

## 自己回帰モデルの係数の推定方法の比較

堀 貞喜\*

国立防災科学技術センター

### Comparison of AR Coefficients Estimated from Time Series Data

By

Sadaki HORI

National Research Center for Disaster Prevention, Japan

#### Abstract

The autoregressive (AR) model is characterized by a set of coefficients which linearly combine successive signals contained in an observed time series. There are two methods to estimate the AR coefficients. The traditional method (method 1) determines the AR coefficients to minimize a *prediction error* which is calculated as a difference between the present value and the value predicted from the past observations. In the other method (method 2), which is recently proposed in the *Sompi* method for an application to spectral analyses, the AR coefficients extract autoregressive signals from the observed time series data so that they are determined to minimize an *extraction residual*. Method 1 gives an unbiased estimate of a linear relationship between the past *noisy* data and the present *noise-free* signal. On the other hand, the AR coefficients obtained by method 2 represent a linear relationship among the successive *noise-free* signals. The difference of the two methods becomes clear through three numerical experiments where the estimates are compared with the true value of the AR coefficients. First, we investigate how the contamination of Gaussian white noise makes worse the estimate of the AR coefficients. The AR coefficients obtained by method 1 are found to be quite sensitive to an increase of the noise power. In contrast, method 2 gives a good estimate of the coefficients for the data with a relatively lower S/N ratio. Then, the effect of the data length is examined for a time series containing non decaying sinusoids. When the number of data points used in calculation amounts to a certain level, the estimate of the AR coefficients obtained by method 1 becomes no longer better with an increase of the data length. But we can recognize a consistent negative correlation between the data length and the deviation of the estimate obtained by method 2 from the true value. Finally, we compare the resolvability of the two methods using a time series containing two sinusoids whose frequencies are

---

\* 第2研究部地震前兆解析研究室

very close to each other and also find the result of method 2 is much better than that of method 1. All the three experiments demonstrate how method 2 is superior to method 1 as for the estimation of AR coefficients. The difference results from the confusion in expression of the autoregressive model for a *noisy* time series. The least squares condition in method 1 implicitly assumes that the past data predicting the present signal are not contaminated by noise. As a result, the present *signal* is, in this expression, independent of the *present noise* but not independent of the *past noises* contained in the *past data*. It should be noticed that any spectral analysis methods based on the traditional autoregressive estimation possibly fail to study an accurate spectral structure of data.

## 1. はじめに

地球物理学では、地震波形や、地殻変動の日値・時間値変化など、時系列データを解析対象とすることが多い。こうした時系列データに対する数学的（パラメトリック）モデルとしては、正規性定常確率過程の見本過程と見なせる場合には、自己回帰（AR）モデル、移動平均（MA）モデル、自己回帰移動平均（ARMA）モデルが、非定常過程を含む場合、積分混合モデル（IAR, ARIMA）などがある（Box and Jenkins, 1970）。これらのモデルのうち、積分混合モデルについては、地殻変動データの予測などに適用された例がある（大内・高橋, 1981；橋爪・三雲, 1983）が、最も単純で、しかも扱い易いのは自己回帰モデルで、例えば地震波初動の自動検出（白井・徳弘, 1979；浜口・鈴木, 1981；森田・浜口, 1981；1984）、最大エントロピー法に基づくスペクトル解析（Burg, 1967；Urlych and Bishop, 1975）、そして地球自由振動の固有振動数や、 $Q$ の観測（Chao and Gilbert, 1980；堀・下平・深尾, 1984；Hori *et al.*, 1989）などに広く適用されている。

一般に、自己回帰モデルの係数を推定する場合、予測誤差を最小にする方法がとられるが、熊沢ら（1983）が提唱した存否法では、非 Toeplitz 型相関行列の最小固有値に対する固有ベ

クトルを用いて係数を推定している。両者の違いは、係数を推定するための過去のデータが、誤差を含んでいると見なすか、そうでないかという点である（大内・奥田，1986）が、その優劣について、詳細な議論はなされていない。本論文では、この2つの係数の推定方法の違いについて、数値実験を中心に解析をし、従来の推定方法の問題点を明らかにしたので報告する。

