Time Series Analysis based on Time-Varying Peak Frequencies of Power Spectrum and Application to Seismic Wave Data

IKOMA, Norikazu
Faculty of Information Science, Hiroshima City University

A new model based on time-varying peak frequencies of power spectrum (TVPP model) has been proposed. The model contains peak frequencies of power spectrum as time-varying parameters and time invariant factor of power spectrum. Nonlinear nonstationary state space modeling proposed by G.Kitagawa is used to estimate time-varying peaks, and quasi-Newton method with BFGS modification formula is used to estimate time invariant parameters. Application result to seismic wave data has been reported. Comparison to time-varying coefficient AR model (TVAR mode) has also been reported.

1 Introduction

Spectral analysis is widely used in many field of researches and statistical models for time series analysis can be used to obtain the estimate of spectrum. Peak frequency of power spectrum is one of the most important feature in the analysis using power spectrum. Although there are many conventional methods to obtain stationary spectrum and several methods to obtain nonstationary spectrum, there is little method directly obtaining such important feature of spectrum.

We propose a new model based on time-varying peak frequencies of power spectrum (TVPP model). The most advantage of this model is the direct estimation of peak frequencies of power spectrum. The basic idea of proposed model, named original model, consists of system equation and observation equation. System equation denotes the smoothness of time changes of peak frequencies, and observation equation consists of time-varying coefficient AR model. Original model has been extended to improve fitness to data by making multiplicative model with stationary AR model.

Since the AR coefficients are parametrized by peak frequencies nonlinearly, estimation method conventionally used for linear models cannot be used. There is an estimation method for nonlinear models, nonlinear nonstationary state space modeling proposed by G.Kitagawa[3], and we have used this method for the estimation of time-varying peaks. The model also contains time invariant parameters, and they are estimated by quasi-Newton method based on likelihood function.

An application to seismic wave data has been reported. In the application, not only TVPP model has been examined, but also time-varying coefficient AR model(TVAR mode)[2] has been applied as the counter example. Application results of both models are compared and discussion about the comparison follows.

2 Time-Varying Peak of Power spectrum model(TVPP model)

We consider time series data, which have stationary mean, denoted by a vector

$$Y = [y_1, y_2, \ldots, y_N]^T,$$

where $\mathbf{x}^T$ denotes a transpose of vector $\mathbf{x}$. Let vector of time-varying peak frequencies of power spectrum at time t be

$$\mathbf{\Theta}_t = [\theta_1(t), \theta_2(t), \ldots, \theta_m(t)]^T,$$

let us assume parameters related to power of each peak are time invariant and denoted by

$$\mathbf{R} = [r_1, r_2, \ldots, r_m]^T.$$

The original idea of a model with time-varying peak frequencies of power spectrum is denoted by

$$\begin{cases} \mathbf{\Theta}_t & = & \mathbf{\Theta}_{t-1} + \mathbf{w}_t, \\ y_t & = & \sum_{j=1}^{2m} a_j(\mathbf{\Theta}_t, \mathbf{R}) y_{t-j} + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma^2), \end{cases} \tag{1}$$

where, system noise

$$\mathbf{w}_t = [w_1(t), w_2(t), \ldots, w_m(t)]^T$$

is a random vector accroding to normal distribution of 0 mean and Q covariance matrix

$$Q = diag\left(\tau_1^2, \tau_2^2, \ldots, \tau_m^2\right).$$

AR coefficient, $a_j(\mathbf{\Theta}_t, \mathbf{R})$, is derived by solving the following equations,

$$1 - \sum_{j=1}^{2m} a_j(\mathbf{\Theta}_t, \mathbf{R}) B^j = \prod_{k=1}^{m} (1 - r_k e^{i\theta_k(t)} B)(1 - r_k e^{-i\theta_k(t)} B). \tag{2}$$

3 Estimation Method

Estimation of the model is done by nonlinear nonstationary state space modeling method proposed by G.Kitagawa [3]. The estimation method can be divided into two phases, forward phase by applying filtering and prediction alternatively, and backward phase of smoothing. The value of likelihood which is used in numerical optimization is obtained in forward phase. After determined the optimal values of time invariant parameters, backward phase is done and we obtain smoothing estimation of state.

The main idea of the estimation method is a numerical representation of non-Gaussian distribution. The domain of state vector is divided into many short intervals, and the values of density function of each interval are stored. Numerical integration is done in estimation process. Since the number of intervals gains exponentially with the dimension of the state vector, the dimension of state vector cannot be taken higher.

