Adaptive Estimation of Power Spectrum by using Genetic Algorithm

Norikazu IKOMA & 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 make an adaptive estimation of nonstationary power spectrum is proposed. The method uses a new model based on time-varying coefficient autoregressive (AR) model in which order of autoregression is also varying with time. Nonstationarity of power spectrum can be obtained by estimating the time-varying coefficients, and abrupt change of the structure of spectrum can be estimated by the time-varying order. The model is written in state space representation with system model that defines smoothness of time-varying parameters and observation model consists of the time-varying parameters AR model. Monte Carlo filter and genetic algorithm, they are very similar except crossover, are used for estimation of the AR coefficients and the order, respectively. Simulational experiment shows the estimation result by the proposed method.

Key Words: Adaptive estimation, Nonstationary power spectrum, Monte Carlo, Genetic algorithm

1. INTRODUCTION

Spectral analysis [12] is widely used in the field of engineering. The methods for stationary spectral analysis have been well established and used in practice. On the other hand, spectral analysis methods for nonstationary time series do not always produce satisfactory results and thus that still remain as a topic to investigate. The use of time-varying coefficients of autoregressive model is one of the possible way to do the investigation of nonstationary spectral analysis.

Time-varying coefficient autoregressive model was proposed by G.Kitagawa and W.Gersch in 1985 [10] in a context of time series analysis by statistical method. Although there are early researches of similar methods in the area of recursive estimation e.g. by R.C.K.Lee in 1964 [11], however, likelihood is effectively used for the estimation of trade-off parameters in the research by G.Kitagawa. The use of likelihood is originally proposed by H.Akaike in 1980 [2] as a problem of penalized least square, but not recursive least square. This is a successful example to introduce the idea of recursive estimation to the objective determination of hyper-parameters(trade-off parameters).

In time-varying coefficient autoregressive model, the order of autoregression can be determined by Akaike information criterion(AIC) [1] when the order can be assumed to be constant through the data. However, when we need to assume the order changing with time in the data, and we have no a priori information about the change of order, we have to examine all combination of changes of the order through the data in order to estimate the AIC best combination of the orders. The combination will easily explode even by taking small order, such as 10, as the maximum candidate order. For the estimation of the autoregressive order changing with time, the idea of recursive estimation is considered as a effective way to lessen the combinational explosion.

Recently, a new method called Monte Carlo filtering has been proposed by G.Kitagawa(1996) [8] for non-Gaussian nonlinear state space modeling. In Monte Carlo filtering, particles that are considered as realizations from non-Gaussian distribution are used to the approximation of non-Gaussian distribution. Since the filtering procedure is done by manipulating each particle, computational cost of Monte Carlo filtering is the order of the number of particles. This is a reduction of computational cost from the exponential order, which is a cost of computation by a method proposed by G.Kitagawa(1987) [7] to approximate the distribution by partially linear function, to the number of particles. Thus non-Gaussian nonlinear filtering is now more practical than the former times.

By using Monte Carlo filter, it becomes possible to investigate a nonlinear state space model that has time-varying autoregressive coefficients and autoregressive order together. In our research, an adaptive method to estimate time-varying autoregressive coefficients and autoregressive order together from data is proposed. The method is based on time-varying coefficients autoregressive model, and the order of autoregression is also varying with time. Estimation of time-varying coefficients is done by Monte Carlo filtering, and time-varying order of autoregression is estimated

by genetic algorithm.

Let characteristic root of AR(2) (second order AR model) be denoted by $re^{\pm i\theta}$, θ and r are equal to peak frequency and bandwidth of power spectrum, respectively. By using this property, and regarding the proposed model as a sequential system consists of AR(2)s, the AR coefficients are parametrized by pairs of peak parameters θ_k and r_k.

State vector consists of P pairs peak parameters and corresponding P-bits gene. Peak parameter is selected and used for the calculation of AR coefficients when the bit corresponding to the peak is equal to 1. Thus the number of value-1-bit is the number of peaks, where twice of number of peaks is equal to the AR order. Smooth change of peak parameters are represented in system model by system noise, and the change of gene is governed by genetic operation such as mutation and crossover in order to select the better combination of peak parameters.

By a simulational experiment, the efficiency of proposed method is checked.

2. MODEL

Problem in this research is firstly defined. Necessity of smoothness prior is mentioned related to the problem definition. After that, a new model to solve the problem is defined, and smoothness prior for the model is shown.

