Ann. Inst. Statist. Math.
39 (1987), Part B, 227-241

AN APPROACH TO THE NONSTATIONARY PROCESS ANALYSIS

YOSHIYASU-HAMADA TAMURA

(Received May 22, 1985; revised July 30, 1986)

Summary

A Bayesian approach to nonstationary process analysis is proposed. Given a set of data, it is divided into several blocks with the same length, and in each block an autoregressive model is fitted to the data. A constraint on the autoregressive coefficients of the successive blocks is considered. This constraint controls the smoothness of the temporal change of spectrum as shown in Section 2. A smoothness parameter, which is called a hyper parameter in this article, is determined with the aid of the minimum ABIC (Akaike Bayesian Information Criterion) procedure. Numerical examples of our procedure are also given.

1. Introduction

The purpose of this paper is to consider the procedure to estimate the power spectral density of nonstationary process. We turn our attention to processes whose statistical properties make gradual or abrupt change such as seismographic records, records of atmospheric turbulence, records of brain wave and time-seriesed clinical records.

Many authors have developed their procedures to deal with such nonstationary processes. Page [11], Priestley [12], [13], Bendat and Piersol [4], Mark [9] and Hino [5] approached this problem by defining their own nonstationary power spectral densities with the aid of the frequency domain time series analysis. It seems to us that their procedures are not practical. For instance, Page's instantaneous spectrum may take negative values for certain processes and so his spectrum is not relevant from the point of view of physical interpretation. Bendat and Piersol's generalized spectrum depends on the behavior of time series over all time and cannot give us the spectral properties of the process at specific time t. Priestley did not give the criterion to select the weight function which is contained in his definition of evolutionary spectrum, though estimated spectrum depends on the weight function.

Key words and phrases: Nonstationary process, Bayesian model, ABIC.

Physical spectrum and developing spectrum defined by Mark and Hino contain the ensemble average in their definitions. In the case only one observed series is given, we cannot apply their spectra without replacing the ensembles average by a time average. These are why their procedures seem to be unpractical.

Other approaches to nonstationary processes were made with the aid of autoregressive type models by Rao [14], Ozaki and Tong [10], Kitagawa and Akaike [6], [7], Akaike [1] and Kitagawa [8].

Ozaki and Tong [10] considered a locally stationary process. Timeseriesed data are divided into several blocks and in each block a stationary autoregressive model is fitted to data. They determined the partition of the data into blocks by the minimum AIC (Akaike Information Criterion) procedure. Kitagawa and Akaike [6] adopted the same approach. They applied the least squares method via Householder transformation for fitting of the autoregressive model of each block. On the other hand Ozaki and Tong solved Yule-Walker equation. It seems to us that Kitagawa and Akaike's procedure is more powerful and more manipulable than Ozaki and Tong's. In these two procedures information which is contained in the previous blocks cannot be used so effectively. A block and its preceding block are considered simultaneously to check homogeneity of data in these procedures. We have to say reverently that these approaches waste information about the temporal development of process. Even if parameters are assumed to be constant in a subinterval, they had better be estimated by using not only data contained in the subinterval in which these parameters are defined but also data contained in the former subintervals. If the system changes abruptly at a point, the point should be found automatically the aid of the appropriate statistical model itself. It seems not to be natural that autoregressive coefficients change suddenly at discrete points which are determined artificially.

Kitagawa and Akaike [7] proposed a Bayesian procedure for the fitting of a locally stationary model. In their procedure information of the previous blocks are utilized effectively in estimating autoregressive coefficients of the present block.

Rao [14] expanded a method of weighted least square to estimate autoregressive coefficients which change with time. We are sorry to say that his ideas is not so bad but his method cannot be adapted to the practical problems for lack of the rule to choose weights.

Kitagawa [8] took approach to nonstationary processes by fitting an autoregressive type model with time-varying coefficients to those processes. He considered a stochastically perturbed difference equation constraint. The k-th order difference of the successive autoregressive coefficients behaves like a random walk in his model. The order of the autoregression and the order of linear difference coefficient are determined by the minimum AIC procedure. His procedure is useful and computationally efficient. It, however, seems to us that for the practical usage autoregressive coefficients may be kept constant in a block (subinterval), supposing that we make use of information contained in data of the preceding blocks as effectively as we can in order to estimate parameters of autoregressive model of the present block and that the number of data contained in each block is put to be as small as possible. Certainly it seems more natural to consider that the parameters change at each point than to consider that they are constant in each subinterval. But in order to cut down computing time, we assume parameters are constant in each subinterval in our model of this paper to our regret.

