A NEW LOOK AT THE MAXIMUM ENTROPY SPECTRUM EXTENSION METHOD AND SOME RELATED PROBLEMS

THESIS

Submitted in Partial Fulfillment

of the Requirements for the

Degree of

MASTER OF SCIENCE (Electrical Engineering)

at the

POLYTECHNIC UNIVERSITY

by

M.Hafed Benteftifa

June 1988

A NEW LOOK AT THE MAXIMUM ENTROPY SPECTRUM EXTENSION METHOD AND SOME RELATED PROBLEMS

THESIS

Submitted in Partial Fulfillment

of the Requirements for the

Degree of

MASTER OF SCIENCE (Electrical Engineering)

at the

POLYTECHNIC UNIVERSITY

by

M.Hafed Benteftifa

June 1988

Approved:

S.Unnikrishna Pillai
Assistant Professor
of Electrical Engineering

Leonard Shaw
Professor and Chairman
Electrical Engineering Department

Copy no:______

Vita

Mr M.Hafed Benteftifa was born on Sept 9,1963 in Blida, Algeria. After graduating from the University of Science and Technology of Bab Ezzouar in 1982, he entered the Ecole Polytechnique of Algiers, Algeria in the Fall of 1982, where he received a "Ingenieur d'etat en Electronique" in 1985.

In the Spring of 1986 he started his graduate studies at Polytechnic University where he is currently a doctoral candidate.

ACKNOWLEDGEMENTS

The author wishes to thank Professor S.Unnikrishna Pillai for the critical comments and support he has provided throughout this research.

Acknowledgement is given to the Algerian ministry of Higher Education and Scientific Research, whose support throughout the years has made this work possible.

AN ABSTRACT

A NEW LOOK AT THE MAXIMUM ENTROPY METHOD AND SOME RELATED PROBLEMS

by

M.Hafed Benteftifa

Advisor:Professor S.Unnikrishna Pillai

Submitted in Partial Fulfillment of the Requirements
for the degree of Master of Science (Electrical Engineering)

June 1988

In this thesis, the problem of spectrum recovery is investigated from the viewpoint of linear prediction. The mean square error of an r-step predictor is the criterion used for the derivation of the spectral representation associated with a finite set of covariance functions.

The Maximum entropy method is developed in parallel with the extension problem. Following Youla [1], it is shown that the maximum entropy method (MEM) is the most robust one among all extensions of the original sequence.

Finally, a new pole-zero representation is developed for the case of a finite sequence of covariance functions. The parameters defining the spectrum estimator are derived following the same approach as for the maximum entropy method.

Table of Contents

Chapter 1

INTRODUCTION

1.1 Introduction:

One of the most challenging problem in modern spectral estimation theory is the problem of spectrum recovery in the presence of incomplete information about the underlying stochastic process.

A vast amount of literature is available about this exciting area [1-14]. Of these, the recent developments of Youla [1] parametrizing the class of all spectrum extensions is especially noteworthy. Our research which goes along the recent trend of the "pole-zero" modeling presents a new approach to the problem of the realization of the power spectrum from a given finite set of error-free covariance functions of the stochastic process.

The main idea is to consider the problem of spectrum recovery as a problem of linear prediction. The mean square error of an r-step predictor is used as the criterion for selecting the appropriate spectrum estimator.

In the first part of this thesis the concept of innovations is introduced . The mean square error of the r-step predictor is easily derived through the use of the whitening transformation.

In chapter 3 we formulate the problem of spectrum estimation from incomplete data .The maximum entropy method is then introduced through the use of the maximization of the minimum value of the prediction error for a 1-step predictor model.

To consider the spectrum extension problem, where we start with a finite covariance sequence and following Youla [1], we show that the maximum entropy method is the most robust among all methods that provide an extension of the original sequence.

In chapter 4 we consider the case of the 2-step predictor. It is shown that the spectral representation in this case is an ARMA(n,1).

Chapter 5 summarizes the work done and make some suggestions regarding future research .

Chapter 2

PREDICTION FILTER

2.1 Factorization and Innovations:

The derivation of the prediction error ϵ_r of the r-step predictor is of prime interest. Its derivation is made easier through the use of the innovations approach which in essence is a whitening transformation that transforms the original observations to a white noise process by means of a causal invertible linear filter.

The original observations x(n) and the innovation process contain the same statistical information. Therefore in deriving ϵ_r the orthonormal process i(n) will be used instead of x(n).

$$x(n)=\sum_{k=0}^{\infty} b(k)i(n-k) \tag{2.1.1}$$

$$i(n)=\sum_{k=0}^{\infty} l(k)x(n-k) \tag{2.1.2}$$

The power spectrum corresponding to x(n) is:

$$S_{xx}(\theta)=S_{ii}(\theta)B(e^{j\theta})B(e^{-j\theta}) \tag{2.1.3}$$

since $S_{ii}(\theta)=1$ then the power spectrum is :

$$S_{xx}(\theta)=B(e^{j\theta})B(e^{-j\theta})=|B(e^{j\theta})|^2 \tag{2.1.4}$$

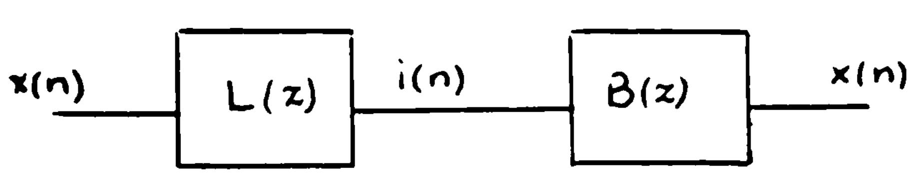

fig 2.1: Whitening and Innovation filter.

L(z): Whitening filter

B(z): Innovation filter

2.2 Derivation of the prediction error:

Given the infinite past of the stochastic process x(i), we want to find the r-step predictor of x(n).

