A Robust-filtering Method for Small Sample Economic Time Series *

Naoto Kunitomo †

October 30, 2018

Key Words

Non-stationary economic time series, Errors-variables models, Measurement Error, trend and seasonality, Robust-filtering, SIML, Large Dimension.

1. Introduction

There has been a vast amount of published research on the use of statistical time series analysis of macro-economic time series. One important feature of macroeconomic time series, which is different from the standard time series analysis, is the fact that the observed time series is an apparent mixture of non-stationary components and stationary components. The second feature is the fact that the measurement errors in economic time series play important roles because macro-economic data are usually constructed from various sources including sample surveys in major official statistics while the statistical time series analysis often ignores measurement errors. There is yet third important issue that the sample size of macro-economic data is rather small and we have 120, say, time series observations for each series when we have quarterly data over 30 years. The quarterly GDP series, which has been the most important data in macro-economy are published since 1994 by the cabinet office of Japan, for instance. Since the sample size is small, it is important to use an appropriate statistical procedure to extract information on trend and noise (or measurement error) components in a systematic way from data.

In this study we will develop a new filtering method to estimate the hidden states of random variables and to handle multiple time series data, and particularly

*A joint work with Seisho Sato (University of Tokyo) and it is a preliminary memorandum. We owe Akihiko Takahashi for a suggestion on our filtering method. The research has been supported by JSPS-Grant JP17H02513.

†School of Potitical Science and Economics, Meiji University, Kanda-Surugadai 1-1, Chiyoda-ku 101-8301, Tokyo, JAPAN, naoto.kunitomo@gmail.com

to deal with small sample economic time series. Kunitomo and Sato (2017), and Kunitomo, Sato and Kurisu (2018) have developed the separating information maximum likelihood (SIML) method for estimating the non-stationary errors-in-variables models. They have discussed the asymptotic properties and finite sample properties of the estimation of unknown parameters. We utilize their results to solve the filtering problem of hidden random variables, which gives a powerful new method of handling macro-economic time series.

Kitagawa (2010) has discussed the standard statistical filtering methods already known including the Kalman-filtering and the particle-filtering methods. Since (i) these methods depend on the underlying distributions such as the Gaussian distributions for the Kalman-filtering and (ii) the procedures essentially depend on the dimension of state variables, there may be some difficulty to extend to the high-dimension cases even when it is fixed, say 100. On the other hand , we can expect that our method has some merits when we handle small sample economic times series with non-stationarity and seasonality with many variables because our method does not depend on the specific distributions as well as the dimensions of random variables. See Kunitomo, Awaya and Kurisu (2017) for a comparison of small sample properties of the ML and SIML methods. The most important feature of the present procedure is that it can be applicable to small sample time series data with large dimension. Also our new method has a solid mathematical and statistical foundation.

2. Non-stationary Errors-in-variables models

Let y_{ji} be the $i-$th observation of the $j-$th time series at i for $i = 1, \cdots, n; j = 1, \cdots, p$. We set $\mathbf{y}_i = (y_{1i}, \cdots, y_{pi})'$ be a $p \times 1$ vector and $\mathbf{Y}_n = (\mathbf{y}_i') \; (= (y_{ij}))$ be an $n \times p$ matrix of observations and denote $\mathbf{y}_0$ as the initial $p \times 1$ vector. We estimate the model when the underlying non-stationary trends $\mathbf{x}_i \; (= (x_{ji})) \; (i = 1, \cdots, n)$, but we have the vector of noise component $\mathbf{v}_i' = (v_{1i}, \cdots, v_{pi})$, which are independent of $\mathbf{x}_i$. We use the non-stationary errors-in-variables representation

$$\mathbf{y}_i = \mathbf{x}_i + \mathbf{v}_i \quad (i = 1, \cdots, n), \tag{2.1}$$

where $\mathbf{x}_i \; (i = 1, \cdots, n)$ are a sequence of non-stationary I(1) process which satisfy

$$\Delta \mathbf{x}_i = (1 - \mathcal{L})\mathbf{x}_i = \mathbf{v}_i^{(x)} \;, \tag{2.2}$$

