

時変係数自己回帰モデル

—— プログラム TVCAR の紹介 ——

統計数理研究所 北 川 源 四 郎

(1986年9月 受付)

1. はじめに

この解説の目的は、時系列解析のためのプログラムパッケージ TIMSAC-84 (Akaike et al., 1985) 中のプログラム TVCAR の原理を簡単に説明し、実例を用いながらその使用法や出力の意味などを示すことにある。

TVCAR は共分散に関し非定常な時系列に対して時間とともに係数が変化する自己回帰モデルをあてはめる。そのために、Kitagawa (1983) で用いられたスペクトル構造の時間変化に関するモデルと Kitagawa-Gersch (1985b) で用いたスペクトルの滑らかさに関するモデルの2つを制約モデルとして採用している。TVCAR では、これらのモデルが状態空間表現により、統一的に記述できることを利用して、効率的な時変係数の推定および尤度計算を実現している。モデル中の未知のパラメータは、非線形最適化を用いて最尤法により求めている。

第2節では TVCAR のモデルおよび推定法を簡単に説明する。第3節では TVCAR の入力の方法と出力の見方を示し、いくつかの注意を与えている。

第2節のモデルおよび計算法は、Kitagawa (1983) および Kitagawa-Gersch (1985a, b) と極めて関連が深い。スペクトルの滑らかさと時間的変化の両方にベイズ的モデルを導入した TVCAR のモデルを紹介した文献は本稿が初めてである。なお、TIMSAC-84 中の LOCCAR および NONSPA も非定常スペクトルの推定を行なうもので TVCAR と関連が深い。LOCCAR の原理は Akaike (1979) に示されている。

2. TVCAR のモデル

2.1 時変分散モデル

$r(n)$ は正規白色雑音でその分散は時間とともに変化するものとする。すなわち、

$$r(n) \sim N(0, \sigma^2(n))$$

とし、簡単のために、 $\sigma^2(2m) = \sigma^2(2m-1)$, $m=1, 2, \dots$ がなりたつものとする。ここで、

$$(1) \quad s(m) = \frac{1}{2}(r(2m)^2 + r(2m-1)^2)$$

という変数を考えると、

$$s(m) \sim \frac{\sigma^2(2m)}{2} \cdot \chi_2^2$$

という形をしていることがわかる。これより、さらに対数変換

$$(2) \quad t(m) = \log s(m)$$

を適用すると

$$(3) \quad t(m) \sim \log \frac{1}{2} \chi^2 + \log \sigma^2(2m)$$

となり、 $\log \sigma^2(2m)$ とそのまわりの確率的な変動という形に分解できる。ここで $\log \frac{1}{2} \chi^2$ の密度関数は $\exp(t - e^t)$ で与えられるが、その平均および分散は、それぞれ $\gamma - \log 2$, $\pi^2/6$ である。(ただし、 γ はオイラー定数、0.57721) 以下では、この分布を正規近似することにする。

ここで、 $\sigma^2(m)$ の推定のために、次のようなモデルを導入することにする。

$$(4) \quad \begin{aligned} \nabla^k u(m) &= \varepsilon(m) \\ t(m) &= u(m) + \delta(m) \end{aligned}$$

ただし、 $\varepsilon(m) \sim N(0, \nu^2)$, $\delta(m) \sim N(0, \pi^2/6)$ 。このモデルは容易に状態空間表現の形に変形できる。したがって、カルマンフィルターおよび固定区間 smoother を用いて $t(m)$ のトレンド、 $\hat{\mu}(m)$ を推定できる。このとき $\hat{u}(m) + \gamma - \log 2$ が $\log \sigma^2(2m)$ の、したがって $\exp\{\hat{u}(m) + \gamma - \log 2\}$ が時変分散 $\sigma^2(2m)$ の推定値となる。

2.2 時変係数 AR モデル

$y(n)$ に対し AR 係数が時間的に変化するモデル

$$(5) \quad y(n) = \sum_{i=1}^m a(i, n) y(n-i) + w(n)$$

を考える。ただし、 $w(n)$ は平均 0、分散 $\sigma^2(n)$ の正規白色雑音とする。

ここで、この AR モデルに対して時変 (逆) 周波数応答関数

$$(6) \quad A(f, n) = 1 - \sum_{j=1}^m a(j, n) \exp(-2\pi i j f) \quad -\frac{1}{2} \leq f \leq \frac{1}{2}$$