First we develop the theory for the case of the 1-step predictor and shall extend it to the general case of the r-step predictor.

2.2.1 Expression of ϵ_1 for the 1-step predictor:

We wish to estimate the present value of x(n) given x(0),x(1),x(2),...,x(n-1)

$$\hat{x}(n)=E[x(n)\mid x(n-k),\ k\geq 1]=\sum_{k=1}^{\infty} a(k)x(n-k) \qquad (2.2.1)$$

The problem is to determine the coefficients a_k which correspond to the coefficients of the prediction filter. The solution is obtained by minimizing the mean-square error $\epsilon_1=x(n)-\hat{x}(n)$. However in our case we are only interested in the expression of ϵ_1.

Let's recall (2.1.1):

$$x(n)=\sum_{k=0}^{\infty} b(k)i(n-k)$$

The estimate of x(n) is $\hat{x}(n)=\sum_{k=1}^{\infty} a(k)x(n-k)$ which can be shown [10] to be equal to $\sum_{k=1}^{\infty} b(k)i(n-k)$. Therefore :

$$\epsilon_1=x(n)-\hat{x}(n)=\sum_{k=0}^{\infty} b(k)i(n-k)-\sum_{k=1}^{\infty} b(k)i(n-k)=b(0)i(n) \qquad (2.2.2)$$

Thus

$$P_1=E[\,|\epsilon_1(n)|^2]=|b(0)|^2 \qquad (2.2.3)$$

2.2.2 Expression of ε_r for an r-step predictor:

The generalization of the previous result to the case of the r-step predictor is straightforward; given x(0),x(1),x(2),... we wish to estimate x(n):

$$\hat{x}(n)=\sum_{k=r}^{\infty} a(k)x(n-k)=\sum_{k=r}^{\infty} b(k)i(n-k) \tag{2.2.4}$$

$$x(n)=\sum_{k=0}^{\infty} b(k)i(n-k)$$

$$\varepsilon_r(n)=x(n)-\hat{x}(n)=\sum_{k=0}^{\infty} b(k)i(n-k)-\sum_{k=r}^{\infty} b(k)i(n-k) \tag{2.2.5}$$

$$\varepsilon_r=\sum_{k=0}^{r-1} b(k)i(n-k) \tag{2.2.6}$$

The mean-square error is therefore:

$$P_r=E[\,|\,\varepsilon_r(n)\,|^2]=\sum_{k=0}^{r-1} |\,b(k)\,|^2 \tag{2.2.7}$$

Chapter 3

Maximum entropy spectral estimator

3.1 Mathematical preliminaries :

For a given zero mean second order stationary discrete time stochastic process x(n) we define the covariance function c(i) by :

$$c(i)=E[x(n).x^*(i+n)] \qquad |i|=0\rightarrow\infty \qquad (3.1.1)$$

where x* denotes the complex conjugate of x. The complex numbers $c(i)]_0^\infty$ possess the representation

$$c(i)=\frac{1}{2\pi}\int_{-\pi}^{+\pi} e^{+ij\theta}dF(\theta) \qquad (3.1.2)$$

where $F(\theta)$ is a bounded monotonic non decreasing function. The derivative $S(\theta)=\frac{dF(\theta)}{d\theta}\geq 0$ defines the spectral density of the process x(n) and exists for almost all θ[1].

If $F(\theta)$ is absolutely continuous then c(i) is given by :

$$c(i)=\frac{1}{2\pi}\int_{-\pi}^{+\pi} S(\theta)e^{+ji\theta}d\theta \qquad |i|=0\rightarrow\infty \qquad (3.1.3)$$

3.2 Spectral realization concept:

The spectral density $S(\theta)$ and the discrete covariance functions $c(i)$, $|i|=0\rightarrow\infty$ form a fourier transform pair [10] :

$$S(\theta)=\sum_{i=-\infty}^{\infty} c(i).e^{-ji\theta} \qquad (3.2.1)$$

$$c(i)=\frac{1}{2\pi}\int_{-\pi}^{+\pi} S(\theta)e^{+ji\theta}d\theta \qquad (3.2.2)$$

The relation (3.2.1) suggests that the computation of $S(\theta)$, in the presence of a finite set of data c_i's $|i|=0\rightarrow n$, requires the knowledge of the remaining covariance functions c(n+1), c(n+2), etc ....

A question arises as to how $c(i)]_0^\infty$ should be specified in order to guarantee that the covariance sequence is positive semi-definite.

One method of interest is the Maximum Entropy Method (MEM), first proposed by Burg [13]. The MEM is characterized by the fact that the completion of the original data c(0),c(1),c(2),...,c(n) is such that the time series corresponding to the complete sequence has maximum entropy.

In the next section, following Youla [1] we show that the power spectrum corresponding to the MEM is the same as the spectrum obtained by maximization of the error variance of the 1-step predictor filter.