where $\mathbf{v}_i^{(x)}$ is a sequence of i.i.d. random vectors with $\mathcal{E}(\mathbf{v}_i^{(x)}) = \mathbf{0}$ and $\mathcal{E}(\mathbf{v}_i^{(x)}\mathbf{v}_i^{(x)'}) = \mathbf{\Sigma}_x$. The random vectors $\mathbf{v}_i \; (i = 1, \cdots, n)$ are a sequence of i.i.d. random variables with $\mathcal{E}(\mathbf{v}_i) = \mathbf{0}$ and $\mathcal{E}(\mathbf{v}_i\mathbf{v}_i') = \mathbf{\Sigma}_v$.

We consider the situation when each pair of vectors $\Delta\mathbf{x}_i$ and $\mathbf{v}_i$ are independently, identically, and normally distributed (i.i.d.) as $N_p(\mathbf{0}, \mathbf{\Sigma}_x)$ and $N_p(\mathbf{0}, \mathbf{\Sigma}_v)$, respectively, and we have the observations of an $n \times p$ matrix $\mathbf{Y}_n = (\mathbf{y}_i^{'})$ and set the $np \times 1$ random vector $(\mathbf{y}_1^{'}, \cdots, \mathbf{y}_n^{'})^{'}$. Given the initial condition $\mathbf{y}_0$, we have

$$\text{vec}(\mathbf{Y}_n) \sim N_{n \times p}\left(\mathbf{1}_n \cdot \mathbf{y}_0^{'}, \mathbf{I}_n \otimes \mathbf{\Sigma}_v + \mathbf{C}_n \mathbf{C}_n^{'} \otimes \mathbf{\Sigma}_x\right) \ , \tag{2.3}$$

where $\mathbf{1}_n^{'} = (1, \cdots, 1)$ and

$$\mathbf{C}_n = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 1 & \cdots & 0 \\ 1 & \cdots & 1 & 1 & 0 \\ 1 & \cdots & 1 & 1 & 1 \end{pmatrix}_{n \times n} . \tag{2.4}$$

We use the K_n^*-transformation that from $\mathbf{Y}_n$ to $\mathbf{Z}_n$ $(= (\mathbf{z}_k^{'}))$ by

$$\mathbf{Z}_n = \mathbf{K}_n^* \left(\mathbf{Y}_n - \bar{\mathbf{Y}}_0\right) \ , \mathbf{K}_n^* = \mathbf{P}_n \mathbf{C}_n^{-1} \ , \tag{2.5}$$

where

$$\mathbf{C}_n^{-1} = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & 0 & \cdots \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}_{n \times n} , \tag{2.6}$$

and

$$\mathbf{P}_n = (p_{jk}^{(n)}) \ , \ p_{jk}^{(n)} = \sqrt{\frac{2}{n + \frac{1}{2}}} \cos\left[\frac{2\pi}{2n+1}(k - \frac{1}{2})(j - \frac{1}{2})\right] . \tag{2.7}$$

By using the spectral decomposition $\mathbf{C}_n^{-1} \mathbf{C}_n^{'-1} = \mathbf{P}_n \mathbf{D}_n \mathbf{P}_n^{'}$ and $\mathbf{D}_n$ is a diagonal matrix with the k-th element $d_k = 2[1 - \cos(\pi(\frac{2k-1}{2n+1}))]$ $(k = 1, \cdots, n)$ and we write

$$a_{kn}^* \ (= d_k) \ = 4 \sin^2 \left[\frac{\pi}{2}\left(\frac{2k-1}{2n+1}\right)\right] \ (k = 1, \cdots, n) \ . \tag{2.8}$$

The separating information maximum likelihood (SIML) estimator of $\hat{\mathbf{\Sigma}}_x$ can be defined by

$$\mathbf{G}_m = \hat{\mathbf{\Sigma}}_{x,SIML} = \frac{1}{m_n} \sum_{k=1}^{m_n} \mathbf{z}_k \mathbf{z}_k^{'} \ . \tag{2.9}$$

