n E_{vk} e^{ki\theta} \rightarrow Y_v(e^{i\theta})$ as $n \rightarrow \infty$ in $L_{2,m}$. From the isomorphism of $L_{2,m}$ and M_{∞} and theorem 23 it follows that $\hat{F}_v = \sum_{k=0}^{\infty} E_{vk} e^{ki\theta}$.

In 1958 Masani [19] proved the following stronger result.

Theorem 26 Let $\{F_n: -\infty < n < \infty\}$ be a regular full rank q -variate process whose spectral density F' only satisfies $F' \in L_{\infty}$ and $(F')^{-1} \in L_1$.

If the matrices E_{vk} are as in the above theorem, then

$$\hat{F}_v = \sum_{k=0}^{\infty} E_{vk} F_{-k}.$$

The proof of this theorem follows readily from the given conditions on F' and lemma 18.

MINIMAL PROCESSES

The characterization of full rank minimal processes due to Masani [19] is the generalization of a theorem on univariate processes due to Kolmogorov.

Recall that a q -variate process is minimal if and only if for some (and hence all) n , $F_n \notin \sigma(F_k)_{k \neq n}$. When this occurs

$$\varphi_n = F_n - (F_n | \sigma(F_k)_{k \neq n}) \neq 0 \text{ for all } n \in \mathbb{N}.$$

It is clear that the sequence $\{\varphi_n: -\infty < n < \infty\}$ is a q -variate process with the same shift operator U as the process $\{F_n: -\infty < n < \infty\}$.

Definition 33 The q -variate process $\{\varphi_n: -\infty < n < \infty\}$ is called the two sided innovation process of the process $\{F_n: -\infty < n < \infty\}$.

Definition 34 The process $\{F_n: -\infty < n < \infty\}$ is full rank minimal if and only if $\Delta(\varphi_0, \varphi_0) > 0$.

Definition 35 Let $\{F_n: -\infty < n < \infty\}$ be a full rank process and let $\{\varphi_n: -\infty < n < \infty\}$ be its two sided innovation process. The q -variate process $\{\psi_n: -\infty < n < \infty\}$ defined by

$$\psi_n = (\varphi_0, \varphi_0)^{-1} \varphi_n$$

is called the normalized two sided innovation process of $\{F_n: -\infty < n < \infty\}$.

The following lemma gives the basic properties of the process

$\{\psi_n: -\infty < n < \infty\}$. Its proof and the proof of theorem 27 can be found in Masani [19].

Lemma 20 Let $\{H_n: -\infty < n < \infty\}$ be the (one sided) normalized innovation

process of the full rank minimal q -variate process $\{F_n: -\infty < n < \infty\}$, and let $\{\psi_n: -\infty < n < \infty\}$ be as above. Then

i) $(\psi_m, \psi_n) = \sigma_{mn} I$,

ii) $\psi_n = \sum_{k=0}^{\infty} D_k^* H_{n+k}$, where D_k is the k 'th Fourier coefficient of $(\Phi_+)^{-1}$, Φ being the generating function of the process $\{F_n: -\infty < n < \infty\}$,

iii) $\Phi^{-1} \in L_2^{0+}$.

Theorem 27 Let $\{F_n: -\infty < n < \infty\}$ be a q -variate process and let F be its spectral distribution function. Then $\{F_n: -\infty < n < \infty\}$ is of full rank minimal if and only if $F'(e^{i\theta})$ is invertible a.e. and $(F')^{-1} \in L_1$ on C .

When this occurs $(F')^{-1}$ is the spectral density of the normalized two sided innovation process $\{\psi_n: -\infty < n < \infty\}$ which is a regular full rank process.

DETERMINATION OF THE GENERATING FUNCTION

It is clear from theorems 23, 25, and 26 that the determination of the generating function Φ of a regular full rank q -variate process $\{F_n: -\infty < n < \infty\}$ is of utmost importance for the solution of the extrapolation problem. In practice it is the spectral density function F' that is known, either through estimation from empirical data or from theoretical considerations. Hence Φ is considered determined if it can be found from the function F' .

In the univariate theory there is no problem, since in this case Φ and F' are merely complex valued and real valued functions respectively. From the equation $|\Phi(e^{i\theta})|^2 = F'(e^{i\theta})$ the function Φ can be found due

to its optimality: see Rozanov [26] and Hannan [7].

