P.R.L.

TECHNICAL NOTE

TN-81-~~89~~ 28

HOW AND WHY ABOUT MEM IN SPECTRAL ANALYSIS - PRACTICAL CONSIDERATIONS

By

D.R. Kulkarni & S.K.Alurkar

December, 1981

PHYSICAL RESEARCH LABORATORY
NAVRANGPURA, AHMEDABAD-380009

CONTENTS

HOW AND WHY ABOUT MEM IN SPECTRAL ANALYSIS - PRACTICAL CONSIDERATIONS

D.R. Kulkarni & S.K. Alurkar
Physical Research Laboratory, Ahmedabad 380 009, India

Abstract

The importance of power spectrum analysis of stationary signals in extracting correct and minute details of spectral contents from a data sequence of a finite length is well-recognized. There has been an impressive development in the methods of spectral analysis due to the ever-increasing demand for precise and highly resolved spectral information. Beginning from the conventional method of spectral analysis by Blackmann-Tukey, which yields low accuracy and low resolution spectra since they are estimated from autocorrelation functions (ACFs) of zero-extended data sequences, high resolution spectral methods, using maximum entropy approach, have recently been developed. The latter is equivalent to least-square fitting of a discrete time all-pole model to the data of a random process. In maximum entropy approach, the ACFs are estimated on the basis of their value at lag zero, and then the subsequent ACFs are computed using the value at zero lag together with the previously estimated ACF using the principle of maximum entropy.

In this report, we first present a qualitative comparison between the Blackmann-Tukey and MEM methods. This is followed by a detailed discussion of the all-pole model with recursive

filter design using an essential and basic set of mathematical formulae.

We show how in the Burg's MEM the ACFs are estimated in a way different from that used in the all-pole model following the Levinson's recursive procedure.

It is very important to accurately choose the order of the filter as it determines the accuracy and resolution of the spectral estimates. We present the various conclusions arrived at by different workers in this field. It is shown that the accuracy of spectral estimates also depends on the data length in comparison with the periodicities of the phenomena and on rounding errors caused due to finite word-size.

To overcome these difficulties a new algorithm, called Least Squares (LS) linear prediction, has been developed by Barrodale & Erickson. We present results of comparative study of the Burg's MEM and LS method made using real data of 10 cm wavelength solar flux values for the last 34 years. It was concluded that it is rather difficult to make correct spectral estimates from a set of 'real data' having length less than the period of phenomenon under consideration.

In Appendix I, we present the Levinson's algorithm for solving linear simultaneous equations. A general description of the computer program for the Burg's method is outlined in Appendix II.

Keywords : Spectral analysis, stationary signals, Maximum Entropy Method, All-pole model, Levinson recursive procedure, Least-square linear prediction.

What is Spectral analysis?

Put simply, it is a process of estimating spectral features contained in a given data sequence emanated from a stationary source. There exist an infinite number of such sequences for a given spectrum, provided these sequences have the same auto-correlation function (ACF). For example, the sequences representing an impulse function and random noise yield an identical flat spectrum since they have the same ACF which itself, is an impulse. In general, any sequence and its complex conjugate time reverse share the same ACF and spectrum. Thus, it is imperative that the ACF of a data sequence should be known precisely to estimate its spectrum which compares well with the true spectrum.

1. Comparison of various spectral methods:

There are two well-known methods of spectral analysis; (a) Blackmann-Tukey and (b) Fast Fourier Transform (FFT).

(a) Blackmann-Tukey Method:

This method requires computation of ACF using a zero-extension of a given data sequence. Consequently, this ad hoc extension of data gives unrealistic estimates of ACF which, in turn, limit the accuracy and resolution of the estimated spectrum. However, an improvement in the resolution can be effected if the ACF is evaluated by dividing the sum of lag-products by the total number of points rather than by the total number of lag-products (Radoski et al., 1975).

b) FFT:

This method is based on a periodogram approach which assumes that the data sequence is periodic, one period being equal to the data length. Here the data extension is made by using the same data; and hence the ACF is not accurately estimated.