We propose a new model to analyse nonstationary process. In our model a set of data is divided into some blocks with the same length. We put a constraint on the autoregressive coefficients of the successive blocks. Our constraint is concerned with the smoothness of the temporal change of power spectrum as shown in the next section. The smoothness parameters, which is called a hyper parameter, is determined by the minimum ABIC (Akaike Bayesian Infomation Criterion) procedure which is attributed to Akaike [2]. Akaike [1] took almost the same approach. He considered constraints concerned with the smoothness of the spectrum in time and frequency domain, the stability of the initial estimate of the spectrum and the mean and the smoothness of the mean value. The order of autoregression is put to be constant. These constraints keep the shape of estimated power spectrum density smooth. Namely unnecessary peaks don't appear owing to constraints in his procedure. On the other hand in our procedure the order of autoregression, which is concerned with the number of hyper parameters as shown later, is determined by the minimum ABIC procedure and so this procedure prevents the order from becoming unnecessarily large.

In Section 2 our model is defined and the likelihood of our model is introduced. In Section 3 our computational scheme is explained. In Section 4 numerical examples are given. Section 5 is devoted to concluding remarks.

2. A Bayesian model and the likelihood of the model

Given a set of data $\{x(1), x(2), \cdots, x(N)\}$, we divide this into P blocks. The length of each block is K and the order of the autoregressive model, which is fitted to each block of data, is put to be M. P, K and M must be chosen under the condition $N \geqq P * K + M$. Data are renumbered as follows for convenience.

$$(1)\qquad y(-M+i)=x(N-P*K-M+i)\,,\qquad i=1,2,\cdots,P*K+M\,.$$

Namely the p-th block ($p=1, 2,\cdots, P$) includes data $y((p-1)K+1)$, $y((p-1)K+2),\cdots, y(pK)$. The 0-th block is the set of data $\{y(-M+1), y(-M+2),\cdots, y(0)\}$. Here we note that K must be equal to or larger than $M+1+\alpha$ in our computer program whose source codes are presented in TIMSAC-84 (Akaike et al. [3]). If the mean value of the data can be put zero then α is put to be 0. Otherwise α is put to be 1. We expand our procedure on the assumption that the mean value of the process is not zero here after. We note that our procedure can be adapted to the process of which the mean value is zero by minor change.

The following autoregressive model is fitted to the p-th block of data ($p=1, 2,\cdots, P$),

$$(2)\qquad y(i)=\sum_{m=1}^{M} a_p(m)y(i-m)+a_p(0)+\varepsilon(i)$$
$$i=(p-1)K+1,\ (p-1)K+2,\cdots,pK$$

where $\varepsilon(i)$ is a Gaussian white noise with mean 0 and variance σ^2. We put a constraint on the temporal change of autoregressive coefficients by assuming that one-step ahead prediction given by autoregressive coefficients of the present block and that given by those of the preceding block are approximately equal. In other words the change of prediction values is penalized. It is assumed that the first order difference between two predictions behaves like a Gaussian white noise with mean 0 and variance σ^2/η^2. The fact that power spectrum can be calculated by using autoregressive coefficients leads us to consider that penalizing the temporal change of autoregressive coefficients is equivalent to penalizing the temporal change of spectrum in the result. Namely it seems to us that we indirectly put a constraint on the smoothness of the temporal change of spectrum. Our constraint is described as the following.

$$(3)\qquad c(p,i)=\sum_{m=1}^{M}(a_{p+1}(m)-a_p(m))y(i-m)+a_{p+1}(0)-a_p(0)$$
$$Ec(p,i)=0,\ \ Ec(p,i)c(p',i')=(\sigma^2/\eta^2)\delta_{pp'}\delta_{ii'}\,.$$

The parameter η controls the smoothness and is called a hyper parameter according to Akaike [1] in this article. It is noted that the larger η becomes, the stronger constraint becomes.

The likelihood of our Bayesian model is deribed as follows. It is to be noted that the definition of the likelihood of Bayesian model is attributed to Akaike [1]. The data distribution of our model is given by

$$(4)\qquad f(y\,|\,a,\sigma^2)=(2\pi\sigma^2)^{-PK/2}\exp\left(-SSR/2\sigma^2\right)\,,$$

where the definition of SSR is given by

$$SSR=\left\| \begin{bmatrix} Y_1 & & & 0 \\ & Y_2 & & \\ & & \ddots & \\ 0 & & & Y_P \end{bmatrix} \begin{bmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \\ \vdots \\ \boldsymbol{a}_P \end{bmatrix} - \begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \\ \vdots \\ \boldsymbol{y}_P \end{bmatrix} \right\|^2 .$$

Here $\| \ \|$ denotes the Euclidean norm and matrix Y_p and column vector $\boldsymbol{y}_p$ and $\boldsymbol{a}_p$ are defined as follows.

