Jour. Korean Inst. Mining Geol.
Vol. 17, No. 4, p. 237-243, 1984

On the Computational Efficiency and Stableness of Burg's Algorithm for Maximum Entropy Spectral Analysis

Hee Joon Kim*

Abstract: Burg's algorithm for maximum entropy spectral analysis is studied with respect to its computational efficiency and stableness. The Burg's method is not only less efficient than the Yule-Walker's method but also sometimes unstable due to its mathematical irrationality. This irrationality is demonstrated by analyzing an artificial time series, and more stable and effective method is proposed. An efficient procedure using Goertzel's algorithm to compute power spectral densities is also proposed.

INTRODUCTION

Various computational methods for estimating the power spectral density of time series are now well known and widely used. In earlier days, available methods were only Fourier analysis or the method of Blackman and Tukey (1959) based on autocorrelation function (ACF). Burg (1967; 1968) introduced maximum entropy method (MEM) for spectral analysis which gives good resolution even for periods comparable to the data length. Its mathematical properties were discussed in detail by Lacoss (1971), Burg (1972) and Ulrych (1972). Baggeroer (1976) derived the probability density and confidence interval for the spectral estimation by the MEM using a Wishart model for the estimated covariance. Ulrych and Bishop (1975) suggested the use of Akaike's final prediction error (FPE) criterion (Akaike, 1969;1970) for determining the optimal length of prediction error filter (PEF). Anderson (1974), on the other hand, gave a fast and simple procedure to compute PEF coefficients in MEM. Based on this procedure, Hino (1977, p. 235-236) and Kanasewich (1981, p. 160-162) reported FORTRAN programs.

* Department of Applied Geology, National Fisheries University of Pusan

In the Burg's algorithm, both PEF and ACF are determined by minimizing the practical prediction error. Thus the ACF is not needed a priori in the Burg's method. It is by now well recognized that the Burg's method has potential superiority over other conventional spectral estimators, in paticular for short data length, but some inherent shortcomings are still remained. The purpose of this paper is to discuss the computational efficiency and the stableness of the Burg's algorithm. After the Burg's algorithm is reviewed briefly and examined by analyzing an artificial time series, a modified Burg's algorithm will be proposed. Furthermore, an efficient procedure to compute the power spectral density for desired frequencies using a estimated PEF will be also proposed.

MAXIMUM ENTROPY SPECTRAL ANALYSIS

Suppose a set of data $x_1, x_2, \cdots, x_N$ with an equal time spacing Δt. The MEM power spectrum $P(f)$ is then estimated by

$$P(f)=P_m\Delta t/|1+\sum_{k=1}^{m}a_{mk}\exp(-2\pi i f k\Delta t)|^2 \quad (1)$$

where f is the frequency, P_m the output power of $m+1$ long PEF, a_{mk} the PEF coefficient and i the imaginary unit. The P_m and a_{mk} are de-

termined by the following normal equation.

$$\begin{bmatrix} \phi_0 & \phi_1 & \cdots & \phi_m \\ \phi_1 & \phi_0 & \cdots & \phi_{m-1} \\ \vdots & \vdots & & \vdots \\ \phi_m & \phi_{m-1} & \cdots & \phi_0 \end{bmatrix} \begin{bmatrix} 1 \\ a_{m1} \\ \vdots \\ a_{mm} \end{bmatrix} = \begin{bmatrix} P_m \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (2)$$

where ϕ_k is the ACF with lag k. If we know all of the ACF, unknowns $(a_{m1}, \cdots, a_{mm}; P_m)$ are easily obtained. Because the matrix of ACF in (2) has a Toeplitz structure, the unknowns can be solved efficiently by Levinson's recursion formula (Levinson, 1947; Robinson and Silvia, 1979, p. 433-436). In this formula equation (2) is solved iteratively by stepwise increase of the matrix dimension $m-1$ to m. For $m=0$, P_0 is estimated by

$$P_0=\phi_0=\frac{1}{N}\sum_{k=1}^{N} x_k^2 \quad (3)$$

Above procedure is called the Yule-Walker's method.