3.3 Spectrum estimator in the case of 1-step predictor:

3.3.1 Preliminary notations:

We have previously shown that the process x(n) can be realized as the output of a causal filter with transfer function B(z) driven by the white noise i(n):

$$x(n)=\sum_{k=0}^{\infty} b(k).i(n-k)$$

$$E[i(n)]=0$$
$$E[i(n).i^*(j)] = \delta_{nj} \qquad (3.3.1)$$

Consider the following coefficients d(k) and f(k) :

$$d_k=\frac{1}{2\pi}\int_{-\pi}^{+\pi} e^{+ij\theta} ln\ S(\theta)d\theta \qquad k=0\rightarrow\infty \qquad (3.3.2)$$

$$\sum_{k=0}^{\infty} f(k)z^k = e^{\sum_{r=1}^{\infty} d_r z^r} \qquad k=1\rightarrow\infty \qquad (3.3.3)$$

$$f(0)=1$$

Then [1]:

$$b(k)=f(k)\exp\frac{d(0)}{2} \qquad k=0\rightarrow\infty \qquad (3.3.4)$$

with the following explicit relations for the f_k's and d_k's

$$f(1)=d(1) \qquad f(2)=d(2)+\frac{d^2(1)}{2} \qquad f(3)=d(3)+d(2).d(1)+\frac{d^3(1)}{6}$$

and $f(k)=d(k)+d(1).d(k-1)+...$ $\quad k>2$

3.3.2 Kolmogoroff's mean-square error formula:

In the case of linear prediction with lead 1 the mean square error (MSE) is :

$$P_1=E[\,|\varepsilon_1|^2]=|b(0)|^2=\exp\frac{1}{2\pi}\int_{-\pi}^{+\pi} ln\; S(\theta)d\theta \qquad (3.3.5)$$

proof [10]: $S(\theta)=|B(e^{j\theta})|^2$ where $B(e^{j\theta})$ is a minimum phase.

$$\int_{-\pi}^{+\pi} ln\; S(\theta)d\theta=\int_{-\pi}^{+\pi} ln\; |B(e^{j\theta}|^2 d\theta \qquad (3.3.6)$$

let $z=e^{j\theta}$ $\qquad dz=jzd\theta$

$$\text{so } \int_{-\pi}^{+\pi} ln\; |B(e^{j\theta}|^2 d\theta=\oint \frac{1}{jz} ln\; B(z).B(z^{-1})dz$$

$$=\oint \frac{1}{jz} ln\; B(z)dz+\oint \frac{1}{jz} ln\; B(z^{-1})dz$$

but

$$\oint \frac{1}{jz} ln\; B(z)dz=\oint \frac{1}{jz} ln\; B(z^{-1})dz$$

which implies :

$$\frac{1}{2\pi}\int_{-\pi}^{+\pi} ln\; S(\theta)d\theta=\frac{2}{\pi}\oint \frac{1}{jz} ln\; B(z)dz \qquad (3.3.7)$$

By using the cauchy integral theorem with the contour of integration the unit cercle we have :

$$\frac{1}{2\pi}\oint \frac{1}{jz} ln\; B(z)dz=ln\; b(0)$$

or

$$\frac{1}{\pi}\oint \frac{1}{jz} ln\; B(z)dz=ln\; |b(0)|^2$$

Finally the mean square error is :

$$P_1 = |b(0)|^2 = exp \frac{1}{2\pi} \int_{-\pi}^{+\pi} ln \ S(\theta) d\theta \qquad (3.3.8)$$

The relations between the coefficients b(k), f(k) and d(k) are the following :

$$b(0) = f(0) \ exp \ d(0)/2$$

Thus $$|b(0)|^2 = exp \ d(0) \qquad (3.3.9)$$

3.3.3 Autoregressive (AR) spectrum estimator:

The c_i's are chosen such that the MSE P_1 is maximized or as we will see later the entropy is maximum. Recalling relation (3.3.8) we see that maximizing $|b(0)|^2$ is the same as maximizing the exponential function.

Since the exponential function is monotonic then it will be sufficient to maximize its argument. Define

$$\Lambda = \frac{1}{2\pi}\int_{-\pi}^{+\pi} \ln\, S(\theta)d\theta \qquad (3.3.9)$$

replacing $S(\theta)$ by (3.2.1) we get :

$$\Lambda = \frac{1}{2\pi}\int_{-\pi}^{+\pi} \ln \sum_{i=-\infty}^{\infty} c(i).e^{-ji\theta}\, d\theta \qquad (3.3.10)$$

Let's differentiate Λ with respect to c_i $\quad |i| = n+1 \rightarrow \infty$

$$\frac{\partial\Lambda}{\partial c_i} = \frac{1}{2\pi}\int_{-\pi}^{+\pi} \frac{e^{-ij\theta}}{\sum_{i=-\infty}^{\infty} c(i).e^{-ji\theta}}\, d\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{e^{-ij\theta}}{S(\theta)}\, d\theta = 0 \qquad (3.3.11)$$

Therefore the coefficients of the fourier series expansion of $\frac{1}{S(\theta)}$ are equal to zero for $|i| \geq n+1$

Thus $S(\theta)$ can be expressed as :

$$S(\theta) = \frac{1}{\sum_{i=-n}^{n} \lambda_i e^{-ij\theta}} \qquad (3.3.12)$$

$\lambda_i = \lambda^*_{-i}$ so as to insure that $S(\theta)$ is positive and real.

The next step is to find the λ_i's corresponding to the expansion of $S(\theta)$. We use the fact that $S(\theta)$ must be consistent with the known fuctions c(i) $|i| = 0 \rightarrow n$:

Consider the polynomial $A(\theta) = \sum_{i=-n}^{n} \lambda_i e^{-ij\theta}$

Since $A(\theta)$ is positive then by Riesz's theorem [11] we can find a unique minimum phase $T(e^{j\theta})$ such that :

$$A(\theta) = T(e^{j\theta}).T(e^{-j\theta}) = |T(e^{j\theta})|^2 \qquad (3.3.13)$$

with $\quad T(e^{j\theta}) = \sum_{i=0}^{n} t(i).e^{-ji\theta}$

Let $z = e^{+j\theta}$ then $\quad T(z) = \sum_{i=0}^{n} t(i).z^{-i}$

But $c(i) = \frac{1}{2\pi} \int_{-\pi}^{+\pi} S(\theta) e^{+ji\theta} d\theta = \frac{1}{2\pi} \int_{-\pi}^{+\pi} \frac{e^{+ij\theta}}{A(\theta)} d\theta \qquad |i| \leq n$

$$= \frac{1}{2\pi} \int_{-\pi}^{+\pi} \frac{e^{+ij\theta}}{T(e^{j\theta}).T(e^{-j\theta})} d\theta = \frac{1}{2\pi j} \oint \frac{z^{i-1}}{T(z).T(z^{-1})} dz \qquad (3.3.14)$$

