SSE-DP-2023-2

# Frequency Regression and Smoothing for Noisy Nonstationary Multivariate Time Series

Naoto Kunitomo
(The Institute of Statistical Mathematics)
and
Seisho Sato
(University of Tokyo)

March 2023

# Frequency Regression and Smoothing for Noisy Nonstationary Multivariate Time Series *

Naoto Kunitomo †
and
Seisho Sato ‡

February 17, 2023

**Abstract**

We develop a new method called frequency regression and smoothing (or the SIML-frequency method) based on the nonstationary errors-in-variables model. It is developed for estimating the relationships among hidden states of random variables and handling noisy nonstationary small sample time series economic data in comparison with data in engineering fields and natural sciences. Many economic time series include not only trend, cycle, seasonal, and measurement error components, but also factors such as abrupt changes, trading-day effects, and institutional changes. The frequency regression and smoothing method can be applied to handle such factors in nonstationary time series. The proposed method is simple and applicable for analyzing nonstationary economic time series and handling seasonal adjustments. Our formulation leads to the asymptotic results on the low frequency method proposed by M*ü*ller and Watson (2018) as a consequence. An illustrative empirical analysis of the macro-consumption in Japan is provided.

**Key words**

Noisy nonstationary time series, Trend-cycle, Measurement error and seasonality, SIML filtering, Frequency Regression, M*ü*ller and Watson (2018), Breaks and institutional changes, Seasonal adjustment

---

*Version 2023-2-17. It is a revised version of the paper (SDS-20, MIMS=RSP Statistics and Data Science Series, MIMS, Meiji University). http://www.mims.meiji.ac.jp/publications/2021-ds/SDS-20.pdf The research has been supported by JSPS-Grant JP17H02513.

†Institute of Statistical Mathematics, Midoricho 10-3, Tachikawa-shi, Tokyo 190-8562, JAPAN, kunitomo@ism.ac.jp

‡Graduate School of Economics, University of Tokyo, Bunkyo-ku,Hongo7-3-1, Tokyo113-0033, seisho@e.u-tokyo.ac.jp

## 1. Introduction

A considerable amount of research has been published on the use of statistical time series analysis of macroeconomic time series data. An important feature of the macroeconomic time series, which is different from the standard time series analysis, is that the observed time series is an apparent mixture of nonstationary and stationary components including apparent seasonality. The nonstationarity may include not only classical trend-cycle components, but also abrupt changes and outliers in the trend-noise components. Recent (vivid) examples are the macro-effects of COVID-19 occurring in 2020-2023 and the financial crisis in 2008-2009. Another feature is the fact that measurement errors in the economic time series play an important role because many macroeconomic data are constructed from various sources including sample surveys from major official statistics, whereas the statistical time series analysis often ignores measurement errors. Further, many official agencies in the world apply the X-12-ARIMA or X-13ARIMA-SEATS programs of the U.S. Census Bureau, which use the univariate reg-ARIMA model to remove seasonality, as the standard filtering procedure to publish the seasonally adjusted data. (See Census Bureau (2020).) The last important feature is that the sample size of the macroeconomic data is rather small in comparison with many data in engineering fields and natural sciences. We obtain 120 time series observations for each series after collecting quarterly data over 30 years, for instance.

The quarterly GDP series and its major components in Japan, which have been the most important data in Japanese macroeconomy, are constructed since 1994 by the cabinet office of Japan at 2023 year. Unlike the U.S.macro-data, both original and seasonally adjusted data have been published from Cabinet Office, and then for us they may give a good opportunity to analyze the appropriateness of official seasonal adjustment from original series. Since the sample size is small, it is important to use an appropriate statistical procedure to extract information on the trend-cycle, seasonal and noise (or measurement error) components in a systematic manner from data.

In this study, we develop a new statistical method called the frequency regression and smoothing (or the SIML filtering or smoothing method) based on the nonstationary errors-in-variables model to estimate hidden states of random variables and handle multiple time series data. In particular, it is based on the frequency domain analysis of nonstationary time series and it can be applicable to small sample economic data. For estimating nonstationary errors-in-variables models, we develop the linear regression methods in the frequency domain for nonstationary multivariate time series. Macroeconomic variables include important factors such as structural breaks, trading-day effects, and institutional changes in addition to trend, cycle, and seasonal components as well as the measurement errors. Since there are many factors in nonstationary time series, statistical method that can cope with them in a systematic and coherent manner yet to be developed. The proposed SIML method for analyzing nonstationary multivariate time series. can be applied to handle these factors systematically. It is simple and applicable to several problems when analyzing a nonstationary multivariate economic time series.

There are several related studies to our method. In statistical multivariate analysis, some studies on the errors-in-variables models are Anderson (1984, 2003) and Fuller (1987); however, they considered multivariate statistical models for independent observations, and the underlying situation is different from ours. As classical time series studies, Granger and Hatanaka (1964), and Brillinger and Hatanala (1969) introduced the spectral and harmonic analysis of economic time series. Engle (1974) proposed the band spectrum regression for stationary economic time series. Also our work is closely related to the problem of Baxter and King (1999), and M*ü*ller and Watson (2018). In particular, our formulation leads to some asymptotic results on the low frequency method proposed by M*ü*ller and Watson (2018) as a consequence, which seem to be new in the literature. Our method of frequency regression could be regarded as extensions of their analyses to nonstationary time series in the sense that we can use not only the trend-cycle components, but also the seasonal components, institutional changes and trading-day effects. The novel feature of this study may be to use the spectral decomposition of non-stationary multivariate time series and it is a generalization of Kunitomo and Sato (2021).