## 2. AR 係数の推定方法

自己回帰過程は、 $\{u_n; -\infty < n < \infty\}$  を、無限に続く時系列信号とすると、個々の信号間に以下の関係式が成り立つものとして定義される：

$$u_n = \sum_{j=1}^M a_j u_{n-j}. \quad (2-1)$$

ここで、 $(a_j; j=1, \dots, M)$  は、AR 係数であり、 $M$  は AR 次数である。この式では、ある時点での値が、それ以前の  $M$  個の値の線形結合で表されることを意味している。熊沢ら（1983）の存否法では（2-1）を、

$$\sum_{j=0}^M a_j u_{n-j} = 0, \quad (2-2)$$

の様に表しているが、（2-2）で  $a_0 = -1$  とおけば（2-1）と同一の式になり、また、AR 係数は絶対値にのみ意味があるため、自己回帰過程の記述という意味では、両者に差はない。

ところで、我々が実際に解析対象とするデータ  $\{x_n\}$  は有限個であり、しかも様々な雑音  $\{e_n\}$  を含んでいる：

$$x_n = u_n + e_n; n = 1, \dots, N. \quad (2-3)$$

一般に、 $\{e_n\}$  は、信号  $\{u_n\}$  とは独立であって、

$$E[e_n] = 0, \quad (2-4)$$

$$E[e_n e_m] = \delta_{nm} \sigma_e^2,$$

という性質を持つ、正規性白色雑音と仮定される。ただし、 $\delta_{nm}$  はクロネッカーのデルタ、 $\sigma_e^2$  は雑音の分散、そして  $E[\cdot]$  はアンサンブル平均を表している。

雑音を含んだ有限個の時系列データから、AR 係数を推定する場合、AR 次数  $M$  を固定して

において、予測誤差  $S_p$  を最小にする方法が一般的である：

$$S_p = \frac{1}{N - M - M + 1} \sum_{n=M+1}^N (x_n - \sum_{j=1}^M a_j x_{n-j})^2 \rightarrow \min. \quad (2-5)$$

(2-5) は、正規方程式、

$$U \vec{a} = \vec{b}, \quad (2-6)$$

ただし、

$$U = (U_{jk}) = (U_{kj}) = \left( \frac{1}{N - M - M + 1} \sum_{n=M+1}^N x_{n-j} x_{n-k} \right); \quad j, k = 1, \dots, M,$$

$$\vec{a} = (a_1, \dots, a_M)^T,$$

$$\vec{b} = \frac{1}{N - M} \left( \sum_{n=M+1}^N x_n x_{n-1}, \dots, \sum_{n=M+1}^N x_n x_{n-M} \right)^T,$$

に帰着されるが、 $N$  が十分大きいときは左辺の係数行列  $U$  の要素  $\frac{1}{N - M - M + 1} \sum_{n=M+1}^N x_{n-j} x_{n-k} / (N - M)$  は、相関関数  $\rho_{j-k} = E [x_{n-j} x_{n-k}]$  に置き換えられるとして、

$$R \vec{a} = \vec{\rho}, \quad (2-7)$$

ただし、

$$R = (R_{jk}) = (R_{kj}) = (\rho_{j-k}),$$

$$\vec{\rho} = (\rho_1, \dots, \rho_M)^T,$$

の様に表すことがある。ここで、(2-7) の係数行列  $R$  は、Toeplitz 行列になっている。(2-7) は Yule - Walker 方程式とも呼ばれ、レビンソン-ダービンアルゴリズムによって、逐次的に解くことができる。また、AR 係数  $M$  は AIC (Akaike, 1973; 1977) 等によって推定される。

ところで、上述した係数の推定方法 (以下、手法1という) は、厳密に云うと、雑音を含む時系列データの中の、完全に自己回帰的な信号部分の線形関係が得られるわけではない。