Time constant parameters,

$$\left[\sigma^2, \tau_1^2, \tau_2^2, \ldots, \tau_m^2, r_1, r_2, \ldots, r_m\right]$$

are obtained by numerical optimization to maximize likelihood function. Quasi-Newton method with BFGS modification formula of Hessian is used as the numerical optimization method.

4 Model Extension

We have seen that low dimension of state vector is required in the estimation method. Low dimension corresponds to small number of peaks of power spectrum. This might cause less of fitness to data. To improve the disadvantage, the original model has been extended as follows. Time invariant factors of power spectrum are introduced to the original model. They can be represented by stationary AR model of order n, and the extended model can be written as follows,

$$\begin{cases} \mathbf{\Theta}_t &= \mathbf{\Theta}_{t-1} + \mathbf{w}_t, \\ v_t &= \sum_{j=1}^{2m} a_j(\mathbf{\Theta}_t, \mathbf{R}) v_{t-j} + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma^2), \\ y_t &= \sum_{j=1}^{n} c_j y_{t-j} + v_t, \end{cases} \tag{3}$$

In the extended model, time constant parameters are rewritten to contain the additional parameters $c_j (j = 1 \sim n)$ as follows,

$$\left[\sigma^2, \tau_1^2, \tau_2^2, \ldots, \tau_m^2, r_1, r_2, \ldots, r_m, c_1, c_2, \ldots, c_n\right].$$

They can be determined by quasi-Newton method.

5 Analysis of Seismic Wave Data

Seismic wave data observed in Hokkaido, Japan has been shown in **fig.** 1. The data consist of east-west direction of vibration, and length of data is 2500 with sampling time 0.02 second. Vibration without earthquake, primaly wave(P-wave), secondary wave(S-wave) and its decay are contained in the data. While the data has nonstationary variance, normalization of variance is required before applying TVPP model or TVAR model[2]. Variance normalized series of seismic wave data is shown in **fig.** 2.

AIC can be used to determine which model and what order is better [1]. Values of AIC obtained by TVPP model and TVAR model for variance normalized series are shown in **tbl.** 1. TVAR model has better fitness than TVPP model in terms of AIC.

Estimated result by TVPP model is as follows, frequency of time-varying peaks have been shown in **fig.** 3, and time invariant factor of power spectrum has been shown in **fig.** 4. Time-varying power spectrum obtained by TVAR model has been shown in **fig.** 6. Time-varying peak frequencies numerically obtained from the spectrum has been shown in **fig.** 5.

By looking at the estimated result of time-varying peaks by TVPP model, different frequencies are confirmed among vibration without earthquake, P-wave, and S-wave. Also we can see that the peak frequency in S-wave is going to back to what of without earthquake as the amplitude of wave going down.

References

[1] H.Akaike: A new look at the statistical model identification, *IEEE Trans* AC-19, **6**, 716/723(1974).

[2] G.Kitagawa and W.Gersch: A Smoothness Priors Time-Varying AR Coefficient Modeling of Nonstationary Covariance Time Series, *IEEE Trans* AC-30, **1**, 48/56(1985).

[3] G.Kitagawa: Non-Gaussian State-Space Modeling of Nonstationary Time Series (with discussion), *Journal of the American Statistical Association* **82**,No.400,1032/1063(1987).

Table 1: Values of AIC for seismic wave data

n of TVPP model	TVPP model		TVAR model	
p of TVAR model	$(m = 1)$	$(m = 2)$	$(k = 1)$	$(k = 2)$
0	5532.13	4920.54	.	
1	5389.31	4896.59	6231.14	6249.99
2	5216.65	4893.50	5275.58	5382.85
3	4877.33	4737.70	4845.18	4890.47
4	4872.92	4729.95	4598.60	4608.77
5	4811.62	4731.02	4628.58	4634.89
6	5034.67	4671.28	4621.63	4616.72
7	4766.36	4636.29	4626.05	4618.25
8	4724.06	4614.22	4601.88	4604.13
9	4724.26	4612.76	4600.42	4618.90
10	4724.99	4629.40	4605.10	4644.37

