Ann. Inst. Statist. Math.
30 (1978), Part B, 351-363

# A PROCEDURE FOR THE MODELING OF NON-STATIONARY TIME SERIES

GENSHIRO KITAGAWA AND HIROTUGU AKAIKE

(Received Jan. 23, 1978; revised Aug. 4, 1978)

## Summary

A minimum AIC procedure for the fitting of a locally stationary autoregressive model is proposed. The least squares computation for the procedure is realized by using the Householder transformation which makes the procedure computationally more flexible and efficient than the one originally proposed by Ozaki and Tong.

## 1. Introduction

Ozaki and Tong [8] extended the autoregressive model fitting procedure developed by Akaike [1], [2], [3] to non-stationary situations. They considered a locally stationary process and fitted a stationary autoregressive model to each stationary block of the data. The goodness of fit of the global model composed of these local stationary models is measured by the corresponding AIC and the partition of the data into blocks which minimizes the AIC defines the best model. The homogeneity of data is checked each time as a block of prescribed number of new data is added and the additional one is pooled to the original one if these two blocks of data are considered to be homogeneous. Otherwise a new process of modeling starts with the new block. The procedure has close connection with the theory of successive process of statistical inference [6], in particular, with that of the estimation after preliminary tests of significance discussed by T. Kitagawa [7]. The main difference is that here the two-step TE (Test-Estimation) type estimation was replaced by the single step minimum AIC procedure.

Ozaki and Tong used the conventional technique of fitting of the autoregressive model described in Akaike [2], which is developed for the analysis of long stationary data and is not quite suitable for the application to a non-stationary situation where the analysis of a short span of data is necessary.

In this paper, we propose a new procedure which utilizes the Householder transformation, a powerful tool for the solution of the least squares problem [5]. This transformation allows the necessary modification of a fitted model due to the addition of observations or the addition and deletion of regressors quite easily. Especially with this procedure even the change of the mean value of the process can be handled very easily.

This paper is organized as follows. In Section 2, the locally stationary autoregressive model is defined and the likelihood and AIC of the model are derived. An on-line type fitting procedure is described in Section 3. In Section 4, the application of the Householder transformation to the solution of the least squares problem of time series is discussed. The minimum AIC procedure for the fitting of locally stationary model is described in Section 5. In Section 6, some numerical examples are given, and the final section is devoted to additional comments on the procedure.

## 2. Locally stationary autoregressive model

Given a set of observations $\{x_1, \cdots, x_N\}$ we consider the situation where the time interval $[1, N]$ is divided into $k$ blocks, each of length $n_i$ $(n_1+n_2+\cdots+n_k=N$; $k$ and $n_i$ are unknown), and the following locally stationary autoregressive model is being fitted to the data:

$$x_n=a_0^i+\sum_{m=1}^{M(i)} a_m^i x_{n-m}+\varepsilon_n^i\ , \qquad \left(\sum_{j=1}^{i-1} n_j+1\leqq n\leqq\sum_{j=1}^{i} n_j;\ i=1,\cdots,k\right)\ ,$$

where $\varepsilon_n^i$ is a Gaussian white noise with $\mathrm{E}\,\varepsilon_n^i=0$, $\mathrm{E}\,(\varepsilon_n^i)^2=\sigma_i^2$ and $\mathrm{E}\,\varepsilon_n^i x_{n-m}=0$ $(m>0)$.

The approximate likelihood of the model to be used in our analysis is defined by

$$\prod_{i=1}^{k}\left(\frac{1}{2\pi\sigma_i^2}\right)^{n_i/2}\exp\left\{-\frac{1}{2\sigma_i^2}\sum_{n=p_i}^{q_i}\left(x_n-a_0^i-\sum_{m=1}^{M(i)} a_m^i x_{n-m}\right)^2\right\}\ ,$$

where $p_1=M(1)+1$, $p_i=1+\sum_{j=1}^{i-1} n_j$, $q_i=\sum_{j=1}^{i} n_j$. Thus by denoting $a^i=(a_0^i, a_1^i,\cdots, a_{M(i)}^i)$, the logarithm of the approximate likelihood is given by

$$\begin{aligned}&l(x;k, n_i, M(i), a^i, \sigma_i^2\ (i=1,\cdots,k))\\&\quad=-\frac{1}{2}\sum_{i=1}^{k}\left\{n_i\log 2\pi\sigma_i^2+\frac{1}{\sigma_i^2}\sum_{n=p_i}^{q_i}\left(x_n-a_0^i-\sum_{m=1}^{M(i)} a_m^i x_{n-m}\right)^2\right\}\ .\end{aligned}$$