3. SIML Filtering

Let an $m \times n$ choice matrix $\mathbf{J}_m = (\mathbf{I}_m, \mathbf{O})$, and let also $n \times p$ matrix

$$\mathbf{Z}_n^* = \mathbf{J}_m^{'} \mathbf{J}_m \mathbf{Z}_n = \mathbf{J}_m^{'} \mathbf{J}_m \mathbf{P}_n \mathbf{C}_n^{-1} (\mathbf{Y}_n - \bar{\mathbf{Y}}_0) \ , \ \mathbf{Z}_n = \mathbf{P}_n \mathbf{C}_n^{-1} (\mathbf{Y}_n - \bar{\mathbf{Y}}_0) \tag{3.1}$$

and an $n \times n$ matrix

(3.2) $$\mathbf{P}_n^* = \mathbf{J}_m^{'}\mathbf{J}_m\mathbf{P}_n \ .$$

We define $\mathbf{Z}_n^{**} = \mathbf{P}_n^{*'}\mathbf{Z}_n^* = \mathbf{P}_n^{'}\mathbf{Z}_n^*$ and then we construct an estimator of $n \times p$ hidden state matrix $\mathbf{X}_n$ by using the inverse transformation of $\mathbf{Z}_n^*$ (by deleting the estimated noise parts) as

(3.3) $$\hat{\mathbf{X}}_n = \mathbf{C}_n\mathbf{P}_n^{*'}\mathbf{Z}_n^* \ .$$

Then we have the relation of $p \times p$ matrices

(3.4) $$\mathbf{Z}_n^{**'}\mathbf{Z}_n^{**} = \mathbf{Z}_n^{*'}\mathbf{Z}_n^*$$

and hence the $p \times p$ variance-covariance matrix of $\mathbf{P}_n^{'}\mathbf{Z}_n^*$ is numerically the same as that of $\mathbf{J}_m\mathbf{Z}_n$.

Let the $[m + (n - m)] \times [m + (n - m)]$ partitioned matrix

(3.5) $$\mathbf{P}_n = \begin{pmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{P}_{21} & \mathbf{P}_{22} \end{pmatrix} \ .$$

Then

(3.6) $$\mathbf{P}_n\mathbf{J}_m^{'}\mathbf{J}_m\mathbf{P}_n = \begin{pmatrix} \mathbf{P}_{11}^{'} \\ \mathbf{P}_{12}^{'} \end{pmatrix} (\mathbf{P}_{11}, \mathbf{P}_{12}) = \mathbf{I}_n - \begin{pmatrix} \mathbf{P}_{21}^{'} \\ \mathbf{P}_{22}^{'} \end{pmatrix} (\mathbf{P}_{21}, \mathbf{P}_{22}) \ .$$

After some calculations, the $(j, j^{'})$-th element of $\mathbf{Q}_n = \mathbf{P}_n\mathbf{J}_m^{'}\mathbf{J}_m\mathbf{P}_n$ $(= (q_{j,j^{'}}))$ is given by

$$\begin{aligned} q_{j,j} &= \frac{2m}{2n+1} + \frac{1}{2n+1}\left[\frac{\sin\frac{2m\pi}{2n+1}(2j-1)}{\sin\frac{\pi}{2n+1}(2j-1)}\right] \ , \\ q_{i,j^{'}} &= \frac{1}{2n+1}\left[\frac{\sin\frac{2m\pi}{2n+1}(j+j^{'}-1)}{\sin\frac{\pi}{2n+1}(j+j^{'}-1)} + \frac{\sin\frac{2m\pi}{2n+1}(j-j^{'})}{\sin\frac{\pi}{2n+1}(j-j^{'})}\right] \quad (j \neq J^{'}) \ . \end{aligned}$$

More generally, let an $m_2 \times [m_1 + m_2 + (n - m_1 - m_2)$ choice matrix $\mathbf{J}_{m_1,m_2,n} = (\mathbf{O}, \mathbf{I}_{m_2}, \mathbf{O})$, and let also $n \times p$ matrix

