Adaptive Order Selection with Aid of Genetic Algorithm

Norikazu IKOMA and *Hiroshi MAEDA*

Department of Computer Science,
Faculty of Engineering, Kyushu Institute of Technology
1-1 Sensui-Cho, Tobata-Ku, Kita-Kyushu 804-8550, JAPAN
(E-mail : ikoma@comp.kyutech.ac.jp)

Abstract

A new method to estimate nonstationary power spectrum with adaptive selection of autoregressive order is proposed. Time-varying PARCOR (partial autocorrelation coefficient) and AR (autoregressive) order are estimated from time series data. The data are assumed to be observations of vibration that contain abrupt change of spectrum due to arrivals of different signal, structural changes of vibrating object, etc. The model that consists of an autoregressive model with time-varying PARCORs and time-varying order is used. The time-varying PARCORs are estimated by Monte Carlo filter, and the time-varying order is estimated by genetic algorithm. An application to analysis of seismic wave data is reported.

Key words: model selection, nonstationary time series, genetic algorithm

1 Introduction

Autoregressive(AR) model is one of the most important tool for spectral analysis of stationary time series data. In case of nonstationary spectral analysis, covariance nonstationary AR models such as time-varying coefficient AR model [7] has been investigated. In these conventional researches of nonstationary spectral analysis, it is usually assumed that the order of AR model is constant through the data unless we know a priori information about the structural change of data. However, in general, many of time series data sets having nonstationary spectrum can be considered to have abrupt change of spectral structure.

For the example of such data sets, we can see an observation of seismic wave. Seismic wave consists of three signals. The first is background noise, the second is P(primary)-wave, and the last is S(secondary)-wave. The arrival of these signals with time delay should be considered in the modeling of spectral analysis. The second example is concerned with an observation of vibrating object. When a break of the object occurred, spectral feature of the vibration will be changed due to the change of degree of freedom. In the situations mentioned above, i.e. abrupt change of data due to the arrival of different signal, the structural change of vibrating object, etc, the analysis of nonstationary spectrum with the assumption of constant order is insufficient.

Motivated by this fact, we propose a new model to have time-varying order of covariance nonstationary AR model with aid of genetic algorithm [3]. In this paper, we firstly define the model after summarizing the properties of covariance nonstationary AR model and PARCOR (partial autocorrelation coefficient). Secondly, an estimation method for the model is shown. The estimation method is based on non-Gaussian nonlinear state space modeling and its state estimation, and that is called Monte Carlo filter [5]. Time-varying PARCORs are estimated by Monte Carlo filter, and time-varying order of AR model is estimated by genetic algorithm. As an application of the method, an analysis of seismic wave data has been reported.

2 Model

We firstly show the aim of analysis, and secondly properties of nonstationary AR model that are required for the model definition are summarized. After that, we will define a new model for the analysis.

2.1 The aim of analysis

Given a set of time series data denoted by

$$Y_N = \{y(1), y(2), \cdots, y(N)\} \tag{1}$$

where ensemble mean of data, $\mathrm{E}\{y(t)\}$ ($\mathrm{E}\{\cdot\}$ denotes expectation), is assumed to be constant in time. We can make the ensemble mean be equal to zero without loss of generality. The nonstationary properties of the data are represented by time-varying autocovariance function denoted by

$$C_t(k) = \mathrm{Cov}\{y(t), y(t+k)\} = \mathrm{E}\{y(t)y(t+k)\}.$$

Since the Fourier transform of autocovariance function is the power spectrum in stationary case, it is natural to assume in nonstationary case that Fourier transform of time-varying autocovariance function for each time is defined as time-varying power spectrum $p(\omega, t)$, where ω denotes angular frequency in $[0, \pi]$.

The aim of analysis is to estimate the time-varying power spectrum from the data. Here, we also assume time-varying power spectra have not only time-varying properties represented by time-varying AR coefficients, but also structural change that is governed by the change of AR order. The time-varying AR order corresponds to the change of degree of freedom of oscillation.

To define a model for this aim of analysis, we will briefly summarize the properties of nonstationary AR model and PARCOR, which are required for the model definition, in the following subsections. After that, we will define the model with time-varying AR order based on the properties.

2.2 Covariance nonstationary AR-model

Covariance nonstationary AR model of order p is defined as follows,

$$y(t) = \sum_{j=1}^{p} a_j^p(t) y(t-j) + \varepsilon_p(t), \tag{2}$$

where $\varepsilon_p(t) \sim N(0, \sigma_p^2)$, i.i.d, and $a_1^p(t), a_2^p(t), \cdots, a_p^p(t)$ are time-varying AR coefficients of order p.