The maximum of the function $l$ for given $a_m^i$'s is attained at

$$\sigma_i^2=\frac{1}{n_i}\sum_{n=p_i}^{q_i}\left(x_n-a_0^i-\sum_{m=1}^{M(i)} a_m^i x_{n-m}\right)^2\ ,$$

and the maximum of log likelihood $l$ is given by

$$\begin{aligned} l^*(x;k,n_i,M(i),\ (i=1,\cdots,k)) &= -\frac{1}{2}\sum_{i=1}^{k}\{n_i \log 2\pi\hat{\sigma}_i^2+n_i\} \\ &= -\frac{N}{2}(1+\log 2\pi)-\frac{1}{2}\sum_{i=1}^{k} n_i \log \hat{\sigma}_i^2 . \end{aligned}$$

where $\hat{\sigma}_i^2$ is the approximate maximum likelihood estimate of $\sigma_i^2$ and is obtained by minimizing $\sigma_i^2$ with respect to $a_m^i$'s.

The AIC of our locally stationary model is then given by

$$\mathrm{AIC}=\sum_{i=1}^{k} n_i \log \hat{\sigma}_i^2+2\sum_{i=1}^{k}(M(i)+2) .$$

The MAICE's (minimum AIC estimates) of the number of stationary blocks, the size of each block and the order of the autoregressive model fitted to each block are defined as those values of $k$, $n_i$ $(i=1,\cdots,k)$ and $M(i)$ $(i=1,\cdots,k)$ which minimize the AIC.

## 3. An on-line type fitting procedure

Instead of considering every possible combination of blocks, the procedure proposed for practical use is as follows [8]:

Consider the situation where an autoregressive model, $\mathrm{AR}_0$, has been fitted to the set of data $\{x_1,\cdots,x_n\}$ and an additional set of $m$ observations $\{x_{n+1},\cdots,x_{n+m}\}$ is newly obtained where $m$ is a prescribed number. We consider two competing models. The first one is defined by connecting two autoregressive models, the model $\mathrm{AR}_0$ and the model $\mathrm{AR}_1$ which is fitted to the newly obtained data $\{x_{n+1},\cdots,x_{n+m}\}$ and which are assumed to be independent. The AIC of this jointed model is given by

$$\mathrm{AIC}_1=n\log\sigma_0^2+m\log\sigma_1^2+2(M_0+M_1+4) ,$$

where $\sigma_0^2$ and $M_0$ are the innovation variance and the order of the autoregressive model $\mathrm{AR}_0$, $\sigma_1^2$ and $M_1$ are those of $\mathrm{AR}_1$, respectively. The second model is an autoregressive model, $\mathrm{AR}_2$, fitted to the whole span of the pooled data $\{x_1,\cdots,x_{n+m}\}$. The AIC of the model is given by

$$\mathrm{AIC}_2=(n+m)\log\sigma_2^2+2(M_2+2) ,$$

where $\sigma_2^2$ is the innovation variance and $M_2$ is the order of the fitted model.

If $\mathrm{AIC}_1$ is less than $\mathrm{AIC}_2$, we switch to the new model $\mathrm{AR}_1$. Otherwise, the two sets of data are considered to be homogeneous and the model $AR_2$ is accepted. The procedure repeats these steps whenever a set of $m$ new observations is given. Hereafter we will call $m$ as the basic span.

The procedure is so designed as to follow the change of the structure of the time series, while if the structure remains unchanged it will improve the model by using the additional observations.

## 4. Least squares computation by Householder transformation

First we assume that the mean value of the process $\{x_n\}$ is zero. The least squares estimates of the $k$th order autoregressive coefficients are obtained by minimizing the sum of squares

$$\frac{1}{N-K}\sum_{n=K+1}^{N}(x_n-a_1x_{n-1}-\cdots-a_Kx_{n-K})^2 .$$

Define the matrix $\boldsymbol{X}$ and the vectors $\boldsymbol{y}$ and $\boldsymbol{a}$ by

$$\boldsymbol{X}=\begin{bmatrix} x_K & x_{K-1}\cdots x_1 \\ x_{K+1} & x_K \quad\cdots x_2 \\ \vdots & \vdots \qquad \vdots \\ x_{N-1} & x_{N-2}\cdots x_{N-K}\end{bmatrix}, \qquad \boldsymbol{y}=\begin{bmatrix} x_{K+1} \\ x_{K+2} \\ \vdots \\ x_N \end{bmatrix}, \qquad \boldsymbol{a}=\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_K \end{bmatrix}.$$

The least squares method minimizes $\|\boldsymbol{y}-\boldsymbol{X}\boldsymbol{a}\|^2$, where $\|\ \|$ denotes the Euclidean norm, and the solution is given by the normal equation

$$\boldsymbol{X}'\boldsymbol{X}\boldsymbol{a}=\boldsymbol{X}'\boldsymbol{y} .$$

As a numerical procedure the direct solution of the normal equation is not quite efficient and the procedure realized by first orthogonalizing the column vectors of $\boldsymbol{X}$ and then solving the resultant equation supersedes the normal equation approach in both manipulability and numerical accuracy [5].