Problem

Let time series data be given by

$$Y_N = \{y_1, y_2, \cdots, y_N\}. \tag{1}$$

Assume that power spectrum of the data has several peaks and the peaks are smoothly changing, and the number of peaks is not constant among the series.

Problem we concern here is to fit the autoregressive model with varying coefficients, $a_{j,t}$, and even order, $2p_t$, denoted by

$$y_t = \sum_{j=1}^{2p_t} a_{j,t}\ y_{t-j} + \varepsilon_t \tag{2}$$

where $\varepsilon_t \sim N(0, \sigma^2)$, i.i.d., and thus we obtain time-varying power spectrum of the model as follows,

$$p(\omega, t) = \frac{\sigma^2}{\left| 1 - \sum_{j=1}^{2p_t} a_{j,t}\ e^{-i\omega j} \right|^2}. \tag{3}$$

Smoothness Prior

In the model fitting, parameters to be estimated from data are time-varying coefficients $a_{j,t}$, time-varying order p_t, and variance σ^2. Note that the number of parameters to be estimated is $\sum_{t=1}^{N} 2p_t + N + 1$ and is, in general, greater than the number of data, N. It is so even in the case the order is constant with time. Consider the case of constant order, p, for a while. To estimate the parameters in this situation, there is a research Time-varying coefficient autoregressive model proposed by G.Kitagawa and W.Gersch in 1985 [10] in a context of time series analysis by statistical method.

Although there are early researches of recursive estimation e.g. by R.C.K.Lee in 1964 [11], however, likelihood is effectively used for the estimation of trade-off parameters (hyper-parameters) in the research by G.Kitagawa and W.Gersch. The use of likelihood is originally proposed by H.Akaike in 1980 [2] as a problem of penalized least square, and developed by G.Kitagawa [6] as recursive least square with Kalman filter.

According to [10], by assuming smoothness prior of time-varying coefficients, in which smoothness is with respect to time t, such as

$$\Delta a_{j,t} \sim N(0, \tau^2), \qquad j = 1, \cdots, p, \tag{4}$$

where Δ is difference operator defined by

$$\Delta y_t \equiv y_t - y_{t-1}, \qquad \Delta\Delta^n = \Delta^{n+1} \tag{5}$$

we can estimate all parameters from data.

Model Definition

We use the following autoregressive model with time-varying coefficients

$$y_t = \sum_{j=1}^{2P} a_j(\mathbf{x}_t)\ y_{t-j} + \varepsilon_t \tag{6}$$

where $\varepsilon_t \sim N(0, \sigma^2)$, i.i.d. . Time-varying coefficients denoted by $a_j(\mathbf{x}_t)$ are parametrized by state vector

$$\mathbf{x}_t = [\ \mathbf{r}_t,\ \theta_t,\ \mathbf{g}_t\ ]^T. \tag{7}$$

The state vector consists of a vector of bandwidth $\mathbf{r}_t$ and frequency θ_t of power spectrum peaks such that

$$\mathbf{r}_t = [\ r_{1,t},\ r_{2,t},\ \cdots,\ r_{P,t}\ ]^T \tag{8}$$

$$\theta_t = [\ \theta_{1,t},\ \theta_{2,t},\ \cdots,\ \theta_{P,t}\ ]^T, \tag{9}$$

and a vector $\mathbf{g}_t$. We call $\mathbf{g}_t$ "gene" since this is denoted by sequence of bit as follows,

$$\mathbf{g}_t = [\ g_{1,t},\ g_{2,t},\ \cdots,\ g_{P,t}\ ]^T. \tag{10}$$

Role of the gene is to specify which peak is active. When j-th bit, $g_{j,t}$, is equal to 1, corresponding peak parameters such that $r_{j,t}$ and $\theta_{j,t}$ are used. Thus actual order of autoregression in (6), denoted by $2p_t$, where p_t is equal to the number of bits of value 1, i.e.

$$p_t = \sum_{j=1}^{P} g_{j,t}. \tag{11}$$

Note that the actual order of autoregression is common to that appeared in (2).

The role of the gene to select the peaks is illustrated in Figure 1.