Time-varying power spectrum of the model (2) is derived as

$$p(\omega, t) = \frac{\sigma_p^2}{\left|1 - \sum_{j=1}^{p} a_j^p(t) e^{-ij\omega}\right|^2}, \tag{3}$$

where ω denotes angular frequency in $[0, \pi]$. Thus, we can obtain the time-varying power spectrum by estimating the time-varying AR coefficients from a set of time series (1).

In the estimation of time-varying AR coefficients, notice that the number of AR coefficients, $N \times p$, is greater than the number of data N. Then a simple application of least squares method or maximum likelihood method will fail, and some restriction among the coefficients are required for the estimation.

In the conventional research, for example, time-varying coefficient AR model [7] assumes the smoothness of AR coefficients

$$\Delta^k a_j^p(t) = v_j(t), \qquad v_j(t) \sim N(0, \tau_p^2), \text{i.i.d} \tag{4}$$

where Δ denotes difference operator with respect to time index t, such that $\Delta a_j^p(t) = a_j^p(t) - a_j^p(t-1)$, and k takes value 1 or 2 in practice. In this formulation, the estimation of time-varying AR coefficients is done by Kalman filter algorithm, where (2) is observation equation and (4) is system equation with state vector defined by time-varying AR coefficients.

The optimal order of time-varying coefficient AR model, $\hat{p}$, is determined by minimizing AIC [1],

$$\text{AIC} = -2l(\hat{\theta}) + 2(p+1), \tag{5}$$

where $l(\cdot)$ is log-likelihood defined by

$$l(\theta) = \sum_{t=1}^{N} \log p(y_t | Y_{t-1}). \tag{6}$$

θ is defined as a vector consists of hyper-parameters such as σ_p^2, τ_p^2, and so on. $\hat{\theta}$ denotes the value of θ that maximize log-likelihood [2].

For the estimation of variance parameters in θ, the use of augmented state vector, incorporated the original state vector and the variance parameters, has been proposed by [6].

2.3 PARCOR

There are relationships between AR coefficients of different order p and $p-1$ as follows,

$$a_j^p = a_j^{p-1} - a_p^p a_{p-j}^{p-1}, \qquad (j = 1, 2, \cdots, p-1). \tag{7}$$

The coefficients, $a_1^1, a_2^2, \cdots, a_P^P$, are called PARCOR (partial autocorrelation coefficient).

Variances also have the relationship

$$\sigma_p^2 = \sigma_{p-1}^2 \left\{1 - (a_p^p)^2\right\}. \tag{8}$$

Thus, by knowing PARCOR and variance σ_0^2, we can obtain all AR coefficients and variance for all orders, $1, 2, \cdots, P$. When we change the AR order, we only need to calculate AR coefficients from PARCOR, i.e. we have no need to store AR coefficients of each order. Figure 1 shows the relationship mentioned here. PARCOR and variance σ_0^2 are shown in box in the figure to emphasize them as the essential factor. In the following section, we use this property in model definition by making state vector denoted by PARCOR and variance.

2.4 Model definition

We propose a new model by assuming the smoothness of time-varying PARCOR,

$$\Delta a_p^p(t) = v_p(t), \qquad v_p(t) \sim N(0, \tau^2), \text{i.i.d.}, \quad , p = 1, 2, \cdots, P$$

observation variance changes with time,

$$\Delta \sigma^2(t) = u(t), \qquad u(t) \sim N(0, \mu^2), \text{i.i.d}, \tag{9}$$

and AR order p is varying with time as follows,

$$\Delta p(t) = \nu(t). \tag{10}$$

	σ_0^2				
AR(1)	a_1^1	σ_1^2			
AR(2)	a_1^2	a_2^2	σ_2^2		
AR(3)	a_1^3	a_2^3	a_3^3	σ_3^2	
$\vdots$	$\vdots$		$\ddots$	$\ddots$	
AR(P)	a_1^P	a_2^P	a_3^P	$\cdots$ a_P^P	σ_P^2

Figure 1: Relationship between PARCOR, AR coefficients, and variances

where $\nu(t)$ is discrete random variable taking the value as follows,

$$\nu(t) = \begin{cases} 0 & \text{with prob.} \quad \alpha, \\ -p(t) \sim P - p(t) & \\ \quad \text{with prob.} & (1-\alpha)/P. \end{cases} \tag{11}$$

Then, the proposed model can be written in state space representation

$$\begin{cases} \mathbf{x}_t = \mathbf{x}_{t-1} + \mathbf{v}_t \\ y_t = h(\mathbf{x}_t, Y_{t-1}) + \varepsilon_t \end{cases} \tag{12}$$

by denoting the state vector, $\mathbf{x}_t$, and system noise vector, $\mathbf{v}_t$, as follows,

$$\mathbf{x}_t = \left[a_1^1(t), a_2^2(t), \cdots, a_P^P(t), \sigma^2(t), p(t)\right]^T, \tag{13}$$

$$\mathbf{v}_t = [v_1(t), v_2(t), \cdots, v_P(t), u(t), \nu(t)]^T. \tag{14}$$

Here, the nonlinear function $h(\cdot,\cdot)$ is as follows,

$$h(\mathbf{x}_t, Y_{t-1}) = \sum_{j=1}^{p(t)} a_j^{p(t)}(t) y_{t-j} \tag{15}$$

and AR coefficients in the above can be obtained from PARCOR by (7). Hyper-parameter vector denoted by θ consists of τ^2 , μ^2 , and α.