The rest of the manuscript is organized as follows. In Section 2, we explain the nonstationary errors-in-variables model and the SIML filtering (or smoothing) method. Then, in Section 3, we introduce the frequency regression method and as an application, we mention to the result obtained by Müller and Watson (2018). In Section 4, we discuss the regression smoothing method based on SIML smoothing. In Section 5, we discuss the likelihood function and in Section 6, we show an illustrative empirical analysis of the macro-consumption of durable goods in Japan. In Section 7, we provide some concluding remarks. Some details of the mathematical derivations of the theoretical results on frequency regression and the corresponding figures are presented in the Appendix.

## 2. Nonstationary errors-in-variables models and SIML Filtering

### 2.1 Nonstationary errors-in-variables models

Let $y_{ji}$ be the $i-$th observation of the $j-$th time series at $i$ for $i = 1, \cdots, n; j = 1, \cdots, p$. Let $\mathbf{y}_i = (y_{1i}, \cdots, y_{pi})'$ be a $p \times 1$ vector and $\mathbf{Y}_n = (\mathbf{y}_i^{'})$ $(= (y_{ij}))$ be an $n \times p$ matrix of observations, further let $\mathbf{y}_0$ be the initial $p \times 1$ vector. We investigate the statistical time series model with several unobservable components when we have the underlying nonstationary component $\mathbf{x}_i$ $(= (x_{ji}))$ $(i = 1, \cdots, n)$, which may be represented an $I(1)$ process, the vector of the noise (or measurement error) component $\mathbf{v}_i^{'} = (v_{1i}, \cdots, v_{pi})$, which may include the seasonal component and be represented an $I(0)$ process. We use the nonstationary errors-in-variables representation in the additive form

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

where $\mathbf{x}_i$ and $\mathbf{v}_i$ $(i = 1, \cdots, n)$ are sequences of a nonstationary I(1) process and a stationary I(0) process. (We assume that $\mathbf{x}_i$ and $\mathbf{v}_i$ $(i = 1, \cdots, n)$ are mutually independent for the simplicity.) The nonstationary state variable satisfies

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

with the lag operator $\mathcal{L}\mathbf{x}_i = \mathbf{x}_{i-1}$, $\Delta = 1 - \mathcal{L}$, and

$$\mathbf{v}_i^{(x)} = \sum_{j=0}^{\infty} \mathbf{C}_j^{(x)} \mathbf{e}_{i-j}^{(x)} \ , \tag{2.3}$$

where $\mathbf{e}_i^{(x)}$ denotes a sequence of i.i.d. random vectors with $\mathbf{E}(\mathbf{e}_i^{(x)}) = \mathbf{0}$ and $\mathbf{E}(\mathbf{e}_i^{(x)} \mathbf{e}_i^{(x)'}) = \mathbf{\Sigma}_e^{(x)}$ (a positive-semi-definite matrix). The $p \times p$ coefficient matrices $\mathbf{C}_j^{(x)}$ $(= c_{kl}^{(x)}(j))$ are absolutely summable and $\|\mathbf{C}_j^{(x)}\| = O(\rho^j)$, where $0 \leq \rho < 1$ and $\|\mathbf{C}_j^{(x)}\| = \max_{k,l=1,\cdots,p} |c_{kl}^{(x)}(j)|$. The initial values $\mathbf{y}_0$ $(=\mathbf{x}_0)$ is fixed. The stationary noise component $\mathbf{v}_i$ satisfies

$$\mathbf{v}_i = \sum_{j=0}^{\infty} \mathbf{C}_j^{(v)} \mathbf{e}_{i-j}^{(v)} \ , \tag{2.4}$$

where the $p \times p$ coefficient matrices $\mathbf{C}_j^{(v)}$ are absolutely summable and $\|\mathbf{C}_j^{(v)}\| = O(\rho^j)$, where $0 \leq \rho < 1$ and $\mathbf{e}_i^{(v)}$ represents a sequence of i.i.d. random vectors with $\mathbf{E}(\mathbf{e}_i^{(v)}) = \mathbf{0}$ and $\mathbf{E}(\mathbf{e}_i^{(v)} \mathbf{e}_i^{(v)'}) = \mathbf{\Sigma}_e^{(v)}$ (positive definite matrix).

We consider the situation when we have the observations of an $n \times p$ matrix $\mathbf{Y}_n = (\mathbf{y}_i^{'})$ and set the $np \times 1$ random vector $(\mathbf{y}_1^{'}, \cdots, \mathbf{y}_n^{'})^{'}$. The non-stationary errors-in-variables formulation in (2.1)-(2.4) includes several important time series decomposition with the trend, cycle, seasonal and irregular components. The important feature is that we have both non-stationary and stationary components. Apparently, the trend component is nonstationary while the measurement error component is stationary. Although the observed process $\mathbf{y}_i$ is an I(1) process and the difference process $\Delta \mathbf{y}_i$ is an I(0) process, it is important to distinguish different components by considering their spectral characteristics in the frequency domain. It is because we often need to detect the trend-cycle component and seasonal component of macro-economic time series although the number of observations is usually small in comparison with data in many engineering fields and natural sciences. The information of time series in different frequencies is important. (The implication of our formulation in frequency domain will be explained in Section 5.)

