

スペクトル解析と最大エントロピー法

小 倉 久 直*

1. ま え が き

測定したデータ列あるいは時系列のスペクトル推定は広範な分野にわたるデータ処理において必要とされる問題であり、単なるスペクトルの推定にとどまらず、システムパラメータの推定または同定、予測、汙波、検出などの多目的なデータ解析の1つの段階としても重要な問題である。スペクトル推定法として従来からペリオドグラム法、あるいは相関関数を用いる Blackman-Tukey 法などがあるが、近年は Burg によって最大エントロピー法 (maximum entropy method 略して MEM) の名のもとに新しい推定法が提案され¹⁾、各方面で効果的に用いられると同時に、これによって多くのスペクトル推定法の研究が誘発された。これは与えられたデータ列が、エントロピーが最大すなわち最もランダム、となるようにスペクトルまたは相関関数を決定する方法である。後に述べるようにこの方法は、定常 Gauss 系列の場合には、自己回帰系列あるいは線形予測理論に基づくスペクトル推定法と等価である。定常系列の予測理論はすでに数学的に完成されており²⁾、多くの著書にも述べられている。この理論は応用上価値のある多くの内容を含むだけでなく、MEM をさらに拡張・変形・改良するための基礎をも与えるので、現在では予測理論に沿ったスペクトル推定法の研究が盛んに行われている。

スペクトル解析に関しては近年多くの論文・著書が出版されている。ここでは紙面の都合上2～3の著書^{3)～5)}、論文集⁶⁾をあげるにとどめる。スペクトル解析の現状については最近の解説論文⁷⁾があり詳しい文献リストがあげられている。本稿で述べる余裕のないスペクトル解析法の歴史、細かい計算技術とアルゴリズム、拡張、各種の応用例についてはそれらを参照さ

れたい。予測理論の文献は非常に多いが、古典的な名著²⁾のほかに、工科系向きの教科書^{8), 9)}およびスペクトル推定に関連して書かれた著書^{10), 11)}のみをあげておく。スペクトル推定の立場から予測理論を解説したものに論文¹²⁾があり、詳しい文献リストが付されている。

スペクトルの推定にあたって、種々の解析法の特長、欠点を理解しないで、計算法をそのままデータ処理に用いるとしばしば誤った結果をひきおこす。特に最大エントロピー法はその名のためか多分に魔術的な方法のごとく受取られがちである。各種の解析法の使用に際し誤りなく推定を行うには、それらの理論的な基礎と共に方法の特徴をよく理解していることが望ましい。本稿ではこのような目的のためにスペクトル推定法の基礎理論を、特に MEM に焦点をあててその基礎に関連する問題を種々の角度から解説する。同時にデータ処理ならびに研究に際し有用な公式をいくつか導き、シミュレーションによる事例によって各種推定法の特徴を説明する。まず2節ではデータ列の数学的モデルである定常系列について、スペクトル解析で用いる必要事項をまとめる。3節ではペリオドグラムおよび相関関数によるスペクトル推定法の基礎を述べる。4節では定常系列の予測理論をスペクトル解析に便利な形で初等的に解説し、実際に用いるいくつかの公式を導出し、さらに自己回帰系列、Gauss 系列のエントロピーなどとの関係を述べる。多くの数式が登場するが MEM を手の内のものとするためにはこの節を十分理解されることを希望する。5節では MEM の代表的な計算法として Yule-Walker 法および Burg 法を説明し、実験例を用いて MEM の特性を説明する。

2. 定常過程の基礎知識

相関関数・スペクトル表現 ここでは定常過程 $X(t)$, $-\infty < t < \infty$, を周期 Δ ごとに標本化した時系列 $X_m = X(m\Delta)$, $m=0, \pm 1, \pm 2, \dots$, を考える。平均を

* 京都工芸繊維大学 工芸学部

キーワード: スペクトル (spectrum), 相関関数 (correlation function), 予測 (prediction), 時系列 (time series), AR モデル (autoregressive model)

$\langle \rangle$ で、複素共役を $\bar{}$ で表わす。 X_m は一般に複素数値をとるものとし (たとえば電波の散乱波), 平均値は 0 としておく: $\langle X_m \rangle = 0$. 定常系列 X_m の相関関数を

$$R_{m-n} \triangleq \langle \bar{X}_n X_m \rangle, \quad n, m=0, \pm 1, \dots \quad (1)$$

とおく. 定義 (1) から導かれる重要な性質として

$$\bar{R}_m = R_{-m} \quad (2)$$

$$\sum_{m,n} R_{m-n} a_m \bar{a}_n \geq 0, \quad (a_m \text{ は任意}) \quad (3)$$

がある. (3) を R_m の非負定値性という. 相関関数は

$$R_m = \int_{-\pi}^{\pi} e^{imA} S(A) dA, \quad m=0, \pm 1, \pm 2, \dots \quad (4)$$

$$S(A) = \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} e^{-imA} R_m \geq 0, \quad |A| \leq \pi \quad (5)$$

の形に表現できる. $S(A)$ をスペクトル密度または電力スペクトルと呼ぶ. ここで周波数パラメータ A は周波数 f (Hz), 標本周期 Δ (s) と, $A = 2\pi f \Delta$ (rad) の関係にある. (4) を相関関数のスペクトル表現または Wiener-Khinchin の定理と呼ぶ. $S(A) \geq 0$ の性質は R_m の非負定値性 (3) の結果として導かれる.

$$F(A) = \int_{-\pi}^A S(A') dA', \quad |A| \leq \pi \quad (6)$$

をスペクトル分布関数という. $F(A)$ は有界な単調非減少関数である. $S(A)$ が存在しない (たとえばデルタ関数 $\delta(A-A_0)$ のような線スペクトルが現れる) 場合でも $F(A)$ は有界である.

定常系列に対してエルゴード性または多少緩い条件のもとで相関関数のエルゴード定理 (文献⁹⁾ 20 節参照)

$$r_m \triangleq \lim_{M \rightarrow \infty} \frac{1}{M} \sum_{n=0}^{M-1} \bar{X}_n X_{n+m} = \langle \bar{X}_n X_{n+m} \rangle = R_m \quad (7)$$

が成り立つ. 左辺は時間平均で定義される見本系列 (データ列) X_m の相関関数で, (1) と区別するために見本相関関数と呼んでおく. 長さ M のデータ列 X_m , $m=0, 1, \dots, M-1$, の見本相関関数を

$$r_m^{(M)} \triangleq \frac{1}{M} \sum_{n=0}^{M-m-1} \bar{X}_n X_{n+m}, \quad r_{-m}^{(M)} \triangleq \bar{r}_m^{(M)} \quad (8)$$

$$0 \leq m < M$$

$$\triangleq 0, \quad |m| \geq M$$

で定義する. $r_m^{(M)}$ も非負定値関数で, 明らかに $r_m^{(M)} \rightarrow r_m (M \rightarrow \infty)$ である.