On the other hand, the Burg's method regards ϕ_m in (2) as an unknown. So an additioal criterion is required before a unique solution is obtained. According to Burg (1968), a_{mm} is determined by minimizing an average output power when the PEF is operated in both forward and backward directions. The average power π_m of the $m+1$ long PEF is given by

$$\pi_m=\frac{1}{2(N-m)}\sum_{l=1}^{N-m}\left[\left(x_l+\sum_{k=1}^{m}a_{mk}x_{l+k}\right)^2 + \left(x_{l+m}+\sum_{k=1}^{m}a_{mk}x_{l+m-k}\right)^2\right] \quad (4)$$

By minimizing π_m with respect to a_{mm}, we can obtain a_{mm}. Once the a_{mm} is determined, remaining unknowns are calculated by the Levinson's recursion:

$$a_{mk}=a_{m-1k}+a_{mm}a_{m-1m-k} \qquad k=1, 2, \cdots, m-1 \quad (5)$$

If we put $a_{m0}=1$ and $a_{mk}=0$ for $k\geq m$, (5) holds for all m. The recursion formula for P_m is derived by inserting (5) in (2), i.e.,

$$P_m=P_{m-1}(1-a_{mm}^2) \quad (6)$$

The unknown ACF ϕ_m, if required, is also derived from (2) as

$$\phi_m=-\sum_{k=1}^{m}a_{mk}\phi_{m-k} \quad (7)$$

Equation (4) can be now rewritten by use of (5) as

$$\begin{aligned}\pi_m&=\frac{1}{2(N-m)}\sum_{l=1}^{N-m}\left[\left(\sum_{k=0}^{m}a_{mk}x_{l+k}\right)^2 + \left(\sum_{k=0}^{m}a_{mk}x_{l+m-k}\right)^2\right] \\ &=\frac{1}{2(N-m)}\sum_{l=1}^{N-m}\left[\left(\sum_{k=0}^{m}a_{m-1k}x_{l+k} + a_{mm}\sum_{k=0}^{m}a_{m-1m-k}x_{l+k}\right)^2 + \left(\sum_{k=0}^{m}a_{m-1k}x_{l+m-k}+a_{mm}\sum_{k=0}^{m}a_{m-1m-k}x_{l+m-k}\right)^2\right] \\ &=\frac{1}{2(N-m)}\sum_{l=1}^{N-m}\left[(b_{ml}+a_{mm}f_{ml})^2 + (f_{ml}+a_{mm}b_{ml})^2\right] \end{aligned} \quad (8)$$

We have here introduced the quantities

$$b_{ml}=\sum_{k=0}^{m}a_{m-1k}x_{l+k}=\sum_{k=0}^{m}a_{m-1m-k}x_{l+m-k} \quad (9)$$

$$f_{ml}=\sum_{k=0}^{m}a_{m-1k}x_{l+m-k}=\sum_{k=0}^{m}a_{m-1m-k}x_{l+k} \quad (10)$$

$$l=1, 2, \cdots, N-m.$$

Since b_{ml} and f_{ml} are independent of a_{mm}, the minimizing condition

$$\frac{\partial\pi_m}{\partial a_{mm}}=0, \quad (11)$$

gives

$$a_{mm}=-2\sum_{l=1}^{N-m}b_{ml}f_{ml}\Big/\sum_{l=1}^{N-m}(b_{ml}^2+f_{ml}^2) \quad (12)$$

Because

$$\frac{\partial^2\pi_m}{\partial a_{mm}^2}=\frac{1}{N-m}\sum_{l=1}^{N-m}(b_{ml}^2+f_{ml}^2)>0$$

the extremum of π_m for a_{mm} given by (8) is a minimum. A fast and simple procedure for estimating the Burg's algorithm was given by

Anderson (1974).

COMPUTATIONAL EFFICIENCY OF MEM

In this section we estimate required computational quantities for both the Yule-Walker and the Burg methods. In the Yule-Walker method the ACF in (2) is usually obtained by

$$\phi_k=\frac{1}{N}\sum_{i=1}^{N-k}x_{i+k}x_i,\quad k=0,1,\cdots,m \tag{13}$$

where $x_i=0$ for $i>N$. Its required computation is approximately

$$(m+1)(N-m/2).$$

To solve (2), the Levinson's recursion formula with respect to a_{mm} is

$$\begin{bmatrix}P_{m-1}\\0\\\vdots\\0\\Q_{m-1}\end{bmatrix}+a_{mm}\begin{bmatrix}Q_{m-1}\\0\\\vdots\\0\\P_{m-1}\end{bmatrix}=\begin{bmatrix}P_m\\0\\\vdots\\0\\0\end{bmatrix} \tag{14}$$