There are several decomposition models of time series based on (2.1)-(2.4).
When the state variables of our interest are the trend-cycle components $\mathbf{TC}_i$ ($p \times 1$ vectors), we take $\mathbf{x}_i = \mathbf{TC}_i$ ($i = 1, \cdots, n$) and $\Delta\mathbf{TC}_i = \mathbf{TC}_i - \mathbf{TC}_{i-1}$ is a stationary process, which has the MA representation of (2.3). In this case $\mathbf{v}_i$ is the stationary error process including the measurement errors except the trend-cycle components. An important example is the seasonal adjustment such as X-13ARIMA-SEATS of the U.S. Census Bureau, which usually estimates the seasonal components to construct seasonal adjusted series. When the state variables of our interest are the trend-cycle components $\mathbf{TC}_i$ and the seasonal component $\mathbf{s}_i$ ($p \times 1$ vectors), we may interpret that $\Delta\mathbf{TC}_i = \mathbf{TC}_i - \mathbf{TC}_i$ and $\mathbf{s}_i$ are stationary processes, and $\Delta\mathbf{x}_i = \Delta\mathbf{TC}_i + \mathbf{s}_i$ has the MA representation of (2.3). In this case, we may have $\Delta\mathbf{TC}_i = \sum_{j=0}^{\infty} \mathbf{C}_j^{(TC)} \mathbf{e}_{i-sj}^{(TC)}$ and $\mathbf{s}_i = \sum_{j=0}^{\infty} \mathbf{C}_{sj}^{(s)} \mathbf{e}_{i-sj}^{(s)}$ , where the lag operator is defined by $\mathcal{L}^s \mathbf{s}_i = \mathbf{s}_{i-s}$ ($s \geq 2$), and $\mathbf{e}_i^{(TC)}$ and $\mathbf{e}_i^{(s)}$ represent sequences of i.i.d. random vectors. That is, $\mathbf{E}(\mathbf{e}_i^{(TC)}) = \mathbf{E}(\mathbf{e}_i^{(s)}) = \mathbf{0}$ and $\mathbf{E}(\mathbf{e}_i^{(TC)} \mathbf{e}_i^{(TC)'}) = \mathbf{\Sigma}_e^{(TC)}$ (non-negative definite matrix) and $\mathbf{E}(\mathbf{e}_i^{(s)} \mathbf{e}_i^{(s)'}) = \mathbf{\Sigma}_e^{(s)}$ (non-negative definite matrix). The $p \times p$ coefficient matrices $\mathbf{C}_j^{(TC)}$ and $\mathbf{C}_j^{(s)}$ are absolutely summable such that $\|\mathbf{C}_j^{(TC)}\| = O(\rho^j)$ and $\|\mathbf{C}_j^{(s)}\| = O(\rho^j)$, where $0 \leq \rho < 1$.
We will explain the filtering procedure to estimate the trend-cycle components as state variables and the frequency regression on trend-cyle components in Sections 2 and 3. Also we will explain the filtering procedure to estimate the seasonal components as the state variables for the seasonal adjustment in Sections 2 and 4.

When there are no cycle and seasonal components, we take that 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, $\mathbf{\Sigma}_x = \mathbf{\Sigma}_e^{(x)}$ and $\mathbf{\Sigma}_v = \mathbf{\Sigma}_e^{(v)}$. Then, given the initial condition $\mathbf{y}_0$ and $\mathbf{x}_0$, $\mathrm{vec}(\mathbf{Y}_n) \sim N_{n \times p}\left(\mathbf{1}_n \cdot \mathbf{y}_0^{'}, \mathbf{I}_n \otimes \mathbf{\Sigma}_v + \mathbf{C}_n \mathbf{C}_n^{'} \otimes \mathbf{\Sigma}_x\right)$, where $\mathbf{1}_n^{'} = (1, \cdots, 1)$ and

$$\mathbf{C}_n = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 1 & \cdots & 0 \\ 1 & \cdots & 1 & 1 & 0 \\ 1 & \cdots & 1 & 1 & 1 \end{pmatrix}_{n \times n} . \tag{2.5}$$

For the non-stationary errors-in-variables model in (2.1)-(2.4), we introduce the $\mathrm{K}_n^*-$transformation from $\mathbf{Y}_n$ to $\mathbf{Z}_n$ $(= (\mathbf{z}_k^{'}))$ using

$$\mathbf{Z}_n = \mathbf{K}_n^* \left(\mathbf{Y}_n - \bar{\mathbf{Y}}_0\right) , \mathbf{K}_n^* = \mathbf{P}_n \mathbf{C}_n^{-1} , \tag{2.6}$$

where $\bar{\mathbf{Y}}_0 = \mathbf{1}_n \mathbf{y}_0^{'}$,

$$\text{(2.7)} \qquad \mathbf{C}_n^{-1} = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & 0 & \cdots \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}_{n\times n} ,$$

and

$$\text{(2.8)} \qquad \mathbf{P}_n = (p_{jk}^{(n)}) \ , \ p_{jk}^{(n)} = \sqrt{\frac{2}{n+\frac{1}{2}}} \cos\left[\frac{2\pi}{2n+1}(k-\frac{1}{2})(j-\frac{1}{2})\right] \ .$$

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

$$\text{(2.9)} \qquad a_{kn}^{*} \ (= d_k) \ = 4\sin^2\left[\frac{\pi}{2}\left(\frac{2k-1}{2n+1}\right)\right] \ (k = 1, \cdots, n) \ .$$

When we have the non-stationary and stationary parts, it is natural to use the $\mathbf{K}_n^{*}-$transformation to decompose time series data with discrete time into two parts. Because we use (2.7), the transformed time series is stationary and it is possible to use the spectral decomposition of stationary time series with discrete time in the frequency domain.

### 2.2 SIML filtering method

We consider the general filtering procedure based on the $\mathbf{K}_n^{*}-$transformation (2.6). It is easy to interpret the role of the elements of the resulting $n \times p$ random matrix $\mathbf{Z}_n$ in the data analysis because they are obtained by the transformation that takes real values in the frequency domain. We consider the inversion of the transformed parts of orthogonal frequency processes.