ペリオドグラム 長さ M のデータ列の Fourier 変換を用いて

$$S^{(M)}(A) \triangleq \frac{1}{M} |\tilde{X}^{(M)}(A)|^2, \quad |A| \leq \pi \quad (9)$$

$$\tilde{X}^{(M)}(A) \triangleq \frac{1}{\sqrt{2\pi}} \sum_{m=0}^{M-1} X_m e^{-imA}, \quad |A| \leq \pi \quad (10)$$

により定義される $S^{(M)}(A)$ をペリオドグラムという. ペリオドグラムについては以下の性質が導かれる.

$$\lim_{M \rightarrow \infty} \langle S^{(M)}(A) \rangle = \lim_{M \rightarrow \infty} \frac{1}{M} \langle |\tilde{X}^{(M)}(A)|^2 \rangle = S(A) \quad (11)$$

すなわち $S^{(M)}(A)$ のアンサンブル平均は $M \rightarrow \infty$ で $S(A)$ に収束する⁹⁾. また (8) ~ (10) より明らかに

$$r_m^{(M)} = \int_{-\pi}^{\pi} e^{imA} S^{(M)}(A) dA \quad (12)$$

が成り立つ. (7) により左辺は $M \rightarrow \infty$ で R_m に収束する. (4) と比べれば右辺は $S^{(M)}(A) \rightarrow S(A)$ のように見えるがこの場合極限と積分が可換でないためこれは成立しない. すなわちペリオドグラムはスペクトル密度に収束しない. 一方, 累積ペリオドグラムを

$$F^{(M)}(A) = \int_{-\pi}^A S^{(M)}(A') dA' \quad (13)$$

で定義すると, Fourier-Stieltjes 積分の Lévy の定理によって $F^{(M)}(A) \rightarrow F(A)$ が証明される. すなわち, 累積ペリオドグラムはスペクトル分布関数に収束する. これは

$$\lim_{M \rightarrow \infty} \int_{-\pi}^A S^{(M)}(A') dA' = \int_{-\pi}^A S(A') dA' \quad (14)$$

が成り立つことを意味する.

$S^{(M)}(A)$ は $S(A)$ に収束しないが, (14) が成り立つ理由は直観的につぎのように説明できる. 長さ M のデータ列はただか $M/2$ 回振動するから, その Fourier 変換 $\tilde{X}^{(M)}(A)$ も, したがって $S^{(M)}(A)$ も $|A| \leq \pi$ の区間上で同程度振動し, M が大きくなるほど激しく振動するため $S(A)$ に収束しない. 一方 $F^{(M)}(A)$ は積分によって $S^{(M)}(A)$ の振動が平滑化されるため, $M \rightarrow \infty$ の極限で $F(A)$ に収束し (14) が成り立つ.

3. ペリオドグラム・相関関数による スペクトル推定

ペリオドグラム法 (9) はデータ列のエネルギースペクトルに相当し, FFT による処理が行えるため簡便なスペクトル推定法としてよく用いられる. (10) をデータ列 (X_0, \dots, X_{M-1}) の離散 Fourier 変換 (DFT) として計算する場合には, 離散周波数における値 $S^{(M)}(A_n)$, $A_n = 2\pi n/M$, $n = \text{整数 (mod. } M)$, が計算される. このとき $r_m^{(M)}$ は $S^{(M)}(A_n)$ の DFT として mod M の周期関数となり, $|m| \geq M$ で 0 にならない. 周波数分割を細かくするためには $A_n = 2\pi n/M'$ ($M' > M$) とすればよいが, これは $M' - M$ 個

の0を付加した長さ M' のデータ列 $(X_0, \dots, X_{M-1}, 0, \dots, 0)$ を用いることと等価である。

上に述べたように $S^{(M)}(A)$ は M 個程度のリップルをもち $S(A)$ の推定値にはなりえないが、(14)の定理が成り立つから、矩形窓または適当な周波数窓 $W(A)$ を用いた平滑化ペリオドグラム

$$\hat{S}(A) = \int_{-\pi}^{\pi} W(A-A') S^{(M)}(A') dA' \quad (15)$$

によってスペクトルの推定値 $\hat{S}(A)$ とすることができ、周波数窓の大きさはスペクトル形、データ長に応じて適当に選ぶ必要がある。もう1つの推定法は(11)の性質によって平均ペリオドグラム $\langle S^{(M)}(A) \rangle$ を用いるものである。アンサンブルを作るために長さ $M = N \times L$ のデータ列を N 個のデータ列に分割し、長さ L のペリオドグラム $S_n^{(L)}(A)$, $n=1, \dots, N$, の算術平均

$$\hat{S}(A) = \frac{1}{N} \sum_{n=1}^N S_n^{(L)}(A) \quad (16)$$

をスペクトルの推定値とする。アンサンブル平均によって $S^{(L)}(A)$ を平滑化するわけであるが、 N を増すと L が減るために DFT の周波数間隔が粗くなる欠点がある。一般的にいて同じ M に対しては(15)のほうがよい推定値が得られる。

データ長 M が有限のために、そのデータ窓の Fourier 変換に相当するサイドローブに基づくスペクトル歪みがある。これを軽減するためにデータに適当な時間窓を用いるが、詳細については文献 5), 7) を参照されたい。

相関関数法 (Blackman-Tukey 法) 相関関数 R_m に適当な推定値 \hat{R}_m を用い、(5)によって

$$\hat{S}(A) = \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} e^{-imA} \hat{R}_m w_m \quad (17)$$

をスペクトルの推定値とすることができる。ここに w_m は適当な時間窓である。相関関数の推定値 \hat{R}_m としては通常、(8)の見本相関関数 $r_m^{(M)}$ を用いる。 $\langle r_m^{(M)} \rangle = (1 - |m|/M) R_m$, つまり平均値は R_m に一致しないから $r_m^{(M)}$ は偏推定量である。 \hat{R}_m として $r_m^{(M)}$ の代りに

$$r_m^{(M)} \triangleq \frac{M}{M-m} r_m^{(M)} = \frac{1}{M-m} \sum_{n=0}^{M-m-1} \bar{X}_n X_{n+m} \quad (18)$$

を用いれば不偏推定量となるが、(18)はもはや非負定値でないから、(17)により計算したスペクトル密度に負値が現れることがある。 M が十分大きいときは両者に大差はない。(17)によるスペクトル推定法を Blackman-Tukey (略して BT) 法と呼ぶ。

(17)で時間窓を $w_m=1$ とし、 $\hat{R}_m=r_m^{(M)}$ とすれば(15)の関係によって(17)はペリオドグラム $S^{(M)}(A)$ に戻ってしまい $S(A)$ の推定値になりえない。 $|R_m| \approx 0$, $|m| > m_0$ とみなしうる最小の m_0 , すなわち相関長を考える。推定値 $r_m^{(M)}$ は $|m| > m_0$ でも $1/M$ に比例するほぼ一定の分散値をもって変動することが示される。これが $S^{(M)}(A)$ のリップルの原因であって、これを抑えるために(17)に時間窓 w_m が必要とされる。 w_m として非負定値な時間窓を用いればスペクトルに負値は現れないが常にスペクトル幅が広がる傾向があり、非負定値でない窓 (たとえば矩形窓) をとれば負のスペクトルが現れることがある。 $\hat{R}_m=r_m^{(M)}$ とおき、 W_m の Fourier 変換を

$$W(A) = \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} e^{-imA} w_m \quad |A| \leq \pi \quad (19)$$

とおけば(17)は(15)の平滑化ペリオドグラムに帰着する。したがってこの意味では2つのスペクトル推定法は同等なものである。計算機処理の点では(8)による $r_m^{(M)}$ の計算は時間がかかるので、FFT を用いるペリオドグラム法のほうが便利である。ペリオドグラムおよび平滑化ペリオドグラムの例は図1(b), 図2(b)に示す。太線は幅 $\pi/16$ の矩形窓で平滑化したものである。

4. 線形予測理論の基礎知識

Yule-Walker 方程式 定常系列 X_m の値を過去の n 個のデータ X_{m-1}, \dots, X_{m-n} , の線形結合

$$\hat{X}_m^{(n)} = - \sum_{j=1}^n a_j^{(n)} X_{m-j} \quad (20)$$

によって推定しよう。ここで予測係数 $a_j^{(n)}$ (定常性のため m によらない) は、予測誤差

$$\eta_m^{(n)} \triangleq X_m - \hat{X}_m^{(n)} = X_m + \sum_{j=1}^n a_j^{(n)} X_{m-j} \quad (21)$$