Since the actual order of autoregression is p_t, autoregressive coefficients that lag is greater than p_t is equal to zero, i.e. $a_j(\mathbf{x}_t) = 0$ for $j > p_t$. How to calculate the remaining coefficients is according to the manner proposed in [5]. Consider a characteristic equation of autoregressive model of order $2p_t$

$$1 - \sum_{j=1}^{2p_t} a_{j,t} B^j = 0 \tag{12}$$

where B corresponds to backward shift operator such that $By_t = y_{t-1}$. Assume that all the roots of characteristic equation are complex, and denoted by $r_j e^{\pm i\theta_j}$, then we regard r_j and θ_j of characteristic roots as the bandwidth and frequency of spectral peak. They correspond to spectral peaks of second order autoregressive models of sequentially connected system, and we estimate the peaks in this meaning. This is illustrated in Figure 2.

Smoothness Prior of The Model

Smoothness prior of the model is defined for spectral peak parameters

$$\Delta r_{j,t} \sim N(0, \tau_r^2), \tag{13}$$

$$\Delta \theta_{j,t} \sim N(0, \tau_\theta^2), \qquad j = 1, \cdots, p. \tag{14}$$

For the gene, mutation and crossover of genetic operation [3] are employed as the smoothness prior. The rates of mutation and crossover are denoted by P_m and P_c, respectively. The reason to use genetic operation as the representation of smoothness prior is based on the similarity of nonlinear state space model and genetic algorithm shown by T.Higuchi [4]. As in the next section, one-step-ahead prediction in state estimation procedure is adding noise vector into state vector. Genetic operations, mutation and crossover, are just this operation of adding noise into gene.

3. Estimation

To estimate time-varying parameters of the proposed model, general state space model is firstly summarized and transform the proposed model into it. Secondly, as a state estimation method, Monte Carlo filtering is summarized. Relationship between Monte Carlo filtering and genetic algorithm is also mentioned.

General State Space Model

General state space model that consists of system model and observation model is as follows.

$$\mathbf{x}_t = \mathbf{g}(\mathbf{x}_{t-1}, \mathbf{w}_t) \tag{15}$$

$$y_t = h(\mathbf{x}_t, \varepsilon_t) \tag{16}$$

In our model defined in previous section, (6) is the observation model (16). System model (15) represents smoothness prior (13), (14), and genetic operations for gene. System model can be written by separately as follows,

$$\mathbf{r}_t = \mathbf{r}_{t-1} + \mathbf{w}_{r,t}, \tag{17}$$

$$\theta_t = \theta_{t-1} + \mathbf{w}_{\theta,t}, \tag{18}$$

$$\mathbf{g}_t = \mathbf{f}_{GA}(\mathbf{g}_{t-1} | P_m, P_c), \tag{19}$$

where $\mathbf{w}_{r,t} \sim N(0, \tau_r^2)$ and $\mathbf{w}_{\theta,t} \sim N(0, \tau_\theta^2)$, and $\mathbf{f}_{GA}(\cdot | P_m, P_c)$ represesnts genetic operation given the mutation rate P_m and the crossover rate P_c.

Monte Carlo Filtering

For the general state space model (15) and (16), state estimation will be done by Monte Carlo filtering [8]. The idea of Monte Carlo filtering to approximate the non-Gaussian distribution by realizations of the distribution and manipulate the realizations instead of distribution itself. The procedure of Monte Carlo filtering consists of the one-step-ahead prediction and the filtering described as follows.

One-step-ahead Prediction: One-step-ahead prediction distribution and filtering distribution are approximated by realizations of their distribution as follows.

$$\mathbf{P}_t = \left\{ \mathbf{p}_1^{(t)}, \mathbf{p}_2^{(t)}, \cdots, \mathbf{p}_M^{(t)} \right\} \sim p(\mathbf{x}_t | Y_{t-1}) \tag{20}$$

$$\mathbf{F}_t = \left\{ \mathbf{f}_1^{(t)}, \mathbf{f}_2^{(t)}, \cdots, \mathbf{f}_M^{(t)} \right\} \sim p(\mathbf{x}_t | Y_t) \tag{21}$$

We call these realizations "particles".