A Householder transformation is an orthogonal transformation defined by a matrix $\boldsymbol{P}=\boldsymbol{I}-2\boldsymbol{u}\boldsymbol{u}'$, where $\boldsymbol{u}$ is a vector with $\|\boldsymbol{u}\|=1$. Let $\boldsymbol{X}^{(1)}=\boldsymbol{X}$ and $\boldsymbol{X}^{(2)},\cdots,\boldsymbol{X}^{(K+1)}$ be defined by $\boldsymbol{X}^{(k+1)}=\boldsymbol{P}^{(k)}\boldsymbol{X}^{(k)}$ $(k=1,\cdots,K)$, where $\boldsymbol{P}^{(k)}$ is a Householder transformation and is chosen so that $x_{i,k}^{(k+1)}=0$ $(i=k+1, k+2,\cdots, N-K)$, where $x_{i,k}^{(k+1)}$ denotes $(i, k)$th element of $\boldsymbol{X}^{(k+1)}$. $\boldsymbol{P}^{(k)}$ is defined by $\boldsymbol{P}^{(k)}=\boldsymbol{I}-\boldsymbol{v}_k\boldsymbol{v}_k'/h_k$ with

$$\boldsymbol{v}_k'=(0,\cdots,0,\ x_{k,k}^{(k)}\pm\tau,\ x_{k+1,k}^{(k)},\cdots,\ x_{N-K,k}^{(k)}) ,$$

where $\tau^2=\sum_{i=k}^{N-K}\{x_{i,k}^{(k)}\}^2$ and $h_k=\tau^2\mp\tau x_{k,k}^{(k)}$. We have

$$\boldsymbol{X}^{(K+1)}=\boldsymbol{Q}\boldsymbol{X}=\boldsymbol{R}=\left[\frac{\boldsymbol{S}}{\boldsymbol{0}}\right] ,$$

where $\boldsymbol{Q}=\boldsymbol{P}^{(K)}\boldsymbol{P}^{(K-1)}\cdots\boldsymbol{P}^{(1)}$ is an orthogonal matrix and $\boldsymbol{S}$ is an upper triangular matrix. The matrix $\boldsymbol{S}$ and the vector $\boldsymbol{Qy}$ keep the complete information for the least squares fitting of autoregressive models up to the $K$th order. Denote $\boldsymbol{S}$ and $\boldsymbol{z}=\boldsymbol{Qy}$ by

$$\boldsymbol{S}=\begin{bmatrix} s_{11} & \cdots & s_{1K} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & s_{KK} \end{bmatrix} \qquad \text{and} \qquad \boldsymbol{z}'=(z_1,\cdots,z_{N-K}) .$$

For each $k\leqq K$, the least squares estimates of the coefficients of the $k$th order autoregressive model

$$x_n=\sum_{i=1}^{k} a_i^k x_{n-i}+\varepsilon_n^k$$

are obtained by solving the linear equation

$$\begin{bmatrix} s_{11} & \cdots & s_{1k} \\ & \ddots & \vdots \\ 0 & & s_{kk} \end{bmatrix}\begin{bmatrix} a_1^k \\ \vdots \\ a_k^k \end{bmatrix}=\begin{bmatrix} z_1 \\ \vdots \\ z_k \end{bmatrix} .$$

The corresponding estimate $d(k)$ of $\sigma_k^2=\mathrm{E}(\varepsilon_n^k)^2$ is given by ([5])

$$d(k)=\frac{1}{N-K}\sum_{i=k+1}^{N-K} z_i^2 .$$

We note that for the fitting of the generalized model with a constant term $a_0$

$$x_n=a_0+\sum_{m=1}^{k} a_m x_{n-m}+\varepsilon_n , \qquad (k=1,\cdots,K)$$

we have only to define $\boldsymbol{X}$ by

$$\boldsymbol{X}=\begin{bmatrix} 1 & x_K & \cdots x_1 \\ 1 & x_{K+1} & \cdots x_2 \\ \vdots & \vdots & \vdots \\ 1 & x_{N-1} & \cdots x_{N-K} \end{bmatrix} .$$

By using the least squares estimates as approximations to the maximum likelihood estimates under the Gaussian assumption, the minimum AIC procedure for the fitting of an autoregressive model is realized as follows.