の分散すなわち2乗平均値

$$\sigma_n^2 \triangleq \langle |\eta_m^{(n)}|^2 \rangle = \left\langle \left| X_m + \sum_{j=1}^n a_j^{(n)} X_{m-j} \right|^2 \right\rangle \quad (22)$$

が最小になるように定めるものとする。(22)を $a_j^{(n)}$ で微分して0とおけば

$$\langle \eta_m^{(n)} \bar{X}_{m-j} \rangle = 0, \quad j=1, 2, \dots, n \quad (23)$$

が得られる。これは予測誤差 $\eta_m^{(n)}$ とデータ X_{m-j} との直交性を意味する。これと(22)から $(\sigma_n^2, a_1^{(n)}, \dots, a_n^{(n)})$ を定める、いわゆる Yule-Walker (Y-W) 方程式

$$R_0 + \sum_{j=1}^n R_{-j} a_j^{(n)} = \sigma_n^2 \quad (24)$$

$$R_m + \sum_{j=1}^n R_{m-j} a_j^{(n)} = 0, m=1, 2, \dots, n \quad (25)$$

が得られる。この解を Cramer 公式で書きおろせば

$$\sigma_n^2 = \frac{D_n}{D_{n-1}}, a_j^{(n)} = \frac{A_j^{(n)}}{D_{n-1}}, j=1, 2, \dots, n \quad (26)$$

$$D_n = |D_n| \triangleq |R_{j-k}|_{n+1} = \begin{vmatrix} R_0 R_{-1} \cdots R_{-j} \cdots R_{-n} \\ R_1 R_0 \cdots \cdots R_{1-n} \\ \cdots \cdots \cdots \cdots \cdots \cdots \\ R_j R_{j-1} \cdots R_0 \cdots R_{j-n} \\ \cdots \cdots \cdots \cdots \cdots \cdots \\ R_n R_{n-1} \cdots R_{n-j} \cdots R_0 \end{vmatrix} \quad (27)$$

となる。ここで D_n は $[R_{j-k}]$ を要素とする $(n+1)$ 次元の Toeplitz 行列 (対角線に沿う成分が等しい行列), $D_n = |D_n|$ は Toeplitz 行列式である。 D_{n-1} および $A_j^{(n)}$ はそれぞれ D_n の最上行の R_0 および R_{-j} の余因子である。

Levinson の漸化式¹⁵⁾ n が大きいときは Cramer 解を直接計算するかわりにつぎの漸化式を用いて能率的に計算できる。これは計算機に適する解法であると同時に $n=1, 2, \dots$ に対し順に解が求まる利点がある。

$$\sigma_0^2 = R_0 \quad (28)$$

$$a_n^{(n)} = -\frac{1}{\sigma_{n-1}^2} \left[R_n + \sum_{j=1}^{n-1} a_j^{(n-1)} R_{n-j} \right] \quad (29)$$

$$a_j^{(n)} = a_j^{(n-1)} + a_n^{(n)} \overline{a_{n-j}^{(n-1)}}, j=1, 2, \dots, n-1 \quad (30)$$

$$\sigma_n^2 = (1 - |a_n^{(n)}|^2) \sigma_{n-1}^2 \quad (31)$$

(28) はデータの無い場合の 2 乗平均予測誤差である。

(28) ~ (31) を Levinson-Durbin のアルゴリズムと呼ぶ^{5), 6), 11)}。この証明は付録に与える。

2 乗平均予測誤差 予測の意味から明らかなように 2 乗平均予測誤差 σ_n^2 はデータ数 n と共に減少する:

$$R_0 = \sigma_0^2 \geq \sigma_1^2 \geq \cdots \geq \sigma_n^2 \geq \cdots \geq 0 \quad (32)$$

$$1 \geq |a_n^{(n)}| \geq 0, n=1, 2, \dots \quad (33)$$

(33) は (31), (32) より得られる。(33) の性質により $a_n^{(n)}$ は反射係数とも呼ばれる。 $\sigma_n^2 \rightarrow 0 (n \rightarrow \infty)$ ならば定常系列 X_m は完全予測可能または決定的という。以下では非決定的な場合のみを考え

$$\lim_{n \rightarrow \infty} \sigma_n^2 = \lim_{n \rightarrow \infty} \frac{D_n}{D_{n-1}} \triangleq e_0^2 > 0 \quad (34)$$

とする。 e_0^2 は最良 2 乗平均予測誤差である。特別な場合としてある次数 $n=N$ で最小値 e_0^2 に到達すれば

$$\sigma_{N-1}^2 > \sigma_N^2 = \sigma_{N+1}^2 = \cdots = e_0^2 \quad (35)$$

$$|a_n^{(n)}| = 0, n=N+1, N+2, \dots \quad (36)$$

が成り立つ。

予測フィルタ (逆フィルタ) 予測を周波数領域で考察しよう。(22) を相関関数で表わし (4) を用いれば

$$\sigma_n^2 = \int_{-\pi}^{\pi} |A_n(e^{-i\Lambda})|^2 S(\Lambda) d\Lambda \quad (37)$$

$$A_n(z) \triangleq 1 + a_1^{(n)} z + \cdots + a_n^{(n)} z^n \quad (38)$$

が得られる。Yule-Walker 方程式 (24), (25) を (4) を用いて書き直すと

$$\sigma_n^2 = \int_{-\pi}^{\pi} A_n(e^{-i\Lambda}) S(\Lambda) d\Lambda \quad (39)$$

$$\int_{-\pi}^{\pi} e^{im\Lambda} A_n(e^{-i\Lambda}) S(\Lambda) d\Lambda = 0, m=1, \dots, n \quad (40)$$

が得られる。これは n 次多項式 $A_n(z)$ を動かして (37) を最小化する条件である。ここで $n \rightarrow \infty$ とし $a_j^{(n)} = a_j, A_\infty(z) = A(z)$ とおく。すなわち

$$A(z) = \sum_{j=0}^{\infty} a_j z^j, a_0 \triangleq 1 \quad (41)$$

このとき (40) の条件は $A(e^{-i\Lambda}) S(\Lambda)$ が $e^{im\Lambda} (m \geq 0)$ の項のみを含むことを示す。 $\overline{A(e^{-i\Lambda})}$ も同様であるからこれらに乗じた $|A(e^{-i\Lambda})|^2 S(\Lambda)$ も $e^{im\Lambda} (m \geq 0)$ の項のみを含む。ところがこれは実であるから $m=0$ の項のみからなる。つまり常数に等しい。結局 $n \rightarrow \infty$ では (37) の被積分関数は常数に等しく、左辺は $\sigma_\infty^2 = e_0^2$ であるから、スペクトル密度は

$$S(\Lambda) = \frac{e_0^2}{2\pi |A(e^{-i\Lambda})|^2}, |\Lambda| \leq \pi \quad (42)$$

の形に書けることが示された。

$n \rightarrow \infty$ の場合の予測誤差を η_m と書けば, (21) は

$$\eta_m \triangleq e_0 W_m = X_m + \sum_{j=1}^{\infty} a_j X_{m-j} \quad (43)$$

となる。(4), (42), (43) より予測誤差の直交性

$$\langle W_m W_n \rangle = \delta_{mn} \quad (44)$$

が導かれる。すなわち W_m は分散 1 の白色雑音である。これをイノベーション系列とも呼ぶ。(43) は X_m を白色化する 1 つのフィルタを表わし, a_j はインパルス応答, $A(e^{-i\Lambda})$ は周波数応答を与える。これを予測フィルタまたは逆フィルタと呼ぶ。