$$Y_p=\begin{pmatrix} 1 & y((p-1)K) & y((p-1)K-1) & \cdots & y((p-1)K+1-M) \\ 1 & y((p-1)K+1) & y((p-1)K) & \cdots & y((p-1)K+2-M) \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & y(pK-1) & y(pK-2) & \cdots & y(pK-M) \end{pmatrix}$$

$$\boldsymbol{y}_p=\begin{pmatrix} y((p-1)K+1) \\ y((p-1)K+2) \\ \vdots \\ y(pK) \end{pmatrix} \qquad \boldsymbol{a}_p=\begin{pmatrix} a_p(0) \\ a_p(1) \\ \vdots \\ a_p(M) \end{pmatrix} .$$

In the case the mean value of process is zero, the first column of Y_p and the first component of $\boldsymbol{a}_p$ are not necessary. The prior distribution for $\boldsymbol{a}_p$ is given by

$$(6) \qquad \begin{aligned} g(a \mid y, \sigma^2, \eta^2) &= (2\pi\sigma^2/\eta^2)^{-(P-1)(M+\alpha)/2} \\ &\times \{\det (Y_1'Y_1)\cdots\det (Y_{P-1}'Y_{P-1})\}^{1/2} \\ &\times \exp(-\eta^2 \,\mathrm{CON}/2\sigma^2) \end{aligned}$$

where CON is defined by

$$\mathrm{CON}=\left\| \begin{bmatrix} Y_1 & & & 0 \\ -Y_2 & Y_2 & & \\ & \ddots & \ddots & \\ 0 & & -Y_{P-1} & Y_{P-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{a}_2 \\ \boldsymbol{a}_3 \\ \vdots \\ \boldsymbol{a}_P \end{bmatrix} - \begin{bmatrix} Y_1\boldsymbol{a}_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \right\|^2 .$$

The estimates of autoregressive coefficients, $\hat{a}_p(m)$ ($m=0, 1,\cdots, M$, $p=2,\cdots, P$) and the estimate of σ^2, $\hat{\sigma}^2$ are obtained by maximizing $f(y \mid a, \eta^2)g(a \mid y, \sigma^2, \eta^2)$ with η^2 and $\boldsymbol{a}_1$ fixed. The estimates of autoregressive coefficients of the first block $\hat{a}_1(m)$ ($m=0, 1,\cdots, M$), which are regarded as hyper parameters, are obtained by searching for the maximum of the likelihood of our Bayesian model.

$$(8)\qquad l(\boldsymbol{a}_1, \sigma^2, \eta^2)=\int \prod_{m=2}^{M} d\boldsymbol{a}_m f(y\,|\,a, \sigma^2) g(a\,|\,y, \sigma^2, \eta^2)\,.$$

The straightforward calculation leads to the following result.

$$(9)\qquad \begin{aligned} l(\boldsymbol{a}_1, \sigma^2, \eta^2)=&(2\pi\sigma^2)^{-PK/2}(\eta^2)^{(P-1)(M+\alpha)/2}\\ &\times\{\det(Y_1'Y_1)\cdots\det(Y_{P-1}'Y_{P-1})\}^{1/2}\\ &\times\{\det(Z'Z)\}^{1/2}\exp(-v(\boldsymbol{a}_1, \eta)/2\sigma^2)\\ &\times\exp\left(-\sum_{i=1}^{K}\left(y(i)-\sum_{m=1}^{M}a_1(m)y(i-m)-a_1(0)\right)^2\Big/2\sigma^2\right)\end{aligned}$$

where

(10)

$$Z=\begin{bmatrix} Y_2 & & & & \\ & Y_3 & & 0 & \\ & & \ddots & & \\ & 0 & & \ddots & \\ & & & & Y_P \\ \eta Y_1 & & & 0 & \\ -\eta Y_2 & \eta Y_2 & & & \\ & \ddots & \ddots & & \\ 0 & & & -\eta Y_{P-1} & \eta Y_{P-1} \end{bmatrix} \qquad v(\boldsymbol{a}_1, \eta)=\left\| Z\begin{bmatrix}\hat{\boldsymbol{a}}_2\\ \hat{\boldsymbol{a}}_3\\ \vdots\\ \hat{\boldsymbol{a}}_P\end{bmatrix}-\begin{bmatrix}\boldsymbol{y}_2\\ \boldsymbol{y}_3\\ \vdots\\ \boldsymbol{y}_P\\ \eta Y_1\boldsymbol{a}_1\\ 0\\ \vdots\\ 0\end{bmatrix}\right\|^2$$

It is noted that $\hat{\boldsymbol{a}}_1$ is the solution of the equation

$$(11)\qquad (1+\eta^2)Y_1'Y_1\hat{\boldsymbol{a}}_1=Y_1'(\boldsymbol{y}_1+\eta^2 Y_1\hat{\boldsymbol{a}}_2)\,\cdot$$