When $q > 1$, however, problems arise due to the failure of the exponential law for matrices. At present no closed form expression for Φ in terms of F' is known: however, infinite series expansions for Φ in terms of F' are available under certain restrictions. Wiener and Masani [30] found such a solution when the boundedness conditions are met. Subsequently Masani [19] extended this result by weakening the restrictions on F' . He showed that only the weaker assumptions

- i) The ratio of the largest to the smallest eigenvalues of F' is in L_1 ,
- ii) F' is invertible a.e. on C_1 and
- iii) $(F')^{-1} \in L_1$

are needed. Since these assumptions are satisfied when $F' \in L_\infty$ and $(F')^{-1} \in L_1$, both the generating function and the autoregressive representation of \hat{F}_v can be found (see theorem 26) in this case.

Let $\{F_n: -\infty < n < \infty\}$ be a regular full rank q -variate process and let F' be its spectral density function. If F' satisfies the boundedness conditions it is not difficult to show that part iii) of the following assumption involves no loss of generality. Hence for the remainder of this section the following assumption is made:

Assumption 1 i) F' is the spectral density function of a regular full rank q -variate process $\{F_n: -\infty < n < \infty\}$,

ii) There exist real numbers λ and λ' such that

$$0 < \lambda I \leq F' \leq \lambda' I < \infty,$$

iii) $F'(e^{i\theta}) = I + M(e^{i\theta})$, where $\mu = \text{ESS. L.U.B. } |M(e^{i\theta})|_B < 1$.
 $0 \leq \theta \leq 2\pi$

Definition 36 Let $G \in L_p$ on C , $p \geq 2$, and let $G(e^{i\theta}) = \sum_{k=-\infty}^{\infty} A_k e^{ki\theta}$.
Then $G_+(e^{i\theta}) = \sum_{k=1}^{\infty} A_k e^{ki\theta}$.

The motivation for studying the following operator stems from Von Neumann's projection theorem. See section 3 of Wiener and Masani [30] for a heuristic discussion of this material.

Definition 37 For all ϕ in L_2 , $B(\phi) = (\phi M)_+$, where $M = F' - I$.

The following properties of B are easily verified.

Lemma 21 i) B is in the Banach algebra of bounded linear operators on L_2 into itself and $|B| \leq \mu < 1$.

ii) Let I be the identity operator on L_2 . Then $I + B$ is invertible and $(I + B)^{-1} = I - B + B^2 - B^3 + \dots$

iii) $B(I) = M_+$, $B^2(I) = (M_+ M)_+$, $B^3(I) = ((M_+ M)_+ M)_+$, etc.

iv) $\|B^k(I)\| \leq \mu^k / q$.

It follows from the above lemma that the series $\sum_{k=0}^{\infty} (-1)^k B^k(I)$ is absolutely convergent. Hence the following definition in fact makes sense.

Definition 38 $\psi = \sum_{k=0}^{\infty} (-1)^k B^k(I) = (I + B)^{-1}(I)$.

Theorem 28 i) $\psi = G^{\frac{1}{2}} \Phi^{-1}$ where G is the prediction error matrix for lag 1 of the process $\{F_n : -\infty < n < \infty\}$,

ii) $\psi F' \psi^* = G$.

Proof i) Since F' satisfies the boundedness conditions, by theorem 24 $\Phi^{-1} \in L_{\infty}^{0+} \subset L_2$, and hence $G^{\frac{1}{2}}\Phi^{-1} \in L_2$. By lemma 21 the operator $I + B$ is 1:1, thus all that is required is to show that $G^{\frac{1}{2}}\Phi^{-1} = (I+B)^{-1}(I)$. This is equivalent to showing $(I+B)(G^{\frac{1}{2}}\Phi^{-1}) = I$. This is readily established by first noting that $G^{\frac{1}{2}}\Phi(0) = I$ and $G^{\frac{1}{2}}\Phi^{-1} \in L_{\infty}^{0+}$ implies that $G^{\frac{1}{2}}\Phi^{-1} = I + (G^{\frac{1}{2}}\Phi^{-1})_+$. If $k > 0$ then the k 'th Fourier coefficient of $G^{\frac{1}{2}}\Phi^*$ is 0. Since $G^{\frac{1}{2}}\Phi^* = G^{\frac{1}{2}}\Phi^{-1}\Phi\Phi^* = G^{\frac{1}{2}}\Phi^{-1}F' = G^{\frac{1}{2}}\Phi^{-1}(I+M) = G^{\frac{1}{2}}\Phi^{-1} + G^{\frac{1}{2}}\Phi^{-1}M$, it follows that $(G^{\frac{1}{2}}\Phi^{-1})_+ + (G^{\frac{1}{2}}\Phi^{-1}M)_+ = 0$. Thus $(I+B)(G^{\frac{1}{2}}\Phi^{-1}) = G^{\frac{1}{2}}\Phi^{-1} + (G^{\frac{1}{2}}\Phi^{-1}M)_+ = I + (G^{\frac{1}{2}}\Phi^{-1})_+ + (G^{\frac{1}{2}}\Phi^{-1}M)_+ = I$. Thus $\psi = G^{\frac{1}{2}}\Phi^{-1}$.