を考えることにする。このとき、 $y(n)$ の瞬間的なスペクトルが

$$(7) \quad p(f, n) = \frac{\sigma^2(n)}{|A(f, n)|^2} \quad -\frac{1}{2} \leq f \leq \frac{1}{2}$$

によって定義できる。また、 $p(f, n)$ を時刻 n の関数と考えると時変スペクトルが得られる。

TVCAR では時変 AR 係数 $a(j, n)$ の推定のために、 $A(f, n)$ すなわちスペクトルの滑らかさに関連した 2 種類の制約モデルを導入している。

スペクトルの時間的変化

$A(f, n)$ の k 階差分の 2 乗積分値を考えると

$$(8) \quad \int_{-\frac{1}{2}}^{\frac{1}{2}} |\nabla^k A(f, n)|^2 df = \sum_{j=1}^m |\nabla^k a(j, n)|^2$$

となる。したがって、AR 係数の k 階差分の 2 乗和を小さくすることにより、(左辺の意味で) スペクトルの時間変化を小さくすることができる。これは

$$(9) \quad \begin{aligned} \nabla_n^k a(j, n) &= v(j, n) \quad (j=1, \dots, m) \\ v(j, n) &\sim N(0, \tau^2) \end{aligned}$$

というモデルを用いることに相当している。モデル (9) は Kitagawa (1983) では AR 係数の時間的変化を小さくするという目的で導入されたものである。

スペクトルの滑らかさ

各時刻 n に対しては Kitagawa-Gersch (1985b) で用いたスペクトルの滑らかさの制約を用いることにする。すなわち

$$(10) \quad \begin{aligned} R_k &= \int_{-\frac{1}{2}}^{\frac{1}{2}} \left| \frac{\partial^k A(f, n)}{\partial f^k} \right|^2 df \\ &= (2\pi)^{2k} \sum_{j=1}^m j^{2k} a(j, n)^2 \end{aligned}$$

は周波数応答関数の k 階微分の意味で、スペクトルの滑らかさを表わしているの、右辺の項を小さくすることにより、スペクトルの滑らかな推定値が得られる。これは

$$\begin{aligned} a(j, n) &= u_k(j, n) \quad j=1, \dots, m \\ u_k(j, n) &\sim N(0, (j^k \lambda_k)^{-2}) \end{aligned}$$

という制約モデルを導入することにより実現できる。 k としては $0, 1, 2, \dots$ などが考えられるが、TVCAR では $k=0$ と 2 の二つを同時に考えて

$$(11) \quad \begin{aligned} a(j, n) &= u(j, n) \\ u(j, n) &\sim N(0, (\lambda_0^2 + j^4 \lambda_2^2)^{-1}) \end{aligned}$$

というモデルを用いている。

時変係数 AR モデル (5) と 2 つの制約モデル (9), (11) は状態空間表現の形で表現できる。簡単のために以下では $k=1$ の場合だけを示す。

$$\begin{aligned} \begin{bmatrix} a(1, n) \\ \vdots \\ a(m, n) \end{bmatrix} &= \begin{bmatrix} I_m \end{bmatrix} \begin{bmatrix} a(1, n-1) \\ \vdots \\ a(m, n-1) \end{bmatrix} + \begin{bmatrix} v(1, n) \\ \vdots \\ v(m, n) \end{bmatrix} \\ \begin{bmatrix} y(n) \\ 0 \\ \vdots \\ 0 \end{bmatrix} &= \begin{bmatrix} y(n-1) \cdots y(n-m) \\ 1 \cdots \cdots \cdots \cdots 1 \end{bmatrix} \begin{bmatrix} a(1, n) \\ \vdots \\ a(m, n) \end{bmatrix} + \begin{bmatrix} w(n) \\ u(1, n) \\ \vdots \\ u(m, n) \end{bmatrix} \end{aligned}$$