It may be noted that both these methods of spectral analysis suffer from the drawback that the accuracy of their spectral estimates is much low when the expected periodicities are comparable to the data length.

c) MEM:

In information theory, the total information content of a system is related to the probability distribution of various events occurring in it. Further, entropy is defined as the average information per unit time. As probability represents the uncertainty of the phenomenon, the entropy is a measure of the uncertainty in the system. According to the maximum entropy principle, the uncertainty (entropy) in the system is maximized w.r.t. the unavailable information, so that only the available information is used to extract various quantities from the system. In the case of spectral analysis the spectral estimates are derived using ACFs obtained from available data from a system in which the uncertainty is maximized w.r.t. the unavailable ACFs (obtained using data extended beyond the available data). Thus, these spectral estimates are independent

of the extended unknown data (Smylie et al., 1973). It has been shown (van den Bos, 1971; Ulrych & Bishop, 1975) that the maximum entropy approach is equivalent to least square fitting of a discrete time all-pole model to the data of a random process. The ACF estimates by this method are made using the knowledge of its value first at zero lag, and then subsequently by using the value at zero lag together with other estimates evaluated previously. In all these successive computations of ACF the entropy function is maximized. This results in higher spectral accuracy and resolution than those obtainable by the other methods.

The all-pole model in reality involves designing a recursive filter (Kulkarni & Alurkar, 1980) which yields a flat spectrum for any input data sequence. As this approach is intuitively more appealing than the entropy maximization approach, we shall adopt the filter design approach of the all-pole model for further discussions.

2. All-pole (autoregressive, AR) model:

In this model, a data point s_n is considered to be the output of a system with some unknown input u_n such that the following relation holds:

$$s_n = - \sum_{k=1}^{p} a_k \cdot s_{n-k} + G \sum_{l=0}^{q} b_l \cdot u_{n-l}, \qquad b_0 = 1 \qquad \ldots\ldots(1)$$

where a_k, $1 \leq k \leq p$; b_l, $1 \leq l \leq q$ and gain G are the parameters of the system (Makhoul, 1975). The output s_n is a

linear combination of past outputs and present and past inputs. In other words, s_n is a predictable signal point from past outputs and present and past inputs. Hence, this model is also known as Linear Prediction model.

In frequency domain, eqn. (1) can be written by taking z-transforms of its two sides. Thus,

$$H(z) = \frac{S(z)}{U(z)} = G \cdot \frac{1 + \sum_{l=1}^{q} b_l \cdot z^{-l}}{1 + \sum_{k=1}^{p} a_k \cdot z^{-k}} \qquad \ldots\ldots(2)$$

where $S(z) = \sum_{n=-\infty}^{\infty} s_n \cdot z^{-n}$ $\qquad \ldots\ldots(3)$

is the z-transform of s_n, $U(z)$ is the z-transform of u_n and $H(z)$ is the general pole-zero model. This is the transfer function of a recursive filter represented by eqn. (1).

The roots of the numerator and denominator of the polynomial in (2) are the zeros and poles of the model respectively. Two special cases of the model are of interest: (i) all-zero model, where $a_k = o$ for $1 \leqslant k \leqslant p$. This model is known as "moving average model" in statistics; (ii) all-pole model, where $b_l = o$, $1 \leqslant l \leqslant q$; this is known as "auto-regressive (AR) model". In the all-pole model, eqn. (1) takes the form:

$$s_n = -\sum_{k=1}^{p} a_k \cdot s_{n-k} + G\, u_n \qquad \ldots\ldots(4)$$

and the transfer function

$$H(z) = \frac{G}{1 + \sum_{k=1}^{p} a_k \cdot z^{-k}} \qquad \ldots\ldots(5)$$

Thus, the problem is to design a linear prediction filter whose transfer function is H(z). In practice, the input u_n is generally unknown and, therefore, the output s_n is only a linear combination of past samples. The co-efficients a_k in (4) can be obtained by the method of least squares.