最小二乗条件 (2-5) から得られる  $\{a_j\}$  は、

$$x_n = \sum_{j=1}^M a_j x_{n-j} + e_n; \quad n=1, \dots, N, \quad (2-8)$$

という線形関係を表す係数の不偏推定値になっている。この式では、現在のデータ  $x_n$  は、雑音の現在値  $e_n$  とは独立であるが、データの過去値  $\{x_{n-j}; j=1, \dots, M\}$  に含まれる雑音とは完全に独立でない。従って、手法1は過去の（雑音を含んだ）データと、現在のデータ中に存在する信号成分の間の線形関係を推定するためのものであると言える。

これに対して、存否法における AR 係数は、（雑音を含んだ）データから自己回帰的な信号成分を除去した残差  $S_e$  を最小にする様に求められる：

$$S_e = \frac{1}{N-M-\sum_{j=1}^M 1} \left( \sum_{j=1}^M a_j x_{n-j} \right)^2 \rightarrow \min. \quad (2-9)$$

ただし、自明な解  $a_0 = \dots = a_M = 0$  以外の解を求めるために、AR 係数を成分とするベクトル  $\vec{a} = (a_0, \dots, a_M)$  の大きさは1に固定される：

$$\sum_{j=0}^M a_j^2 = 1. \quad (2-10)$$

このような条件付き極値問題は、ラグランジュの未定乗数法によって、

$$\frac{\partial}{\partial a_k} \left[ \frac{1}{N-M-\sum_{j=1}^M 1} \left( \sum_{j=1}^M a_j x_{n-j} \right)^2 - \lambda \left( \sum_{j=0}^M a_j^2 - 1 \right) \right] = 0; \quad k=0, \dots, M, \quad (2-11)$$

という方程式を解くことになり、結局、

$$\sum_{j=0}^M (P_{jk} - \lambda \delta_{jk}) a_j = 0; \quad k=0, \dots, M, \quad (2-12)$$

ただし、

$$P_{jk} = \frac{1}{N-M-\sum_{j=1}^M 1} \sum_{n=k+1}^N x_{n-j} x_{n-k}; \quad j, k=0, \dots, M,$$

という固有値問題に帰着する。行列  $\mathbf{P} = (P_{jk})$  の  $(M+1)$  個の固有値及び固有ベクトルをそれぞれ  $\lambda^{(r)}$ ,  $\vec{a}^{(r)}$  ( $r=0, \dots, M$ ) とすると、 $\vec{a}^{(r)}$  の成分を (2-9) に代入して得られる  $S_e = S_e^{(r)}$  は、(2-12) から、 $\lambda^{(r)}$  に等しくなるので、結局、 $\lambda^{(r)}$  のうちで最小の固有値  $\lambda^{(0)}$  に対する固有ベクトル  $\vec{a}^{(0)}$  が、AR 係数の推定値となる。この手法（以下、手法2という）では、従来の方法とは異なって、連続する  $(M+1)$  個のデータに含まれる信号成分の線形関係を記述する係数の不偏推定値が得られることになる。

### 3. 数値実験

偶数次の自己回帰モデルで表現される信号は、次数の半数の（増大または減衰する）正弦波の重ね合わせと等価である（Fröberg, 1969）. 今、信号  $u_n$  が、 $L$  個の正弦波によって構成されているものとする.

$$u_n = \sum_{l=1}^L A_l \exp(\alpha_l n \Delta t) \cos(\omega_l n \Delta t + \varphi_l). \quad (3-1)$$

ここで、 $A_l$ 、 $\alpha_l$ 、 $\omega_l$  及び、 $\varphi_l$  はそれぞれ、 $l$  番目の正弦波の初期振幅、増大率、角振動数及び初期位相を表しており、 $\Delta t$  はサンプリング間隔である. この時の AR 係数は、 $a_0 = 1$  として残りは  $M (= 2L)$  元連立一次方程式、

$$\begin{aligned} & \sum_{k=1}^M \sum_{l=1}^L A_l \exp[(j-k)\alpha_l \Delta t] \cos[(j-k)\omega_l \Delta t + \varphi_l] a_k \\ & = \sum_{l=1}^L A_l \exp(j\alpha_l \Delta t) \cos(j\omega_l \Delta t + \varphi_l); \quad j=1, \dots, M, \end{aligned} \quad (3-2)$$

を解くことによって厳密に求めることができる.