The state vector, $\mathbf{x}_t$, takes fundamental role in the estimation mentioned the following section. Elements of the state vector are illustrated in Figure 2.

a_1^1	a_2^2	$\cdots$	a_P^P	σ_0^2	p

Figure 2: Elements of state vector

3 Estimation

We propose to estimate the state by Monte Carlo filter [5] for continuous variables such as PARCOR and variance, by genetic algorithm [3] for discrete variable, the AR order.

3.1 Monte Carlo filter

Since the proposed model contains nonlinear formula in observation equation (2) with respect to the state $\mathbf{x}_t$, non-Gaussian nonlinear filtering is required for the state estimation. As the filtering method, we use Monte Carlo filter(MCF) [5]. It is summarized in the followings.

The key idea of MCF is an approximation of non-Gaussian distribution on state vector by its sample particles as,

$$\left\{\mathbf{p}_1^{(t)}, \mathbf{p}_2^{(t)}, \cdots, \mathbf{p}_M^{(t)}\right\} \sim p(\mathbf{x}_t|Y_{t-1}) \tag{16}$$

$$\left\{\mathbf{f}_1^{(t)}, \mathbf{f}_2^{(t)}, \cdots, \mathbf{f}_M^{(t)}\right\} \sim p(\mathbf{x}_t|Y_t) \tag{17}$$

where, M denotes the number of particles.

By using these particles, filtering and one-step-ahead prediction are done as follows. For the one-step-ahead prediction, calculate the particles by

$$\mathbf{p}_i^{(t)} = \mathbf{f}_i^{(t-1)} + \mathbf{v}_i^{(t)} \tag{18}$$

where $\mathbf{v}_i^{(t)}$ is a vector $\mathbf{v}_i^{(t)} \sim \mathbf{v}_t$, i.e. the realized value of random vector $\mathbf{v}_t$ in equation (14).

Next, for the filtering, likelihood of each particle is firstly calculated as follows,

$$\alpha_i^{(t)} = p(y_t|\mathbf{p}_i^{(t)}). \tag{19}$$

Since observation noise ε_t is Gaussian, $\alpha_i^{(t)}$ can be calculated by Gaussian pdf of $y_t - h(\mathbf{p}_i^{(t)}, Y_{t-1})$ with zero mean and variance $\sigma^2(t)$

Secondly for the filtering, resample the particles by according to the probabilities

$$\mathbf{f}_i^{(t)} = \begin{cases} \mathbf{p}_1^{(t)} & \text{with prob.} \quad \alpha_1^{(t)}/\sum\alpha, \\ \vdots & \\ \mathbf{p}_M^{(t)} & \text{with prob.} \quad \alpha_M^{(t)}/\sum\alpha, \end{cases} \tag{20}$$

where, $\sum\alpha = \sum_{j=1}^M \alpha_j^{(t)}$.

Log-likelihood of the model is approximately obtained by

$$\begin{aligned} l(\theta) &= \sum_{t=1}^N \log p(y_t|Y_{t-1}) \\ &\simeq \sum_{t=1}^N \log\left(\sum_{j=1}^M \alpha_j^{(t)}\right) - N\log M. \end{aligned} \tag{21}$$

Table 1: Relationship between GA and MCF

GA	MCF
individual	particle
mutation	one-step-ahead prediction
fitness	likelihood
selection	resampling

3.2 Genetic algorithm

By regarding the filtering and the one-step-ahead prediction of MCF as the selection and the mutation operations of genetic algorithm(GA)[3] respectively, the calculation of MCF is equivalent to the process of GA except the crossover operation. The relationship between MCF and GA are shown in Table 3.2. The research [4] has been shown this relationship, and proposed to incorporate the genetic operation, such as crossover, to MCF calculation. We use this idea into our model as follows.

The state vector shown in Figure 2 can be considered as "gene" in the interpretation of genetic algorithm. Among items of state vector $\mathbf{x}_t$, only the order $p(t)$ is a discrete random variable, and all other items are continuous. Since GA is intended to the optimization of discrete variables, we incorporated GA operations to the order $p(t)$. Other items of state vector $\mathbf{x}_t$ is calculated by MCF. For the coding of $p(t)$, several method, such as binary code, Gray code, and $0-1$ coding, can be considered.