ただし

$$\begin{aligned} \begin{pmatrix} v(1, n) \\ \vdots \\ v(m, n) \end{pmatrix} &\sim N \left(\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} \tau^2 & & \\ & \ddots & \\ & & \tau^2 \end{pmatrix} \right) \\ \begin{pmatrix} w(n) \\ u(1, n) \\ \vdots \\ u(m, n) \end{pmatrix} &\sim N \left(\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & & & \\ \hline & (\lambda_0^2 + \lambda_2^2)^{-1} & & \\ & & \ddots & \\ & & & (\lambda_0^2 + m^4 \lambda_2^2)^{-1} \end{pmatrix} \right) \end{aligned}$$

である。このように、 m 次の状態ベクトル、 $(m+1)$ 次の観測ベクトルを用いて、3つのモデルが表現できる。したがって、カルマンフィルター、固定区間スモザーを用いて、 $a(j, n)$ ($j=1, \dots, m; n=1, \dots, N$) が推定できる。状態空間表現に含まれるパラメータ、 $\tau^2, \lambda_0^2, \lambda_2^2$ はカルマンフィルターで求められる尤度の値を最大化することによって定められる。

2.3 TVCAR の推定手続き

TVCAR では次の順にモデル推定を行なう。

- (1) 原データの変動分散を2.1の方法で推定する。この平方根をプログラムでは ENVELOPE FUNCTION と呼んでいる。原データを ENVELOPE FUNCTION で割ることにより、NORMALIZED DATA が得られる。このデータは近似的に分散1の分散定常時系列と見なせる。
- (2) NORMALIZED DATA に2.2の方法で変動係数 AR モデルをあてはめる。
- (3) 変動係数 AR モデルの残差系列にさらに2.1を適用して、その変動分散を推定する。

3. TVCAR の使用法

3.1 TVCAR への入力

TVCAR を実行させるためには、NAMELIST 文による制御パラメータと解析の対象となる原データの二種類の入力が必要である。

〈制御パラメータの入力〉

プログラムの制御パラメータとしての次のようなものがある。制御パラメータには標準値が設定されているので、変更する必要があるものだけを NAMELIST 文を用いて入力すればよい。

| 制御パラメータ | 標準値 | 説明 |
|---------|-----------|---|
| MT | 10 | 原データの入力機番 =5: カードリーダー |
| MAR | 8 | AR 次数 |
| KAR | 2 | AR 係数の変化を表わす差分モデルの階差 |
| NOBS | 10 | 何ステップ毎に係数を変化させるか |
| IDIF | 1 | Gradient 推定のための数値差分 =1 片側差分 =2 両側差分 |
| MESH | 1 | グラフの座標上に mesh を描く =0 描かない |
| NOUT | 0 | システムの急激な変化の回数 |
| TAU 1 | 10^{-4} | τ_1^2 |
| TAU 2 | 0.3 | τ_2^2 |
| TAU 3 | 10^{-5} | τ_3^2 |

ただし、NOUT に正の値を入れたときは、NAMELIST 文のあとに、急激な変化がおこった場所を示す NOUT 個の整数をフリーフォーマットで入力しなければならない。

後で示す計算例の場合には

△&PARAM △ NOUT=2, &END
635△1030

の 2 行を入力した。ただし △ は空白を表わす。

〈原データの入力〉

原データは機番 10 の入力装置 (NAMELIST 文により変更された場合はその機番) から入力される。原データは DECOMP と同様次のフォーマットに従っているものと仮定する。

| レコード番号 | 内 容 | フォーマット |
|--------|---------------------|----------------|
| 1 | データのタイトル | (20A4) |
| 2 | データ数 | (I5) |
| 3 | データのフォーマット | (20A4) |
| 4 以降 | データ Y(1), ..., Y(N) | 3 で指定されたフォーマット |

以下の計算例では次の様なデータが用いられた。

```
EARTHQUAKE DATA: MYE1F
2600
(20F4.0)
-1 -1 -1 -1 0 -1 0 -1 -1 0 -1 -1 0 -1 -1 -2 -2 -2 -1
-2 -2 -2 -1 -1 -2 -1 -1 -2 -1 -1 -3 -1 -1 -1 -1 -1 1 1
 0  1  0 -1 -1 -1  0  0 -2 -2 -3 -1 -2 -2 -3 -3 -2 -1 -2
-3 -1  1 -1 -1 -2  1  0  1  0 -3 -1  2  0 -1 -1  0  3  0
 2  0  0  0  1 -1  0  1  0  0  0 -2  0  0 -1 -3  0 -1 -4
-2 -3 -4 -2 -4 -5 -4 -2 -2 -4 -7 -3 -3 -3 -4 -5 -6 -1 -2
-5 -4  0  0 -2 -2  0  0  0  1 -1  0  3  2  2  1 -3 -3 -2
 1  1  2  2  0  0 -1  0  0  1 -1 -1 -2 -1 -2 -3 -3 -3 -3
-5 -4 -3 -3 -2 -3 -4 -3 -1 -1 -3 -1  0 -1  0 -1 -1  1  0
 0  1  0  0  1  0  2  0  1  0  0  1  2  2  0  0  1  2  0 -1
```