移動平均・システムフィルタ (42) の関係を

$$S(\Lambda) = |G(e^{-i\Lambda})|^2 \quad (45)$$

$$G(z) \triangleq \frac{e_0}{\sqrt{2\pi} A(z)} = \frac{1}{\sqrt{2\pi}} \sum_{n=0}^{\infty} g_n z^n \quad (46)$$

$$g_0 = \sqrt{2\pi} G(0) = e_0 \quad (47)$$

の形に書き表わす。ここで g_n は (40), (42) により

$$g_n \triangleq \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} G(e^{-i\Lambda}) e^{-in\Lambda} d\Lambda = 0, n < 0 \quad (48)$$

となるから (46) の右辺のべき展開が成り立つ。したがって $A(z), G(z)$ は共に $|z| < 1$ で解析的で極も零ももたない。つぎに $\sqrt{2\pi} G(z) A(z) = e_0$ の両辺の z^n の

係数を比べると, $\{g_n\}, \{a_j\}$ の間の関係式

$$g_n + \sum_{j=0}^n g_{n-j} a_j = e_0 \delta_{n0}, \quad n=0, 1, 2, \dots \quad (49)$$

が得られる. これを $n=0, 1, \dots$ の順に解くことにより g_n, a_j の相互の変換が可能である.

(49) の関係を用いると (43) の逆の変換

$$X_m = g_0 W_m + \sum_{n=1}^{\infty} g_n W_{m-n} \quad (50)$$

が得られる. (50) の右辺を移動平均 (moving average 略して MA) という. これは白色雑音 W_n を X_m に変換するフィルタであって g_n がインパルス応答, $G(e^{-iA})$ が周波数応答を表わす. これをシステムフィルタと呼んでおく. (43) と比較すればわかるように (50) の第 1 項は予測誤差 η_m , 第 2 項は予測値 \hat{X}_m に等しい.

スペクトルの因子分解・ケプストラム (45) は $S(A)$ が (46) の形の関数 G および \bar{G} の積で与えられることを示す. これをスペクトルの因子分解という.

(46) の $G(z)$ は, Y-W 方程式の解で $n \rightarrow \infty$ として得られる予測フィルタ (41) によって定義されたが, これは $\log S(A)$ の Fourier 級数を用いて $S(A)$ から直接求めることもできる (これはいわゆる Wiener-Hopf 法 の原形となるものである). すなわち

$$S(A) = \exp[\log S(A)] \\ = \exp \left[c_0 + \sum_{n=1}^{\infty} c_n e^{-inA} + \sum_{n=1}^{\infty} \bar{c}_n e^{inA} \right] \quad (51)$$

$$c_n = \bar{c}_{-n} = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{inA} \log S(A) dA, \quad n=0, 1, 2, \dots \quad (52)$$

$$c_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log S(A) dA > -\infty \quad (53)$$

と置く. ここで $\log S(A)$ が可積分すなわち (53) を仮定すれば Fourier 係数 c_n は有界である. ここで

$$C(z) \triangleq \frac{c_0}{2} + \sum_{n=1}^{\infty} c_n z^n \quad (54)$$

を定義すればこれは $|z| < 1$ で解析的であって, (45), (51) を比較すれば直ちにつきの関係が得られる:

$$G(z) = e^{C(z)} \quad (55)$$

(34), (47), (53)-(55) よりさらにつきの重要な関係式

$$e_0^2 = \lim_{n \rightarrow \infty} \frac{D_n}{D_{n-1}} = 2\pi \exp \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} \log S(A) dA \right] \quad (56)$$

が導かれる. したがって (34) の $e_0^2 > 0$ の条件と $|c_0| < \infty$ つまり (53) の条件とは同等である.

$\{c_n\}$ または $C(z)$ は 1 つのフィルタを定義する.

$C(z)$ と他の 2 つのフィルタ $G(z), A(z)$ との関係は (55), (46) によって与えられているが, ここではスペクトル解析のうで都合のよい $\{c_n\}$ と $\{g_n\}, \{a_n\}$ との間の相互変換を可能にする漸化式を導こう. $G'(z) = C'(z)G(z)$ の両辺の z^m の係数を比較すれば

$$g_0 = \sqrt{2\pi} e^{c_0/2}, \quad m=0 \quad (57)$$

$$g_m = c_m g_0 + \frac{1}{m} \sum_{n=1}^{m-1} n c_n g_{m-n}, \quad m=1, 2, \dots \quad (58)$$

が得られる. (57) は (56) にほかならない. (58) を順に解けば $g_m(c_m)$ を $c_n(g_n), 0 \leq n \leq m$, によって表わす式が得られる. 同様に $A'(z) = -C'(z)A(z)$ の関係から, $a_0=1$, および

$$a_m = -c_m - \frac{1}{m} \sum_{n=1}^{m-1} n c_n a_{m-n}, \quad m=1, 2, \dots \quad (59)$$

が得られ, $a_m(c_m)$ を $c_n(a_n), 0 \leq n \leq m$, によって表わすことができる.

このフィルタの特長はシステムフィルタが多段に分解する, すなわち $G(z)$ が積に分解するとき, $C(z)$ は対応する和に分解することである. これはたとえば

$$G_1(z)G_2(z) = e^{C_1(z)+C_2(z)} \quad (60)$$

の関係から明らかである. スペクトル解析では $\{c_n\}$ をケプストラム (cepstrum) と呼ぶ.

自己回帰系列 上で述べた予測の特別の場合として, ある次数 $n=N$ で最良予測が得られるものとする. このとき (35), (36) が成り立つから, (30), (38) により

$$a_j = a_j^{(N)}, \quad j=1, \dots, N; \quad a_j=0, \quad N < j < \infty \\ AN(z) = A_{N+1}(z) = \dots = A(z) \quad (61)$$

の関係が得られる. これは予測が過去の N 個の値のみに依存し, それ以前の過去の値によらないことを意味する. これを広義 N 重 Markov 性という.

このときスペクトル密度 (42) および相関関数は

$$S_N(A) = \frac{e_0^2}{2\pi |A_N(e^{-iA})|^2}, \quad |A| \leq \pi \quad (62)$$

$$R_m = \frac{e_0^2}{2\pi} \int_{-\pi}^{\pi} \frac{e^{imA}}{|A_N(e^{-iA})|^2} dA, \quad m=0, \pm 1, \pm 2, \dots \quad (63)$$

$$A_N(z) \triangleq 1 + a_1 z + \dots + a_N z^N \\ = a_N(z-z_1) \dots (z-z_N), \quad |z_*| > 1 \quad (64)$$

で与えられる. ここに z_* は $A_N(z)=0$ の N 個の根であって, 上に述べたようにすべて $|z_*| > 1$ である.

(62) の形のスペクトル密度を次数 N の自己回帰形 (autoregressive 略して AR) あるいは全極形 (all-pole) スペクトルと呼び, そのような定常系列を AR 系列, 予測係数 $\{a_j\}$ を AR 係数とも呼ぶ.