4 Data analysis

As an example of application of the proposed model, seismic wave data shown in Figure 3 have been analized. Arrivals of P-wave and S-wave can obviously seen in the data. Variance of the data is also changing with time.

According to the proposed method, we have estimated the time-varying AR order as shown in Figure 4. Also time-varying variance has been estimated and shown in Figure 5. By looking Figure 4, the order increases at the time where P-wave arrives, and also increases more at S-wave arrives. The order is slowly decreasing as amplitude of S-wave going small. From Figure 5, we can see that the estimated variance follows the change of data variance.

Estimated result of power spectra varying with time are shown in Figure 6. They are plotted from the estimated AR coefficients and AR order both varying with time. By looking the result, simple structure of power spectrum (represented by low order) can be seen at background noise, and for P-wave and S-wave, more complicated spectra (represented by high order) are obtained.

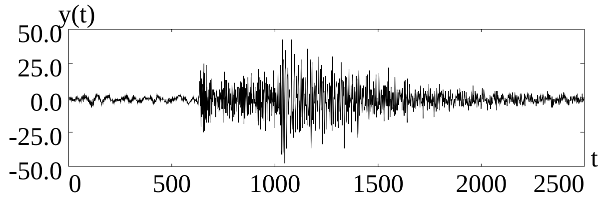

Figure 3: Seismic wave data

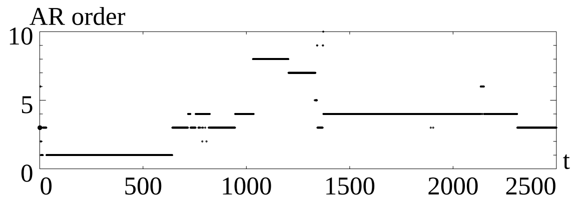

Figure 4: Estimated AR order for seismic wave data

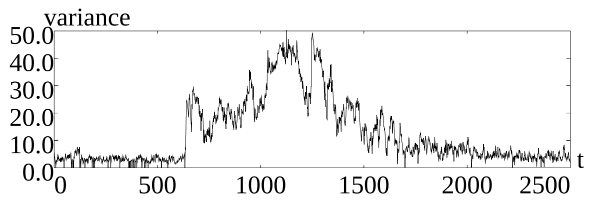

Figure 5: Estimated variance for seismic wave data

5 Conclusion

A new method to analyze covariance nonstationary AR process data has been proposed. In the method, a model based on nonstationary AR model with time-varying AR coefficients parametrized by PARCOR is proposed. AR order is also varying with time. The model can be written in state space representation with state vector that consists of PARCOR, observation variance and AR order. Estimation of the state is done by Monte Carlo filtering for PARCOR and variance, by genetic algorithm for AR order. An application of the proposed method to the analysis of seismic wave data has been demonstrated.

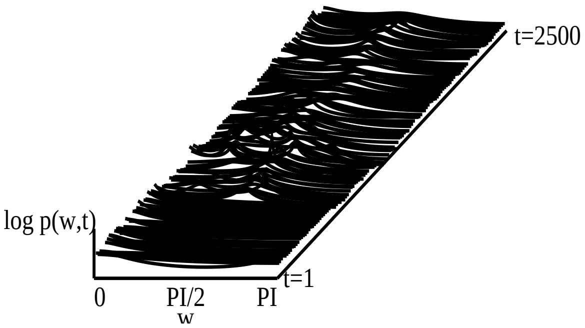

Figure 6: Estimated power spectra for seismic wave data

Acknowledgement

This research was supported by the Grant-in-Aid for Encouragement of Young Scientists by the Japanese Ministry of Education, Science, Sports and Culture.

References

[1] H.Akaike: A new look at the statistical model identification, *IEEE Trans.* AC-19, No.6, pp.716-723, 1974.

[2] H.Akaike: Likelihood and Bayesian procedure, *Bayesian Statistics*, University Press, Valencia, Spain, pp.143-166, 1980.

[3] D.E.Goldberg: *Genetic algorithm in search, optimization and machine learning*, Addison-Wesley MA, 1989.

[4] T.Higuchi: Monte Carlo filter using the Genetic algorithm operators, *Journal of Statistical Computation and Simulation*, Vol.59, No.1, pp.1-23, 1997.

[5] G.Kitagawa: Monte Carlo filter and smoother for non-Gaussian nonlinear state space model, *Journal of Computational and Graphical Statistics*, 5, pp.1-25, 1996.

[6] G.Kitagawa: A self-organizing stat-space model, *Journal of the American Statistical Assoc.*, Vol.93, No.443, pp.1203-1215, 1998.

[7] G.Kitagawa and W.Gersch: A smoothness priors time-varying AR coefficient modeling of nonstationary covariance time series, *IEEE Trans.* AC-30, No.1, pp.48-56, 1985.