Let an $n \times p$ matrix

$$\text{(2.10)} \qquad \hat{\mathbf{X}}_n = \mathbf{C}_n\mathbf{P}_n\mathbf{Q}_n\mathbf{P}_n\mathbf{C}_n^{-1}(\mathbf{Y}_n - \bar{\mathbf{Y}}_0) \ ,$$

where $\mathbf{Z}_n = \mathbf{P}_n\mathbf{C}_n^{-1}(\mathbf{Y}_n - \bar{\mathbf{Y}}_0)$ and $\mathbf{Q}_n$denotes an $n \times n$ filtering matrix.
The stochastic process $\mathbf{Z}_n$ represents the orthogonal decomposition of the original time series $\mathbf{Y}_n$. Then, the time series $\hat{\mathbf{X}}_n$ represents the realization of some frequency parts of the original time series.
We provide explicit forms of useful examples including the trend-cycle filtering procedure and seasonal filtering procedure for economic time series. For the filtering (or smoothing) method in the form of (2.10), we provide two examples.

**Example 1 : Trend Smoothing** : Let an $m \times n$ choice matrix $(0 < m < n)$ $\mathbf{J}_m = (\mathbf{I}_m, \mathbf{O})$, and let $n \times p$ matrix

$$\hat{\mathbf{X}}_n = \mathbf{C}_n \mathbf{P}_n \mathbf{J}_m^{'} \mathbf{J}_m \mathbf{P}_n \mathbf{C}_n^{-1} (\mathbf{Y}_n - \bar{\mathbf{Y}}_0) \tag{2.11}$$

and an $n \times n$ matrix

$$\mathbf{Q}_n = \mathbf{J}_m^{'} \mathbf{J}_m \ . \tag{2.12}$$

We construct an estimator of the $n \times p$ hidden state matrix $\mathbf{X}_n$ in the frequency domain using the inverse transformation of $\mathbf{Z}_n$. We can recover the trend-cycle components by deleting the estimated noise parts in the high-frequency and using only low frequency parts.

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

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

and

$$\mathbf{P}_n \mathbf{J}_m^{'} \mathbf{J}_m \mathbf{P}_n = \begin{pmatrix} \mathbf{P}_{11}^{'} \\ \mathbf{P}_{12}^{'} \end{pmatrix} (\mathbf{P}_{11}, \mathbf{P}_{12}) = \mathbf{I}_n - \begin{pmatrix} \mathbf{P}_{21}^{'} \\ \mathbf{P}_{22}^{'} \end{pmatrix} (\mathbf{P}_{21}, \mathbf{P}_{22}) \ . \tag{2.13}$$

Then the $(j, j^{'})$-th element of $\mathbf{A}_n = \mathbf{P}_n \mathbf{J}_m^{'} \mathbf{J}_m \mathbf{P}_n$ $(= (a_{j,j^{'}}))$ is given by

$$\begin{aligned} a_{j,j} &= \frac{2m}{2n+1} + \frac{1}{2n+1} \left[ \frac{\sin \frac{2m\pi}{2n+1}(2j-1)}{\sin \frac{\pi}{2n+1}(2j-1)} \right] , \\ a_{i,j^{'}} &= \frac{1}{2n+1} \left[ \frac{\sin \frac{2m\pi}{2n+1}(j+j^{'}-1)}{\sin \frac{\pi}{2n+1}(j+j^{'}-1)} + \frac{\sin \frac{2m\pi}{2n+1}(j-j^{'})}{\sin \frac{\pi}{2n+1}(j-j^{'})} \right] \quad (j \neq J^{'}) \ . \end{aligned} \tag{2.14}$$

**Example 2 : Band Smoothing** : We consider the band filtering based on the $\mathbf{K}_n^*-$ transformation in (2.6) and use the inversion of only a band of frequency parts from the random matrix $\mathbf{Z}_n$. A leading example is the seasonal frequency in the discrete time series, we consider $s$ $(> 1)$ to be a positive integer in this case. 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 the $n \times p$ matrix

$$\hat{\mathbf{X}}_n = \mathbf{C}_n \mathbf{P}_n \mathbf{J}_{m_1,m_2,n}^{'} \mathbf{J}_{m_1,m_2,n} \mathbf{P}_n \mathbf{C}_n^{-1} (\mathbf{Y}_n - \bar{\mathbf{Y}}_0) \tag{2.15}$$

and the $n \times n$ matrix

$$\mathbf{Q}_n = \mathbf{J}_{m_1,m_2,n}^{'} \mathbf{J}_{m_1,m_2,n} \ . \tag{2.16}$$

As an example, when we have the seasonal frequency $\lambda_s$ $(0 \leq \lambda_s \leq \frac{1}{2})$, we take $m_1 = [2n/s] - h$ and $m_2 = 2h+1$. The $(j, j^{'})$-th element of $\mathbf{A}_n = \mathbf{P}_n \mathbf{J}_{m_1,m_2,n}^{'} \mathbf{J}_{m_1,m_2,n} \mathbf{P}_n$ $(= (a_{j,j^{'}}))$ is given by

$$a_{j,j} = \frac{2m_2}{2n+1} + \frac{1}{2n+1} \left[ \frac{\sin \frac{2(m_1+m_2)\pi}{2n+1}(2j-1) - \sin \frac{2(m_1)\pi}{2n+1}(2j-1)}{\sin \frac{\pi}{2n+1}(2j-1)} \right] , \tag{2.17}$$

$$a_{i,j'} = \frac{1}{2n+1}\left[\frac{\sin\frac{2(m_1+m_2)\pi}{2n+1}(j+j'-1) - \sin\frac{2(m_1)\pi}{2n+1}(j+j'-1)}{\sin\frac{\pi}{2n+1}(j+j'-1)} + \frac{\sin\frac{2(m_1+m_2)\pi}{2n+1}(j-j') - \sin\frac{2(m_1)\pi}{2n+1}(j-j')}{\sin\frac{\pi}{2n+1}(j-j')}\right] \quad (j \neq j') .$$