where

$$Q_{m-1}=\sum_{k=0}^{m}a_{m-1k}\phi_{m-k} \tag{15}$$

From (14), a_{mm} is obtained by

$$a_{mm}=-Q_{m-1}/P_{m-1} \tag{16}$$

Once a_{mm} is determined, remaining unknowns are obtained by (5) and (6). In this case, since we have already known Q_{m-1}, (6) can be rewritten as

$$P_m=P_{m-1}+a_{mm}Q_{m-1} \tag{17}$$

Thus required computations in (5), (15), (16) and (17) are about

$$2[m(m-1)/2]+2m=m(m+1)$$

Consequently total computations in the Yule-Walker method are about

$$(m+1)(N+m/2)$$

In the Burg method, if ACF is not needed, we must evaluate equations (5), (6), (9), (10) and (12) for estimating a power spectrum. Hence the total computations in the Burg method are

$$(m+1)[5(N-m)+3m]=(m+1)(5N-2m)$$

Consequently the ratio of required computations between the Yule-Walker and the Burg methods is

$$(5N-2m)/(N+m/2)$$

If we assume $N\gg m$, then the Burg method requires about five times more computations than the Yule-Walker method.

STABLENESS OF BURG'S ALGORITHM

Another shortcoming of the original Burg method is that equation (12) has a mathematical irrationality. Based on (12), the estimated coefficients of PEF sometimes diverge. In fact, we fail to obtain reasonable power spectra for the following time series

$$x_k=\exp[-c(k-1)]\sin[2\pi f_0\Delta t(k-1)]+n$$
$$k=1,2,\cdots,N \tag{18}$$

where c is a positive constant, f_0 a certain frequency and n the Gaussian noise. Fig. 1, for example, when $c=0.05$, $f_0=0.05$(Hz), $\Delta t=1$ (sec) and $N=101$, shows the Akaike's FPE (Akaike, 1969;1970) for the noise-free time series (the amplitude of noise is 0) in the range of $m=1$ to 9. The Akaike's FPE is the variance of prediction error in applying PEF to a time

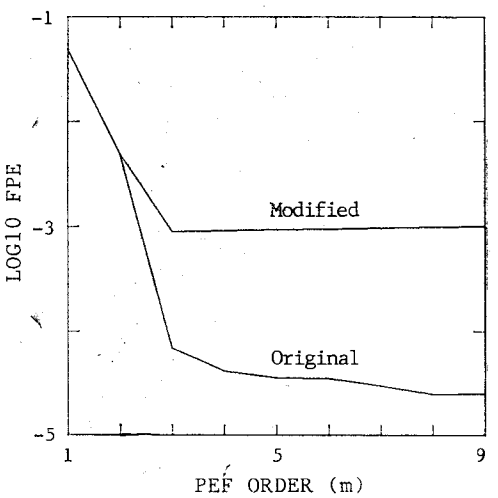

Fig. 1 Akaike's final prediction errors estimated by the original and the modified Burg methods. The tested time series is $x_k=\exp[-0.05(k-1)]\sin[0.1\pi(k-1)]$, $k=1,2,\cdots,101$.

series. That is,

$$FPE=E[|x_k-\hat{x}_k|^2] \tag{19}$$

where $E[\ ]$ represents the ensemble mean and $\hat{x}_k$ is the prediction of x_k. When we design $m+1$ long PEF using N-point time series x_k ($k=1, 2, \cdots, N$), the FPE is given by

$$FPE_{m+1}=P_m(N+m+1)/(N-m-1) \tag{20}$$

The FPE's in Fig. 1 (marked "Original") have no local minimum in the range of $m=1$ to 9, because P_m in (20) diverges (i.e., decreses too rapidly). The minimum FPE occurs in $m=26$.

To avoid such a divergence, the first term of (4) or (8) should be ignored. Then the alternative output power π_m' when the PEF is operated in only forward direction is

$$\pi_m'=\frac{1}{N-m}\sum_{l=1}^{N-m}\left(\sum_{k=0}^{m}a_{mk}x_{l+m-k}\right)^2$$

$$=\frac{1}{N-m}\sum_{l=1}^{N-m}(f_{ml}+a_{mm}b_{ml})^2 \tag{21}$$

Using the condition (11), we get

$$a_{mm}=-\sum_{l=1}^{N-m}b_{ml}f_{ml}/\sum_{l=1}^{N-m}b_{ml}^2 \tag{22}$$

The algorithm using (22) is refered as modified Burg method in this paper. The line marked "Modified" in Fig. 1 shows the Akaike's FPE's estimated by the modified Burg method for the same time series as in the original Burg method. A minimum value appears at $m=3$, so we can regard the optimal length of PEF as 4.