The initial distribution denoted by $\mathbf{F}_0$ is given as follows.

$$\mathbf{F}_0 = \left\{ \mathbf{f}_1^{(0)}, \mathbf{f}_2^{(0)}, \cdots, \mathbf{f}_M^{(0)} \right\}, \sim p(\mathbf{x}_0 | Y_0) \tag{22}$$

One-step-ahead prediction is done by the following procedure. $\mathbf{w}_i^{(t)}$

$$\mathbf{p}_i^{(t)} = \mathbf{g}(\mathbf{f}_i^{(t-1)}, \mathbf{w}_i^{(t)}) \tag{23}$$

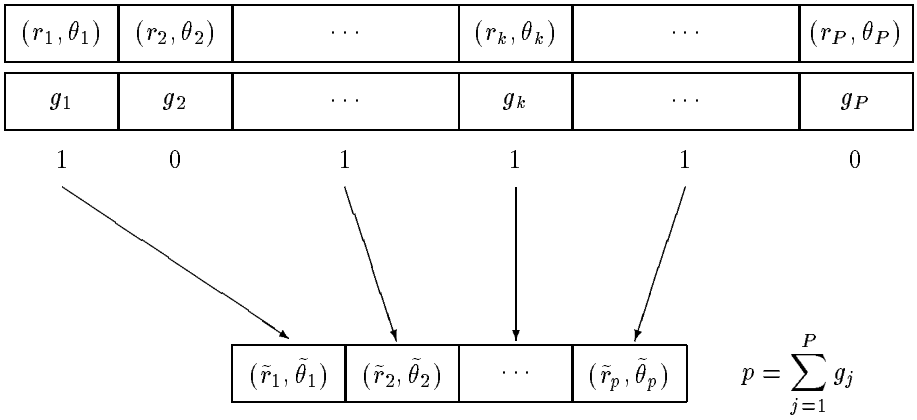

Figure 1: Selection of peaks(characteristic roots) by gene(bit sequence)

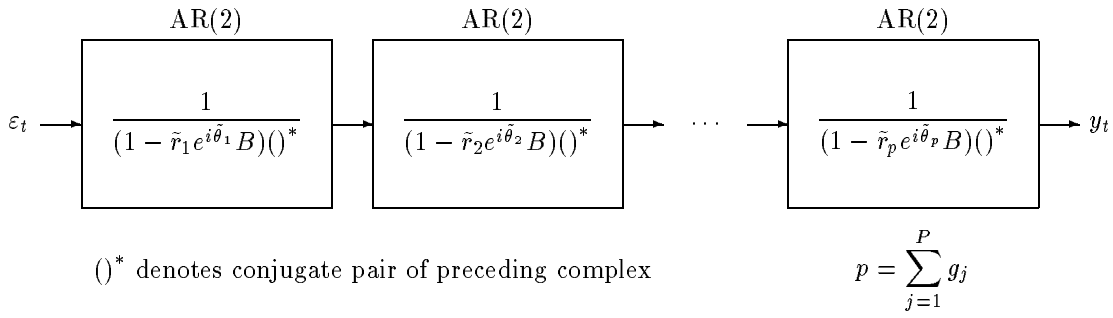

Figure 2: Sequential system of AR(2) consists of selected peaks

Filtering: Filtering is done by the following procedures. Firstly, likelihood of each particle is calculated by

$$\alpha_i^{(t)} = p(y_t|\mathbf{p}_i^{(t)}) = r\left(h^{-1}(y_t, \mathbf{p}_i^{(t)}) \right), \quad (24)$$

where we assume for $y = h(x, z)$ in (16) that there exist a function $h^{-1}(\cdot,\cdot)$ such that $z = h^{-1}(y, x)$. $r(\cdot)$ is density function of ε_t.

Secondly, filtering particles are obtained by doing resampling according to the following probabilities.

$$\mathbf{f}_j^{(t)} = \begin{cases} \mathbf{p}_1^{(t)} & \text{with prob.} \quad \alpha_1^{(t)} / \sum_{j=1}^{M} \alpha_j^{(t)} \\ \vdots & \vdots \\ \mathbf{p}_k^{(t)} & \text{with prob.} \quad \alpha_k^{(t)} / \sum_{j=1}^{M} \alpha_j^{(t)} \\ \vdots & \vdots \\ \mathbf{p}_M^{(t)} & \text{with prob.} \quad \alpha_M^{(t)} / \sum_{j=1}^{M} \alpha_j^{(t)} \end{cases} \quad (25)$$

Likelihood: Likelihood of the model is approximately obtained from likelihood values of each particle in filtering procedure, $\alpha_{j,t}$, for all time $t = 1, \cdots, N$ as follows.

$$\begin{aligned} l(\vartheta) &= \sum_{t=1}^{N} \log p(y_t|Y_{t-1}) \\ &\simeq \sum_{t=1}^{N} \log \left(\sum_{j=1}^{M} \alpha_j^{(t)} \right) - N \log M \end{aligned} \quad (26)$$

Where, ϑ is a vector of hyperparameters τ_r^2, τ_θ^2, P_c, and P_m.