以下略

3.2 TVCAR の出力の説明

TVCAR の出力には数値と図がある。数値はラインプリンタ(又は端末)、図はレーザープリンタ(又はプロッタ)から出力される。

〈数値出力〉

TVCAR が実行され、NAMELIST 文が入力されると、まず次の様な出力が得られる。

```
--- PROGRAM TVCAR ---
MT = 8
IDIF = 1
TAU1 = 0.0001000
TAU2 = 0.3000000
TAU3 = 0.0000100
NOBS = 10
NOUT = 2
MESH = 1

*** KOUT(I), I=1,...,NOUT ? ***
```

これによって NAMELIST 文および標準値によって設定されたパラメータを確認できる。NOUT に正の値を設定した場合には、ここでその個数だけ構造変化の場所 KOUT(i) を入力する必要がある。

次に、機番 MT から原データが入力されると次のような出力が得られる。

```
<< ORIGINAL DATA X(I) (I=1,N) >> N = 2600 FORMAT = (20F4.0)
MYE1F
```

これにより、入力されたデータのデータ長、入力フォーマット、タイトル等が確認できる。次にサブルーチン EPARA 2 が非線形最適化法により τ_1^2 , τ_2^2 および τ_3^2 の最尤推定値を求める。この過程で出力される情報は次のようなものである。

PARAMETER: 非線形最適化中のパラメータベクトル。ただし、適当な変数変換がほどこされている。
 GRADIENT : 関数値のパラメータに関する偏微分
 RAM : 線形探索のステップ幅
 E : そのときの関数値

最尤推定値が求まると改ページされて、次のように、データ名、プログラム名、モデル、計算日時、初期推定値、最尤推定値がまとめられた出力が得られる。

```

--- DATA ---
MYE1F                                     N = 2600

--- MODEL ---
MAR = 8      KAR = 2

--- PROGRAM TVAR ---
DATE: 85-03-04      TIME: 14 48 06

<<< INITIAL ESTIMATES >>>
-- PARAMETER VECTOR --
-0.921034D+01  0.120397D+01  0.115129D+02
-- GRADIENT --
0.220927D+00  0.330481D+02 -0.898171D-03
LIKELIHOOD = -2727.768
AIC = 5463.536

I      TAU2(I)
1  0.10000000D-03
2  0.30000000D+00
3  0.10000000D-04

<<< FINAL ESTIMATES >>>
-- PARAMETER VECTOR --
-0.114116D+02  0.937851D+00  0.115162D+02
-- GRADIENT --
0.226783D-02  0.828236D-02 -0.315199D-03
LIKELIHOOD = -2723.253
AIC = 5454.506

I      TAU2(I)
1  0.11066335D-04
2  0.39146833D+00
3  0.99674242D-05

```

〈図による出力〉

図 1-1 は原データである。この地震データの場合には $n=630, 1030$ の付近で急激に分散が増大し、その後、徐々にもとの分散にもどっていく様子が見られる。急激な分散の増大は P 波と S 波の到着に対応している。

図 1-2 は (1), (2) により変換したデータと分散の対数の推定値である。(2) の変換により分散の増大は平均値の上昇に対応していることがわかる。

図 1-3 は推定された時変分散により規格化されたデータである。この規格化により、非定常時系列がほぼ分散一様の時系列に変換されていることがわかる。時変係数 AR モデルはこのデータから推定される。