Assume that a set of data $\{x_n;\ n=1,\cdots,N\}$ is given.

1) Replace $x_n$ by $\tilde{x}_n=x_n-\bar{x}$, where $\bar{x}=\frac{1}{N}\sum_{n=1}^{N} x_n$.

2) Determine the upper limit $K$ of the order of the autoregressive models to be fitted to the data.

3) Define the $(N-K)\times(K+1)$ matrix

$$\boldsymbol{X}=\begin{bmatrix} x_K & x_{K-1}\cdots x_1 & x_{K+1} \\ x_{K+1} & x_K \cdots x_2 & x_{K+2} \\ \vdots & \vdots \quad \ddots \quad \vdots & \vdots \\ x_{N-1} & x_{N-2}\cdots x_{N-K} & x_N \end{bmatrix}.$$

4) Reduce the matrix $\boldsymbol{X}$ to an upper triangular matrix

$$\boldsymbol{S}=\begin{bmatrix} s_{11}\cdots s_{1K} & s_{1,K+1} \\ \ddots \quad \vdots & \vdots \\ s_{KK} & s_{K,K+1} \\ 0 \qquad & s_{K+1,K+1} \end{bmatrix}$$

by the successive application of the Householder transformations described in the preceding section.

5) Define the AIC of the autoregressive model of order $m$ by

$$\mathrm{AIC}(m)=(N-K)\log(d(m))+2(m+2),$$

where $d(m)$, the estimate of the innovation variance, is in this case given by

$$d(m)=\frac{1}{N-K}\sum_{i=m+1}^{K+1} s_{i,K+1}^2 .$$

6) Adopt the $m$ which gives the minimum of $\mathrm{AIC}(m)$ $(m=0, 1, \cdots, K)$ as the order $L$ of the model.

7) The minimum AIC estimates of the autoregressive coefficients $a_i^L$ $(i=1,\cdots, L)$ are obtained by

$$a_L^L=s_{L,L}^{-1}s_{L,K+1}, \qquad a_i^L=s_{i,i}^{-1}\left(s_{i,K+1}-\sum_{j=i+1}^{L} a_j^L s_{i,j}\right), \quad (i=L-1,\cdots, 1).$$

## 5. Minimum AIC estimation of locally stationary models

In this section we will discuss an implementation of the minimum AIC procedure for the fitting of autoregressive models to locally stationary time series.

Let the matrix $\boldsymbol{X}$ be defined by

$$\boldsymbol{X}=\begin{bmatrix} 1 & x_K & \cdots x_1 & x_{K+1} \\ 1 & x_{K+1} & \cdots x_2 & x_{K+2} \\ \vdots & \vdots & \ddots \quad \vdots & \vdots \\ 1 & x_{N_0-1} & \cdots x_{N_0-K} & x_N \end{bmatrix}$$

and is reduced to the upper triangular matrix $\boldsymbol{S}$ by the Householder transformation. When an additional set of observations $\{x_{N_0+1},\cdots, x_{N_0+M}\}$ is obtained the matrix $\boldsymbol{Z}$ is constructed as

$$\boldsymbol{Z}=\begin{bmatrix} 1 & x_{N_0} & \cdots x_{N_0-K+1} & x_{N_0+1} \\ 1 & x_{N_0+1} & \cdots x_{N_0-K+2} & x_{N_0+2} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{N_0+M-1} & \cdots x_{N_0+M-K} & x_{N_0+M} \end{bmatrix}.$$

By repeated applications of Householder transformations the matrix $\boldsymbol{Z}$ is reduced to an upper triangular matrix $\boldsymbol{S}_1$ by

$$\boldsymbol{Q}_1\boldsymbol{Z}=\overset{\overleftrightarrow{K+2}}{\left[\frac{\boldsymbol{S}_1}{\boldsymbol{0}}\right]}\begin{matrix}\updownarrow K+2 \\ \updownarrow M-K-2\end{matrix}.$$

Again by applying the Householder transformations to the matrix

$$\left[\frac{\boldsymbol{S}}{\boldsymbol{S}_1}\right]\begin{matrix}\updownarrow K+2 \\ \updownarrow K+2\end{matrix}$$

we obtain an upper triangular matrix $\boldsymbol{S}_2$ by

$$\boldsymbol{P}\left[\frac{\boldsymbol{S}}{\boldsymbol{S}_1}\right]=\left[\frac{\boldsymbol{S}_2}{\boldsymbol{0}}\right].$$

Obviously the triangular matrix $\boldsymbol{S}_2$ is one and the same as the one obtained by reducing the augmented matrix $\left[\frac{\boldsymbol{X}}{\boldsymbol{Y}}\right]$. This means that the least squares estimates of the coefficients of an autoregressive model obtained by pooling two consecutive time series can be obtained quite easily. The procedure by Ozaki and Tong [8] assumes the zero initial and end conditions for each block. It is one of the advantages of the present procedure that the fitting of an autoregressive model is realized with the initial condition given by its preceding block. Another advantage is that it can be applied to the situation where the mean value of the process varies between blocks.