The hyperparameters are estimated to maximize the likelihood. However, exact numerical optimization to obtain the maximum of likelihood is useless in the situation of Monte Carlo method. Currently, make a grid search in the hyperparameter space and plot contour to obtain the optimal combination of them is practical. As an another method to avoid this problem, involving

Table 1: Relationship between GA and MCF

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

hyperparameters into a state vector and estimate them simultaneously is proposed [9].

Similarity to Genetic Algorithm

Although, Monte Calro filtering is developed in the context of non-Gaussian nonlinear state space modeling and its state estimation method, and it is not motivated by nature of biological system, there are similarities to operations of genetic algorithm. By considering both Monte Calro filtering and genetic algorithm are developed in the literature of adaptation, the similarity is not a coincidence, rather it can be considered as the necessity.

According to the Higuchi's research [4], by regarding particles appear in Monte Carlo filtering as individuals of genetic algorithm, we can see similarities between Monte Calro filtering and genetic algorithm as shown in Table 1. One-step-ahead prediction of Monte Carlo filtering corresponds to mutation operation in genetic algorithm, likelihood of each particles to fitness of each individual, and resampling to selection. There is no operation in Monte Carlo filtering corresponds to crossover in genetic algorithm directly, however, by considering a operation (19), crossover operation can be incorporated into Monte Calro filtering procedure.

4. SIMULATIONAL EXPERIMENT

By numerical simulation, time series data shown in Figure 3 are generated. The data have nonstationary power spectrum that peaks are changing as follows,

$$\begin{aligned} \theta_{1,t} &= 0.310 + 0.004 \times t, \\ r_{1,t} &= 0.8, \\ \theta_{2,t} &= 2.944 - 0.004 \times t, \\ r_{2,t} &= 0.5 + 0.001 \times (t-1), \end{aligned} \tag{27}$$

where the second peak denoted by $r_{2,t}$ and $\theta_{2,t}$ is used in $t = 151 \sim 300$, i.e., the number of peaks p is 1 $t = 1 \sim 150$, and $p = 2$ in $t = 151 \sim 300$. The peak frequencies denoted by $\theta_{1,t}$ and $\theta_{2,t}$ are shown in Figure 5 by solid line. AR coefficients used for data generation to have those characteristic roots are calculated by

$$\begin{aligned} a_{1,t} &= 2r_{1,t}\cos\theta_{1,t} \\ a_{2,t} &= -{r_{1,t}}^2 \end{aligned} \tag{28}$$

for $t = 1 \sim 150$, and by

$$\begin{aligned} a_{1,t} &= 2r_{1,t}\cos\theta_{1,t} + 2r_{2,t}\cos\theta_{2,t} \\ a_{2,t} &= -{r_{1,t}}^2 - {r_{2,t}}^2 \\ &\quad -4r_{1,t}r_{2,t}\cos\theta_{1,t}\cos\theta_{2,t} \\ a_{3,t} &= 2r_{1,t}{r_{2,t}}^2\cos\theta_{1,t} + 2{r_{2,t}}^2 r_{2,t}\cos\theta_{2,t} \\ a_{4,t} &= -{r_{1,t}}^2{r_{2,t}}^2 \end{aligned} \tag{29}$$

for $t = 151 \sim 300$ [5].

Conditions of estimation by proposed model are as follows. The number of particles is set to 1000. As the hyperparameters, genetic operation rates are $P_c = 0.05$ and $P_m = 0.01$, variances of system noise are $\tau_r^2 = 0.01$ and $\tau_\theta^2 = 0.01$.

Since bandwidth r_k takes value in $(0, 1)$, and frequency θ_k takes value in $(0, \pi)$, numerical treatment of state variables not to be take out of range is required. Use of logistic function is better mathematically, however, it takes computational cost in calculation of logarithm and exponential. In this experiment, we use a method to make boundary to the state variables, i.e. it takes boundary value when it is out of range in the procedure of one-step-ahead prediction.

The estimated number of peaks varying with time is plotted in Figure 4. In this figure, the mode of distribution of estimated peak number is plotted at each time. Peak frequencies estimated by the model are plotted in Figure 5. In the plot, median of estimated distribution for the peak number equal to the mode is used. Bandwidths are also estimated and by combining with estimated peak frequencies, nonstationary power spectrum can be obtained. Nonstationary power spectra estimated by the proposed method are plotted in Figure 6.