(3.7) $$\mathbf{Z}_n^* = \mathbf{J}_{m_1,m_2,n}^{'}\mathbf{J}_{m_1,m_2,n}\mathbf{Z}_n = \mathbf{J}_{m_1,m_2,n}^{'}\mathbf{J}_{m_1,m_2,n}\mathbf{P}_n\mathbf{C}_n^{(s)-1}(\mathbf{Y}_n - \bar{\mathbf{Y}}_0)$$

and an $n \times n$ matrix

$$\mathbf{P}_n^{**} = \mathbf{J}_{m_1,m_2,n}^{'}\mathbf{J}_{m_1,m_2,n}\mathbf{P}_n \ ,$$

where $\mathbf{C}_n^{(s)} = \mathbf{C}_n \otimes \mathbf{I}_s$ $(s \geq 2)$, $m_1 = [2n/s] - [m/2]$ and $m_2 = m$. (s is the seasonal frequency and we have used the seasonal differencing.)

Then we construct an estimator of $n \times p$ hidden (seasonal factor) matrix $\mathbf{S}_n$ by using the Fourier-inversion of $\mathbf{Z}_n^*$ (by deleting other parts), which is given as

(3.8) $$\hat{\mathbf{S}}_n = \mathbf{C}_n^{(s)}\mathbf{P}_n^{**'}\mathbf{Z}_n^* \ .$$

After some calculations, the $(j, j^{'})$-th element of $\mathbf{Q}_n = \mathbf{P}_n \mathbf{J}_{m_1,m_2,n}^{'} \mathbf{J}_{m_1,m_2,n} \mathbf{P}_n$ $(= (q_{j,j^{'}}))$ is given by

$$
\begin{aligned}
q_{j,j} &= \frac{2m_2}{2n+1} + \frac{1}{2n+1}\left[\frac{\sin\frac{2(m_1-1+m_2)\pi}{2n+1}(2j-1) - \sin\frac{2(m_1-1)\pi}{2n+1}(2j-1)}{\sin\frac{\pi}{2n+1}(2j-1)}\right] , \\
q_{i,j^{'}} &= \frac{1}{2n+1}\left[\frac{\sin\frac{2(m_1-1+m_2)\pi}{2n+1}(j+j^{'}-1) - \sin\frac{2(m_1-1)\pi}{2n+1}(j+j^{'}-1)}{\sin\frac{\pi}{2n+1}(j+j^{'}-1)}\right. \\
&\qquad \left. + \frac{\sin\frac{2(m_1-1+m_2)\pi}{2n+1}(j-j^{'}) - \sin\frac{2(m_1-1)\pi}{2n+1}(j-j^{'})}{\sin\frac{\pi}{2n+1}(j-j^{'})}\right] \quad (j \neq j^{'}) .
\end{aligned}
$$

We note that when $m_1 = 1$ and $m_2 = m$, the resulting formulae become to those in the standard case.

However, Sato (2018, a personal communication) has reported that the differencing may be better than the seasonal differencing to estimate the state of seasonal factors. This observation can be represented as

$$
\mathbf{Z}_n^{(s)} = \mathbf{J}_{m_1,m_2,n}^{'} \mathbf{J}_{m_1,m_2,n} \mathbf{Z}_n = \mathbf{J}_{m_1,m_2,n}^{'} \mathbf{J}_{m_1,m_2,n} \mathbf{P}_n \mathbf{C}_n^{-1} (\mathbf{Y}_n - \bar{\mathbf{Y}}_0) \tag{3.9}
$$

and

$$
\hat{\mathbf{S}}_n = \mathbf{C}_n \mathbf{P}_n^{**'} \mathbf{Z}_n^{(s)} . \tag{3.10}
$$

The resulting formura for MSE is the same except $\mathbf{C}_n$ instead of $\mathbf{C}_n^{(s)}$.
There may be some reason for this empirical observation. Since we have small observations in macro-economic time series (, say 100), the finite sample properties of estimation method should be important. The resulting formula is the same by putting $s = 1$. When we had an infinite number of data, we could have recovered the spectral density, which diverges at seasonal frequencies, and it helps us to identify the seasonal factor. We need to investigate the effects of differencing, seasonal differencing and periodogram in a systematic way. See Kunitomo (2018) for the estimation problem of unknown structural parameters.