一般に任意に相関関数 $R_m(0 \leq m < \infty)$ が与えられ

たとき, 2 乗平均予測誤差に関しては常に不等式 (32), すなわち

$$\sigma_{N^2} = D_N / D_{N-1} \geq e_0^2 \quad (65)$$

が成り立つ. N 次の AR 系列の予測誤差に対しては, (35) すなわち (65) で等式が成り立つ. このことを逆にいえば, (R_0, R_1, \dots, R_N) を指定した定常系列のうちで N 次の AR 系列が最大の 2 乗平均予測誤差をもつことになる. そのスペクトル密度は (62) で与えられ, (63) で計算される相関関数の最初の $N+1$ 個の値 R_0, R_1, \dots, R_N , は指定した値に一致する.

AR 系列に対しては, 予測誤差を与える (43) 式は

$$X_m + \sum_{j=1}^N a_j X_{m-j} = e_0 W_m \quad (66)$$

と書ける. ここに W_m は (44) の性質をもつ白色雑音である. 逆に (66) は白色雑音 W_m で励振される X_m の確率差分方程式を表す. これを AR 方程式と呼ぶ. この定常な解, すなわち AR 系列は (50) の移動平均によって与えられ, その係数 $\{g_n\}$ は, (46), (48) から, あるいは (49) を漸次的に解いて求められる. 実際 (49) は方程式 (66) の Green 関数 g_n の方程式にほかならない. AR 方程式 (66) を相関関数の満たす関係式として書き直しておこう. (66) の両辺に X_0 を乗じて平均をとり, (50), (44) を用いれば

$$R_m + \sum_{j=1}^N R_{m-j} a_j = e_0 g_m, \quad m=0, -1, -2, \dots \quad (67)$$

$$R_m + \sum_{j=1}^N R_{m-j} a_j = 0, \quad m=1, 2, 3, \dots \quad (68)$$

が得られる. (67) の $m=0$, および (68) の最初の N 個の式が Y-W 方程式 (24), (25) にほかならない.

ARMA 系列 有理形のスペクトル密度

$$S_N^M(\lambda) = \frac{e_0^2 |B_M(e^{-i\lambda})|^2}{2\pi |A_N(e^{-i\lambda})|^2} \quad (69)$$

$$B_M(z) = \sum_{k=0}^M b_k z^k, \quad b_0=1 \quad (70)$$

$$A_N(z) = \sum_{j=0}^N a_j z^j, \quad a_0=1 \quad (71)$$

を自己回帰-移動平均 (ARMA) 形のスペクトルとも呼ぶ. ここで多項式 $A_N(z)$, $B_M(z)$ はいずれも $|z| < 1$ で極, 零点をもたないものとする. (69) は (42), (45) いずれの形にも書き直すことができる. $N=0$ の場合を MA または全零形 (all-zero) スペクトルといい, (46) が M 次多項式となった場合に相当する. $M=0$ の場合は AR すなわち全極形である.

(69) のスペクトルをもつ定常系列は AR 方程式 (66) に類似した確率差分方程式

$$\sum_{j=0}^N a_j X_{m-j} = e_0 \sum_{k=0}^M b_k W_{m-k} \quad (72)$$

の定常解として表わしうる. ここで W_m は 1 つの白色雑音である. これを (M, N) 次の ARMA 方程式という. (72) の解も Green 関数 g_n を用いて (50) と同様な形に求められるが, これ以上はふれないでおく.

定常 Gauss 系列のエントロピー 実の強定常系列 X_m の長さ $M+1$ の系列 $(X_0, X_{-1}, \dots, X_{-M})$ の確率密度関数を $p(x_0, x_1, \dots, x_M)$ で表わそう. そのとき

$$H_\infty = \lim_{M \rightarrow \infty} H_M(x_0 | x_1, \dots, x_M) \\ \triangleq - \lim_{M \rightarrow \infty} \left(\log \frac{p(X_0, X_{-1}, \dots, X_{-M})}{p(X_{-1}, \dots, X_{-M})} \right) \quad (73)$$

を定常系列のエントロピーと呼ぶ. ここで $H_M(x_0 | x_1, \dots, x_M)$ は過去の M 個の値 (X_{-1}, \dots, X_{-M}) を条件とする X_0 の条件付エントロピー, したがって H_∞ は過去のすべての値を条件とするエントロピーを表わす. 右辺の平均は p 中の確率変数 X_m についての平均を意味する. X_m がエルゴード的ならば, McMillan の定理¹³⁾により単位時間当りの情報量は確率 1 で収束し, H_∞ に等しい:

$$- \lim_{M \rightarrow \infty} \frac{1}{M+1} \log p(X_0, X_{-1}, \dots, X_{-M}) \\ = H_\infty \quad (a. s.) \quad (74)$$

相関関数 R_m をもつ長さ $M+1$ の定常 Gauss 系列の確率密度はつぎの $M+1$ 次元 Gauss 分布,

$$p(x_0, x_1, \dots, x_M) \\ = \frac{1}{(2\pi)^{M+1/2} D_M^{1/2}} \exp \left[-\frac{1}{2} \sum_{m,n=0}^M [D_M^{-1}]_{mn} x_m x_n \right] \quad (75)$$

で与えられる. ここで D_M^{-1} は Toeplitz 行列 D_M の逆行列を表わす. (75) を (73) に代入して Gauss 系列のエントロピーを計算すると容易に

$$H_\infty = \log \sqrt{2\pi e} + \frac{1}{2} \lim_{M \rightarrow \infty} \log \frac{D_M}{D_{M-1}} \\ = \log \sqrt{2\pi e} + \frac{1}{2} \log e_0^2 \quad (76)$$

が得られる. ここで e_0^2 は (56) で与えられる最良 2 乗平均予測誤差である. もし (73) の代りに (74) の左辺を計算すれば, (76) で D_M/D_{M-1} を $[D_M]^{1/M+1}$ に置き換えたものが得られるが, これらは $M \rightarrow \infty$ で一致する. ちなみに (76) は予測誤差, すなわち (50) の第 1 項, のエントロピーにも等しい. さて (76) により H_∞ は e_0^2 について単調増加であるから, (R_0, \dots, R_N) を指定した Gauss 系列でエントロピーが最大のものは予測誤差 e_0^2 が最大のものである. それは, (65) で述べたことにより, N 次の (Gauss) AR 系列にほかな

らない。

5. 予測理論によるスペクトル推定法 (最大エントロピー法)

Burg の提案した MEM すなわちエントロピーが最大となるようにスペクトルを推定する方法は、前節の終りに述べたように、 (R_0, \dots, R_N) を指定した Gauss 系列の場合には、予測誤差が最大となるスペクトルすなわち AR スペクトル (62) を推定することにはほかならない。いいかえればスペクトルを AR モデルにあてはめることである。これは次数 N を定めた場合 $N+1$ 個の AR パラメータ (e_0^2, a_1, \dots, a_N) を推定することに帰着するが、その方法にいくつかの場合が考えられる。

Yule-Walker 法 長さ M のデータ (X_0, \dots, X_{M-1}) が与えられたとき、 n 個のデータ $(X_{m-n}, \dots, X_{m-1})$ で 1 つ先の値 X_m を予測する前向き予測誤差を

$$\eta_m^{(n)} \triangleq \sum_{j=0}^n a_j^{(n)} X_{m-j}, m=n, n+1, \dots, M-1 \quad (77)$$

$(a_d^{(n)} \triangleq 1)$ とおく。これは (21) にほかならない。 $M-n$ 個の予測誤差 $\eta_m^{(n)}$ の見本分散値を

$$\delta_n \triangleq \frac{1}{M-n} \sum_{k=n}^{M-1} |\eta_k^{(n)}|^2 = \sum_{m,j=0}^n \bar{a}_m^{(n)} a_j^{(n)} Q_{mj} \quad (78)$$

$$Q_{mj} \triangleq \frac{1}{M-n} \sum_{k=n}^{M-1} \bar{X}_{k-m} X_{k-j}, 0 \leq m, j \leq n \quad (79)$$