The estimate of σ^2 is derived easily.

$$(12)\qquad \hat{\sigma}^2=V/PK$$

where $V=v(\hat{\boldsymbol{a}}_1, \eta)$ is

$$(13)\qquad V=\left\|\begin{bmatrix} Y_1 & & & & \\ & Y_2 & & 0 & \\ & & \ddots & & \\ & 0 & & \ddots & \\ & & & & Y_P \\ -\eta Y_1 & \eta Y_1 & & & \\ & -\eta Y_2 & \eta Y_2 & 0 & \\ & & \ddots & \ddots & \\ 0 & & & -\eta Y_{P-1} & \eta Y_{P-1} \end{bmatrix}\begin{bmatrix}\boldsymbol{a}_1\\ \boldsymbol{a}_2\\ \vdots\\ \boldsymbol{a}_P\end{bmatrix}-\begin{bmatrix}\boldsymbol{y}_1\\ \boldsymbol{y}_2\\ \vdots\\ \boldsymbol{y}_P\\ 0\\ 0\\ \vdots\\ 0\end{bmatrix}\right\|^2.$$

Now we must determine the appropriate smoothness parameter η and the number of autoregressive coefficients of the first block. To do so we calculate ABIC (Akaike Bayesian Information Criterion) defined by Akaike [2].

$$\text{ABIC} = -2\log(\text{maximum likelihood of the Bayesian model}) + 2(\text{the number of estimated hyper parameters}).$$

In our model the likelihood of the Bayesian model $l(\boldsymbol{a}_1, \sigma^2, \eta^2)$ is described by $a_1(0), a_1(1), \cdots, a_1(M)$, σ^2 and η^2 (if $\alpha=0$ then $a_1(0)$ is not included). But independent variables of $l(\boldsymbol{a}_1, \hat{\sigma}^2, \eta^2)$ are $\boldsymbol{a}_1$ and η^2 as long as the maximum likelihood estimate is concerned, since $\hat{\sigma}^2$ is expressed by $\hat{\boldsymbol{a}}_1$ and η^2. In our model $\boldsymbol{a}_1$ and η^2 are regarded as hyper parameters. Namely the number of estimated hyper parameters is $M+\alpha+1$. Here "1" is common to all models and so we ignore this term. Then ABIC of our model is given by

$$\begin{aligned}(14)\quad \text{ABIC}(M, \eta) = {} & PK(\ln 2\pi + 1 - \ln(PK) + \ln V) - (P-1)(M+\alpha)\ln \eta^2 \\ & -\ln(\det(Y_1'Y_1)\cdots\det(Y_{P-1}'Y_{P-1})) \\ & +\ln(\det(Z'Z)) + 2(M+\alpha)\,.\end{aligned}$$

The hyper parameter η is chosen by searching for minimum value of ABIC with the aid of a nonlinear optimization method. In our computer program (Akaike et al. [3]) grid search of η is done to perform an approximate nonlinear optimization with less computational time.

The number of autoregressive coefficients of the first block, which are hyper parameters of our model, is determined so that the value of ABIC becomes minimum. In Section 1 we stated that the order of autoregression is determined by the minimum ABIC procedure. But it is more accurate to phrase that the number of hyper parameters $(a_1(0), a_1(1), \cdots, a_1(M))$ is determined by the minimum ABIC procedure. It is noted that M is the order of autoregressive models in all blocks as well as in the first block. And so to determine the order of autoregressions can be regarded as paraphrase of to determine the number of hyper parameters.

3. Computational scheme and power spectrum

The minimum ABIC procedure for the fitting of our nonstationary time series model is summarized as follows.

(1) Given a set of data $\{x(1), x(2), \cdots, x(N)\}$, set the upper limit $M_{\max}$ of the order of autoregression and choose the length of a block K and the number of blocks P. It is noted that the conditions $N \geqq PK + M_{\max}$ and $K \geqq M_{\max}+2$ must be satisfied.