The minimum AIC procedure for the fitting of a locally stationary autoregressive model is summarized as follows. Assume that a set of data $\{x_n;\ n=1,\cdots,N\}$ is given. Let AIC$(\boldsymbol{S})$ and AR$(\boldsymbol{S})$ denote the minimum of AIC$(m)$ among $m=0,1,\cdots,K$ and the minimum AIC estimate of the autoregressive model obtained through the matrix $\boldsymbol{S}$, respectively.

1) Set the upper limit $K$ of the order of autoregressive models and choose $M$, the length of the basic span of data.

2) (i) Construct the $M\times(K+2)$ matrix

$$\boldsymbol{X}=\begin{bmatrix} 1 & x_K & \cdots x_1 & x_{K+1} \\ 1 & x_{K+1} & \cdots x_2 & x_{K+2} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{K+M-1} & \cdots x_M & x_{K+M} \end{bmatrix}.$$

(ii) Reduce the matrix $\boldsymbol{X}$ to an upper triangular matrix $\boldsymbol{S}_0$.

(iii) Determine the minimum AIC autoregressive model $\mathrm{AR}(\boldsymbol{S}_0)$ and put $\mathrm{AIC}_0 \equiv \mathrm{AIC}(\boldsymbol{S}_0)$.

(iv) Set $N_0 = M$.

3) (i) Construct the $M \times (K+2)$ matrix

$$\boldsymbol{Y}=\begin{bmatrix} 1 & x_{N_0+K} & \cdots x_{N_0+1} & x_{N_0+K+1} \\ 1 & x_{N_0+K+1} & \cdots x_{N_0+2} & x_{N_0+K+2} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{N_0+K+M-1} & \cdots x_{N_0+M} & x_{N_0+K+M} \end{bmatrix}.$$

(ii) Reduce the matrix $\boldsymbol{Y}$ to an upper triangular matrix $\boldsymbol{S}_1$.

(iii) Determine the minimum AIC autoregressive model $\mathrm{AR}(\boldsymbol{S}_1)$ and put $\mathrm{AIC}_1 \equiv \mathrm{AIC}(\boldsymbol{S}_1) + \mathrm{AIC}_0$.

4) (i) Construct the $2(K+2) \times (K+2)$ matrix

$$\boldsymbol{Z}=\left[\frac{\boldsymbol{S}_0}{\boldsymbol{S}_1}\right].$$

(ii) Reduce the matrix $\boldsymbol{Z}$ to an upper triangular matrix $\boldsymbol{S}_2$.

(iii) Determine the minimum AIC autoregressive model $\mathrm{AR}(\boldsymbol{S}_2)$ and put $\mathrm{AIC}_2 = \mathrm{AIC}(\boldsymbol{S}_2)$.

5) If $\mathrm{AIC}_1$ is less than $\mathrm{AIC}_2$, replace the current model $\mathrm{AR}(\boldsymbol{S}_0)$ by $\mathrm{AR}(\boldsymbol{S}_1)$, overwrite $\boldsymbol{S}_1$ on $\boldsymbol{S}_0$ and set $\mathrm{AIC}_0 = \mathrm{AIC}_1$. If $\mathrm{AIC}_1$ is greater than or equal to $\mathrm{AIC}_2$, replace the current model $\mathrm{AR}(\boldsymbol{S}_0)$ by $\mathrm{AR}(\boldsymbol{S}_2)$, overwrite $\boldsymbol{S}_2$ on $\boldsymbol{S}_0$ and set $\mathrm{AIC}_0 = \mathrm{AIC}_2$.

6) If $N_0 + M$ equals to $N$ stop the procedure, otherwise replace $N_0$ by $N_0 + M$ and go back to step 3).