ii) This is clear from i), since $\psi F' \psi^* = G^{\frac{1}{2}}\Phi^{-1}F'(\Phi^{-1})^*G^{\frac{1}{2}} = G$.

Since M is obtainable from the spectral density F' so is ψ . Thus Φ can be found when F' satisfies assumption 1.

For the remainder of this chapter the following assumption is needed. It makes possible the factorization $F' = f'\tilde{F}'$ of the spectral density F' , where $f' \in L_1$, $\frac{1}{f'} \in L_1$, $\tilde{F}' \in L_{\infty}$, and $(\tilde{F}')^{-1} \in L_1$. Such a factorization is required to make use of the theory of minimal processes.

Assumption 2 The regular full rank process $\{F_n: -\infty < n < \infty\}$ has a spectral density F' which satisfies the following:

- i) $(F')^{-1} \in L_1$ on C .
- ii) If $\lambda(e^{i\theta})$ and $\mu(e^{i\theta})$ are the smallest and largest eigenvalues of $F'(e^{i\theta})$, then $\frac{\mu(e^{i\theta})}{\lambda(e^{i\theta})} \in L_1$, on C .

Lemma 22 If $F' \in L_2$ and $(F')^{-1} \in L_2$ assumption 2 is satisfied.

Proof If $F', (F')^{-1} \in L_2$ then $F', (F')^{-1} \in L_1$. For any $\theta \tau F'(e^{i\theta})$ equals the sum of the eigenvalues of $F'(e^{i\theta})$ and similarly for $(F')^{-1}$. Thus, by theorem 2, $0 \leq \mu \leq \tau F' \in L_2$ and $0 \leq \frac{1}{\lambda} \leq \tau(F')^{-1} \in L_2$. Hence $\frac{\mu(e^{i\theta})}{\lambda(e^{i\theta})} \in L_1$.

Definition 39 i) $f'(e^{i\theta}) = \frac{1}{2}(\lambda(e^{i\theta}) + \mu(e^{i\theta}))$.

ii) $M(e^{i\theta}) = \frac{1}{f'(e^{i\theta})} F'(e^{i\theta}) - I$ a.e.

$M(e^{i\theta})$ is actually well defined, since $f'(e^{i\theta}) \geq \frac{1}{2} \lambda(e^{i\theta}) > 0$ a.e.

Lemma 23 Under the conditions of assumption 2

i) $F' = f'(I+M)$ a.e.,

ii) $|M(e^{i\theta})|_B \leq 1$ for $0 \leq \theta \leq 2\pi$ and $|M(e^{i\theta})|_B < 1$ a.e.,

iii) $I + M$ is the spectral density of a full rank minimal q -variate process,

iv) f' is the spectral density of a regular minimal univariate process.

The proof of this lemma along with that of the following theorem can be found in Masani [19].

Theorem 29 Let Φ, Φ_1 , and ϕ be the generating functions of the processes with spectral densities $F', I + M$, and f' respectively. Let G and \tilde{G} be the prediction error matrices with lag 1 of the first two

processes, and let g be the innovation function of the third process.

Then

$$i) \quad \Phi^{-1}, \Phi_1^{-1} \in L_2^{0+}, \quad \frac{1}{\Phi} \in L_2^{0+},$$

$$ii) \quad \Phi = \Phi \Phi_1,$$

$$iii) \quad G = |g|^2 \tilde{G}.$$

As in the case under assumption 1, the following operator on L_2 is examined. In this case, however, its norm is not necessarily strictly less than 1.