とおく。 M が大きくなればエルゴード性により

$$\delta_n \rightarrow \sigma_n^2, Q_{mj} \rightarrow r_{m-j} \quad (M \rightarrow \infty) \quad (80)$$

が成り立つ。ここで σ_n^2 は (22), r_{m-j} は見本相関関数 (7) である。 δ_n を最小にする条件により $a_j^{(n)}$ を定めるものとすれば、Y-W 方程式と同形の方程式

$$Q_{00} + \sum_{j=1}^n Q_{0j} a_j^{(n)} = \delta_n \quad (81)$$

$$Q_{m0} + \sum_{j=1}^n Q_{mj} a_j^{(n)} = 0, m=1, \dots, n \quad (82)$$

が得られる。この Cramer 解は直ちに求められるが、 $[Q_{mj}]$ は Toeplitz 行列でないので Levinson 漸化式は使えない。Y-W 法では Q_{mj} を長さ M のデータによる見本相関関数、すなわち (8) により近似して

$$Q_{mj} \cong r_{m-j}^{(M)}, \delta_n \cong \sigma_n^2 \quad (83)$$

とおく。これは (80) により M が十分大きければ良い近似である。次数 N を定め、 $a_j^{(N)} = a_j, \sigma_n^2 = e_0^2$ とおけば、(81), (82) は $n=N$ の場合として

$$r_0^{(M)} + \sum_{j=1}^N r_{-j}^{(M)} a_j = e_0^2 \quad (84)$$

$$r_m^{(M)} + \sum_{j=1}^N r_{m-j}^{(M)} a_j = 0, j=1, \dots, N \quad (85)$$

を与える。これは Y-W 方程式 (24), (25) において $N+1$ 個の (R_0, \dots, R_N) に対し、推定値 $(r_0^{(M)}, \dots, r_N^{(M)})$ を用いたものであって、これを解いて (e_0^2, a_1, \dots, a_N) を定める。この解は Levinson の漸化式 (28) - (31) を $n=1, \dots, N$ まで用いれば容易に求められる。以上を (62) に用いれば AR スペクトルが推定される。相関関数の推定値 \hat{R}_m は (63) によって計算できるが、特に最初の $N+1$ 個の値は

$$\hat{R}_0 = r_0^{(M)}, \dots, \hat{R}_N = r_N^{(M)} \quad (86)$$

に一致する。残りの推定値 $\hat{R}_m (m > N)$ も (63) によって求めるが、AR 系列に対して成り立つ関係 (68),

$$\hat{R}_m = - \sum_{j=1}^N \hat{R}_{m-j} a_j, m=1, 2, \dots \quad (87)$$

に (86) を用いれば漸化的に簡単に計算できる。特に (85) と (87) を比較すれば (86) は明らかである。したがって Y-W 法では (86) から $\hat{R}_m (m > N)$ へ一種の相関関数の外挿が行われることになるので、小さい次数 N でよく外挿ができる場合には、この方法は効果的な推定法となる。これが BT 法と大きく異なる特長であって、BT 法では (17) のように時間窓を乗ずるために一般にスペクトル幅は広がる欠点がある。逆に次数 N を極端に大きくしてデータ数 M に近づければ (86) によって $\hat{R}_m = r_m^{(M)}$ となるから、(12) の関係により推定スペクトル $\hat{S}(A)$ はペリオドグラム $S^{(M)}(A)$ に近づく。したがって N の増加と共にペリオドグラムの振動に対応するリップルが $\hat{S}(A)$ に現れ、 N が大きくなるほど起伏は激しくなる。

Burg 法 n 個のデータ $(X_{m-n}, \dots, X_{m-1})$ により X_m を予測する前向き予測誤差 (77) に対して、 n 個のデータ $(X_{m-n+1}, \dots, X_{m-1}, X_m)$ により 1 つ手前の値 X_{m-n} を予測する後向き予測誤差を

$$\zeta_m^{(n)} \triangleq \sum_{j=0}^n \bar{a}_j^{(n)} X_{m-n+j}, m=n, n+1, \dots, M-1 \quad (88)$$

とおく。予測係数が $\bar{a}_j^{(n)}$ であるのは、(2) によって後向き相関関数が $R_{-n} = \bar{R}_n$ となるためである。

Burg 法では、(78) の代りに、前向きおよび後向き予測誤差の見本分散値の和を算術平均したもの

$$\varepsilon_n \triangleq \frac{1}{2(M-n)} \sum_{m=n}^{M-1} [|\eta_m^{(n)}|^2 + |\zeta_m^{(n)}|^2] \quad (89)$$

$$= \frac{1}{2} \sum_{m,j=0}^n \bar{a}_m^{(n)} a_j^{(n)} [Q_{mj} + Q_{n-j, n-m}] \quad (90)$$

を取り扱う。 $M \rightarrow \infty$ ならば (80) と同様 $\varepsilon_n \rightarrow \sigma_n^2$ となり (78) と (89) の差はなくなる。もしすべての $a_j^{(n)}$ に

ついて (90) を最小化すれば, (81), (82) において $Q_{mj} \rightarrow P_{mj} \triangleq (Q_{mj} + Q_{n-j, n-m})/2$ の置換をした方程式が得られる. この Cramer 解を求める代りに, (82) に相当する式で, (83) のように P_{mj} がある Toeplitz 行列 \hat{P}_{m-j} で近似されるものとすれば, Levinson の漸化式 (30) を使うことができる. すなわち

$$a_j^{(n)} = a_j^{(n-1)} + a_n^{(n)} \bar{a}_{n-j}^{(n-1)}, \quad j=1, \dots, n-1 \quad (91)$$

が成り立つものとする. これは予測フィルタの安定性, つまり (64) が成り立つことを保証するのみならず, $\eta_m^{(n)}, \xi_m^{(n)}$ に対するつぎの漸化式を導く利点がある:

$$\eta_m^{(n)} = \eta_m^{(n-1)} + a_n^{(n)} \xi_{m-1}^{(n-1)} \quad (92)$$

$$\xi_m^{(n)} = \xi_{m-1}^{(n-1)} + \bar{a}_n^{(n)} \eta_m^{(n-1)} \quad (93)$$

実際 (77) に (91) を代入すれば

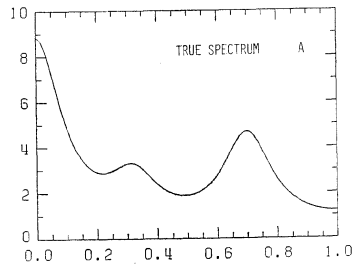
$$\begin{aligned} \eta_m^{(n)} = & X_m + \sum_{j=1}^{n-1} (a_j^{(n-1)} \\ & + a_n^{(n)} \bar{a}_{n-j}^{(n-1)}) X_{m-j} \\ & + a_n^{(n)} X_{m-n} \end{aligned} \quad (94)$$

となるが, これを並べかえて (88) を用いれば (92) が得られる. まったく同じことを (88) に対して行えば (93) が得られる.