## 6. Numerical examples

The procedure of fitting autoregressive models to non-stationary time series was applied to two artificially generated non-stationary time series. The first series is generated by adjoining the realizations of three types of autoregressive models

$$x_n = 1.6x_{n-1} - 1.25x_{n-2} + 0.35x_{n-3} + \varepsilon_n ,$$

$$x_n = 1.1x_{n-1} - 1.00x_{n-2} + 0.35x_{n-3} + \varepsilon_n$$

and

Table 1. Comparison of two models by AIC's at each block

| Block | New data | $AIC_1$ | $AIC_2$ | Decision |
|---|---|---|---|---|
| 1 | $\{x_1,\cdots, x_{105}\}$ | $-9.34$ | — | |
| 2 | $\{x_{106},\cdots, x_{205}\}$ | 0.02 | $-4.51$ | New data pooled |
| 3 | $\{x_{206},\cdots, x_{305}\}$ | $-6.02$ | $-11.26$ | New data pooled |
| 4 | $\{x_{306},\cdots, x_{405}\}$ | 16.24 | 47.67 | **Model switched** |
| 5 | $\{x_{406},\cdots, x_{505}\}$ | 30.83 | 25.21 | New data pooled |
| 6 | $\{x_{506},\cdots, x_{605}\}$ | 9.17 | 7.66 | New data pooled |
| 7 | $\{x_{606},\cdots, x_{705}\}$ | 11.28 | 29.83 | **Model switched** |
| 8 | $\{x_{706},\cdots, x_{805}\}$ | 3.26 | $-0.68$ | New data pooled |
| 9 | $\{x_{806},\cdots, x_{900}\}$ | 1.32 | $-3.26$ | New data pooled |

$$x_n=0.8x_{n-1}-0.82x_{n-2}+0.40x_{n-3}+\varepsilon_n ,$$

each of length 300, where $\varepsilon_n$ denotes Gaussian white noise with zero mean and variance 1.0. The length of the basic span was chosen to be 100 and the upper limit of the order was arbitrarily set to 5. Table 1 shows AIC's of two competing models, $AIC_1$ and $AIC_2$, and the decision made by our procedure. First, third order autoregressive model was obtained for the interval [6,105]. When another 100 data were obtained two AIC's ($AIC_1$ and $AIC_2$) were compared. The former is the AIC of the succession of two local models fitted to [6,105] and [106,205] and the latter is that of the autoregressive model fitted to the whole interval [6,205]. Since $AIC_2$ was less than $AIC_1$, the latter was adopted

Table 2. Selected model at each interval

| Intervals | | Fitted models | Residual variances | Entropies |
|---|---|---|---|---|
| from | to | | | |
| 6 | 105 | $x_n=1.581x_{n-1}-1.189x_{n-2}+0.277x_{n-3}+\varepsilon_n$ | 0.8408 | $-0.01447$ |
| 6 | 205 | $x_n=1.532x_{n-1}-1.112x_{n-2}+0.233x_{n-3}+\varepsilon_n$ | 0.9394 | $-0.01012$ |
| 6 | 305 | $x_n=1.565x_{n-1}-1.139x_{n-2}+0.267x_{n-3}+\varepsilon_n$ | 0.938 | $-0.00869$ |
| 306 | 405 | $x_n=1.091x_{n-1}-0.839x_{n-2}+0.304x_{n-3}+\varepsilon_n$ | 1.215 | $-0.03241$ |
| 306 | 505 | $x_n=1.123x_{n-1}-0.891x_{n-2}+0.282x_{n-3}+\varepsilon_n$ | 1.153 | $-0.02153$ |
| 306 | 605 | $x_n=1.098x_{n-1}-0.896x_{n-2}+0.285x_{n-3}+\varepsilon_n$ | 1.037 | $-0.01190$ |
| 606 | 705 | $x_n=0.601x_{n-1}-0.713x_{n-2}+0.240x_{n-3}+\varepsilon_n$ | 0.9572 | $-0.02715$ |
| 606 | 805 | $x_n=0.728x_{n-1}-0.744x_{n-2}+0.304x_{n-3}+\varepsilon_n$ | 0.9215 | $-0.00820$ |
| 606 | 900 | $x_n=0.738x_{n-1}-0.816x_{n-2}+0.341x_{n-3}+\varepsilon_n$ | 0.9379 | $-0.00670$ |

as the current model, i.e., the data were considered to be homogeneous and the two blocks were pooled. The third block was also pooled. When the fourth block was added, the $AIC_1$ became less than $AIC_2$. Thus the interval [6,405] was divided into two sub-intervals, [6,305] and [306,405], and the current model was replaced by the one fitted to the sub-interval [306,405]. This means that the procedure detected the change of the model at the 305th data point. Continuing further, the procedure also detected the change at the 605th data point. The minimum AIC of the stationary model fitted to the whole span [6,900] is 176.80, while that of the locally stationary model adopted by our procedure is $-3.36$. This shows that the fit of the model was greatly improved by the adoption of the locally stationary model. Table 2 shows the selected partition of the data and the model fitted to each sub-interval. The table also contains the value of the entropy of the true model with respect to the fitted model defined by