4. A Mathematical Foundation

At the first glance, the SIML filtering method might be seen as an *ad-hoc* statistical procedure without any mathematical foundation. However, on the contrary, there is a rather solid statistical foundation.
Let $\theta_{jk} = \frac{2\pi}{2n+1}(j - \frac{1}{2})(k - \frac{1}{2})$,

$$
p_{jk}^{(n)} = \frac{1}{\sqrt{2n+1}}(e^{i\theta_{jk}} + e^{-i\theta_{jk}})
$$

and we write

$$\Delta_\lambda \mathbf{z}^{(n)}(\lambda_k^{(n)}) = \sum_{j=1}^{n} p_{jk}^{(n)} \mathbf{r}_j^{(n)} \ , \ \mathbf{r}_j^{(n)} = \mathbf{y}_j - \mathbf{y}_{j-1} \ , \tag{4.1}$$

which is actually a (real-valued) form of Fourier-transformation. Then $\Delta_\lambda \mathbf{z}^{(n)}(\lambda_k^{(n)})$ $(k = 1, \cdots, n)$ are a sequence of (real-valued) form of transformations of data at the frequency $\lambda_k^{(n)}$, which are the estimates of the underlying orthogonal incremental process, say, $\mathbf{z}(\lambda)$.

For the development of statistical inferences, we have the next result by using the CLT for dependent variables, which may be useful for applications.

Theorem 1 : Let $\mathbf{r}_j$ $(j = 1, \cdots, n)$ be an ergodic stationary stochastic process with $\mathbf{\Gamma}(h) = \mathcal{E}(\mathbf{r}_j \mathbf{r}_{j-h}^{'})$ and

$$\sum_{h=0}^{\infty} \|\mathbf{\Gamma}(h)\| < \infty \ . \tag{4.2}$$

(i) Let $\Delta_\lambda \mathbf{z}^{(n)}(\lambda_k^{(n)}) = \sum_{j=1}^{n} p_{jk}^{(n)} \mathbf{r}_j^{(n)}$ and $\mathbf{r}_j^{(n)}$ be an ergodic stationary sequence with $\mathcal{E}[\mathbf{r}_j] = \mathbf{0}$ and

$$\mathbf{f}(\lambda) = \mathbf{\Gamma}(0) + \sum_{h=1}^{\infty} \cos(2\pi h \lambda)[\mathbf{\Gamma}(h) + \mathbf{\Gamma}(-h)] \ , \tag{4.3}$$

is the positive definite and bounded (real-valued and symmetrized) spectral density matrix. Also assume that $\lambda_k^{(n)} \to s$, $\lambda_{k'}^{(n)} \to t$ and $0 < s < t < \frac{1}{2}$. Then as $n \longrightarrow \infty$

$$\begin{bmatrix} \Delta_\lambda \mathbf{z}^{(n)}(\lambda_k^{(n)}) \\ \Delta_\lambda \mathbf{z}^{(n)}(\lambda_{k'}^{(n)}) \end{bmatrix} \xrightarrow{w} N_{2p} \left[\mathbf{0}, [\begin{array}{cc} \mathbf{f}(s) & \mathbf{0} \\ \mathbf{0} & \mathbf{f}(t) \end{array}] \right] \ . \tag{4.4}$$

(ii) Let $\mathbf{Z}_n(t) - \mathbf{Z}_n(s) = \sum_{k=[sn]}^{[tn]} \frac{1}{\sqrt{n}} \sum_{j=1}^{n} p_{jk}^{(n)} \mathbf{r}_j^{(n)}$ for $0 < s < t < 1$. Then as $n \longrightarrow \infty$

$$\mathbf{Z}_n(t) - \mathbf{Z}_n(s) \xrightarrow{w} [\frac{1}{t-s} \int_{\frac{2}{2}}^{\frac{t}{2}} \mathbf{f}(\lambda) d\lambda]^{1/2} [\mathbf{B}(t) - \mathbf{B}(s)] \ , \tag{4.5}$$

where $\mathbf{B}(t)$ is the vector of (standard) Brownian motions, which is the continuous time vector process with independent increments.