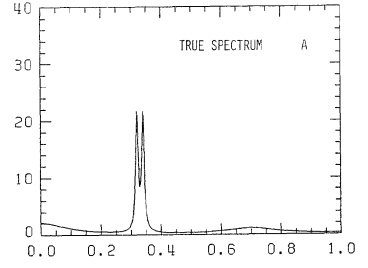
つぎに残る反射係数 $a_n^{(n)}$ を直接に (89) を最小にする条件により定める. すなわち (89) に (92), (93) を代入し, $\bar{a}_n^{(n)}$ で微分して 0 とおけば容易に漸化式

$$a_n^{(n)} = \frac{-2 \sum_{m=n}^{M-1} \xi_{m-1}^{(n-1)} \eta_m^{(n-1)}}{\sum_{m=n}^{M-1} [|\xi_{m-1}^{(n-1)}|^2 + |\eta_m^{(n-1)}|^2]} \quad (95)$$

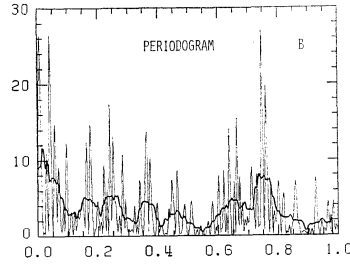
が得られる. (33) すなわち $|a_n^{(n)}| \leq 1$ が成り立つことは Schwarz の不等式を用いて容易に確かめられる. (95) が Levinson 漸化式 (29) に代るものである. 以上により Burg 法では (91) ~ (93), (95) および (31) の漸化式によってパラメータ $(\sigma_n^2, a_1^{(n)}, \dots, a_n^{(n)})$ を逐



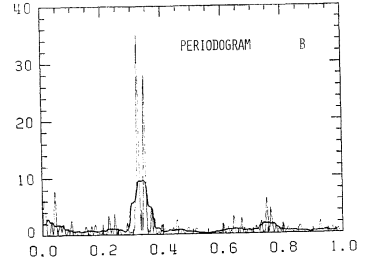
1-(a)



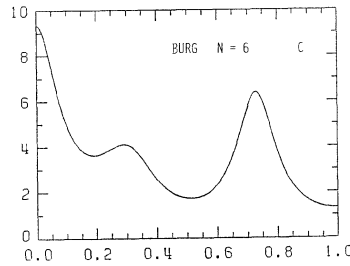
2-(a)



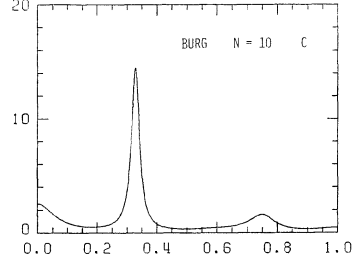
1-(b)



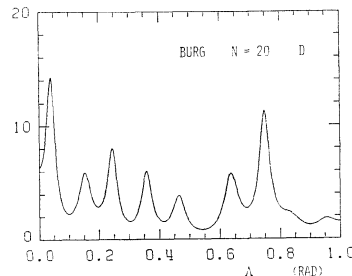
2-(b)



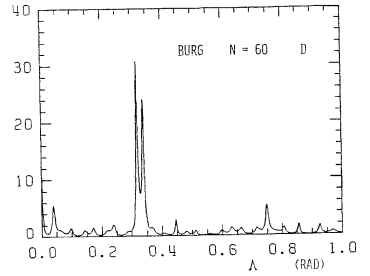
1-(c)



2-(c)



1-(d)



2-(d)

図 1 スペクトルの推定例

図 2 スペクトルの推定例

次求め AR スペクトルおよび相関関数を推定する.

Y-W 法の $a_n^{(n)}$ を求める漸化式 (29) では近似 (83) を用いるが, Burg 法では $a_n^{(n)}$ は直接に ϵ_n を最小化する条件から求め, 仮定 (91) 以外は用いない. この意味で, M が有限の場合には Burg 法は Y-W 法よりもデータの利用率が高く優れているが, M が増加すればこの差は小さくなる. 次数 N の増加と共にスペクトルにリップルを生ずることも Y-W 法と同様である.

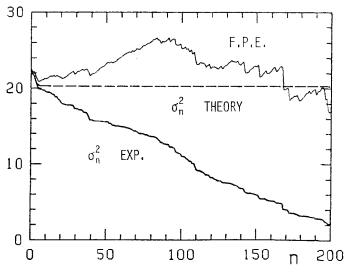


図3 2乗平均予測誤差

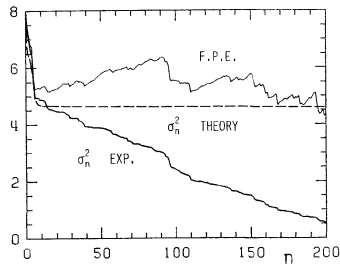


図4 2乗平均予測誤差

図1, 図2の(c), (d)にBurg法による推定例を示す。用いた時系列は、単峰スペクトルをもつ独立なGauss系列を、電力・中心周波数・スペクトル幅を変えて発生いくつか加えたもので、全体としてのスペクトルは(69)の有理形である。図の例のデータ数は $M=256$ であり、この M では同じ次数 N に対するY-W法とBurg法の推定スペクトルの差は小さい。

次数の決定 スペクトルが未知の場合に次数 N を定めることは重要な問題である。 N を定める1つの手懸りは(31)から得られる2乗平均予測誤差

$$\sigma_n^2 = R_0 \prod_{m=1}^n (1 - |a_n^{(m)}|^2) \quad (96)$$

である。これは n と共に単調に減少し理論的には(56)の最終値 e_0^2 に近づく。しかしながら実際の $a_n^{(m)}$ の推定値は M が有限のため必ず誤差を伴うので、(96)の右辺の積として求まる実験値 $\hat{\sigma}_n^2$ は n と共に0に収束し(図3, 図4), N を定める規準にはならない。 N の推定法としてはいくつかの提案があるが、一例としてよく知られている赤池の最終予測誤差¹⁴⁾をあげておく:

$$\text{FPE} \triangleq \hat{\sigma}_n^2 \frac{M+n+1}{M-n-1} \quad (97)$$

これは $\hat{\sigma}_n^2$ と n に関する増加因子の積により生じたFPEの最小位置 $n=N$ によって次数を定めようとするものである。しかしながら、通常は $\hat{\sigma}_n^2$ の減少が増加因子よりも速く、 n を大きくすれば減少して最小値をもたないので、最初の極小値を $n=N$ と定めておく。ほかにFPEと似た振舞いをする赤池の情報量規準AIC, ParzenのAR伝達規準などが提案されているが、いずれも決定的な有効性はない。以下に述べるようにこれらの方法の有効性はスペクトル形に大きく左右される。

連続スペクトルと線スペクトルの推定 理論的には、予測理論の適用は純非決定的^{9), 10)}すなわち $S(A)$ が条件(53), つまり条件 $e_0^2 > 0$, を満たす場合に限り、しかも線スペクトルのような完全予測可能な成分

は予測誤差、予測公式に寄与しないので除外される。まず緩い変化をする連続スペクトルをもつ場合を取り扱う(図1(a)). $S(A)=0$ となる区間はないから条件(53)を満たす。理論的には σ_n^2 は最終値 e_0^2 まで急速に下降して以後一定値を保つが、実験値 $\hat{\sigma}_n^2$ は下降点よりさらに0まで緩く下降する。