Let $$\tilde{s}_n = - \sum_{k=1}^{p} a_k \cdot s_{n-k} \qquad \ldots\ldots(6)$$

be an approximation to s_n. Then the error between the actual and predicted s_n is given by

$$e_n = s_n - \tilde{s}_n = s_n + \sum_{k=1}^{p} a_k \cdot s_{n-k} \qquad \ldots\ldots(7)$$

Let $$E = \sum_n e_n^2 = \sum_n \left(s_n + \sum_{k=1}^{p} a_k \cdot s_{n-k}\right)^2 \qquad \ldots\ldots(8)$$

be the total squared error. According to the least squared method, we minimize E w.r.t. the coefficients a_k by setting

$$\frac{\delta E}{\delta a_i} = 0 \quad , \quad \text{for} \quad 1 \leqslant i \leqslant p \qquad \ldots\ldots(9)$$

From (8) and (9), we obtain a set of p normal equations with p unknowns:

$$\sum_{k=1}^{p} a_k \sum_n s_{n-k}\, s_{n-i} = - \sum_n s_n\, s_{n-i} \,, \quad 1 \leqslant i \leqslant p \qquad \ldots\ldots(10)$$

Solving these equations, the coefficients a_k are obtained.

The minimum squared error, E_p, will be

$$E_p = \sum_n s_n^2 + \sum_{k=1}^{p} a_k \sum_n s_n \cdot s_{n-k} \qquad \ldots\ldots(11)$$

Depending on the range of s_n used in (8), (10) and (11), two methods can be used to determine the coefficients a_k:

a) Autocorrelation method:-

Here the range of n is given by

$-\infty < n < \infty$. Hence, eqn. (10) and (11) become

$$\sum_{k=1}^{p} a_k . R(i-k) = -R(i) \quad , \quad 1 \leq i \leq p \qquad .....(12)$$

$$\text{and} \quad E_p = R(o) + \sum_{k=1}^{p} a_k \cdot R(k) \qquad .....(13)$$

$$\text{where } R(i) = \sum_{n=-\infty}^{\infty} s_n \cdot s_{n+i} \qquad .....(14)$$

is the ACF of the signal s_n.

b) Covariance method:-

Here the error is minimized over a finite interval $(0 \leq n \leq N-1)$ of n. Therefore, eqns. (10) and (11) reduce to

$$\sum_{k-1}^{p} a_k \cdot \Phi_{ki} = -\Phi_{oi} \qquad ....(15)$$

$$\text{and} \quad E_p = \Phi_{oo} + \sum_{k=1}^{p} a_k \cdot \Phi_{ok} \qquad ....(16)$$

$$\text{where } \Phi_{ik} = \sum_{n=o}^{N-1} s_{n-i} \cdot s_{n-k}$$

is the covariance of the signal s_n in the given interval.

The filter parameters a_k can be obtained by solving eqns. (12) & (13) or (15) & (16). The stability of the resulting filter transfer function H(z) is guaranteed in the auto-

correlation method, whereas it is not so in the case of the covariance method. Further development of MEM by Burg (Burg, 1967) is also based on autocorrelation method. Hence we shall confine our discussion to the autocorrelation method.

As shown earlier, parameters a_k can be evaluated using eqn. (10). Thus, in order to design a filter with H(z) as its transfer function as shown in eqn. (5), we need to estimate the gain parameter G. Comparing eqns. (4) and (7), it is seen that the input signal u_n that will result in the signal s_n as output, is that where $G\, u_n = e_n$. Thus, the input signal u_n is proportional to the error signal e_n. Therefore, the total energy in the input signal Gu_n is equal to the total energy in the error signal e_n given by E_p in eqn. (13). Assuming that the system model is energized by an input signal of the form of either stationary white noise or of an impulse function, we get,