Definition 40 For all $\phi \in L_2$ $(\phi) = (\phi M)_+$, where M is as in lemma 23.

Lemma 24 i) \mathcal{D} is a bounded linear operator on L_2 , and $|\mathcal{D}|_B \leq 1$.

ii) $\mathcal{D}(I) = M_+$, $\mathcal{D}^2(I) = (M_+ M)_+$, $\mathcal{D}^3(I) = ((M_+ M)_+ M)_+$, etc.

Proof i) follows from part ii) of lemma 23, and ii) is obvious.

Lemma 25 Let Φ , and \tilde{G} be as in theorem 29. Then, if I is the identity operator on L_2 ,

$$(I + \mathcal{D})(G^{\frac{1}{2}} \Phi_1^{-1}) = I.$$

Proof The proof is exactly the same as that of theorem 28, part i) once it is noted that $\Phi_1^{-1} \in L_2^{0+}$ due to the full rank minimality of the process with spectral density $I + M$.

If it were possible to invert the operator $I + \mathcal{D}$, the generating function Φ_1 , could then be found in a manner similar to that of theorem 28. The generating function ϕ of the univariate process with spectral density f' can easily be found as noted on page 58. Then by theorem 29 the generating function Φ of the process $\{F_n: -\infty < n < \infty\}$ could be obtained.

The proof of the following theorem rests upon the fact that $|\mathcal{D}|_B < 1$ on some set of positive measure. For its proof see Masani [19].

Theorem 30 i) \mathcal{D} is a strict contraction operator on L_2^{0+} , i.e. if $\psi \neq 0 \in L_2^{0+}$ then $\|\mathcal{D}(\psi)\| < \|\psi\|$.

ii) $I + \mathcal{D}$ is 1:1 on L_2 into itself.

By the above theorem the operator $I + \mathcal{D}$ is invertible on its range. This fact, together with lemma 26, shows that $(I + \mathcal{D})^{-1}(I)$ exists. The next theorem shows that the geometric series for $(I + \mathcal{D})^{-1}$ converges strongly on the range of $I + \mathcal{D}$. Its proof can be found in Masani [19].

Theorem 31 i) $\mathcal{D}^n \rightarrow 0$ strongly on L_2 as $n \rightarrow \infty$; i.e. for any $\psi \in L_2$

$$\lim_{n \rightarrow \infty} \|\mathcal{D}^n(\psi)\| = 0.$$

ii) If ψ is in the range of $I + \mathcal{D}$, then

$$\lim_{n \rightarrow \infty} \sum_{k=0}^n (-1)^k \mathcal{D}^k(\psi) = (I + \mathcal{D})^{-1}(\psi).$$

Theorem 32 If Φ_1 , and \tilde{G} are as in theorem 29, then

i) The series $I - M_+ + (M_+ M_+)^- \dots$ is mean convergent.

ii) If ψ is its sum, then $\psi = G^{\frac{1}{2}} \Phi_1^{-1}$.

iii) $\tilde{G} = \psi(I+M)\psi^*$.

Proof i) is clear from the last theorem. ii) follows from the previous theorem and lemma 25. Since $I + M = \Phi_1 \Phi_1^*$, $\psi(I+M)\psi^* = G^{\frac{1}{2}} \Phi_1^{-1} \Phi_1 \Phi_1^* (\Phi_1^{-1})^* G^{\frac{1}{2}} = \tilde{G}$.

This proves iii).

Thus when the spectral density F' of a regular full rank q -variate process $\{F_n; -\infty < n < \infty\}$ satisfies assumption 2, the generating function Φ of $\{F_n; -\infty < n < \infty\}$ can be found.

CHAPTER V

DEGENERATE RANK PROCESSES

There are several problems that arise in the study of degenerate rank q -variate processes that do not exist in the full rank theory. In the full rank case the determinant of the spectral density matrix F' is often investigated: the analogue of this in the degenerate theory is the examination of various proper minors of F' . This naturally leads to the utilization of the theory of the Nevanlinna class N_σ , a relevant discussion of which can be found in Wiener and Masani [31].