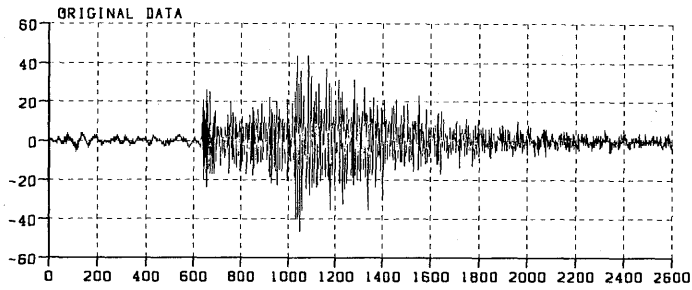


図 1-1. 地震波データ

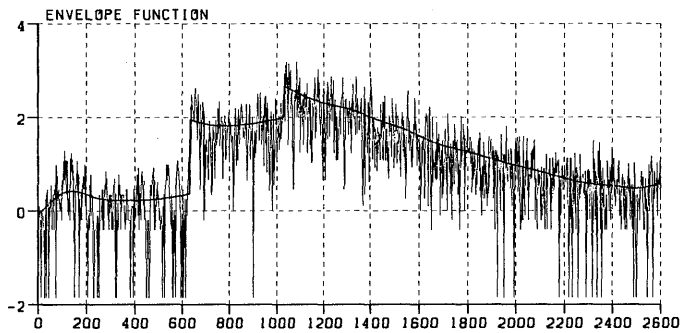


図 1-2. 変換したデータの分散の対数

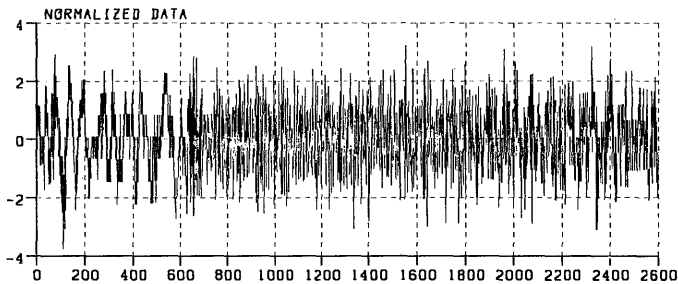


図 1-3. 規格化したデータ

図 2 は推定された偏自己回帰係数の時間的変化を図示したものである。P 波および S 波が到着した $n=630, 1030$ の付近では、偏自己回帰係数が急激に変化し、変動の分散だけでなく共分散構造自体も大きく変化したことがわかる。また $n=2600$ では、ほぼ P 波到着前の状態にもどっていることがわかる。

図 3 は、(7) によって求めた時変スペクトルである。P 波および S 波の到着による分散の変化、スペクトルの変化がよくわかる。とくに、S 波到着後、二つのスペクトルのピークが高周波側へ移動しながら、徐々に常時微動のスペクトルにもどっていく様子がよくわかる。このような、周波数による到着時刻の変化は表面波については知られているが、S 波に関してはよく判っていない。