When $m_1 = 0$ and $m_2 = m$, (2.16) becomes (2.12) in Example 1. When we have seasonality, however, there is a complication in the data analysis to be considered. Because we have discrete observations such as quarterly or monthly data, we need to use several frequency parts of the transformed process. For quarterly data, a 1 year (4 quarters) cycle cannot be distinguished from the 2 quarters cycle. For monthly data, the 1 year cycle cannot be distinguished from the 6, 4, 3, 2.4, and 2 months cycles. (We shall discuss examples in Section 4 in details.)

## 3. Frequency Regression

In this section, we partition the transformed random variables in the frequency domain and consider a linear regression model based on observations of $q \times p$ matrix $\mathbf{Z}_m^*$ by

(3.1) $$\mathbf{Z}_m^* = \mathbf{F}_q \mathbf{P}_n \mathbf{C}_n^{-1}(\mathbf{Y}_n - \bar{\mathbf{Y}}_0) = [\mathbf{z}_{1m}^*, \mathbf{Z}_{2m}^*] ,$$

where $\mathbf{F}_q$ denotes a $q \times n$ matrix and $q$ $(> p)$ depends on $n$ as $q = q_n$. In this notation $\mathbf{z}_{1m}^*$ is a $q \times 1$ vector and $\mathbf{Z}_{2m}^*$ is a $q \times (p-1)$ matrix.

There are several interesting examples. Since we consider the case when the rank of $\mathbf{F}_q$ is $p$ $(p < q)$, let us investigate this case.

When we have nonstationary time series, we often have trend, cycle, seasonal, and noise components. To handle these components, we can use a more complicated transformation $\mathbf{F}_q$. Further, there are trading-day components, leap year effects, structural changes such as the 2008 financial crisis and the 2020-2022 COVID-19 crisis, and institutional changes such as the consumption tax in Japan. When we use seasonal adjusted data, which are published by official agencies, it is important to handle these effects in meaningful ways. It is important to understand that the many official agencies use the X-12-ARIMA or X-13ARIMA-SEATS programs of the U.S. Census Bureau, which utilized the Reg-ARIMA modelling to deal with the problems. Thus, from the standard statistical view, it is known that an ad hoc method may be followed to handle these effects in official statistics partly because there are many factors to be considered and/or interpreted.

We can consider the simple the case of the transformation when $\mathbf{F}_q = \mathbf{J}_m$. We first investigate this case and assume that the rank of $\mathbf{F}_q$ is $p$ $(p < q)$. We define $p \times p$ matrices

(3.2) $$\mathbf{G}_m^* = \frac{1}{m}\mathbf{Z}_m^{*'}\mathbf{Z}_m^* \ , \mathbf{G}_n = \frac{1}{n}\mathbf{Z}_n^{'}\mathbf{Z}_n$$

and their probability limits as $m = m_n \to \infty$ $(n \to \infty, m_n/n \to 0)$

$$\text{plim}_{n\to\infty}\mathbf{G}_m^* = \mathbf{\Sigma}_x \ , \ \text{plim}_{n\to\infty}\mathbf{G}_n = \mathbf{\Sigma}_{\Delta y} \ , \tag{3.3}$$

where

$$\mathbf{\Sigma}_x = (\sum_{j=0}^{\infty}\mathbf{C}_j^{(x)})\mathbf{\Sigma}_e^{(x)}(\sum_{j=0}^{\infty}\mathbf{C}_j^{(x)'}) \ (= \mathbf{f}_{\Delta x}(0)) \ , \tag{3.4}$$

where $\mathbf{\Sigma}_x$ represents the spectral density matrix of $\Delta\mathbf{x}_i$ at zero frequency. (See (5.1) in Section 5 and Theorem A.1 in the Appendix.) The $p\times p$ matrix $\mathbf{\Sigma}_{\Delta y}$ is the spectral density matrix of $\Delta\mathbf{y}_i$ at zero frequency, whch is different from $\mathbf{\Sigma}_x$ (the long-run variance-covariance matrix of $\Delta\mathbf{x}_i$) in the errors-in-variables models of (2.1)-(2.4). Let $\mathbf{\Sigma}_v = (\sum_{j=0}^{\infty}\mathbf{C}_j^{(v)})\mathbf{\Sigma}_e^{(v)}(\sum_{j=0}^{\infty}\mathbf{C}_j^{(v)'})$, which is a $p \times p$ positive definite matrix because we assumed that $\mathbf{\Sigma}_e^{(v)}$ is positive definite in (2.4). This assumption on the errors-in-variables models has an important role because we have both the signal and noise terms in the observed time series.