Let $\{F_n: -\infty < n < \infty\}$ be a degenerate rank q -variate process, and let $F(\theta)$ be its spectral distribution function. The following problems are still receiving investigation.

i) Regular processes. By theorem 19 the process $\{F_n: -\infty < n < \infty\}$ is regular if and only if $F(\theta)$ is absolutely continuous with respect to Lebesgue measure and $F'(e^{i\theta}) = \phi(e^{i\theta})\phi^*(e^{i\theta})$ where $\phi \in L_2^{0+}$ on C . Matveev [23] has found necessary and sufficient conditions for such a factorization to exist. One part of his conditions, however, requires that certain functions on C be the boundary values of functions in the N_σ class. As noted in Rozanov [26] there is no general method known for determining when a function on C is in fact such a boundary value.

In the special case of a rational spectral density F' (i.e., the elements of F' are rational functions of $e^{i\theta}$) the process $\{F_n: -\infty < n < \infty\}$ is easily shown to be regular. A method for actually performing the factorization $F' = \Phi\Phi^*$, $\Phi \in L_2^{0+}$, can be found in Rozanov [26].

ii) Deterministic processes. Although the definition of a deterministic q -variate process is quite simple and basic to the theory, at present no practical spectral characterization of such processes is known. Matveev [22] has found a sufficient condition for the process $\{F_n: -\infty < n < \infty\}$ to be deterministic: he has also shown that this condition is not necessary. In [10] and [11] Jang Ze Pei has developed necessary and sufficient conditions for $\{F_n: -\infty < n < \infty\}$ to be deterministic, but his criteria have the same problem as those of Matveev in i) above. Only in certain bivariate cases is it actually practical.

iii) Concordance In [17] Masani has shown that the conclusion of theorem 22 does not hold in general by exhibiting a process having both a spectral density F' and a non zero remote past. Robertson [24] showed that a necessary and sufficient condition for concordance of the Wold decomposition in the time domain and the Lebesgue Cramer decomposition in the spectral domain is that $R(F'(\theta)) = \rho$ a.e. Leb., where ρ is the rank of the process $\{F_n: -\infty < n < \infty\}$. However, there is no method known of determining ρ except in certain special cases. This problem is closely related to that of determining the spectral density F'_u of the regular process $\{U_n: -\infty < n < \infty\}$ in the Wold decomposition of $\{F_n: -\infty < n < \infty\}$.

By utilizing the theory of subordination (this is called the theory of linear transformations in translations of the Russian papers) Jang Ze Pei [10] and [11] has been able to find F'_u in theory for any process $\{F_n: -\infty < n < \infty\}$. but his method is only practical in some bivariate cases.

iv) Prediction theory. In [21] Masani gives the extension of theorem 23 to processes of arbitrary rank. However, the new expression for $Y_V(e^{i\theta})$ involves a certain matrix J^\perp when $\{F_n: -\infty < n < \infty\}$ is not of full rank. Since the matrix J^\perp is determined from the prediction error matrix G (thought of as a linear operator on C^q) and the calculation of G is still an open problem, this new expression cannot be put to practical use.

No studies have yet been made on the determination of an autoregressive representation for \hat{F}_V when the process $\{F_n: -\infty < n < \infty\}$ is of degenerate rank, except in the case of a rational spectral density F' (see Yaglom [32]).

BIBLIOGRAPHY

- [1] Cox, D.R. and Miller, H.D. The Theory of Stochastic Processes. John Wiley and Sons Inc., London, 1965.
- [2] Cramér, H. On the Theory of Stationary Random Processes. Ann.Math. 41 (1940) pp. 215-230.
- [3] Cramér, H. and Leadbetter, M.R. Stationary and Related Stochastic Processes. John Wiley and Sons Inc., New York, 1967.
- [4] Doob, J.L. Stochastic Processes. John Wiley and Sons Inc., New York, 1953.
- [5] Gantmacher, F.R. Applications of the Theory of Matrices, Vol. I. Chelsea Publishing Company, New York, 1959.
- [6] Halmos, P. Introduction to Hilbert Space. Chelsea Publishing Company, New York, 1957.
- [7] Hannan, E.J. Multiple Time Series. John Wiley and Sons Inc., New York, 1970.
- [8] Hewitt, E. and Stromberg, K. Real and Abstract Analysis. Springer-Verlag, New York, 1965.
- [9] Hoffman, K. Banach Spaces of Analytic Functions. Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1962.
- [10] Jang Ze-Pei. The Prediction Theory of Multivariate Stationary Processes I. Chinese Math. 4 (1963) pp. 291-322.
- [11] Jang Ze-Pei. The Prediction Theory of Multivariate Stationary Processes II. Chinese Math. 5 (1964) pp. 471-484.
- [12] Jenkins, G. and Watts, D. Spectral Analysis and its Applications. Holden-Day, San Francisco, 1968.
- [13] Koopmans, L.H. On the Coefficient of Coherence for Weakly Stationary Stochastic Processes. Ann.Math.Statist. 35 (1964) pp. 532-549.
- [14] Masani, P. Sur la Fonction Génératrice d'un Processus Stochastique Vectoriel. C.R. Acad.Sci.Paris 249 (1959) pp. 360-362.
- [15] Masani, P. Isomorphie Entre les Domaines Temporel et Spectral d'un Processus Vectoriel, Régulier. C.R. Acad.Sci.Paris 249 (1959) pp. 496-498.