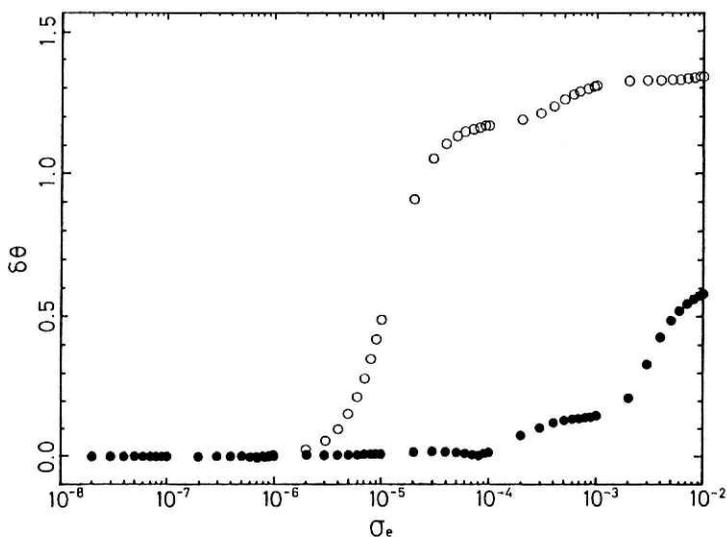
以下の数値実験では、(3-2) を解いて厳密解として得られた AR 係数と、前述した2つの手法で得られた AR 係数の推定値を比較することにする. 厳密解と推定値の差は、AR 係数を  $(M+1)$  次元ベクトルと見なしたときの両者の方向のなす角  $\delta \theta$  で表すことにする:

$$\delta \theta = \sum_{j=0}^M a_j^{(0)} / |\vec{a}^{(0)}| |\vec{a}|. \quad (3-3)$$

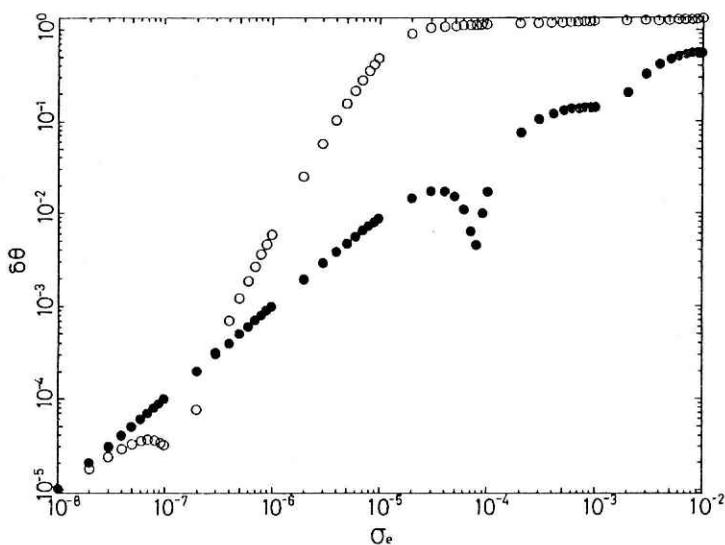
ただし、 $\vec{a}^{(0)} = (a_0^{(0)}, \dots, a_M^{(0)})$  及び、 $\vec{a} = (a_0, \dots, a_M)$  は、それぞれ厳密解より得られる AR 係数、及び、手法1または手法2から得られる AR 係数の推定値である.