Using this relation we next form the sum $\sum_{i=0}^{n} t(i).c(m-i)$:

$$\sum_{i=0}^{n} t(i).c(m-i) = \frac{1}{2\pi j} \oint \frac{z^{m-1}}{T(z).T(z^{-1})} . \sum_{i=0}^{n} t(i) z^{-i} \, dz = \frac{1}{2\pi} \oint \frac{z^{m-1}}{T(z^{-1})} dz \qquad (3.3.15)$$

$T(z)$ is minimum phase therefore $T(z^{-1})$ is a maximum phase (all its zeros are located outside the unit circle). By recalling the Cauchy integral theorem we conclude that :

$$\frac{1}{2\pi j}\oint \frac{z^{m-1}}{T(z^{-1})}dz = \begin{cases} \dfrac{1}{t^*(0)} & m=0 \\ 0 & m=1,2,3,\ldots,n \end{cases} \qquad (3.3.16)$$

Therefore :

$$\sum_{i=0}^{n} t(i).c(m-i) = \begin{cases} \dfrac{1}{t^*(0)} & m=0 \\ 0 & m=1,2,3,\ldots,n \end{cases} \qquad (3.3.17)$$

which is equivalent to the system of equations:

$$\begin{cases} t(n).c(0) & +t(n-1).c(1) & +\ldots & +t(0).c(n) & = & 0 \\ t(n).c(-1) & +t(n-1).c(0) & +\ldots & +t(0).c(n-1) & = & 0 \\ \cdots & \cdots & +\ldots & \cdots & & 0 \\ t(n).c(1-n) & +t(n-1).c(-n) & +\ldots & +t(o).c(1) & = & 0 \\ t(n).c(-n) & +t(n-1).c(-n+1) & +\ldots & +t(0).c(0) & = & \dfrac{1}{t^*(0)} \end{cases} \qquad (3.3.18)$$

and $\qquad T(z)=t(n)z^{-n} + t(n-1)z^{-(n-1)} + \ldots + t(0)$

These standard Yule-Walker equations can be solved in a variety of ways [9-10]; Solving for T(z) we have [1]:

$$T(Z) = \frac{\begin{vmatrix} 0 & c(0) & c(1) & \cdots & c(n) \\ 0 & c^*(1) & c(0) & \cdots & c(n-1) \\ . & . & . & \cdots & . \\ . & . & . & \cdots & . \\ 0 & c^*(n-1) & c^*(n) & \cdots & c(1) \\ \frac{1}{t^*(0)} & c^*(n) & c^*(n-1) & \cdots & c(0) \\ 0 & z^{-n} & z^{-(n-1)} & \cdots & 1 \end{vmatrix}}{\begin{vmatrix} 0 & c(0) & c(1) & \cdots & c(n) \\ 0 & c^*(1) & c(0) & \cdots & c(n-1) \\ . & . & . & \cdots & . \\ . & . & . & \cdots & . \\ 0 & c^*(n-1) & c^*(n) & \cdots & c(1) \\ 0 & c^*(n) & c^*(n-1) & \cdots & c(0) \\ -1 & z^{-n} & z^{-(n-1)} & \cdots & 1 \end{vmatrix}} \qquad (3.3.19)$$

Let Δ_n be the determinant of order n :

$$\Delta_n = \begin{vmatrix} c(0) & c(1) & \cdots & c(n) \\ c^*(1) & c(0) & \cdots & c(n-1) \\ \cdots & \cdots & \cdots & \cdots \\ c^*(n-1) & c^*(n) & \cdots & c(1) \\ c^*(n) & c^*(n-1) & \cdots & c(0) \end{vmatrix} \qquad (3.3.20)$$

Then

$$T(Z)=\frac{1}{t^*(0)}\frac{\begin{vmatrix} c(0) & c(1) & \cdots & c(n) \\ c^*(1) & c(0) & \cdots & c(n-1) \\ \cdots & \cdots & \cdots & \cdots \\ c^*(n-1) & c^*(n) & \cdots & c(1) \\ z^{-n} & z^{-(n-1)} & \cdots & 1 \end{vmatrix}}{\Delta_n} \qquad (3.3.21)$$

From (3.3.19) we get:

$$t(0)=\frac{\begin{vmatrix} c(0) & c(1) & \cdots & c(n-1) & 0 \\ c^*(1) & c(0) & \cdots & c(n-2) & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ c^*(n-1) & c^*(n) & \cdots & c(0) & 0 \\ c^*(n) & c^*(n-1) & \cdots & c^*(1) & \frac{1}{t^*(0)} \end{vmatrix}}{\Delta_n}=\frac{1}{t^*(0)}.\frac{\Delta_{n-1}}{\Delta_n} \qquad (3.3.22)$$

where Δ_{n-1} is the determinant of order n-1, thus :

$$|t(0)|^2=\frac{\Delta_{n-1}}{\Delta_n} \qquad (3.3.23)$$

and from (3.3.19) we get [1]:

$$T(Z)=\frac{1}{\sqrt{\Delta_n.\Delta_{n-1}}}.\begin{vmatrix} c(0) & c(1) & \cdots & c(n) \\ c^*(1) & c(0) & \cdots & c(n-1) \\ \cdots & \cdots & \cdots & \cdots \\ c^*(n-1) & c^*(n) & \cdots & c(1) \\ z^{-n} & z^{-(n-1)} & \cdots & 1 \end{vmatrix}$$

Therefore by replacing $T(z)$ in (3.3.12) we get the final expression for $S(\theta)$. The spectrum $S(\theta)$ is completely determined by the known covariance functions $c(i)]_0^n$. An interesting remark at this point is that the maximization procedure doesn't tell us anything about the location of the unknown c_i's.