Fig. 2 shows power spectral densities obtained by the original and the modified Burg methods with each PEF length of 4. From this illustration we see that the modified Burg method gives a peak at 0.05 Hz, but the original Burg method has a some frequency shift. Such inaccurate frequency estimate for sinusoidal data in the original Burg method has been pointed out by many investigators (e.g., Barradale and Erickson, 1980). Moreover, the original Burg method requires about 20% more computations

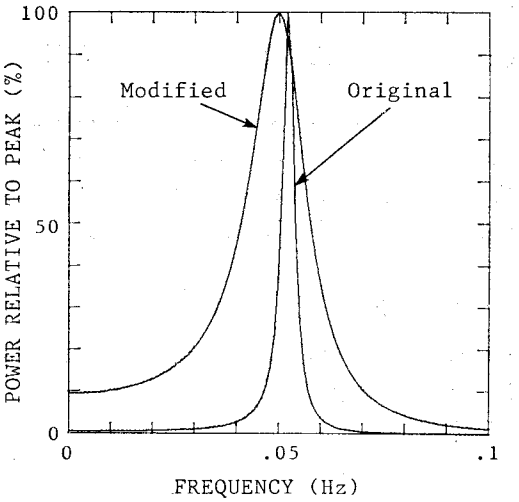

Fig. 2 Power spectra estimated by the original and the modified Burg methods. The tested time series is $x_k=\exp[-0.05(k-1)]\sin[0.1\pi(k-1)]$, $k=1, 2, \cdots, 101$.

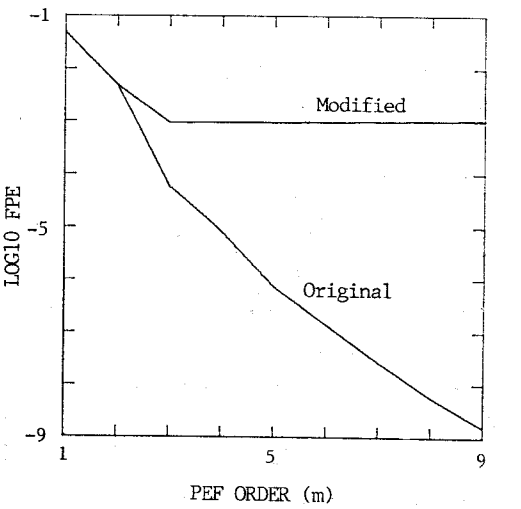

Fig. 3 Akaike's final prediction errors estimated by the original and the modified Burg methods. The tested time series is $x_k=\exp[-0.05(k-1)]\sin[0.1\pi(k-1)]+n$, $k=1, 2, \cdots, 101$ where n is the Gaussian noise with amplitude 0.001.

than the modified Burg method (compare (22) and (12)). Because of short PEF, the Yule-Walker's spectrum (not shown in Fig. 2) is almost the same as the modified Burg's one.

Fig. 3 shows the Akaike's FPE for the time series with a small Gaussian noise (the ampli-

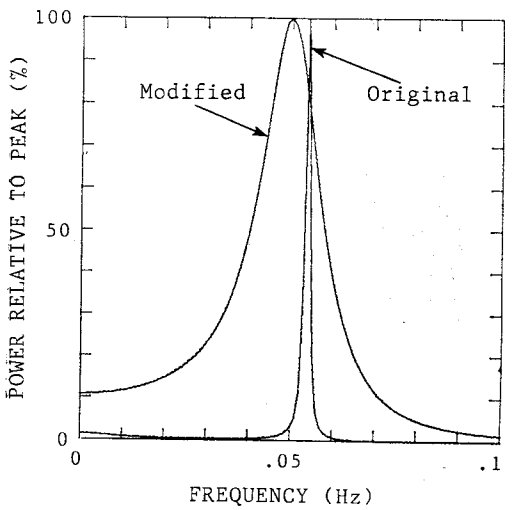

Fig. 4 Power spectra estimated by the original and the modified Burg methods. The tested time series is $x_k=\exp[-0.05(k-1)]\sin[0.1\pi(k-1)]+n$, $k=1, 2, \ldots, 101$, where n is the Gaussian noise with amplitude 0.001.

tude of noise is 0.001) in the range of $m=1$ to 9. The original Burg method also fails to have a minimum FPE in the range of $m=1$ to 9, and the minimum FPE appears in $m=40$. The modified Burg method, however, is stable for the time series with the small Gaussian noise, and the minimum FPE also appears in $m=3$. Fig. 4 shows power spectral densities for the time series with the small Gaussian noise. From Figs. 2 and 4, we see that the modified Burg spectrum is not affected with the small Gaussian noise.