We partition $\mathbf{G}_m^*$, $\mathbf{\Sigma}_x$ and $\mathbf{\Sigma}_v$ into $(1+k)\times(1+k)$ $(k = p-1)$ submatrices as

$$\mathbf{G}_m^* = \begin{bmatrix} g_{11}^* & \mathbf{g}_{12}^* \\ \mathbf{g}_{21}^* & \mathbf{G}_{22}^* \end{bmatrix} \ , \ \mathbf{\Sigma}_x = \begin{bmatrix} \sigma_{11}^{(x)} & \boldsymbol{\sigma}_{12}^{(x)} \\ \boldsymbol{\sigma}_{21}^{(x)} & \mathbf{\Sigma}_{22}^{(x)} \end{bmatrix} \ , \ \mathbf{\Sigma}_v = \begin{bmatrix} \sigma_{11}^{(v)} & \boldsymbol{\sigma}_{12}^{(v)} \\ \boldsymbol{\sigma}_{21}^{(v)} & \mathbf{\Sigma}_{22}^{(v)} \end{bmatrix} \ . \tag{3.5}$$

Then, we will investigate statistical properties of the least squares estimator in the frequency domain

$$\hat{\boldsymbol{\beta}}_m = \mathbf{G}_{22}^{*-1}\mathbf{g}_{21}^* \ , \tag{3.6}$$

which is an estimator of vector $\boldsymbol{\beta}_m = \mathbf{\Sigma}_{22}^{(x)-1}\boldsymbol{\sigma}_{21}^{(x)}$ under the assumption that the inverse matrices of $\mathbf{G}_{22}^*$ and $\mathbf{\Sigma}_{22}^{(x)}$ exist. (We need to assume that $\mathbf{\Sigma}_{22}^{(x)}$ has a full rank.)

We write

$$\hat{\boldsymbol{\beta}}_m - \boldsymbol{\beta} = [\mathbf{Z}_{2m}^{*'}\mathbf{Z}_{2m}^*]^{-1}\mathbf{Z}_{2m}^{*'}\mathbf{Z}_m^*\begin{pmatrix} 1 \\ -\boldsymbol{\beta} \end{pmatrix} \ , \tag{3.7}$$

where we partitioned $\mathbf{Z}_m^*$ into $q\times(1+k)$ submatrices $\mathbf{Z}_m^* = (\mathbf{z}_{1m}^*, \mathbf{Z}_{2m}^*)$.
Then we have the next result on the asymptotic properties of the least squares estimator and the proof is presented in the Appendix.

**Theorem 3.1** : Let $m_n = n^{\alpha}$, $m = [m_n]$ and $m \to \infty$ (as $n \to \infty$). In (2.1)-(2.4), assume that the fourth-order moments of $\mathbf{e}_i^{(x)}$ and $\mathbf{e}_i^{(v)}$ are bounded.
(i) For $0 < \alpha < 1$, $\boldsymbol{G}_m^*$ is a consistent estimator of $\mathbf{\Sigma}_x$ as $n \to \infty$.
(ii) Assume that the rank of $\mathbf{\Sigma}_{22}^{(x)}$ is $k$ $(= p-1)$ and $0 < \alpha < 0.8$. Then when $m \to \infty$ $(n \to \infty)$, $\sqrt{m_n}[\hat{\boldsymbol{\beta}}_m - \boldsymbol{\beta}]$ is asymptotically and normally distributed as $N(\mathbf{0}, \sigma_{11.2}\mathbf{\Sigma}_{22}^{(x)-1})$ and $\sigma_{11.2}^{(x)} = \sigma_{11}^{(x)} - \boldsymbol{\sigma}_{12}^{(x)}\mathbf{\Sigma}_{22}^{(x)-1}\boldsymbol{\sigma}_{21}^{(x)}$.

Then, we can rewrite $\mathbf{u}_m = \mathbf{z}^*_{1m} - \mathbf{Z}^*_{2m}\boldsymbol{\beta}$, that is,

$$\mathbf{z}^*_{1m} = \mathbf{Z}^*_{2m}\boldsymbol{\beta} + \mathbf{u}_m \tag{3.8}$$

and $\mathbf{E}[\mathbf{u}_m] = \mathbf{0}$. This is a linear regression equation, however, the error term of $\mathbf{u}_m$ has a specific form of heteroscedasticity.
Theorem 3.1 is valid when we estimate the covariance matrix $\boldsymbol{\Sigma}_x$ of the hidden state variables, which is different from the observed covariance of the differenced data $\boldsymbol{\Sigma}_{\Delta y}$. We can delete the effects of noisy parts of nonstationary time series by using $\mathbf{G}_m$ instead of $\mathbf{G}_n$ with the condition $m_n/n \to 0$ and $m_n \to \infty$. By the condition $0 < \alpha < 0.8$, we can recover the asymptotic normality of the least squares estimator without noises.

One direct application of Theorem 3.1 is Müller and Watson (2018), who proposed the *so-called* long-run co-variability of macroeconomic time series. They investigated many nonstationary time series using their method and obtained some interesting findings. We can interpret their method as the relationships among long-run trends in our framework when $p = 2$. Let $2 \times 2$ matrices $\boldsymbol{\Sigma}_e^{(x)} = (\sigma_{ij}^{(x)})$; then, we define the regression coefficient $\boldsymbol{\beta} = [\sigma_{22}^{(x)}]^{-1}\sigma_{21}^{(x)}$ under the assumption that $\sigma_{22}^{(x)}$ $(= \boldsymbol{\Sigma}_{22}^{(x)}) > 0$. Further, let $\boldsymbol{G}^*_m = (\hat{g}_{ij}^{(x)})$, and an $n \times 2$ matrix

$$(\mathbf{a}_{1n}, \mathbf{a}_{2n}) = \mathbf{C}_n^{-1}(\mathbf{Y}_n - \mathbf{Y}_0)\ . \tag{3.9}$$

For estimating $\boldsymbol{\beta}$, we define the estimated regression coefficient as

$$\hat{\boldsymbol{\beta}} = [\hat{g}_{22}^{(x)}]^{-1}\hat{g}_{21}^{(x)} = [\mathbf{a}_{2n}^{'}\mathbf{P}_n\mathbf{J}_m\mathbf{J}_m^{'}\mathbf{P}_n\mathbf{a}_{2n}]^{-1}[\mathbf{a}_{2n}^{'}\mathbf{P}_n\mathbf{J}_m\mathbf{J}_m^{'}\mathbf{P}_n\mathbf{a}_{1n}]\ . \tag{3.10}$$