#### 3・1 雑音のパワーとの関係

図1aと図1bはそれぞれ、 $\delta \theta$  と雑音のパワーの関係を直線スケールと対数スケールで示したものである. ここでは、時系列データとして2つの正弦波の重ね合わせを用いており、それぞれの初期振幅、初期位相、角振動数、増大率は、 $A_1 = A_2 = 1$ 、 $\varphi_1 = \varphi_2 = 0$ 、 $\omega_1 = -\alpha_1 = 0.1$ 、 $\omega_2 = -\alpha_2 = 0.2$  とした. また、データ長  $N$  及びサンプリング間隔  $\Delta t$  はそれぞれ1000、1とした. そして、このデータに標準偏差  $\sigma_e$  の正規性白色雑音を加え、 $\sigma_e$  を  $10^{-8}$  から  $10^{-2}$  まで、増大させていった. 図1中白丸と黒丸は、それぞれ手法1と手法2の結果である. 予想されるように、



(a)



(b)

図 1. 雑音のパワーの増大に伴う AR 係数の推定値の変化を直線スケール (a) と対数スケール (b) で示したもの. 白丸と黒丸はそれぞれ手法 1 と手法 2 の結果である.  $\sigma_e$ : 正規性白色雑音の標準偏差.  $\delta \theta$ : AR 係数の推定値及び真の値からなる 2 つのベクトルの方向の差 (ラジアン).

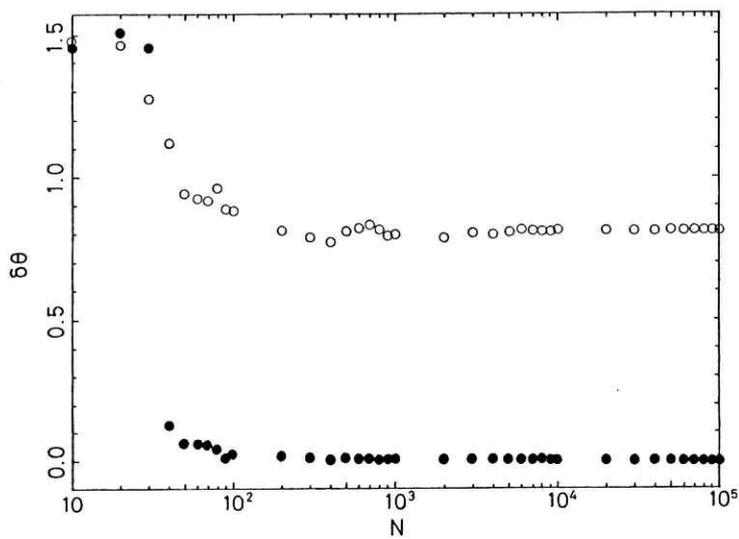
Fig. 1 Variation of the estimate of the AR coefficients with an increase of the noise power shown in the linear scale (a) and in the log scale (b). Open and filled circles represent the results of method 1 and method 2 respectively.  $\sigma_e$ : Standard deviation of Gaussian white noise.  $\delta \theta$ : Difference in directions of the estimated vector and the true vector constituted by AR coefficients (radian).

雑音のパワーが小さい場合 ( $\sigma_e < 10^{-6}$ ), 両者とも  $\delta\theta < 10^{-2}$  で良好な結果が得られている。しかし, 手法1の結果は, 直線スケールで見ると  $\sigma_e = 2 \times 10^{-6}$  から  $\delta\theta$  が急激に増大し,  $\sigma_e = 10^{-5}$  では約0.5 (約30°) にも達しており, 真の値からかなりずれている。これに対し, 手法2の結果は, 対数スケールで見ると, ほぼ  $\delta\theta$  が直線的に増大しているが,  $\sigma_e = 10^{-3}$  でも  $\delta\theta$  は約0.1 (約6°) と十分小さい。以上から, 手法1で得られるAR係数の推定値は, 手法2によるものと比べると, 雑音の混入に対してはるかに敏感であることが分かる。