$$-\frac{1}{2}\log 2\pi\hat{\sigma}^2-\frac{1}{2\hat{\sigma}^2}\left(\rho_0-2\sum_{i=1}^{m}\hat{a}_i\rho_i+\sum_{i=1}^{m}\sum_{j=1}^{m}\hat{a}_i\hat{a}_j\rho_{|i-j|}\right),$$

where $m$, $\hat{\sigma}^2$ and $\hat{a}_i$ $(i=1,\cdots,m)$ are the order, the innovation variance and the autoregression coefficients of the fitted model and $\rho_i$ $(i=0,1,\cdots,m)$ is the theoretical autocorrelation function defined by the true model [4]. The entropy is non-positive and we consider that the greater

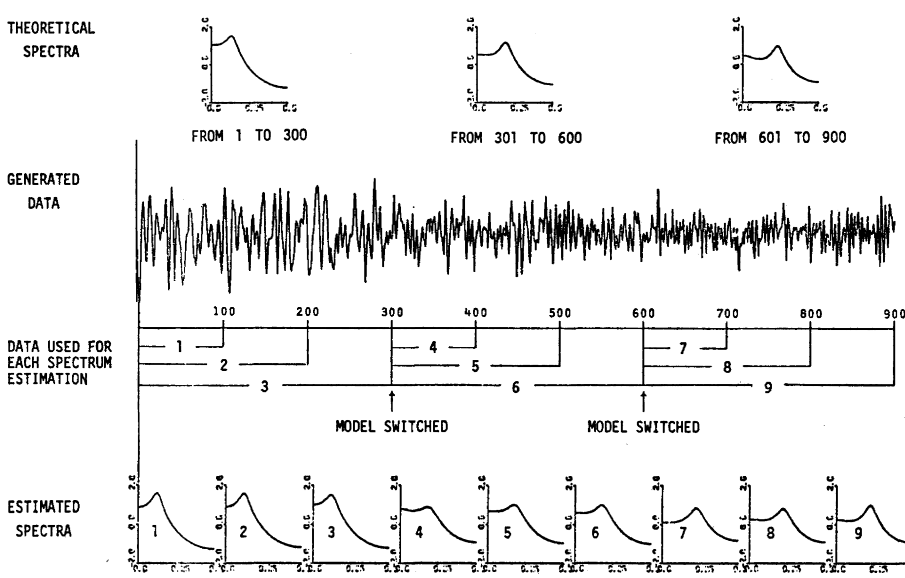


Fig. 1. Power Spectral Density Functions Estimation
MAICE within the locally stationary autoregressive models at each block is compared with the theoretical spectrum. Data length $N=900$, basic span $M=100$, highest order $K=5$.

the value of entropy, the better the fit of the estimated model. From the table, we know that the amount of entropy was reduced when the additional data were pooled.

Fig. 1 illustrates the record of the generated data, the theoretical spectra of three models and the estimated spectra through the fitted models. The figure shows two significant merits of our procedure: If the additional data are obtained from the same model as that of the preceding data then the new data will be pooled and the current model will be refined. On the other hand, if the statistical structure of the process is significantly changed, the procedure will detect and follow the change of the structure by switching the model.

The second series was generated by a seventh order autoregressive model with time-variant coefficients. The autoregressive coefficients $a_m(t)$, $(m=1,\cdots,7)$ were determined so that the characteristic roots of the equation $s^7-a_1(t)s^6-\cdots-a_6(t)s-a_7(t)=0$ were given by 0.4, $-0.6 \pm 0.6989i$, $0.86 \pm 0.4005i$ and $0.35\{1+(1.4+c(t_n))\cos t_n \pm (3+\sin t_n)i\}$, where $t_n$ and $c(t)$ were defined by $t_n=0$, for $n<500$, $m\pi/60$, for $30(m-1)$