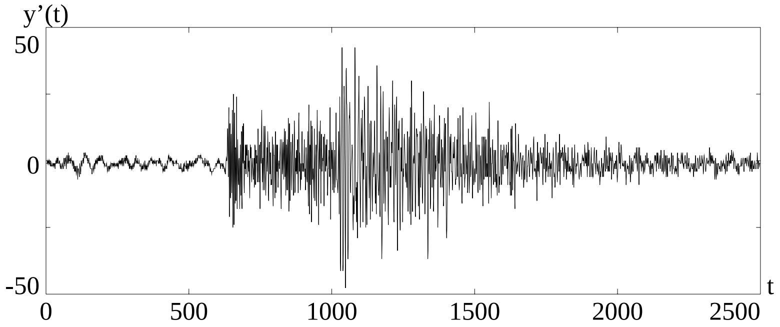

Figure 1: Seismic wave data

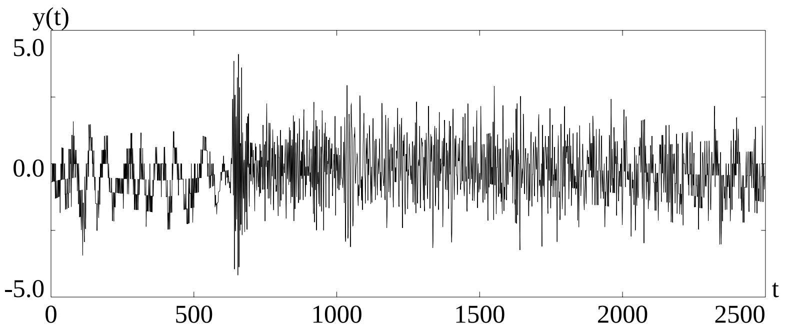

Figure 2: Variance normalized series of seismic wave data

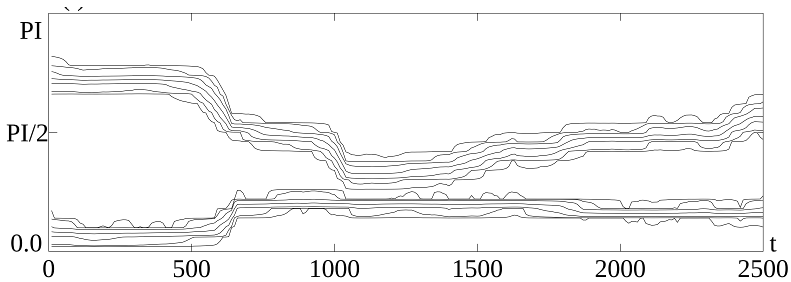

Figure 3: Estimated result of time-varying peak frequencies of power spectrum by TVPP model(50%,68%,95%,and 99% points are plotted)

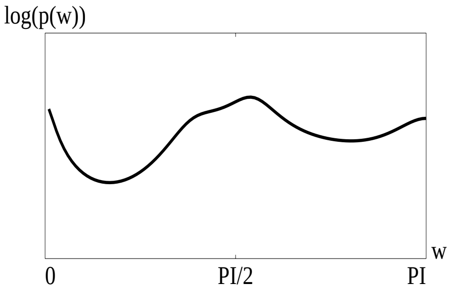

Figure 4: Estimated result of time invariant power spectrum by TVPP model

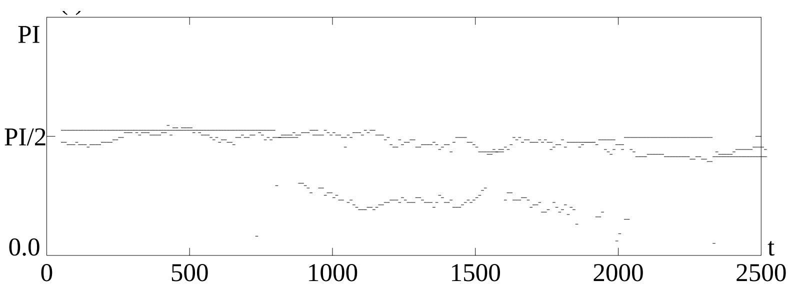

Figure 5: Estimate result of time-vaying peak frequencies of power spectrum by TVAR model

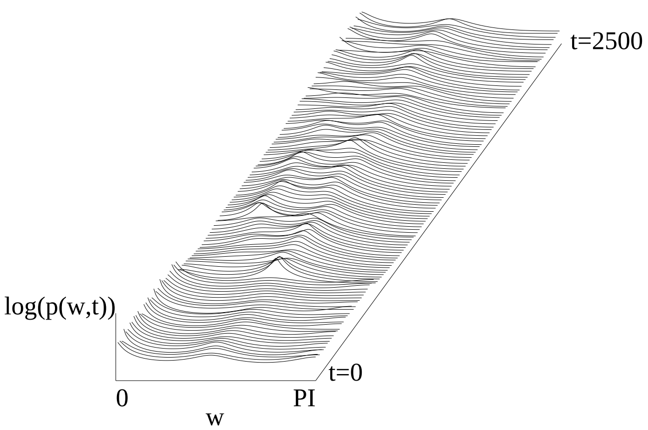

Figure 6: Estimated result of time-varying power spectrum by TVAR model