By looking the estimated result for time-varying peak number in Figure 4, the number of peaks started from 1 and changed to 2 after $t = 150$. Estimated peaks in Figure 5 are placed around the true peaks plotted by solid line in this figure. In Figure 6, nonstationary power spectra estimated by the model are corresponding to the result of peaks estimation. That started from one peak at low frequency, and gradually increasing the frequency. After $t = 150$, second peak appeared and moving to lower frequency.

5. CONCLUSION

A new model to estimate nonstationary power spectrum adaptively is proposed. The model is defined by spectral peaks of AR(2) denoted by characteristic root

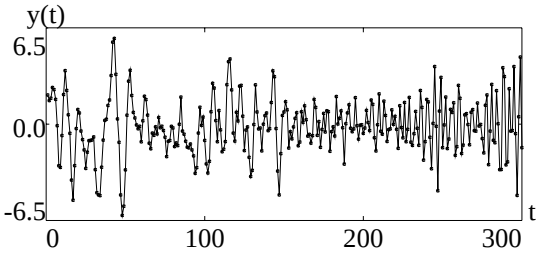

Figure 3: Simulation data

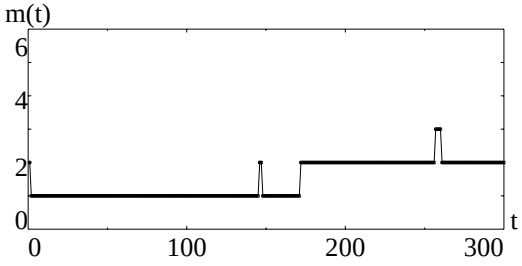

Figure 4: Estimated number of peaks

and sequence of bits called "gene". By selecting the p-peaks corresponding to the value-1-bit of gene, AR(2p) is constructed by the peaks to represent power spectrum at the time. Applying Monte Carlo filter and genetic operations of mutation and crossover, we can estimate these time-varying parameters. From the estimated parameters, we can obtain the time-varying power spectrum. The performance of the method is shown by simple numerical simulation.

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.

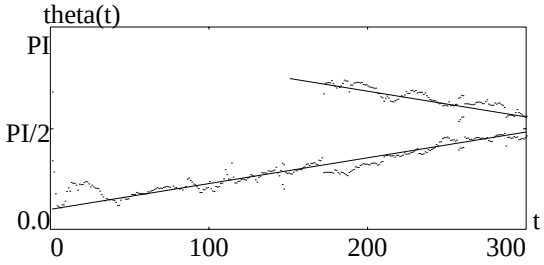

Figure 5: Estimated frequency of the peaks

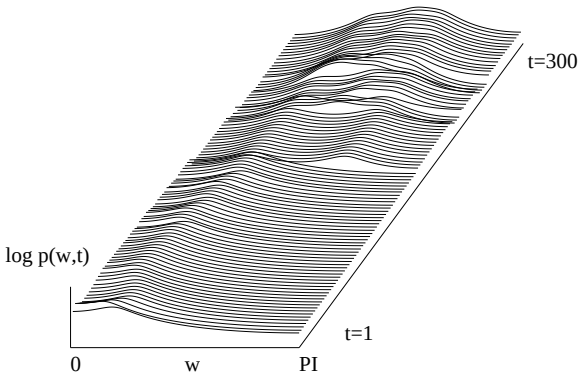

Figure 6: Estimated power spectrum

[5] N.Ikoma: Estimation of time varying peak of power spectrum based on non-Gaussian nonlinear state space modeling, *Signal Processing*, Vol.49, pp.85-95(1996).

[6] G.Kitagawa: A nonstationary time series model and its fitting by a recursive filter, *Journal of Time Series Analysis*, Vol.2, pp.103-116, 1981.

[7] G.Kitagawa: Non-Gaussian state-space modeling of nonstationary time series (with discussion), *Journal of the American Statistical Association*, Vol.82, No.400, pp.1032-1063, 1987.

[8] 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.

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

[10] 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.

[11] R.C.K.Lee: Optimal estimation, identification, and control, The MIT Press, Cambridge, MA, 1964.

[12] M.B.Priestley: Spectral analysis and time series, Academic Press, 1981.