However by exploiting the positive definiteness of the covariance sequence we show in the next section that the unknown c_i's in the MEM case are located at the center of successive circles whose center and radii at stage k depend only on the sequence of order k-1.

For a detailed derivation of these standard results see Youla [1] and Geronimus [5].

3.4 Extension problem :

Given the finite set of covariance functions $c(i)]_n^{\infty}$

Let's form the determinant of order n+1 with $\xi = c(n+1)$ the unknown:

$$\Delta_{n+1}(\xi) = \begin{vmatrix} c(0) & c(1) & \cdots & \xi \\ c^*(1) & c(0) & \cdots & c(n) \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ c^*(n) & c^*(n\text{-}1) & c(0) & c(1) \\ \xi^* & c^*(n) & c^*(1) & c(0) \end{vmatrix} \qquad (3.4.1)$$

By the vertue of positive definitess of the sequence c(n):

$$\Delta_{n+1}(\xi) > 0 \qquad (3.4.2)$$

(3.4.2) represents the equation for the interior of a circle, whose radius r_n and center ρ_n has been shown to be [1]:

$$r_n = \frac{\Delta_n}{\Delta_{n-1}} \qquad (3.4.3)$$

$$\rho_n = (-1)^{n+1} \frac{\begin{vmatrix} c(1) & \cdots & c(n) & 0 \\ c(0) & \cdots & c(n-1) & c(n) \\ & & & \\ c^*(n-1) & \cdots & c(0) & c(1) \end{vmatrix}}{\Delta_{n-1}} \tag{3.4.4}$$

Our next step is to show that the sequence of radii $r_n, r_{n+1}, \ldots etc$ is monotone non increasing.

The minors of $\Delta_{n+1}(\xi)$ at any stage k obey the following relation:

$$\Delta_{k+1} \cdot \Delta_{k-1} = \Delta_k^2 - \Delta^2(\xi) \tag{3.4.5}$$

Dividing both sides by Δ_{k-1}^2 we get :

$$\frac{\Delta_{k+1}}{\Delta_{k-1}} = \frac{\Delta_k^2 - \Delta^2(\xi)}{\Delta_{k-1}^2} = \frac{\Delta_k^2}{\Delta_{k-1}^2} - \frac{\Delta^2(\xi)}{\Delta_{k-1}^2} \tag{3.4.6}$$

But $\dfrac{\Delta^2(\xi)}{\Delta_{k-1}^2} = (\xi - \rho_k)^2$

Which implies that :

$$\frac{\Delta_{k+1}}{\Delta_{k-1}} = \frac{\Delta_k^2}{\Delta_{k-1}^2} - (\xi - \rho_k)^2 =$$

$$= \frac{\Delta_k^2}{\Delta_{k-1}^2} \cdot \left[1 - \left((\xi - \rho_k) \cdot \frac{\Delta_{k-1}}{\Delta_k} \right)^2 \right] \tag{3.4.7}$$

But from (3.4.2) we have

$$\left\{(\xi - \rho_k) \cdot \frac{\Delta_{k-1}}{\Delta_k}\right\}^2 \leq 1 \tag{3.4.8}$$

Therefore we have the inequality $\dfrac{\Delta_{k+1}}{\Delta_k} \leq \dfrac{\Delta_k}{\Delta_{k-1}}$

But since $r_{k+1} = \dfrac{\Delta_{k+1}}{\Delta_k}$ and $r_k = \dfrac{\Delta_k}{\Delta_{k-1}}$

Then $\quad r_{k+1} \leq r_k$

By using the same argument for stage k+1 we can prove that $r_{k+2} \leq r_{k+1}$ The inequality is then true for any consecutive stages k and k+1 .Therefore the sequence of radii of the given circles ς_k is montone non decreasing.

Special case: For the case where the unknown functions $c(i)]_{n+1}^{\infty}$ are chosen at the center of the circles $\varsigma_i]_n^{\infty}$ we have $r_n = r_{n+1} = r_{n+2} = \dots = r_k = r_{k+1} = \cdots$

with $r_n = \dfrac{\Delta_n}{\Delta_{n-1}}$

proof: Consider $\xi_k = c(k+1)$

Then if ξ is taken at the center of the circle ς_k;$\Delta(\xi_k) = 0$,we have from (3.4.5) :

$$\frac{\Delta_{k+1}(\xi_k)}{\Delta_k} = \frac{\Delta_k}{\Delta_{k-1}} \tag{3.4.13}$$

Now let's consider $\xi_{k+1} = c(k+2)$ with c(k+1) already chosen at the center of the circle ς_k then :

$$\frac{\Delta_{k+2}(\xi_{k+1})}{\Delta_{k+1}(\xi_k)} = \frac{\Delta_{k+1}(\xi_k)}{\Delta_k} = \frac{\Delta_k}{\Delta_{k-1}} \tag{3.4.14}$$

Thus $\qquad r_{k+2} = r_{k+1} = r_k$

The same argument is valid for any stage k.

3.5 Relation between the MEM and the extension problem:

We have previously seen that while the maximum entropy method provides an autoregressive representation for the spectrum $S(\theta)$ it doesn't tell us very much about the location of the remaining unknown covariance functions. By contrast, from the extension problem we can conclude that given a finite covariance sequence the remaining covariances are located at the center of the respective circles ς_k, k=n+1,n+2,....