(1)′ the 0-th block: $\{y(-M_{\max}+1), y(-M_{\max}+2), \cdots, y(0)\}$

the p-th block: $\{y((p-1)K+1), y((p-1)K+2), \cdots, y(pK)\}$

$p=1, 2, \cdots, P$

where

$$y(-M_{\max}+i)=x(N-PK-M_{\max}+i)\ , \qquad i=1, 2, \cdots, PK+M_{\max}\ .$$

(2) (i) Construct $K*(M_{\max}+2)$ matrix for $p=1, 2, \cdots, P$

$$Y_p=\begin{pmatrix} 1 & y((p-1)K) & \cdots & y((p-1)K+1-M_{\max}) & y((p-1)K+1) \\ 1 & y((p-1)K+1) & \cdots & y((p-1)K+2-M_{\max}) & y((p-1)K+2) \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & y(pK-1) & \cdots & y(pK-M_{\max}) & y(pK) \end{pmatrix}.$$

(ii) By applying the Householder transformation the matrix Y_p is reduced to an upper triangular matrix R_p

$$R_p=\begin{pmatrix} r_{11}^{(p)} & \cdots & r_{1,M_{\max}+1}^{(p)} & r_{1,M_{\max}+1}^{(p)} \\ & \ddots & \vdots & \vdots \\ \text{\Large 0} & & r_{M_{\max}+1,M_{\max}+1}^{(p)} & r_{M_{\max}+1,M_{\max}+2}^{(p)} \\ & & & r_{M_{\max}+2,M_{\max}+2}^{(p)} \\ 0 & & & 0 \end{pmatrix}=\begin{bmatrix} R_p(M_{\max}) & r_p(M_{\max}) \\ 0 & 0 \end{bmatrix}.$$

(iii) Calculate $\det(Y_p' Y_p)$ for $p=1, 2, \cdots, P-1$ and for each $M \leqq M_{\max}$

$$\det(Y_p' Y_p)=\prod_{i=1}^{M}(r_{ii}^{(p)})^2\ .$$

(iv) Sum of squared residuals of each block ($p=1, 2, \cdots, P$) is given by

$$\mathrm{SSR}(p)=\sum_{i=M+2}^{M_{\max}+2}(r_{i,M_{\max}+2}^{(p)})^2\ .$$

(3) Set the matrix

$$
\begin{bmatrix}
R_2(M) & & & & \\
& R_3(M) & & 0 & \\
0 & & \ddots & & \\
& & & R_P(M) & \\
\eta R_1(M) & & & & \\
-\eta R_2(M) & \eta R_2(M) & & & 0 \\
0 & \ddots & \ddots & & \\
& & -\eta R_{P-1}(M) & \eta R_{P-1}(M) &
\end{bmatrix}
$$

and reduce this matrix to the upper triangular matrix Q via the Householder transformation. We can derive $\det(Z'Z)$ by using diagonal components of Q, q_{ii} $(i=1,\cdots,(P-1)(M+1))$,

$$
\det(Z'Z)=\prod_{i=1}^{(P-1)(M+1)} q_{ii}^2 .
$$

(4) Construct the matrix

$$
\begin{bmatrix}
R_1(M) & & & & 0 & r_1(M) \\
& R_2(M) & & & & r_2(M) \\
0 & & \ddots & & & \vdots \\
& & & & R_P(M) & r_P(M) \\
-\eta R_2(M) & \eta R_2(M) & & & & 0 \\
& -\eta R_3(M) & \eta R_3(M) & & 0 & 0 \\
0 & & \ddots & \ddots & & \vdots \\
& & & -\eta R_P(M) & \eta R_P(M) & 0
\end{bmatrix}
$$

and reduce this matrix to the upper triangular matrix by applying the Householder transformation

$$
\begin{bmatrix}
S_{11}(M) & S_{12}(M) & & & & s_1(M) \\
& S_{22}(M) & S_{23}(M) & & 0 & s_2(M) \\
& & \ddots & \ddots & & \vdots \\
& & & S_{P-1,P-1}(M) & S_{P-1,P}(M) & s_{P-1}(M) \\
& 0 & & & S_{P,P}(M) & s_P(M) \\
& & & & & s_{P+1}(M)
\end{bmatrix}
$$

where $S_{ii}(M)$ $(i=1, 2,\cdots, P)$ are upper triangular matrices. Autoregressive coefficients are given by

$$\hat{\boldsymbol{a}}_P = S_{PP}(M)^{-1}\boldsymbol{s}_P(M) ,$$

$$\hat{\boldsymbol{a}}_p = S_{pp}(m)^{-1}(\boldsymbol{s}_p(M) - S_{p,p+1}(M)\hat{\boldsymbol{a}}_{p+1}) , \qquad p = P-1, \cdots, 1 .$$

The maximum likelihood estimate of the innovation variance $\hat{\sigma}^2$ is given by

$$\hat{\sigma}^2 = \left(\sum_{p=1}^{P} \mathrm{SSR}\,(p) + s_{P+1}(M)^2\right) \Big/ PK .$$

(5) Calculate ABIC (M, η) $(M=0, 1, \cdots, M_{\max})$ for several values of the hyper parameter η and select the number of the autoregressive coefficients of the first interval M and the hyper parameter η so that ABIC (M, η) has the minimum value.