This quantity can be interpreted as the least squares slope of the transformed vector from $\mathbf{y}_{1n}$ on the transformed vector from $\mathbf{y}_{2n}$ for a $n \times 2$ matrix $\mathbf{Y}_n = (\mathbf{y}_{1n}, \mathbf{y}_{2n})$; that is, essentially the same as the estimation method proposed by Müller and Watson (2018) [1]. However, there is an important difference between the SIML method and their method, that is, we consider the situation when $m = [m_n] \to \infty$ as $n \to \infty$ while $m_n/n \to 0$. This is a natural framework for the asymptotic theory on their method.
We fix $m$, which is independent from $n$ and we investigate the case when $\Delta\mathbf{x}_i$ and $\mathbf{v}_i$ $(i = 1, \cdots, n)$ are mutually independent with $\mathbf{y}_0 = \mathbf{x}_0 = \mathbf{0}$ for the simplicity. Because (2.1)-(2.4), we find that $\boldsymbol{\Sigma}_{\Delta y} = \boldsymbol{\Sigma}_x + 2\boldsymbol{\Sigma}_v$. Then, as $n \to \infty$,

$$\hat{\boldsymbol{\beta}} \stackrel{p}{\longrightarrow} [\boldsymbol{\Sigma}_{22}^{(x)} + 2\boldsymbol{\Sigma}_{22}^{(v)}]^{-1}[\boldsymbol{\sigma}_{21}^{(x)} + 2\boldsymbol{\sigma}_{21}^{(v)}]\ . \tag{3.11}$$

[1] In their notation, $m$ corresponds to $q$, which is fixed. They did use (differenced) stationary data, and thus, we could interpret that they calculated the linear regression from the filtered data $\hat{\mathbf{X}}^*_n = \mathbf{P}_n^{'}\mathbf{J}_m^{'}\mathbf{J}_m\mathbf{P}_n\mathbf{C}_n^{-1}(\mathbf{Y}_n - \mathbf{Y}_0)$ as a modification of (2.11) in our notation.

This corresponds to the fact that the least squares estimator is not consistent when the sample size is large in the classical errors-in-variables models. (See Anderson (1984) in details.) When $\mathbf{\Sigma}_v$ is relatively small, the probability limit of $\hat{\boldsymbol{\beta}}$ is close to $\boldsymbol{\beta}$ because the error terms $\mathbf{v}_i$ are negligible. If there are no measurement errors as the standard time series analysis, $\mathbf{\Sigma}_v = \mathbf{O}$ and $\hat{\boldsymbol{\beta}} \xrightarrow{p} \boldsymbol{\beta}$ as $n \to \infty$. From Theorem 3.1 (and Theorem A.1 in the Appendix), we btain the following result.

**Corollary 3.1** : When $p = 2$, we assume that $\mathbf{\Sigma}_e^{(x)}$ is positive semi-definite, $\mathbf{\Sigma}_e^{(v)}$ is positive definite, and the fourth-order moments of $\mathbf{e}_i^{(x)}$ and $\mathbf{e}_i^{(v)}$ $(i = 1, \cdots, n)$ are bounded.
(i) Fix $m$, which is independent of $n$. Then $\hat{\beta}$ in (3.10) is not consistent when $n \to \infty$.
(ii) Set $m_n = n^{\alpha}$ and $0 < \alpha < 1$ and construct $\hat{\boldsymbol{\beta}}_m$ in (3.6). Then, as $n \longrightarrow \infty$, $\hat{\boldsymbol{\beta}}_m - \boldsymbol{\beta} \xrightarrow{p} \mathbf{0}$ .
(iii) Set $m_n = n^{\alpha}$ and $0 < \alpha < 0.8$, then, as $n \longrightarrow \infty$, $\sqrt{m_n}[\hat{\boldsymbol{\beta}}_m - \boldsymbol{\beta}]$ is asymptotically normal with $N(0.\sigma_{11.2}^{(x)})$.

Since our method can be generalized to other situations beyond the trend-cycle components of non-stationary time series, it is a generalization of M$\ddot{u}$ller and Watson (2018). For instance, it is rather straight-forward to incorporate the regression effects of dummy variables in trend relations such as structural breaks and the seasonal frequency parts.

## 4. Regression Smoothing

When we have noisy-nonstationary time series, we often need to remove the seasonality and/or low frequency component. However, in some applications of official statistics, we need to construct the seasonally adjusted data after removing additional effects such as trading-day components including the leap year effect, structural changes such as the 2008 financial crisis and institutional changes such as the introduction of consumption tax in Japan. These effects are can be defined in deterministic ways.