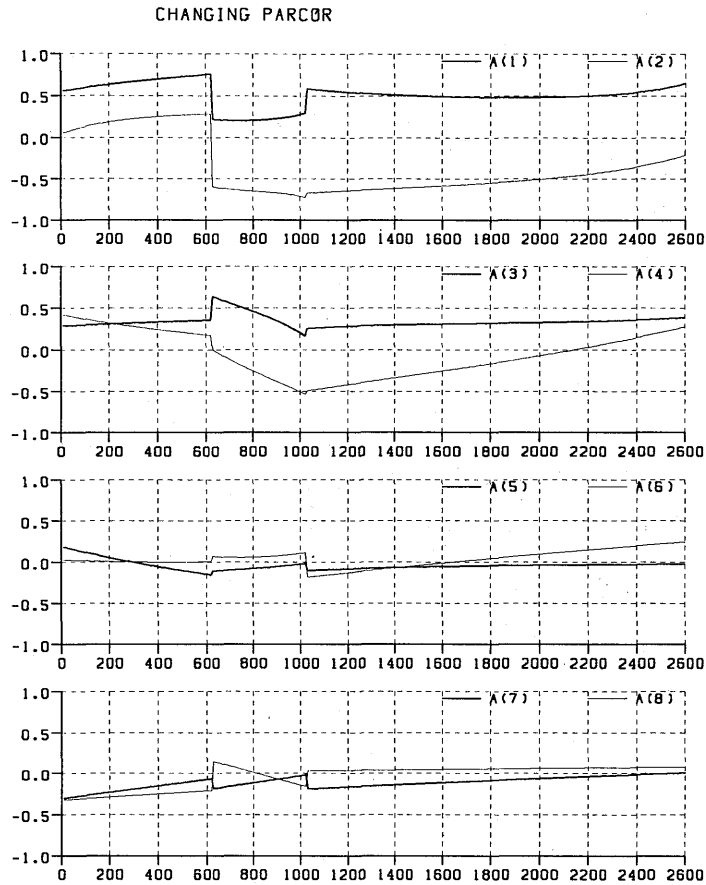


図2. 偏自己回帰係数の時間変化

3.3 パラメータの急激な変化

TVCAR で用いたモデルは本来、パラメータのゆっくりした変化に追従するためのものである。したがって、本稿で示した地震波の様に急激な変化をともなう場合には特別な取扱いが必要である。TVCAR では、この様な場合を想定して、システム雑音に異常値の存在を許している。この異常値により、前後のパラメータ間の制約条件は無視されることになり、したがって、パラメータの急激な変化が可能となる。この異常値の処理を全く行わずに推定した図4と図3を比較するとこの効果がはっきりと判る。ただし、現在のプログラムでは、異常値の数およびその場所はユーザーが指定しなければならない。異常値の場所が簡単には判らない場合には、原データおよび残差系列の異常値の自動検出を行なってみるのも一つの方法である。浪花(1986)は異常値の個数と場所に対応する多くのモデルを求め AIC によって、その場所を特定することを試みている。