It has been known that in the statistical time series analysis for a stationary discrete (vector) process $\mathbf{r}_k^*$ with the spectral distribution F, there exists a right-continuous orthogonal increment (vector, complex-valued) process $\mathbf{z}^*(\lambda)$ $(-1/2 \leq \lambda \leq 1/2)$ such that

$$\mathbf{r}_k^* = \int_{(-1/2,1/2]} e^{i2\pi i k \nu} d\mathbf{z}^*(\nu) \ (k = 1, \cdots, n). \tag{4.6}$$

(The topic here goes back to Doob (1953), but see Hannan (1971) or Brockwell and Davis (1990).)

The trend component and seasonal component of (real-valued) time series in our setting can be defined by

$$\text{(4.7)} \qquad \mathbf{r}_k^{(u)} = \int_{(0,1/2]} \cos(2\pi ik\nu)h(\cos(2\pi ik\nu)d\mathbf{z}(\nu) \ (k=1,\cdots,n)$$

for $u = x$ or $u = s$, where $h(\cdot)$ is the indicator function of some frequencies around zero (for trend) and seasonal frequency (for seasonality), respectively, and $\mathbf{z}(\nu)$ $0 < \nu \leq 1/2$ is the right-continuous orthogonal increment (real-valued) process, which is the limiting continuous process of (4.1) .

Since $\mathbf{z}(\nu)$ is not observed with finite data, the (real-valued) estimate of $\mathbf{r}_k^{(u)}$ (i.e. the hidden components of $\mathbf{r}_k$) from data can be represented as

$$\text{(4.8)} \mathbf{r}_k^{(u,n)} = \int_{(0,1/2]} \cos(2\pi ik\nu)h_n(\cos(2\pi ik\nu))d\mathbf{z}^{(n)}(\nu) \ (\text{at } \nu = \frac{k}{2n}, k=1,\cdots,n),$$

where $h_n(\cdot)$ is a measurable function and we have abused some notations.

There may be an interesting representation problem of (discrete time and continuous time) stationary processes and orthogonal incremental stochastic processes.

5. Model Selection

When we have estimates of the state variables $\mathbf{x}_i$ $(i = 1, \cdots, n)$, the estimates of noise components are $\hat{\mathbf{v}}_i = \mathbf{y}_i - \hat{\mathbf{x}}_i$ $(i = 1, \ldots, n)$.

Then an estimated MSE of the one-step ahead prediction errors based on the SIML-filtering is given by

$$\text{(5.1)} \qquad \boldsymbol{M}\hat{S}E(1) = \hat{\mathbf{\Sigma}}_x + c\hat{\mathbf{\Sigma}}_v \ .$$

where c is a constant.

Then one may try to minimize the estimated one-step prediction MSE by choosing an appropriate m.

References

[1] Brockwell, P. and R. Davis (1990), *Time Series : Theory and Methods*, 2nd Edition, Wiley.

[2] Doob, J.L (1953), *Stochastic Processes*, Wiley.

[3] Durrett, R. (1991), *Probability : Theory and Examples*, Duxbury Press.

[4] Hannan, E.J. (1970), *Multiple Time Series*, Wiley.