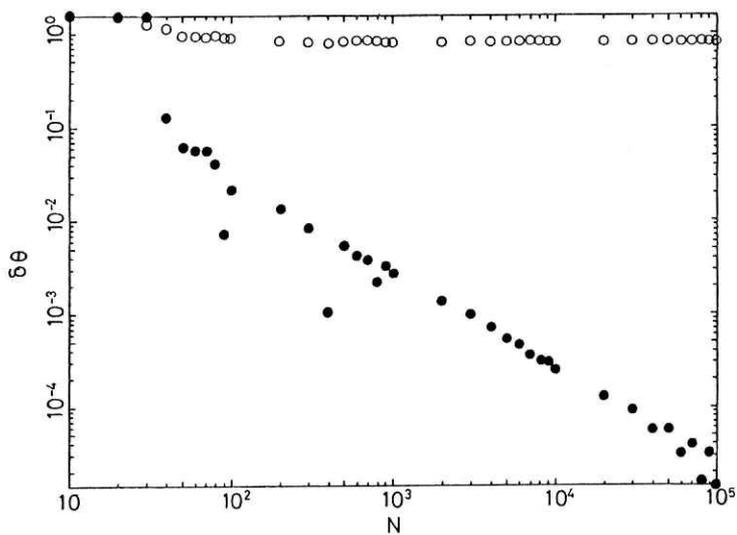
### 3. 2 データ長との関係

次に, 雑音のパワーを固定しておいて, データ長  $N$  を変えたときのAR係数の推定値の変化を調べる。図2aと図2bは  $N$  を  $10^1$  から  $10^5$  まで増大させたときの  $\delta\theta$  の値を, それぞれ直線スケールと対数スケールで示したものである。ここで用いたデータも2つの正弦波と正規性白色雑音の重ね合わせであるが, データ長の影響だけを見るため, 増大率は2つとも0にして, 全データを通してS/N比を一定にした。振幅, 位相など他のパラメータは3.1と同一である。また雑音の標準偏差  $\sigma_e$  は  $10^{-4}$  に固定した。

データ長が有限の場合, 雑音は程度の差こそあれ, 「有色」となる。しかし, データ長が大きくなれば(2-4)で表される「正規性」と「白色性」の傾向が強くなっていくと考えられる。従って, 減衰しない信号ならば, データを長くすればとるほど雑音と分離されてAR係数の推定値については良好な結果が得られるはずである。ところが図2をみると, 手法1ではデータ長が200ポイントになるまでは  $\delta\theta$  は減少していくが, それ以上はいくらデータを増やしていても  $\delta\theta$  はほぼ一定で, よい結果は得られてこない。これに対して手法2では, 直線スケールで見ると, データ長を100ポイントまで増やした時点で  $\delta\theta$  の値はほぼ0となって, 真の値に近いAR係数を推定していることが分かる。また, 対数スケールで見ると, 手法1の結果は, データ長を大きくしても全く向上していないのに対して, 手法2の結果では, 多少のばらつきはあるものの, データ長を増やした分, 結果が向上するという,  $N$  と  $\delta\theta$  の間の一貫した負の相関が認められる。



(a)



(b)

図 2. データ長の増大に伴う AR 係数の推定値の変化.  $N$ : 計算に用いたデータのポイント数.  
 Fig. 2 Variation of the estimate of the AR coefficients with an increase of the number of data points shown in the linear scale (a) and in the log scale. Open and filled circles represent the results of method 1 and method 2 respectively.  $N$ : Number of data points used in calculation.