$$G^2 = E_p = R(o) + \sum_{k=1}^{p} a_k \cdot R(k) \qquad \ldots.(17)$$

Further, to obtain the coefficients a_k, we rearrange eqns. (12) & (13) thus,

$$R(o) + \sum_{k=1}^{p} a_k \cdot R(k) = E_p \qquad \ldots.(18)$$

and $$R(i) + \sum_{k=1}^{p} a_k \cdot R(i-k) = 0 \qquad \ldots.(19)$$

These can be written in the matrix form, called Toeplitz matrix, as

$$\begin{bmatrix} R_o & R_1 & R_2 \cdots R_p \\ R_1 & R_o & R_1 \cdots R_{p-1} \\ \vdots & \vdots & \vdots \\ R_p & R_{p-1} & R_{p-2} \cdots R_o \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ \vdots \\ a_p \end{bmatrix} = \begin{bmatrix} E_p \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad \ldots\ldots(20)$$

where the coefficients $1, a_1, \ldots\ldots a_p$ can be considered as the constants of the "prediction error filter". These equations can be solved very efficiently by the Levinson's recurssive procedure described in Appendix I.

3. Relation of the filter constants to the input spectrum:

From eqn. (2), we have

$$S(z) = H(z) \cdot U(z)$$

For the assumed form of the input signal, its spectrum $U(z)$ is constant which can be made equal to unity.

Hence, $S(z) = H(z) = \frac{G}{A(z)}$ $\qquad \ldots\ldots(21)$

where $A(z) = 1 + \sum_{k=1}^{p} a_k \cdot z^{-k}$ is called an "inverse filter".

$$\therefore \quad S^2(\omega) = \tilde{P}(\omega) = \frac{G^2}{\left|A(e^{j\omega})\right|^2} \qquad \ldots\ldots(22)$$

Taking z-transform of eqn. (7), we get

$$E(z) = A(z) \cdot S(z) \qquad \ldots\ldots(23)$$

where $\mathbf{E}(z)$ is the z-transform of e_n.

$$\therefore \quad P(\omega) = S^2(\omega) = \frac{|E(e^{j\omega})|^2}{|A(e^{j\omega})|^2} \qquad \ldots\ldots (24)$$

$\therefore$ comparing (22) and (24), we have

$G^2 \simeq |E(e^{j\omega})|^2$ for all ω . This means that the output error spectrum is approximated by a flat spectrum. Hence, the filter $A(z)$ is known as a "whitening filter".

As seen from eqn. (22), the estimated power spectrum is inversely proportional to $A^2(z)$. As a result, even very narrow spectral peaks can be easily resolved. Hence, the edge of this spectral method over other methods for obtaining high resolution spectra. However, even a small change in the filter coefficients may change the parameters of the peaks considerably. It is therefore absolutely necessary that the computations are performed with high precision.

According to Parceval's theorem, the total error to be minimized in frequency domain is given by

$$E = \sum_{-\infty}^{\infty} e_n^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |E(e^{j\omega})|^2 \, d\omega \qquad \ldots\ldots (25)$$

Using eqns. (22), (24) and (25) we get

$$E = \frac{G^2}{2\pi} \int_{-\pi}^{\pi} \frac{P(\omega)}{\tilde{P}(\omega)} \, d\omega \qquad \ldots\ldots (26)$$

Thus, minimization of E amounts to minimization of the integrated ratio of the signal spectrum $P(\omega)$ to $\tilde{P}(\omega)$.

As $p \to \infty$, $\tilde{P}(\omega) \to P(\omega)$. In other words, any spectrum can be closely approximated by an all-pole model. In general, it can be shown that the estimated spectrum by an all-pole model fits the peaks in the signal spectrum better than the zeros in that spectrum. Thus, a model spectrum can be a good estimate of the spectral envelop of the signal spectrum.