COMPUTATION OF POWER DENSITY

A power density for density frequencies can be computed by equation (1). Some authors (e.g., Hino, 1977; Barrodale and Erickson, 1980) use equation (1) itself, direct calculation. However, since equation (1) involves a complex exponential function, large computations are required. Required computations at the summation of (1) are approximately m times of sine and cosine functions, respectively, and $2m$ times of multiplication.

If we rearrange (1)

$$P(f)=P_m\Delta t/|\sum_{k=0}^{m} a_{mk}\exp(-2\pi ifk\Delta t)|^2 \quad (23)$$

where $a_{m0}=1$, then we can use fast Fourier transforms (FFT). In this case, Goertzel's algorithm (e.g., Robinson, 1967, p. 279) is convenient as the FFT. In the Goertzel method, the summation in (23), $S(f)$, is calculated by

$$\begin{aligned} S(f)&=\sum_{k=0}^{m} a_{mk}\exp(-2\pi ifk\Delta t)\\ &=(1+S_0\cos\omega-S_1)+iS_1\sin\omega \end{aligned} \quad (24)$$

where $\omega=2\pi f\Delta t$ is the non-dimensional angular frequency, and S_0 and S_1 are obtained by the following recursion formula.

$$S_{k-1}=a_{mk}+2S_k\cos\omega-S_{k+1},\quad k=m, m-1, \cdots, 1 \quad (25)$$

Initial values for (25) are

$$S_{m+1}=S_m=0$$

When we use the recursion formula, required computations at the summation of (23) are approximately 1 time of sine and cosine functions, respectively, and $m+2$ times of multiplication. Since sine or cosine functions require about four times more computations than multiplications, then the ratio of computations between (23) and (1) is

$$\frac{2m\times 4+2m}{2\times 4+(m+2)}=\frac{10m}{m+10}$$

Therefore, for large m the procedure using the Goertzel method is about ten times faster than the conventional one.

DISCUSSION AND CONCLUSIONS

The Burg methd has higher resolution for short sample sinusoidal data than the Yule-Walker method (Radoski et al., 1975). In the Yule-Walker method ACF should be known before computing PEF coefficients (solving(2)). In computing the ACF we usually assume $x_i=0$ for $i>N$ in (13). Thus the ACF in (2) involves some error, and the error increases with

time lag. This means that the Yule-Walker method using PEF long compared with data length may give unreasonable power spectra. In the Burg method, on the other hand, the ACF is not needed a priori and is obtained from PEF coefficients (see equation (7)). This is one of the superiority of Burg method over the Yule-Walker method. In the method of Blackman and Tukey, the appropriate length of ACF is known to be shorter than about 10% of data length (e.g., Hino, 1977, p. 186). This criterion seems to be also useful in the Yule-Walker method. In the case of the tested time series of (18), in fact, the Yule-Walker's spectrum with $m=3$ (not shown in Fig. 2) was almost the same as the modified Burg's spectrum. Furthermore, the Yule-Walker method is about 4 times more efficient than the modified Burg one.

Ulrych and Bishop (1975) suggested that the optimal length of PEF can be determined by a local minimum of Akaike's FPE. In equation (20), P_m usually decreases monotonically with m, but $(N+m+1)/(N-m-1)$ increases monotonically. Thus, if the P_m decreases moderately with m, then the FPE has a minimum value at a certain m. The original Burg method, however, sometimes fails to have a minimum of FPE (Figs. 1 and 3), partly because of too much decreasing rate or divergence of P_m which is resulted from the mathematical irrationality in (12). In the case of the tested data of (18), the original Burg method gives the unreasonable spectrum with some frequency shift, and the spectrum is affected by the amplitude of added noise (Figs. 2 and 4). This shortcoming can be avoided by using the modified Burg method,