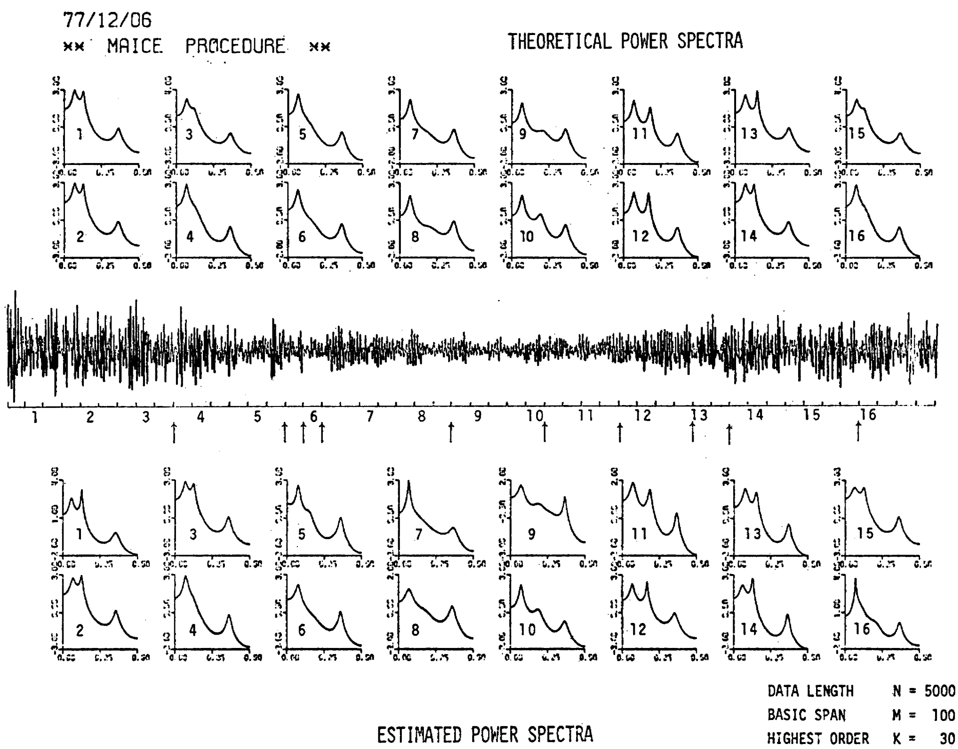


Fig. 2. Power Spectral Density Functions Estimation
MAICE within the locally stationary autoregressive models at each block is compared with the theoretical spectrum. ↑ shows that the model was switched there.

$\leqq n-500 \leqq 30m$ and $c(t)=0.5\sin(2t+\pi/6)$. The process is stationary up to the 500th data point, and after that the model gradually changes a pair of its characteristic roots at every 30 points with the period of 3600. Fig. 2 illustrates a realization of the process, the theoretical spectrum and the spectrum estimated through the fitted autoregressive model at each interval, where the theoretical spectra are defined as the spectra of the stationary time series with coefficients of the characteristic equations. The basic span was chosen to be 100 and the upper limit of the order was set to 30. The procedure detected the change of the model at $n=910$, 1510, 1610, 1710, 2410, 2910, 3310, 3710, 3910 and 4610. The figure shows that the procedure detected the significant changes of the spectrum.

## 7. Remarks

The present procedure will be useful for the automatic detection or monitoring of the change of spectral characteristics of industrial or natural processes. The generalization of the present procedure to the multidimensional case is straightforward. This extension will be useful for the implementation of adaptive control of non-stationary processes.

## Acknowledgements

Our interest in the use of the Householder transformation for time series analysis was inspired by a lecture on least squares methods given by Mr. M. Sibuya of Japan IBM. For this we are very grateful. Thanks are also due to Miss F. Tada of the Institute of Statistical Mathematics for her assistance during the preparation of the necessary programmings. The authors also thank the referees for their comments on the manuscript.

The present work was partly supported by a grant from the Ministry of Education, Science and Culture.

THE INSTITUTE OF STATISTICAL MATHEMATICS

## REFERENCES

[1] Akaike, H. (1969). Power spectrum estimation through autoregressive model fitting, *Ann. Inst. Statist. Math.*, **21**, 407-419.

[2] Akaike, H. (1970). Statistical predictor identification, *Ann. Inst. Statist. Math.*, **22**, 203-223.

[3] Akaike, H. (1976). Canonical correlation analysis of time series and the use of an information criterion, in *System Identification: Advance and Case Studies*, R. K. Mehra and D. G. Lainiotis, eds., Academic Press, New York.

[4] Akaike, H. (1976). On entropy maximization principle, *Symposium on Applications of*

*Statistics*, Dayton, Ohio.

[5] Golub, G. H. (1965). Numerical methods for solving linear least squares problems, *Numer. Math.*, **7**, 206-216.

[6] Kitagawa, T. (1963). Estimation after preliminary tests of significance, *Univ. Calif. Pub. Statist.*, **3**, 147-186.

[7] Kitagawa, T. (1950). Successive process of statistical inferences, *Mem. Fac. Sci. Kyushu University*, Series A, **15**, 139-180.

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