4. Burg's Maximum Entropy Method:

Unlike the all-pole model, the estimates of ACF by this method are made in a different way, but are consistent with the maximum entropy principle. However, these estimates are not usually computed directly in the evaluation of filter coefficients. An ingeneous use is made of the Levinson's recursive procedure in the Burg's method. For example, to solve a (3 x 3) equation such as

$$\begin{bmatrix} r_o & r_1 & r_2 \\ r_1 & r_o & r_2 \\ r_2 & r_1 & r_o \end{bmatrix} \begin{bmatrix} 1 \\ a_1' \\ a_2' \end{bmatrix} = \begin{bmatrix} v' \\ 0 \\ 0 \end{bmatrix} \qquad \text{.....(27)}$$

We need to know according to the Levinson's method solutions to equations

$$\begin{bmatrix} r_o & r_1 \\ r_1 & r_o \end{bmatrix} \begin{bmatrix} 1 \\ a \end{bmatrix} = \begin{bmatrix} v \\ o \end{bmatrix} \qquad \text{.....(28)}$$

so that

$$\begin{bmatrix} 1 \\ a_1' \\ a_2' \end{bmatrix} = \begin{bmatrix} 1 \\ a \\ o \end{bmatrix} - c \begin{bmatrix} o \\ a \\ 1 \end{bmatrix} \quad \ldots\ldots(29)$$

In the Levinson's method, the estimate of r_2 is required to find the value of c (vide Appendix I). On the contrary, Burg does not assume the knowledge of r_2. Consequently, there will be four unknowns and three equations from (27). Burg solves this difficulty by minimizing the error by considering both forward and backward predictions. Thus, he minimizes

$$E(a_1, a_2) = \sum_{t=2}^{N} \left| x_t + a_1 \cdot x_{t-1} + a_2 \cdot x_{t-2} \right|^2 + \left| \bar{x}_{t-2} + a_1 \cdot \bar{x}_{1-1} + a_2 \bar{x}_t \right|^2 \quad \ldots\ldots(30)$$

where the first term in the summation is the forward prediction error and the second term is the backward prediction error. Using (29) and (30), one can rewrite (30) as

$$E(c) = \sum_{t=2}^{N} \left| x_t + ax_{t-1} - c(ax_{t-1} + x_{t-2}) \right|^2 + \left| \bar{x}_{t-2} + a\bar{x}_{t-1} - c(a\bar{x}_{t-1} + \bar{x}_t) \right|^2 \quad \ldots\ldots(31)$$

It can be shown that after minimization of $E(c)$ w.r.to c, we get

$$c = \frac{2 \sum_t (\bar{x}_{t-2} + a.\bar{x}_{t-2})\,(x_t + a.x_{t-1})}{\sum_t (\bar{x}_{t-2} + a\,\bar{x}_{t-2})^2 + (x_t + a\,x_{t-1})^2} \qquad \ldots\ldots(32)$$

It is easy to see from (32) that the coefficient $c \leq 1$ and uses only the available data. Knowing c, the solutions for (3 x 3) set of equations can be obtained from (29). In his method, Burg has made use of the important fact that the Levinson's procedure guarantees a stable (minimum phase) filter. The evaluation of c based only on the available data points is equivalent to estimating the ACF. In the case discussed above the estimated r_2 will be

$$r_2 = \frac{c\,v - r_1 a}{v} \qquad \ldots\ldots(33)$$

In conclusion, the Burg's method can be used to evaluate a general filter with n coefficients $(1, a_1, a_2 \ldots\ldots a_n)$. The general formulae for evaluating these coefficients are given by Anderson (Anderson, 1974).

5. General discussion about the number of filter coefficients and resolution:

It is very crucial to make a correct choice of the order of the filter for accurate spectral estimates. In general, the spectral resolution increases with the order of the filter.

There seem to exist two schools of thought regarding the actual order of the filter. According to one (Ulrych & Bishop, 1975), there is an optimum size of the filter beyond which spectral estimates will not be accurate. Various criteria have been given to find the optimum size of the filter. According to the second (Radoski et al., 1975) there is no harm in having a filter with large number (80 to 90% of the data points) of coefficients, provided that an integrated power spectrum is computed. It may be noted that the variance of the integrated power spectrum is much smaller than that of the single power spectrum estimate.

In principle a synthetic signal with N number of peaks should require a filter of order N. But in this case, the experience shows that the peaks are not very precisely located in the MEM spectrum. However, in the presence of noise in the signal the same peaks are accurately located. This is due to the fact that the MEM is in reality a statistical fitting procedure. This is also true for real signals with high signal-to-noise ratios.