In this section we will show that the maximum entropy method is that particular extension where the new covariances are located at the center of the circles ς_k ,k=n+1,n+2,....

proof: Given $c(i)]_0^{k-1}$ we know that :

$$S(\theta) = |B(e^{j\theta})|^2 = \frac{1}{|T(e^{j\theta}|^2} \tag{3.5.1}$$

But from (3.3.21) $|t(0)|^2 = \dfrac{\Delta_{k-1}}{\Delta_k}$

Using the fact that $|b(0)|^2 = \dfrac{1}{|t(0)|^2}$ we conclude that $|b(0)|^2 = \dfrac{\Delta_k}{\Delta_{k-1}}$

$|b(0)|^2$ is maximum if the $c(i)]_n^{\infty}$ are at the center of the circle $\varsigma_k]_{n+1}^{\infty}$. In that case all the radii are equal and their maximum is $\dfrac{\Delta_n}{\Delta_{n-1}}$, therefore $|b(0)|^2 = \dfrac{\Delta_n}{\Delta_{n-1}}$ is maximum.

Hence, it implies that the entropy is maximum. Since the unknown covariance functions are consistenly chosen at the center of the circles ς_k we may conclude that the maximum entropy method presents a robust approach to spectrum estimation.

chapter 4

POLE-ZERO MODELING

4.1 introduction :

Youla in [1], in addition to deriving a closed form expression for the entire class of spectrum extensions, also suggested that it is possible to select a spectral density estimator which maximizes a function defined on the impulse response coefficients b_r's. The maximum entropy as we have seen in chapter 3 is a special case where the parameter was taken to be $|b(0)|^2$.

In this section we develop a new spectral estimator based on the first two coefficients of the impulse response B(z). We define the function or criterion for selecting the spectral estimator as the mean square error of the predictor filter.

In the first part we review some previous result on the predictor filter and then derive the expression for the spectral estimator.

In the second part we derive the exact expressions for the parameters of $S(\theta)$.

4.2 Pole-zero model :

The prediction error for the r-step predictor is:

$$\epsilon_r = \sum_{k=0}^{r-1} b(k) i(n-k) \tag{4.2.1}$$

$$P_r = E[\,|\epsilon_r(n)|^2] = \sum_{k=0}^{r-1} |b(k)|^2 \tag{4.2.2}$$

For the case of a two-step predictor P_r becomes :

$$P_2 = \sum_{k=0}^{1} |b(k)|^2 = |b(0)|^2 + |b(1)|^2 \tag{4.2.3}$$

recalling (3.3.3):

$$b(0) = f(0) . \exp\frac{d(0)}{2} = \exp\frac{d(0)}{2}$$

$$b(1) = f(1) . \exp\frac{d(0)}{2} = d(1) . \exp\frac{d(0)}{2} \qquad (4.2.4)$$

$$\text{with} \begin{cases} d(0) = \dfrac{1}{2\pi} . \displaystyle\int_{-\pi}^{+\pi} \ln\ S(\theta) d\theta \\ \\ d(1) = \dfrac{1}{2\pi} . \displaystyle\int_{-\pi}^{+\pi} e^{+j\theta} . \ln\ S(\theta)\, d\theta \end{cases}$$

Define $\Delta_1 = |b(0)|^2 + |b(1)|^2 = \left[1 + |d(1)|^2\right] . \exp\ d(0)$

$$= \left[1 + |d(1)|^2\right] . \exp\frac{1}{2\pi} \int_{-\pi}^{+\pi} \ln\ S(\theta) d\theta$$

Differentiating with respect to the unknown $c(i)]_{n+1}^{\infty}$:

$$\frac{\partial \Delta_1}{\partial c(i)} = \left[\frac{1}{2\pi} . [1 + |d(1)|^2] . \int_{-\pi}^{+\pi} \frac{e^{-ij\theta}}{S(\theta)} d\theta + \frac{d^*(1)}{2\pi} . \int_{-\pi}^{+\pi} \frac{e^{-j\theta(i-1)}}{S(\theta)} d\theta \right.$$

$$\left. + \frac{d(1)}{2\pi} . \int_{-\pi}^{+\pi} \frac{e^{-j\theta(i+1)}}{S(\theta)} d\theta \right] \exp\ \frac{d(0)}{2} = 0 \qquad (4.2.5)$$

Let $1 + |d(1)|^2 = \alpha$ and $\beta = d(1)$ then:

$$\frac{\partial \Lambda_1}{\partial c(i)} = \frac{1}{2\pi} . \int_{-\pi}^{+\pi} . e^{-ij\theta} \frac{[\alpha + \beta^* . e^{+j\theta} + \beta . e^{-j\theta}]}{S(\theta)} d\theta = 0 \qquad (4.2.6)$$

which implies that the fourier series for the real function $\dfrac{[\alpha + \beta^* . e^{+j\theta} + \beta . e^{-j\theta}]}{S(\theta)}$

truncates.

Thus $$\frac{[\alpha + \beta^* . e^{+j\theta} + \beta . e^{-j\theta}]}{S(\theta)} = \sum_{i=-n}^{n} \lambda(i) . e^{-ij\theta}$$

and $$S(\theta) = \frac{\alpha + \beta^* . e^{+j\theta} + \beta . e^{-j\theta}}{\sum_{i=-n}^{n} \lambda(i) . e^{-ij\theta}} \qquad (4.2.7)$$

$\lambda_i = \lambda_{-i}{}^*$ to insure that $S(\theta)$ is real;

$d(1) = \dfrac{b(1)}{b(0)}$ so $\beta = d(1) = \dfrac{b(1)}{b(0)}$ and $\alpha = 1 + \left| \dfrac{b(1)}{b(0)} \right|^2$

Finally $S(\theta)$ can be written as :

$$S(\theta) = \frac{b^*(1) . b(0) . e^{+j\theta} + |b(0)|^2 + |b(1)|^2 + b(1) . b^*(0) . e^{-j\theta}}{|b(0)|^2 . \sum_{i=-n}^{n} \lambda(i) . e^{-ij\theta}} \qquad (4.2.8)$$

Again by Riesz's theorem $\sum_{i=-n}^{n} \lambda(i) . e^{-ij\theta}$ can be expressed as the product

$W(e^{j\theta}) . W(e^{-j\theta})$ where $W(e^{j\theta})$ is a minimum phase.