したがって $\hat{\sigma}_n^2$ の急下降の終わった

位置はFPEの最初の極小点に近い(図3)。この附近の位置 $n=N$ をARの次数に定めれば連続スペクトルに対してはよいスペクトル推定値を与える(図1(c)). それ以上に N を増加すればスペクトルにリップルが生じ(図1(d)), さらに N を大きくとればリップルはスペクトル状になって上に述べたように、ペリオドグラム(図1(b)の細線)に近づく。つぎに線スペクトルまたは狭帯域スペクトルをもつデータを取り扱うものとする。このとき σ_n^2 も $\hat{\sigma}_n^2$ も急速に $e_0^2 \approx 0$ に下降し、以後0に近い値をとる。この下降地点の次数では小さすぎるため狭帯域のスペクトルは推定できない。 ΔA (rad)の分解能を得るためには、次数は少なくとも $N \approx \pi/\Delta A$ またはそれ以上の大きさが必要である。この N ではもはや $\hat{\sigma}_n^2$ は誤差にすぎないから、スペクトルの高さはまったく正しくないが、線スペクトルの位置または分解能は得られる。 N をあげていくときデータ数 M が小さい場合には、Burg法のほうがY-W法より速く所要の分解能を与えることが示される。線スペクトルの推定例は多く報告されている⁶⁾⁻⁷⁾。図2は連続スペクトルの成分と、接近した2つの狭帯域スペクトルの成分とを等電力比で混合したデータ列に対する推定例で、この場合には上に述べた2つの事情が重なって現れる。この場合最終値 e_0^2 はほとんど連続部分によるものである(図4)。図2(c)($N=10$)では連続部分がよく推定されているが狭帯域部分の分解能は得られていない。図2(d)($N=60$)では狭帯域部分の分解能は得られているが、連続部分にリップルが生じ推定が劣化している。さらに N を増加するとピークはいずれも尖鋭化して図2(b)のペリオドグラムに近づく。

上の例からわかるようにMEMは有力なスペクトル推定法であるが、まだ完全ではなく種々の問題がある。これを改良・変形する多くの方法、ケプストラム法、ARMAスペクトルの推定法などを紹介する余白はないが、いずれも4節で与えた諸公式をもとに行われる。さらにこの方法は多次元スペクトル、画像など

の確率場のスペクトルの推定にも拡張されることを付記しておく。

謝 辞

本稿の図のために多くの計算機実験をしていただいた本学電子工学教室の吉田靖夫氏にお礼申し上げます。

(昭和 57 年 10 月 9 日受付)

参 考 文 献

- 1) J. P. Burg: Maxmum Entropy Spectral Analysis, Proceedings of the 37th Meeting of the Society of Exploration Geophysicists, (1967) 文献⁹⁾所収
- 2) U. Grenander and G. Szegő: Toeplitz Forms and Their Applications, Univ. of California Press (1958)
- 3) G. M. Jenkins and D. Watts: Spectral Analysis and Its Applications, Holden Day (1968)
- 4) M. B. Priestley: Spectral Analysis and Time Series, **1, 2**, Academic Press (1981)
- 5) 日野幹雄: スペクトル解析, 朝倉書店 (1977)
- 6) D. G. Childers (editor): Modern Spectrum Analysis, IEEE Press (1978)
- 7) S. M. Kay and S. L. Marple, Jr.: Spectrum Analysis—A Modern Perspective, Proc. of IEEE **69-12**, 1380/1419 (1981)
- 8) A. M. Yaglom: Introduction to the Theory of Stationary Random Functions, Dover Publications (1973)
- 9) 小倉久直: 確率過程論, コロナ社 (1978)
- 10) 藤井光昭: 時系列解析, コロナ社 (1974)
- 11) 添田, 中溝, 大松: 信号処理の基礎と応用, 日新出版 (1979)
- 12) J. Makhoul: Linear Prediction: A Tutorial Review, Proc. of IEEE, **63-4**, 561/580 (1975) 文献⁹⁾所収
- 13) R. G. Gallager: Information Theory and Reliable Communication, John Wiley and Sons, Chap. 3 (1968)
- 14) H. Akaike: Fitting Autoregressive Models for Prediction, Ann. Inst. Statist. Math., **21**, 243/247 (1969)
- 15) N. Levinson: The Wiener (Root Mean Square) Error Criterion in Filter Design and Application, J. Math. Phys. **25**, 261/278 (1947)

《附録 Levinson 漸化式の証明》

本文(29)～(30)の漸化式は最初 Levinson によって導かれ, 後に Durbin によって改良されたものである¹⁵⁾. Levinson 漸化式の導出法は種々可能である. 著者が試みた限りでは最も統一的方法は $S(A)$ に関する直交多項式・核多項式の漸化式²⁾を用いるもので, (29)～(30)のほかに逆向き漸化式も自然に導かれ

る. ここでは紙面の都合上, 別の短い証明を与える.

Y-W 方程式の Cramer 解(26)に対し, 余因子展開を用いて漸化式を導こう. まず $A_n^{(n)}$, すなわち(27)の行列式 D_n の最上行 R_{-n} の余因子をさらに余因子の最左列の R_j で展開する. R_j の余因子は D_{n-1} の最右列 R_{j-n} の余因子に等しいが, D_n が Toeplitz 行列式であるから, それは $A_{n-j}^{(n-1)}$ に等しい. したがって,

$$A_n^{(n)} = - \sum_{j=1}^n R_j A_{n-j}^{(n-1)} = - \sum_{j=0}^{n-1} R_{n-j} A_j^{(n-1)} \quad (\text{A. 1})$$

が成り立つ. この両辺を D_{n-1} で割って(26)を用いれば最初の漸化式(29)が得られる.

つぎに Y-W 方程式(25)で最後の $m=n$ の式を除いた $(n-1)$ 次元の方程式を

$$R_m + R_{m-n} a_n^{(n)} + \sum_{j=1}^{n-1} R_{m-j} a_j^{(n)} = 0 \\ m=1, 2, \dots, n-1 \quad (\text{A. 2})$$

と書く. 第1, 2項を既知の項とみなして $a_1^{(n)}, \dots, a_{n-1}^{(n)}$ を Cramer 解の形で求めよう. $a_j^{(n)}$ に対する第1項の寄与は $A_j^{(n-1)}/D_{n-2}$ である. ここで $A_j^{(n-1)}$ は D_n の左上の小行列式 D_{n-1} の第1行の R_{-j} の余因子として得られる. 第2項の寄与は $(\bar{A}_{n-j}^{(n-1)}/D_{n-2})a_n^{(n)}$ である. ここで $\bar{A}_{n-j}^{(n-1)}$ は D_n の右下の小行列式 D_{n-1} の最下行の要素 R_{n-j} の余因子として得られるが, $\bar{R}_n = R_{-n}$ および D_{n-1} が Toeplitz 行列式であることから, これは D_{n-1} の最上行 R_{j-n} の余因子 $A_{n-j}^{(n-1)}$ の複素共役に等しい. これらの2つの寄与を合せると, (A. 2)の解は

$$a_j^{(n)} = \frac{1}{D_{n-2}} [A_j^{(n-1)} + a_n^{(n)} \bar{A}_{n-j}^{(n-1)}] \\ = a_j^{(n-1)} + a_n^{(n)} \bar{a}_{n-j}^{(n-1)}, j=1, \dots, n-1 \quad (\text{A. 3})$$

となる. これが第2の漸化式(30)である.

つぎに(24)に(A. 3)を代入し(29)を用いれば

$$\sigma_n^2 = R_0 + \sum_{j=1}^{n-1} R_{-j} (a_j^{(n-1)} + a_n^{(n)} \bar{a}_{n-j}^{(n-1)}) \\ + R_{-n} a_n^{(n)} \\ = \sum_{j=0}^{n-1} R_{-j} a_j^{(n-1)} + a_n^{(n)} \sum_{j=0}^{n-1} R_{j-n} \bar{a}_j^{(n-1)} \\ = \sigma_{n-1}^2 - |a_n^{(n)}|^2 \sigma_{n-1}^2 \quad (\text{A. 4})$$

($a_0^{(n)} \triangleq 1$) を得る. これが漸化式(31)である.