To summarize, as $p \to \infty$, $\tilde{P}(\omega) \to P(\omega)$. However, in practice, it is necessary to decide the optimum value of p without losing important spectral features. A large value of p increases the computations together with the possibility of ill-conditioning of the normal equations. A practical test for getting an optimum p_o is to see that the error curve becomes almost flat for $p > p_o$ (Makhoul, 1975).

Our experience with the MEM shows that the criteria of getting an optimum filter size often are not effective. In fact, they give too small a filter size for the spectral resolution to be adequate. On the other hand, the spectrum gets stabilized as the filter size increases in agreement with Radoski's finding (Radoski et al., 1975). We found that a filter size of 30 to 50% of the data points is adequate to yield satisfactory spectral resolution in many practical cases.

It is also shown that the accuracy of the spectrum depends on the phase of the input signal (Fougere et al., 1976). If this is so, it may introduce an element of doubt regarding the accuracy of spectral estimates in case of real data, wherein the phase of the signal is often unknown.

As stated earlier, most of the methods of spectral analysis do not give accurate spectral estimates when the data length of real signal is comparable to the periodicities of the phenomena under consideration. Although the MEM gives relatively better results for a given length of data, this is restricted to the data length of the order of one period of the phenomenon. For shorter data lengths, the spectral estimates tend to be inaccurate. In general, for smaller data lengths the order of the filter required is relatively quite large.

Another advantage of MEM is that it does not require windowing of the data sequence to be analyzed. In conventional methods of spectral analysis, any truncated data length calls

for both windowing and extension in some ad hoc manner. On the one hand, windowing reduces the spectral resolution, whereas, on the other hand, extension of data reduces the accuracy of the spectrum. In MEM, there is no ad hoc extension of data since the required additional data points are predicted. Hence, MEM does not require windowing as well.

Recently a new algorithm called Least Squares (LS) linear prediction for spectral analysis has been suggested (Barrodale & Erickson, 1980a) to overcome the shortcomings of the Burg's method, viz. inaccurate spectral estimates of short data samples and contamination of results by rounding errors.

We have tested this method using the FORTRAN Program (Barrodale & Erickson, 1980b) together with the sample data provided by these authors. For this data, Burg's MEM also was tried. It was found that both these methods could accurately resolve the periodicity for which the data length corresponded to 1.5 times the periodicity; whereas, for another periodicity for which the data length was only 0.23 times the periodicity, the MEM gave a wrong estimate and the LS method gave a correct spectral estimate.

Such a comparative study of these two methods was made using the real data of 10 cm wavelength solar flux values over the last 34 years. It is well known that this quantity has a periodicity of 11 years and thus our test data contained three cycles of the

solar activity. Both the methods gave accurate periodicity of 11 years for the entire data length. This exercise was carried out for various data lengths. When the latter became shorter than one cycle of the phenomenon, both the methods failed to yield accurate periodicity.

In conclusion, it is rather difficult to extract correct spectral information from a set of real data which contains less than one cycle of the expected periodicity.

The general description of the computer program of the Burg algorithm and a broad outline of its various modules are presented in Appendix II.

Acknowledgements

We like to sincerely acknowledge the useful discussions with Dr.K. Dobes of the Geophysical Institute of Prague. We are thankful to Mr.D. Stephen and Mr.P. Raghavan for the neat typing of this technical note. We appreciate the assistance rendered by Shri S.K. Shah and Shri K.J. Shah in compiling the 10 cm solar flux data and obtaining their spectra.