### 3.3 分解能

最後の雑音パワーとデータ長を一定にしたときの分解能について、手法1と手法2を比較してみた。時系列データとしては、2つの正弦波と正規性白色雑音を合成したものを用いた。振幅、位相、増大率は2つとも同一で、それぞれ、1, 0, 0とした。また、データ長は10000ポイント、サンプリング間隔は1、雑音の標準偏差は $10^{-3}$ にした。そして、角振動数は $\omega_1 = 0.25/2\pi$  (ナイキスト周波数の1/2),  $\omega_2 = \omega_1 + \Delta\omega/2\pi$ として、 $\Delta\omega$ と $\delta\theta$ の関係を調べた。その結果を図3に示す。周波数の接近した2つの正弦波を合成すると、「うなり」が発生し、その周期は、2つの周波数の差の逆数になることが知られているが、今回の実験に用いたデータ長は、 $\Delta\omega/2\pi = 10^{-4}$ でちょうどうなり1周期分に相当することになる。図3を見るとここでも手法2の方が手法1よりも優れていることが明らかになっているが、特に、手法2の方は $\Delta\omega/2\pi = 10^{-5}$ 、つまり、うなりの周期が全データ長の10倍にも達する時でさえ、 $\delta\theta$ が $10^{-2}$ まで減少しているのは注目に値する。

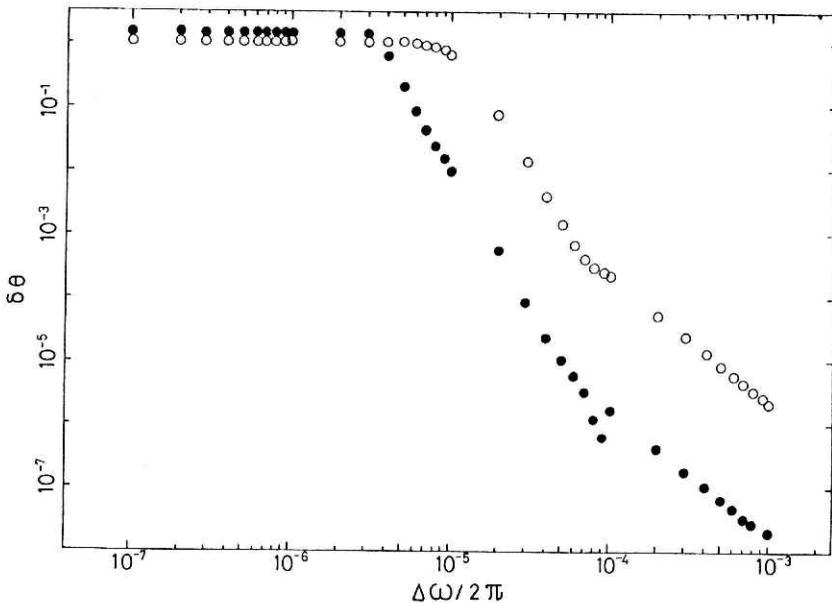


図3. 2つの正弦波の振動数の差の増大に伴うAR係数の推定値の変化。 $\Delta\omega$ ：角振動数の差(サイクル毎サンプリング間隔)

Fig. 3 Variation of the estimate of the AR coefficients with an increase of the frequency difference of the two sinusoids. Open and filled circles represent the results of method 1 and method 2 respectively.  $\Delta\omega$ : Frequency difference between two wave elements (cycle per one sampling interval).

#### 4. 議論と結論

今回行った単純な3つの数値実験では、AR係数を時系列データから推定する場合、いずれも、従来の予測誤差を最小にする方法より、信号を除去した残差を最小にする存否法の推定方法の方が優れていることが分かった。従来の方法が、雑音の混入に対して敏感で、データ長を大きくとることの利点が生かされにくいのは、データ中に含まれる信号成分と雑音を完全に独立なものとして扱っていないためであることは明らかである。つまり、自己回帰モデルを記述する場合、本来分けて記述すべき(2-1)ないし(2-2)と(2-3)を、暗黙のうちに1つにまとめて(2-8)で表現することの、本質的な問題点が今回の研究で判明した。