- [16] Masani, P. Sur les Fonctions Matricielles de la Classe de Hardy H_2 . C.R. Acad.Sci.Paris 249 (1959) pp. 873-875; 906-907.
- [17] Masani, P. Cramer's Theorem on Monotone Matrix-valued Functions and the Wold Decomposition. Probability and Statistics (U. Grenander, editor) John Wiley and Sons Inc., New York, 1959. pp. 175-189.
- [18] Masani, P. Une Généralisation pour les Fonctions Matricielles de la Classe de Hardy H_2 d'un Théorème de Nevanlinna. C.R. Acad.Sci.Paris 251² (1960) pp. 318-320.
- [19] Masani, P. The Prediction Theory of Multivariate Stochastic Processes, III. Acta Math. 104 (1960) pp. 141-162.
- [20] Masani, P. Shift Invariant Spaces and Prediction Theory. Acta Math. 107 (1962) pp. 275-290.
- [21] Masani, P. Recent Trends in Multivariate Prediction Theory. Multivariate Analysis (P.R. Krishnaiah, editor) Academic Press, New York, 1966. pp. 351-382.
- [22] Matveev, R.F. On Singular Multidimensional Stationary Processes. Theory Probability Appl. (U.S.S.R.) English Transl. 5 (1960) pp. 33-39.
- [23] Matveev, R.F. On Multidimensional Regular Stationary Processes. Theory Probability Appl. (U.S.S.R.) English Transl. 6 (1961) pp. 149-165.
- [24] Robertson, J.B. Orthogonal Decompositions of Multivariate Weakly Stationary Stochastic Processes. Canad.J.Math. 20 (1968) pp. 368-383.
- [25] Rozanov, Yu.A. Spectral Properties of Multivariate Stationary Processes and Boundary Properties of Analytic Matrices. Theory Probability Appl. (U.S.S.R.) English Transl. 5 (1960) pp. 362-376.
- [26] Rozanov, Yu.A. Stationary Random Processes, Moscow (Russian). English translation by A. Feinstein. Holden-Day Inc., San Francisco, 1967.
- [27] Rosenberg, M. The Square-integrability of Matrix-valued Functions with respect to a Non-negative Hermitian Measure. Duke Math.J. 31 (1964) pp. 291-298.
- [28] Urbanik, K. Lectures on Prediction Theory. Springer-Verlag, Berlin, 1967.
- [29] Wiener, N. and Masani, P. The Prediction Theory of Multivariate Stochastic Processes I. Acta Math. 98 (1957) pp. 111-150.
- [30] Wiener, N. and Masani, P. The Prediction Theory of Multivariate Stochastic Processes II. Acta Math. 99 (1958) pp. 93-137.

- [31] Wiener, N. and Masani, P. On Bivariate Stationary Processes and the Factorization of Matrix-valued Functions. Theory Probability Appl. (U.S.S.R.) English Transl. 4 (1959) pp. 300-308.
- [32] Yaglom, A.M. Effective Solutions of Linear Approximation Problems for Multivariate Stationary Processes with a Rational Spectrum. Theory Probability Appl. (U.S.S.R.) English Transl. 5 (1960) pp. 239-264.
- [33] Yaglom, A.M. An Introduction to the Theory of Stationary Random Functions. Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1962.
- [34] Youla, D.C. On the Factorization of Rational Matrices. IRE Trans. Inform. Theory 7 (1961) pp. 172-189.