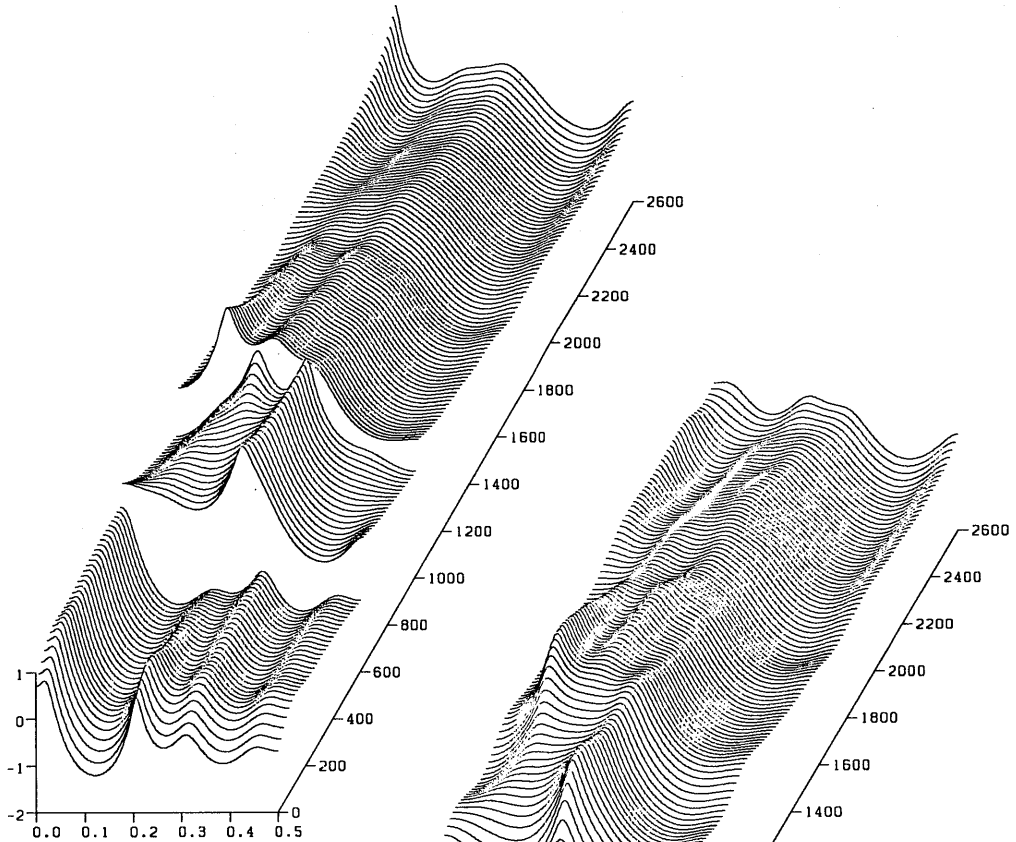


図3. 時変スペクトル

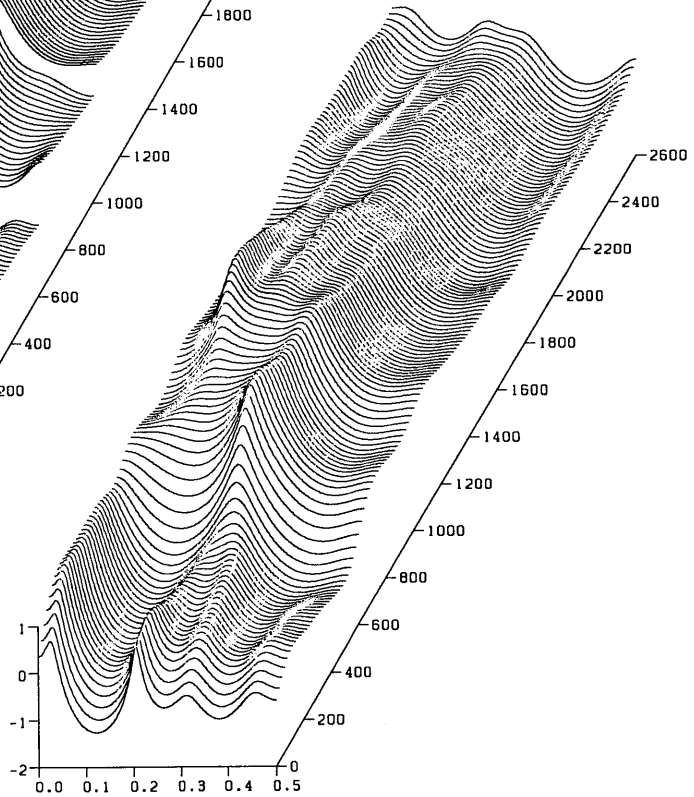


図4. 異常値処理しないで求めた時変スペクトル

謝 辞

地震波のデータを提供していただいた北海道大学理学部の高波鉄夫氏に感謝いたします。

参 考 文 献

- Akaike, H. (1979). On the Construction of Composite Time Series Models, *Proc. 42nd Session Int. Statist. Inst.*
- Akaike, H. et al. (1985). TIMSAC-84 Part 1, *Computer Science Monographs*, **22**, The Institute of

Statistical Mathematics, Tokyo.

Kitagawa, G. (1983). Changing Spectrum Estimation, *Journal of Sound and Vibration*, **89**, No. 3, 433-445.

Kitagawa, G. and Gersch, W. (1985a). A Smoothness Priors Time-Varying AR Coefficient Modeling of Nonstationary Covariance Time Series, *IEEE Trans. A.C.*, **30**, No. 1, 48-56.

Kitagawa, G. and Gersch, W. (1985b). A Smoothness Priors Long AR Model Method for Spectral Estimation, *IEEE Trans. A.C.*, **30**, No. 1, 57-65.

浪花貞夫 (1986). トレンドを除去した経済時系列の非定常性について——構造的変化の統計的検討——, *金融研究*, **5**, 第4号, 97-139.

Time-Varying Autoregressive Model
— An Introduction to the Program TVCAR —

Genshiro Kitagawa

(The Institute of Statistical Mathematics)

A procedure for the identification and spectral estimation of nonstationary time series is described here. The basic model is an autoregressive model with time-varying coefficients. Two types of constraint models are introduced to get reasonable estimates of the coefficients. The first is the model for the time evolution of the coefficients and the second for the smoothness of the spectrum at each time. The basic model and these two constraint models are incorporated into a state space model form. Computationaly efficient Kalman filter and the fixed interval smoother are used to obtain the time-varying coefficients and the likelihood of the model. The values of the parameters which control the smoothness of the estimated spectrum are determined so as to maximize the likelihood of the model. Time varying variance of the process is also estimated by a similar state space model.

The usage of the computer program TVCAR which realizes the procedure is exemplified by explaining the sample inputs and the sample outputs for an actual seismic data. Some remarks on the abrupt changes of the parameters are also given.