(6) The power spectrum of the p-th block is given by

$$q_p(g) = \frac{\hat{\sigma}^2}{\left|1 - \sum_{m=1}^{M} \hat{\boldsymbol{a}}_p(m) \exp(-i2\pi g m)\right|^2} .$$

4. Numerical examples

Our procedure for fitting a Bayesian autoregressive model to nonstationary process was applied to two artificially generated nonstationary time-seriesed data and one time-seriesed clinical data. The artificial data is generated by the following equation.

$$x(n) + \sum_{i=1}^{m} a(i, n) x(n-i) = \varepsilon(n) \qquad n = -209, \cdots, -1, 0, 1, \cdots, 501$$

where $\varepsilon(n) \sim N(0, 1)$ and autoregressive coefficients $a(i, n)$ are determined so that characteristic equation

$$s^m + a(1, n)s^{m-1} + \cdots + a(m-1, n)s + a(m, n) = 0$$

has roots given by

$$\text{model 1 } (m=4): \quad 0.83 \exp(\pm w_1(n)\sqrt{-1}), \quad 0.88 \exp(\pm w_2(n)\sqrt{-1})$$

$$\text{model 2 } (m-4): \quad 0.83 \exp(\pm w_3(n)\sqrt{-1}), \quad 0.88 \exp(\pm w_4(n)\sqrt{-1}) .$$

Here

$$w_1(n) = 5\pi/9 + (\pi/6) \sin(7\pi(n-100)/2700)$$

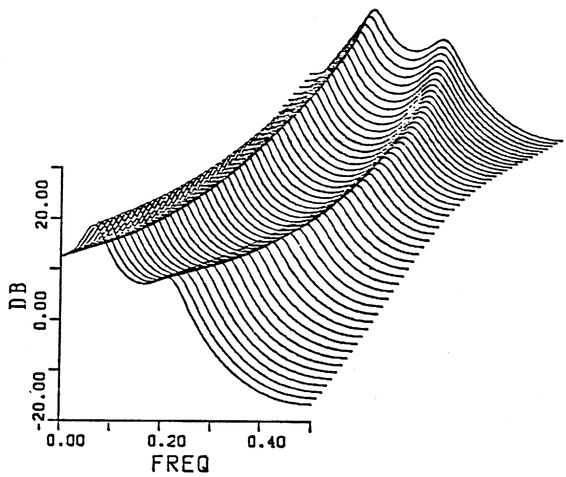
$$w_2(n) = \pi/6 + (\pi/12) \sin(7\pi(n-100)/5400)$$

$$w_3(n) = 5\pi/9 + (\pi/6) \sin(\pi(n-100)/135)$$

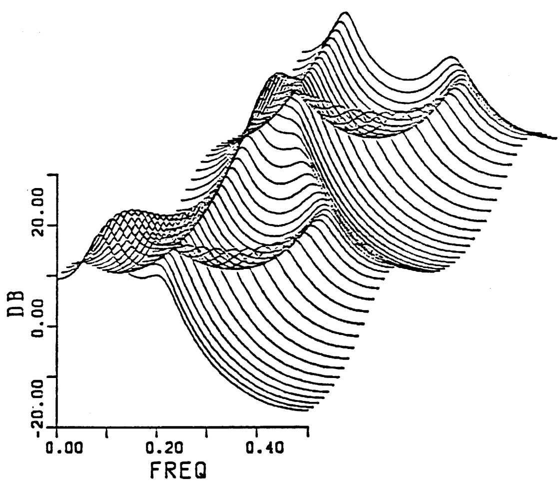
$$w_4(n)=2\pi/9+(\pi/18)\sin(\pi(n-100)/270) .$$

These models are the same models that Kitagawa [8] proposed in his paper.

The sets of data $\{x(-9),\cdots, x(-1), x(0), x(1),\cdots, x(500)\}$ of the above two models are used. The number of blocks, the number of data in each block and the maximum value of the order of autoregression were put to be 50, 10 and 9, respectively. ("The length of each block" must be equal to or larger than "the maximum order of autoregression+1" in the case the mean value is zero.) The hyper param-



(a)



(b)

Fig. 1. Artificially generated changing spectrum.
(a) model 1 (b) model 2

eter η is determined by searching for minimum value of ABIC over a set of η. The change of the theoretical spectrum is illustrated in Fig. 1 and the change of the estimated spectrum is illustrated in Fig. 2. The model whose order of autoregression is 4 was the minimum ABIC model for the first series and the model of which order of autoregression is 4 was the minimum ABIC's for the second one. Namely our procedure could select true order for these examples. We note that our procedure always selected true order for 50 time-seriesed data generated by using the same coefficients with model 1. It seems that our