In this paper the computational efficiency and the stableness of Burg's algorithm were discussed and the modified Burg method was proposed. The computational efficinency of the original Burg method is only about 1/5 and 4/5 compared with that of the Yule-Walker and the modified Burg methods, respectively. Besides, the original Burg method is sometimes unstable becauue it involves mathematical irrationality. For the data lacking in line spectrum, which is frequently encountered in geophysical problems, the original Burg and Yule-Walker methods usually give similar results (e.g., Hino, 1977, p. 213). Therefore the maximum entropy spectral analysis should be carried out either by the Yule-Walker method or by the modified Burg one. The choice of which method is useful can be made by the length of PEF. When the length of PEF is shorter than about 10% of data length, the Yule-Walker method will be useful due to its computational efficiency. When the length of PEF is longer than about 10% of data length, the modified Burg method will be useful due to its superior resolution. Finally, I recommed the use of the Goertzel's algorithm in computing the power density for densired frequencies. For long PEF, this method has an efficiency about ten times more than the conventional one.

ACKNOWLEDGEMENTS

I wish to thank to Dr. Yong Q. Kang, National Fisheries University of Pusan, for his helpul comments on the manuscript.

REFERENCES

Akaike, H. (1969) Fitting autoregressive models for prediction. Ann. Inst. Statist. Math., v.21, p. 243-247.

Akaike, H. (1970) Statistical predictor identification. Ann. Inst. Statist. Math., v. 22, p. 203-217.

Anderson, N. (1974) On the calculation of filter coefficients for maximum entropy spectral analysis. Geophysics, v. 39, p. 69-72.

Baggeroer, A.B. (1969) Confidence intervals for regression (MEM) spectral estimates. IEEE trans. Inform. Theory, IT-22, p. 534-545.

Barrodale, I. and Erickson, R.E. (1980) Algorithms for least-squares linear prediction and maximum

entropy spectral analysis, Part II:Fortran program. Geophysics, v. 45, p. 433-446.

Blackman, R.B. and Tukey, J.W. (1959) The Measurement of Power Spectra. New York, Dover, 190pp.

Burg, J.P. (1967) Maximum entropy spectral analysis. Presented at the 37th Annual International SEG Meeting, Oklahoma, USA.

Burg, J.P. (1968) A new analysis technique for the time series data. Presented at the NATO Advanced Study Institute on Signal Processing with Emphasis on Underwater Acoustics, Eschede, Netherlands.

Burg, J.P. (1972) The relationship between maximum entropy spectra and maximum likelihood spectra. Geophysics, v. 37, p. 375-376.

Hino, M. (1977) Spectral Analysis. Asakura Shoten, 300pp. (in Japanese)

Kanasewich, C.R. (1981) The Time Sequence Analysis in Geophysics, 3rd Ed. Univ. of Alberta Press, 450pp.

Lacoss, R.T. (1971) Data adaptive spectral analysis method. Geophysics, v. 36, p. 661-675.

Levinson, N. (1947) The Wiener RMS error criterion in filter design and prediction. J. Math. Phys., v. 25, p. 261-278.

Radoski, H.R., Fougere, P.F. and Zawalick, E.J. (1975) A comparison of power spectral estimates and applications of the maximum entropy method. J. Geophys. Res., v. 80, p. 619-625.

Robinson, E.A. (1969) Multichannel Time Series Analysis with Digital Computer Programs. Holden-Day, 298pp.

Robinson, E.A. and Silvia, M.T. (1979) Digital Foundation of Time Series Analysis, vol. 1. Holden-Day, 451pp.

Ulrych, T.J. (1972) Maximum entropy power spectrum of truncated sinusoids. J. Geophys. Res., v. 77, p. 1396-1400.

Ulrych, T.J. and Bishop, T.N. (1975) Maximum entropy spectral analysis and autoregressive decomposition. Rev. Geophys. and Space Phys., v. 13, p. 183-200.

최대엔트로피 스펙트럼 분석에 관한 Burg알고리즘의 계산효율과 안정성에 대하여

김 희 준*

요약 : 최대 엔트로피 스펙트럼 분석(Maximum entropy spectral analysis)에 관한 Burg알고리즘(algorithm)의 계산효율과 안정성에 대하여 검토하였다. Burg의 방법은 Yule-Walker의 방법보다 계산효율이 낮을 뿐만 아니라 그의 수학적 불합리성으로 인하여 간혹 불안전할 때가 있다. 이러한 불합리성을 인공 시계열의 해석을 통하여 증명한 후 보다 안정하고 효과적인 방법을 소개하였다. 또한 파워 스펙트럼밀도(Power spectral density)의 계산에서 Goertzel알고리즘을 이용하는 효율적인 방법도 소개하였다.

* 부산수산대학 응용지질학과