References

Anderson, M.	1974	Geophys., 39, 69.
Barrodale, I. & Erickson, R.E.	1980a	Geophys., 45, 420.
Barrodale, I. & Erickson, R.E.	1980b	Geophys., 45, 433.
Burg, J.P.	1967	Maximum entropy spectral analysis, Proc. 37th Annual International Meeting, Oklahama City.
Fougere, P.F., Zawalick, E.J. & Radoski, H.R.	1976	Phys. Earth Planet. Int., 12, 201.
Kulkarni, D.R. & Alurkar, S.K.	1980	PRL Tech. Note, TN-80-03.
Levinson, H.	1947	J. Math. Phys., 25, 261.
Makhoul, J.	1975	Proc. IEEE, 63, 561.
Radoski, H.R., Fougere, P.F. & Zawalick, E.J.	1975	JGR, 80, 619.
Smylie, D.E., Clarke, G.K.C. & Ulrych, T.J.	1973	Methods in Computational Physics, 13, 391, Acad. Press.
Ulrych, T.J. & Bishop, T.N.	1975	Rev. Geophys. Space Phys., 13, 183.
Van den Bos, A.	1971	IEEE Trans. on Information Theory, 17, 493.

Bibliography

Markel, J.D. & Gray, A.H.	1976	Linear Prediction of Speech, Springer-Verlag.
Radoski, H.R., Zawalick, E.J. & Fougere, P.F.	1976	Phys. Earth Planet. Int., 12, 208.
Ulrych, T.J. & Clayton, R.W.	1976	Phys. Earth Planet. Int., 12, 188.

Appendix I

Levinson's algorithm for solving linear simultaneous equations:

Standard computer algorithms for solving N simultaneous equations require time proportional to N^3 and computer memory proportional to N^2. Levinson computer algorithm for Toeplitz matrix requires time proportional to N^2 and memory proportional to N. The Toeplitz matrix is a symmetric matrix with all diagonal elements identical. Levinson utilizes this special symmetry to develop a fast method for solving linear simultaneous equations. This is a recursive method so that given a solution to the (k x k) set of equations, it is possible to find a solution to (k + 1) x (k + 1) set. First a solution to k = 1 is obtained. The solutions for higher k are obtained by increasing k by 1 at each stage using a set of formulae to be derived here. Before dealing with the general case of (n x n) equations, we shall show how the recursion works in going from (2 x 2) to (3 x 3) set of equations. The given (2 x 2) equations and their solutions are represented as follows:

$$\begin{bmatrix} r_o & r_1 \\ r_1 & r_o \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \end{bmatrix} = \begin{bmatrix} v \\ o \end{bmatrix} \qquad \ldots\ldots(1)$$

Now, given r_2 we define a quantity e as follows:

$$\begin{bmatrix} r_o & r_1 & r_2 \\ r_1 & r_o & r_1 \\ r_2 & r_1 & r_o \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ o \end{bmatrix} = \begin{bmatrix} v \\ o \\ e \end{bmatrix} \quad \ldots\ldots(2)$$

The last row in eqn. (2) defines the quantity e.

Our aim is to find a solution to the set of (3 x 3) equations given below:

$$\begin{bmatrix} r_o & r_1 & r_2 \\ r_1 & r_o & r_1 \\ r_2 & r_1 & r_o \end{bmatrix} \begin{bmatrix} 1 \\ a_1' \\ a_2' \end{bmatrix} = \begin{bmatrix} v' \\ o \\ o \end{bmatrix} \quad \ldots\ldots(3)$$

The important trick is to reverse the order of rows and columns in all the three matrices in eqn. (2). Thus,

$$\begin{bmatrix} r_o & r_1 & r_2 \\ r_1 & r_o & r_1 \\ r_2 & r_1 & r_o \end{bmatrix} \begin{bmatrix} o \\ a_1 \\ 1 \end{bmatrix} = \begin{bmatrix} e \\ o \\ v \end{bmatrix} \quad \ldots\ldots(4)$$

Note that the square matrix remains unchanged.