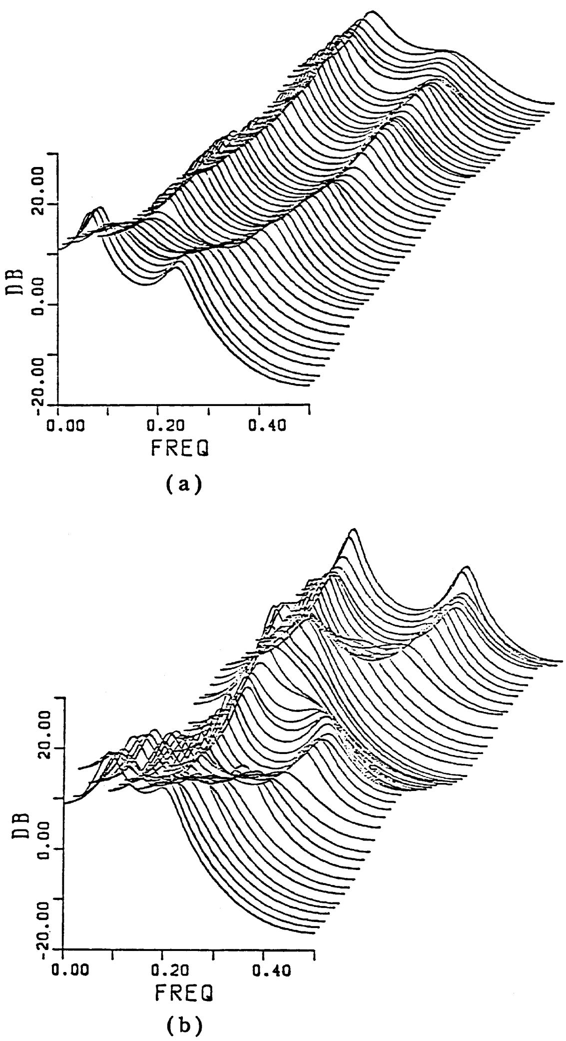


Fig. 2. Estimated changing spectrum obtained with
(a) $M=4$, $\eta=4.2$ for model 1,
(b) $M=4$, $\eta=2.2$ for model 2.

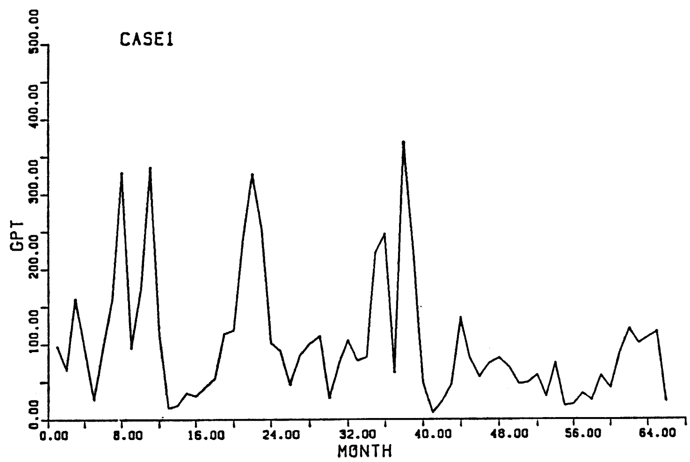


Fig. 3. Time-seriesed clinical data (GPT).

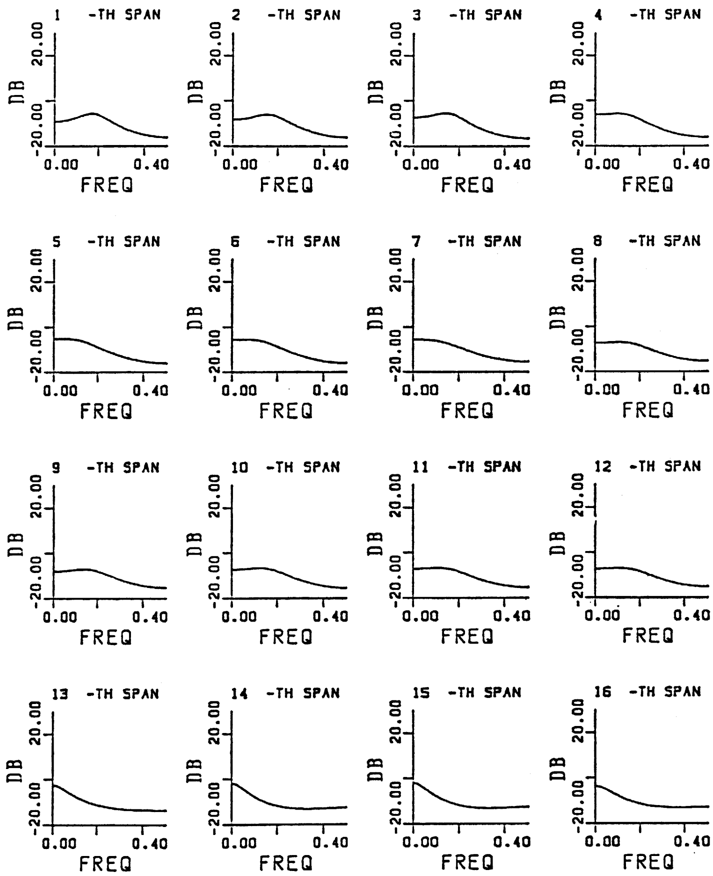


Fig. 4. Estimated changing spectrum for log-transformed GPT data.

procedure can select the reasonable model.