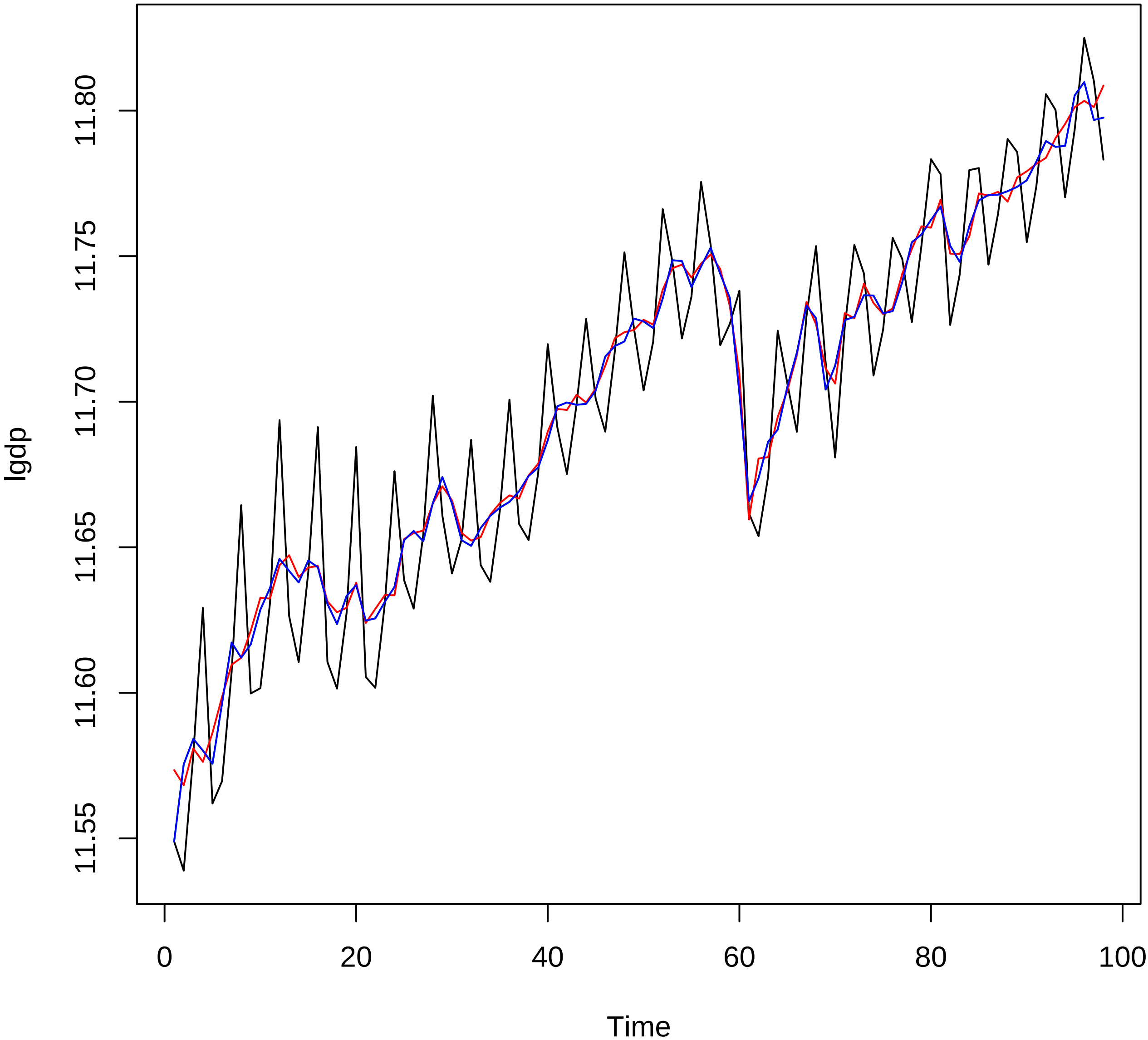
[5] Kitagawa, G. (2010), *Introduction to Time Series Analysis*, CRC Press.

[6] Kunitomo, N. (2018), "Common Seasonal Factors," in preparation.