The Levinson recursion consists of subtracting an unknown portion c_2 of eqn. (4) from eqn. (2) so as to get the results (3):

$$\begin{bmatrix} r_o & r_1 & r_2 \\ r_1 & r_o & r_1 \\ r_2 & r_1 & r_o \end{bmatrix} \left\{ \begin{bmatrix} 1 \\ a_1 \\ o \end{bmatrix} - c_2 \begin{bmatrix} o \\ a_1 \\ 1 \end{bmatrix} \right\} = \left\{ \begin{bmatrix} v' \\ o \\ e \end{bmatrix} - c_2 \begin{bmatrix} e \\ o \\ v \end{bmatrix} \right\} \quad \ldots(5)$$

Comparing RHS of eqn. (3) and (5), we get

$$v' = v - c_2 e \qquad \ldots\ldots(6a)$$

and $e = c_2 v$ $\qquad \ldots\ldots(6b)$

$\therefore c_2 = e/v$

and $v' = v-(e^2/v) = v(1-e^2/v^2)$ $\qquad \ldots\ldots(7)$

Thus, the solutions to (3 x 3) equations are

$$a_1' = a_1 \ (1 - e/v)$$

$$a_2' = -e/v$$

It may be noted that as v and v' be always positive for all values of the order of the filter so that the resultant filter is always stable. We see from eqn. (7) that

$$-1 < e/v < +1$$

Hence, $|c_2|$ is always less than 1 for any intermediate value of $|c_k|$, where k varies from 1 to the order of the filter.

In fact, these values of c provide a built-in check for the stability of the prediction error filter. Incidentally, a small word-length and other round-off errors in the computer cause inaccuracies in r_i which, in turn, make the filter unstable.

The above procedure can be generalized by a set of formulae given below:

$$v_o = r_o$$

$$c_i = \left[r_i + \sum_{j=i}^{i-1} a_j^{(i-1)} \cdot r(i-j) \right] / \; v(i-1)$$

$$a_i^{(i)} = -c_i$$

$$a_j^{(i)} = a_j^{(i-1)} - c_i \, a_{i-j}^{(i-1)}, \quad 1 \leqslant j \leqslant (i-1)$$

$$v_i = (1-c_i^2) \, v_{i-1}$$

where i = 1, 2, ...n and j = 1, 2, ...n. In general, n recursions are necessary to solve the (n+1) equations. The output is the series of coefficients a_i (i = 1 to n) and v_n (Levinson, 1947).

Appendix II

General description of the computer package for obtaining spectral estimates by Burg's method:

A software package has been developed based on the recursive procedure given by Anderson (Anderson, 1974). The package consists of various modules listed below:

i) SUB. BURGSP

ii) SUB. STDVAR

iii) SUB. ORDER

iv) SUB. BPEF

v) SUB. LPEF

vi) SUB. FINDPK

vii) SUB. PWDINT

viii) SUB. PLOTT

ix) FUN. FPW

x) SUB. INTGRL

xi) FUN. SIMSON

xii) SUB. AUTCOR

xiii) SUB. FACTAC

A given data sequence is first converted into a sequence of standard variates so that the mean total input energy is equal to unity (SUB. STDVAR). According to the Parseval's theorem, it then follows that the output power or the area under the spectral density curve must also be unity. This serves as a check on the overall computation of the spectral estimates.

The order of the filter can be specified by the user or he can estimate it by using the Akaike criterion (Ulrych & Bishop, 1975) through the subroutine ORDER and subroutine PLOTT.

Then the Burg's method is used to estimate a spectrum with the specified resolution (SUB.BPEF). This is followed by locating various peaks that may be present in the spectrum (SUB. FINDPK, FUN. FPW). A search for locating the exact position of a peak is then carried out following Radoski's procedure (SUB. PWDINT, SUB. INTGRAL, FUN. SIMSON).

This package also includes the modules to be used for evaluating spectral estimates by the all-pole model (MEM without Burg's modification) using Levinson's algorithm (SUB.LPEF, SUB. AUTCOR, SUB. FACTAC).

The subroutine BURGSP calls the above modules to make spectral estimates either by the Burg's method or by the all-pole model. The main program which calls the subroutine BURGSP, requires the following main inputs:

a) the number of data points and their values
b) the time interval used for sampling
c) the order of the filter
d) the spectral resolution required
e) the control parameter to decide whether the Akaike criterion is to be used.