近年、自己回帰モデルはスペクトル解析法にも適用されるようになってきて、例えば Chao and Gilbert (1980) は、地球の自由振動の固有周波数を複素領域で求める手法を提唱している。同じように自己回帰モデルを自由振動の解析に適用する手法は、例えば、存否法を応用したものとして堀・下平・深尾(1984)や Hori *et al.* (1989) も考案している。地球自由振動の場合、そのスペクトル構造は比較的単純で、自己回帰モデルを適用するのにふさわしいデータであるといえる。しかし、Chao and Gilbert (1980) の AR 係数の推定方法は基本的には手法1に基づいており、地震記象のような、かなり雑音を含んだ時系列データに対する効果には疑問が持たれる。さらに、自由振動を解析する場合、時には、1週間を超えるようなデータを用いる場合があるが、このようなデータに対しては、長いデータほど雑音と信号の分離がうまくいく手法2に基づくスペクトル解析法、つまり存否法の方がより精度の高い結果が得られると考えられる。

今回明らかにした、従来の AR 係数の推定方法における問題点は、予測フィルターや、ARMA モデルの推定にも関係があると思われるが、それについては今後さらに検討すべきであろう。

#### 参 考 文 献

- 1) Akaike, H. (1973): Information theory and an extension of the maximum likelihood principle. *2nd International Symposium on Information Theory*, eds. B. N. Petrov and F. Csaki, pp. 267-281. Akademia Kiado, Budapest.
- 2) Akaike, H. (1977): On entropy maximization principle. *Proceedings of Symposium on Amsterdam*, Norm-Holland, 27-41.
- 3) Box, G. E. and Jenkins, G. M. (1970): Time series analysis. *Forecasting and Control*, Holden Day, San Francisco.
- 4) Burg, J. P. (1967): Maximum entropy spectral analysis. *Paper presented at the 37th Annual International Meeting*, Soc. Explor. Geophysics, Oklahoma City Okla., Oct. 31.

- 5) Chao, B. F. and Gilbert, F. (1980) : Autoregressive estimation of complex eigenfrequencies in low frequency seismic spectra. *Geophys. J. R. Astron. Soc.*, **63**, 641-657.
- 6) Fröberg, C. E. (1969) : Introduction to Numerical Analysis, 2nd edn., pp. 347-349, Addison-Wesley, Reading, Massachusetts.
- 7) 浜口博之・鈴木次郎 (1979) : 微小地震 *P* 波初動の自動検測とその信頼性について. 自然災害特別研究成果, **A-54-2**, 62-83.
- 8) 橋爪道郎・三雲 健 (1983) : 地殻変動観測データの多変量解析—ARIMA モデルによる予測の試み—. 地震 **2**, **36**, 515-530.
- 9) 堀 貞喜・下平 慶・深尾良夫 (1984) : 存否法によるマルチプレットの解析. 地震学会講演予稿集, 1984, No 1, B59.
- 10) Hori, S., Fukao, Y., Kumazawa, M., Furumoto, M., Yamamoto, A. (1989) : A new method of spectral analysis and its application to the Earth's free oscillations : the *Sompi* method. *J. Geophys. Res.*, in press.
- 11) 熊沢峰夫・古本宗充・深尾良夫・山本明彦・水谷 仁 (1983) : 存否スペクトル法—原理と特長, 地震学会講演予稿集, 1983, No. 1, B33.
- 12) 森田裕一・浜口博之 (1981) : 2次元自己回帰過程による *S* 波初動の自動検測. 地震 **2**, **34**, 223-240.
- 13) 森田裕一・浜口博之 (1984) : 自己回帰過程による地震波初動の自動検測とその信頼区間. 地震 **2**, **37**, 281-293.
- 14) 大内 徹・奥田 暁 (1986) : 地震観測データの統計解析 (1) —最小二乗法と主成分分析法—. 地震学会講演予稿集, 1986, No. 2, A95.
- 15) 大内 徹・高橋 享 (1981) : 地震学で扱う非定常時系列データの解析—予測と異常検出の試み—. 地震 **2**, **34**, 455-464.
- 16) 白井克彦・徳弘一路 (1979) : 地震波初動の自動検出. 地震 **2**, **32**, 141-148.
- 17) Urlych, T. J. and Bishop, T. N. (1975) : Maximum entropy spectral analysis and autoregressive decomposition. *Rev. Geophys. Space. Phys.*, **13**, 183-200.

(1988年6月20日原稿受理)