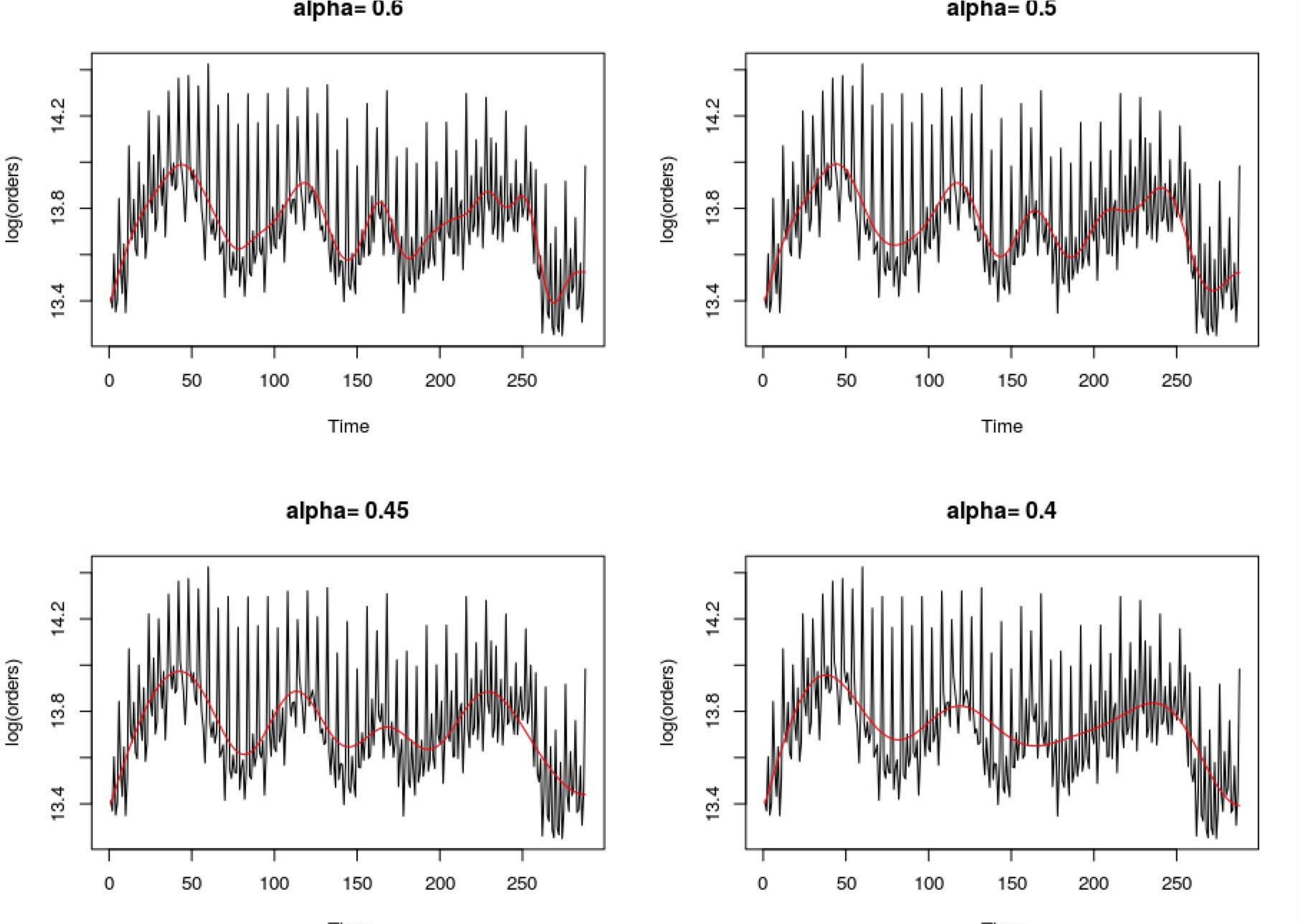
[7] Kunitomo, N. , Sato and D. Kurisu (2018), *Separating Information Maximum Likelihood Estimation for High Frequency Financial Data*, Springer.

[8] Kunitomo, N. and S. Sato (2017), "Trend, Seasonality and Economic Time Series : the Non-stationary Errors-in-variables Models," SDS-4, MIMS, Meiji University, http://www.mims.meiji.ac.jp/publications/2017-ds.

[9] Kunitomo, N., N. Awaya and D. Kurisu (2017), "Some Properties of Estimation Methods for Structural Relationships in Non-stationary Errors-in-Variables Models," SDS-3, MIMS, Meiji University.



lgdp
11.55
11.60
11.65
11.70
11.75
11.80
0
20
40
60
80
100
Time


alpha= 0.6
alpha= 0.5
alpha= 0.45
alpha= 0.4
log(orders)
13.4
13.8
14.2
0
50
100
150
200
250
Time