Our procedure was applied to time-seriesed clinical data. Original data is illustrated in Fig. 3 and the change of the estimated spectrum is illustrated in Fig.4. It is noted that we modified original data by applying log transformation and fitted our model to the modified data. Whole data (the number of data is 66) was divided into 16 blocks. The length of each block was put to be 4 and the maximum order of autoregression was put to be 2. ("The length of each block" must be equal to or larger than "the maximum value of autoregression+2" in the case the mean value is not zero.) In this case the variance of data suddenly changes about the 40th month. We, however, assume that the variance is constant and so we must weaken constraint by using the small value of the hyper parameter in order to alleviate the influence of the change of the variance. The value of the hyper parameter used was $\eta=5$. Our procedure gave the result that the shape of spectrum changes at the 13th span. This result seems to be reasonable, because this data shows that hepatic disorder state alters about the 40th month from hepatitis to cirrhosis of the liver.

5. Concluding remarks

A Bayestian procedure to estimate nonstationary power spectral density is proposed. The number of autoregressive coefficients of the first block a_1 and the value of the hyper parameter η which is related with the strength of constraint on time-varying autoregressive coefficients are determined with the aid of the minimum ABIC procedure. In Section 2 it was noted that our constraint is concerned with the smoothness of one-step ahead prediction. Considering that power spectrum of each block is derived by using estimated autoregressive coefficients, we can say that our constraint is concerned with the smoothness of spectral density in time domain. Numerical examples show that our procedure is sufficiently practicable.

Extension of our procedure to multi-variate time-seriesed data is discussed in the separate paper. The case where the variance of data σ^2 is not constant will be also discussed in near future.

Acknowledgements

The author would like to express his thanks to Professor Akaike for his continual encouragement.

THE INSTITUTE OF STATISTICAL MATHEMATICS

REFERENCES

[1] Akaike, H. (1979). On the construction of composite time series model, *Bull. 42nd Session I.S.I.*, **48**, 411-422.

[2] Akaike, H. (1980). Likelihood and the Bayes procedure, Bayesian Statistics (eds. Bernardo, J. S., Degroot, M H., Lindley, D. V. and Smith, A. F. M.).

[3] Akaike, H. et al. (1985). TIMSAC-84 (A Time Series Analysis and Control Program Package), *Computer Science Monographs*, **22**, The Institute of Statistical Mathematics, Tokyo.

[4] Bendat, J. S. and Piersol, A. G. (1966). *Measurement and Analysis of Random Data*, John Wiley & Sons, New York.

[5] Hino, M. (1977). *Spectral Analysis*, Asakura Shoten, Tokyo (in Japanese).

[6] Kitagawa, G. and Akaike, H. (1978). A procedure for the modeling of non-stationary time series, *Ann. Inst. Statist. Math.*, **30**, 351-363.

[7] Kitagawa, G. and Akaike, H. (1981). *On* TIMSAC-78, *Applied Time Series Analysis* II (ed. Findley, D.), 449-548, Academic Press, New York.

[8] Kitagawa, G. (1983). Changing spectrum estimation, *J. Sound and Vib.*, **89**, 433-445.

[9] Mark, W. D. (1970). Spectral analysis of the convolution and filtering of non-stationary stochastic processes, *J. Sound and Vib.*, **11**, 19-63.

[10] Ozaki, T. and Tong, H. (1975). On the fitting of non-stationary autoregressive models in time series analysis, *Proceeding of 8th Hawaii International Conference on System Sciences*, 224-226.

[11] Page, C. H. (1952). Instantaneous power spectra, *J. Appl. Phys.*, **23**, 103-106.

[12] Priestley, M. B. (1965). Evolutionary spectra and non-stationary processes, *J. R. Statist. Soc.*, B**27**, 221-224.

[13] Priestley, M. B. (1967). Power spectral analysis of non-stationary random processes, *J. Sound and Vib.*, **6**, 86-97.

[14] Subba Rao, T. (1970). The fitting of non-stationary time-series model with time-dependent parameters, *J. R. Statist. Soc.*, B**32**, 312-332.