The numerator can be written as $r(-1) \cdot e^{j\theta} + r(0) + r(1) \cdot e^{-j\theta}$

with : $\begin{cases} r(+1) = b(1) \cdot b^*(0) \\ r(-1) = b(0) \cdot b^*(1) \\ r(0) = |b(0)|^2 + |b(1)|^2 \end{cases}$

Finally :

$$S(\theta) = \frac{|b(0) + b(1) \cdot e^{-j\theta}|^2}{|b(0)|^2 \cdot |W(e^{j\theta})|^2} \qquad \textbf{(4.2.9)}$$

4.3 Derivation of the spectrum parameters:

Now that we have the expression of $S(\theta)$ the next step is to derive the relations for its parameters.

Let's form the sum $\sum_{k=0}^{n} \lambda(k).c(m-k)$:

$$\sum_{k=0}^{n} \lambda(k).c(m-k) = \frac{1}{2\pi} \sum_{k=0}^{n} \lambda(k) \int_{-\pi}^{\pi} \frac{\sum_{i=-1}^{1} r(i).e^{j\theta(m-i)}.e^{-j\theta k}}{|b(0)|^2 \sum_{i=-n}^{n} \lambda(i).e^{-ij\theta}} d\theta \qquad (4.3.1)$$

$$= \frac{1}{2\pi j} \oint \frac{z^{m-1}.\sum_{i=-1}^{1} r(i)z^{-i}}{|b(0)|^2.W(z)} dz = \frac{1}{2\pi j} \oint \frac{z^{m-1}.(r(+1)z^{-1}+r(0)+r(-1)z)}{|b(0)|^2(\lambda^*(0)+\lambda^*(1)z+...+\lambda^*(n)z^n)} dz =$$

$$= \frac{1}{2\pi j} \oint \frac{z^{m-2}.(r(+1)+r(0)z+r(-1)z^2)}{|b(0)|^2(\lambda^*(0)+\lambda^*(1)z+...+\lambda^*(n)z^n)} dz \qquad (4.3.2)$$

$W(z^{-1}) = \lambda^*(0) + \lambda^*(1)z + \cdots + \lambda^*(n)z^n$ is a maximum phase; all its zeros are located outside the unit circle, therefore by using Cauchy's integral relation we have:

$$\sum_{k=0}^{n} \lambda(k).c(m-k) = \begin{cases} A & m=0 \\ \dfrac{r(+1)}{|b(0)|^2.\lambda^*(0)} & m=1 \\ 0 & m=2,3,...,n \end{cases} \qquad (4.3.3)$$

where A is defined by :

$$A = \frac{\partial \left[\dfrac{r(+1)+r(0)z+r(-1)z^2}{\lambda^*(0)+\lambda^*(1)z+\cdots\lambda^*(n)z^n} \right]}{\partial z} \Bigg|_{z=0} \qquad (4.3.4)$$

$$= \frac{\lambda^*(0).r(0) - r(1).\lambda^*(1)}{|b(0)|^2.\lambda^*(0)^2}$$

with: $\begin{cases} r(+1) = b(1).b^*(0) \\ r(-1) = b^*(1).b(0) \\ r(0) = |b(0)|^2 + |b(1)|^2 \end{cases}$

For a 2-step predictor the equations are the following [3]:

$$\sum_{k=2}^{n} a(k).c(m-k) = c(m) \qquad 2 \le m \le n \qquad (4.3.5)$$

$$P_2 = E[\,|\epsilon_2|^2] = c(0) - \sum_{k=2}^{n} a(k).c(-k) \qquad (4.3.6)$$

The coefficients $a_k]_n^2$ are easily obtained from the matrix equation :

$$\mathbf{a} = \mathbf{R}^{-1}.\mathbf{c} \qquad (4.3.7)$$

where $\mathbf{a}^T = [\,a(2), a(3), \ldots, a(n)\,]$, $\mathbf{c}^T = [\,c(2), c(3), \ldots, c(n)\,]$

$$\text{and } \mathbf{R} = \begin{bmatrix} c(0) & \cdots & c(2-n) \\ \cdots & \cdots & \cdots \\ c(n-2) & \cdots & c(0) \end{bmatrix}$$

The above equations can be combined together to form the following system :

$$\sum_{k=0}^{n} \psi(k) \,.\, c(m-k) = \begin{cases} P_2 & m=0 \\ \\ 0 & m=2,3,4,\ldots,n \end{cases} \qquad (4.3.8)$$

where the set of coefficients ψ_k is defined as:

$$\psi(k)\big]_2^n = -a(k)\big]_2^n \qquad (4.3.9)$$

and $\psi(0) = 1, \psi(1) = 0$.

Now it remains to find b(0) and b(1) to completely characterize $S(\theta)$.

The system of equations (4.3.3) and (4.3.8) can be written in the following :

$$\sum_{k=0}^{n} \frac{\psi(k)}{P_2} \,.\, c(m-k) = \begin{cases} 1 & m=0 \\ \displaystyle\sum_{k=0}^{n} \frac{\psi(k).}{P_2} . c(1-k) & m=1 \\ 0 & m=2,3,4,\ldots,n \end{cases} \qquad (4.3.10)$$

$$\sum_{k=0}^{n} \frac{\lambda(k)}{A} . c(m-k) = \begin{cases} 1 & m=0 \\ \dfrac{r(1)}{A\,|b(0)|^2 \lambda^*(0)} & m=1 \\ 0 & m=2,3,\ldots,n \end{cases} \qquad (4.3.11)$$

By comparing the two sets of equations we may choose : $\dfrac{\lambda(k)}{A} = \dfrac{\psi(k)}{P_2}$

for k=0

$$\lambda(0) = \frac{\psi(0)A}{P_2} = \frac{\lambda^*(0)r(0) - r(1)\lambda^*(1)}{P_2 \,|\, b(0) \,|^2 \lambda^*(0)^2} \tag{4.3.12}$$

which is equal to $\dfrac{r(0)}{P_2 \,|\, b(0) \,|^2 \lambda^*(0)}$ because $\lambda(1) = \psi(1)\dfrac{A}{P_2} = 0$

For m=1 we have the equality :

$$\frac{r(1)}{A \,|\, b(0) \,|^2 \lambda^*(0)} = \frac{P_1}{P_2} \tag{4.3.13}$$

where $\quad P_1 = \sum_{k=0}^{n} \psi(k).c(1-k)$

$$r(1) = \frac{r(0).P_1}{P_2} \tag{4.3.14}$$

from (4.3.12) and (4.3.14) we have the system:

$$r(0) = |\lambda(0)|^2 \,|\, b(0) \,|^2 P_2 = |\, b(0) \,|^2 + |\, b(1) \,|^2 \tag{4.3.15}$$

$$r(1) = \frac{r(0).P_1}{P_2} = b(1)b^*(0) \tag{4.3.16}$$

The ratio $\dfrac{b(1)}{b(0)}$ is then the solution of the following equation:

$$1 + \left|\frac{b(1)}{b(0)}\right|^2 = \left(\frac{P_2}{P_1}\right)\left(\frac{b(1)}{b(0)}\right) \tag{4.3.17}$$

where the ratio $\frac{P_2}{P_1}$ is fixed by the known covariance functions.

Finally:

$$S(\theta) = \frac{\left| 1 + \frac{b(1)}{b(0)} . e^{-j\theta} \right|^2}{\left| \lambda(0) + \lambda(1) . e^{-j\theta} + \ldots + \lambda(n) . e^{-jn\theta} \right|^2} \qquad (4.3.18)$$

where $\frac{b(1)}{b(0)}$ is determined from the previous quadratic equation.

These solutions are at best only partial for two reasons:

First, for real data one must demonstrate that the filter parameters $b(0), b(1), \lambda(0), \lambda(1), \ldots, \lambda(n)$ are real. More importantly, the denominator polynomial $\lambda(0) + \lambda(1) z^{-1} + \ldots + \lambda(n) z^{-n}$ must be shown to be free of zeros inside the unit circle. These are important features and must be demonstrated before one could claim the realization of the desired Wiener factor.

Chapter 5

Conclusion and Recommendations

5.1: Conclusion :

The research presented in this thesis has concentrated on using the mean square error of a 1 and 2 step predictor as the criterion for the derivation of the spectrum estimator.

The work has proceeded from the case of the 1-step predictor, where it is shown that the spectral representation of the time series is an AR model, to the case of the 2-step predictor where the the spectral representation is shown to be an ARMA(n,1) model.

5.2: Recommendations for future research :

In this thesis we considered the simple case of the 1-step and 2-step predictor. The 2-step predictor filter solutions developed here are only partial for two reasons: First, for real real data one must demonstrate that the filter parameters are real. More importantly, the denominator polynomial $\lambda(0)+\lambda(1)z^{-1}+\ldots+\lambda(n)z^{-n}$ must be shown to be free of zeros inside the unit circle. These are important features and must be demonstrated before one could claim the realization of the desired Wiener factor.

Investigation of these problems together with the more general r-step predictor forms a potential area for future research.

REFERENCES

[1] D.C Youla ,"The FEE: a new tunable high-resolution spectral estimator," *Polyt Inst of NY, Brooklyn , Tech note 3, 36 pp, 1978.*

[2] M.R Schroeder ,"Linear prediction entropy and signal analysis," *IEEE ASSP Mag,* vol1, no.3 ,pp 3-11, 1984

[3] S.Haykin , *Non linear methods of spectral analysis,* Berlin: Springer Verlag, 1979

[4] T.Kailath, M.Morf, A.Vieira , "Inversion of Toeplitz operators, innovations and orthogonal polynomials," , *SIAM Rev,* vol 20, no.1, pp 106-119, 1978.

[5] Y.L Geronimus, "Polynomials orthogonals on a circle and their applications," *Amer Math.Soc.Transl,* vol 2, pp 253-304, 1954.

[6] U.Grenander, G.Szego , *Toeplitz forms and their applications,* Berkeley, CA:Univ.California Press, 1958.

[7] T.Georgiou ,"Realization of power spectra from partial covariance sequences," *IEEE Trans ASSP,* vol 35, no.4, April 87, pp 438-449.

[8] S.M Kay, S.Marple ,"Spectrum analysis- a modern perspective," *IEEE proc,* vol 69, no.11, Nov 1981.

[9] J.Makhoul, "Linear prediction: A tutorial review," *IEEE proc,* vol 53, pp 561-580, April 1975.

[10] A.Papoulis , *Probability, Random variables and stochastic processes,* 2nd ed. New York :Mc Graw Hill 1984.

[11] ________ ,*Signal analysis,* New York :Mc Graw Hill 1977.

[12] ________ "Maximum entropy and spectral estimation: a review," *IEEE Trans.Acoust.Speech,Signal Processing,* vol ASSP -29, pp 1176-1186, Dec 1982

[13] J.P Burg, "Maximum entropy spectral analysis,"
presented at the 37 th Intl Meet Soc Explor Geophys, Orlando, FL, 1976.

[14] T.Kailath, "Three decades of linear filtering theory,"
IEEE Trans.Inform.Theory, vol IT-20, pp 680-695, Mar 1974.