Let the observed vector times series $\mathbf{y}_i$ be decomposed as

$$\mathbf{y}_i = \mathbf{x}_i + \mathbf{SCO}_i + \mathbf{v}_i \ \ (i = 1, \cdots, n), \tag{4.1}$$

and $\mathbf{SCO}_i = \mathbf{SC}_i + \mathbf{O}_i$, where $\mathbf{x}_i$ denotes the trend-cycle component, $\mathbf{SC}_i$ denotes the structural break component, $\mathbf{v}_i$ denotes the noise component, and $\mathbf{O}_i$ represents the outlier component.
In this section we consider the case where in $\mathbf{SC}_i$ and $\mathbf{O}_i$ can be expressed as $\mathbf{SC}_i + \mathbf{O}_i = \mathbf{SCO}_i(w)$, where $\mathbf{w}$ denotes the set of instrumental variables. If these terms can be expressed as linear relationships, we write

$$\mathbf{y}_i = \mathbf{B}^{'}\mathbf{w}_i + \mathbf{u}_i \ \ (i = 1, \cdots, n), \tag{4.2}$$

where $\mathbf{B}'$ denotes a $p \times r$ matrix, $\mathbf{w}_i$ denotes a $r \times 1$ vector of instrumental variables, $\mathbf{z}_i$ and $\mathbf{u}_i = \mathbf{x}_i + \mathbf{v}_i$ represents a sequence of I(1) process. Hence, the model is a multivariate regression model when the noise terms are $I(1)$ process with stationary noise term and seasonal terms. We incorporate extraneous information such as dummy variables to extract or delete some components from the observed time series based on (4.1).
To find the regression and smoothing procedure of trend and seasonal components, we use the $K_n^*-$transformation of data and rewrite (4.2) as

$$\mathbf{Y}_n^* = \mathbf{W}_n^* \mathbf{B} + \mathbf{U}_n^* \ , \tag{4.3}$$

where $\mathbf{Y}_n^* = \mathbf{P}_n \mathbf{C}_n^{-1}(\mathbf{Y}_n - \mathbf{Y}_0)$ and $\mathbf{W}_n^* = \mathbf{P}_n \mathbf{C}_n^{-1} \mathbf{W}_n$ ($\mathbf{W}_n = (\mathbf{w}_t'))$ represent $n \times p$ and $n \times r$ matrices of the explained variables and explanatory variables, respectively, and $\mathbf{U}_n^* = \mathbf{P}_n \mathbf{C}_n^{-1} \mathbf{U}_n$ and $\mathbf{U}_n = (\mathbf{u}_i')$ are $n \times p$ disturbance matrices. (We fix the initial condition $\mathbf{y}_0$ $(= \mathbf{x}_0)$ and the state variables $\mathbf{x}_i^* = \mathbf{x}_i - \mathbf{x}_0$.
As a consequence of the $K_n^*-$transformation, we have the disturbance terms in (4.3), that are stationary processes.

Because (4.2) is a linear regression equation, it is possible to apply Theorem 3.1 by defining a $(p + r) \times 1$ vector

$$\mathbf{y}_i^* = \begin{bmatrix} \mathbf{y}_i \\ \mathbf{w}_i \end{bmatrix} .$$

Then we can estimate the regression coefficients and calculate the residuals from the regression equations. When vectors $\mathbf{w}_i$ $(i = 1, \cdots, n)$ are deterministic, we assume that

$$\lim_{m \to \infty} \frac{1}{m} \mathbf{W}_m^{*\prime} \mathbf{W}_m^* = \mathbf{\Sigma}_{w^*} \ , \tag{4.4}$$

where $\mathbf{\Sigma}_{w^*}$ denotes a positive definite matrix and $\mathbf{W}_m^* = \mathbf{J}_m \mathbf{P}_n \mathbf{C}_n^{-1} \mathbf{W}_n$ represents an $m \times r$ matrix.
When the $r \times 1$ instrumental variables $\mathbf{w}_i$ $(i = 1, \cdots, n)$ are exogenous or deterministic, we have the following result from Theorem 3.1 and the proof is in the Appendix.

**Theorem 4.1** : In (2.1)-(2.4), assume that the fourth-order moments of $\mathbf{e}_i^{(x)}$ and $\mathbf{e}_i^{(v)}$ are bounded. Let $\mathbf{Y}_m^* = \mathbf{W}_m^* \mathbf{B} + \mathbf{U}_m^*$ and $\mathbf{U}_m^* = \mathbf{J}_m \mathbf{U}_n^*$, where $\mathbf{Y}_m^* = \mathbf{J}_m \mathbf{P}_n \mathbf{C}_n^{-1} \mathbf{Y}_n$ and $\mathbf{W}_m^* = \mathbf{J}_m \mathbf{P}_n \mathbf{C}_n^{-1} \mathbf{W}_n$. We also assume the nonsingularity condition (4.4) and $\mathbf{w}_i$ $(i = 1, \cdots, n)$ is exogenous or deterministic function. We denote $\hat{\mathbf{B}}_m = (\mathbf{W}_m^{*\prime} \mathbf{W}_m^*)^{-1} \mathbf{W}_m^{*\prime} \mathbf{Y}_m^*$ is the least squares estimator of $\mathbf{Y}_m^*$ on $\mathbf{W}_m^*$. Let $m_n = n^{\alpha}$ $(0 < \alpha < 0.8)$ and $m = [m_n]$. Then, as $n \longrightarrow \infty$, we have the asymptotic normality

$$\sqrt{m_n}[\hat{\mathbf{B}}_m - \mathbf{B}] \xrightarrow{w} N(\mathbf{0}, \mathbf{\Sigma}_{w^*}^{-1} \otimes \mathbf{\Sigma}_x) \ . \tag{4.5}$$

Define the general transformed instrumental variables

$$\hat{\mathbf{W}}_n = \mathbf{J}_W \mathbf{P}_n \mathbf{C}_n^{-1} \mathbf{W}_n \ , \tag{4.6}$$

where $\mathbf{J}_W$ represents a $q \times n$ choice matrix, and we denote the idempotent matrix ($q \times q$ matrix)

$$\mathbf{Q}_W = \hat{\mathbf{W}}_n (\hat{\mathbf{W}}_n^{'} \hat{\mathbf{W}}_n)^{-1} \hat{\mathbf{W}}_n^{'} \ . \tag{4.7}$$

We utilize the regression information on smoothing by utilizing the projection matrix $\mathbf{Q}_W$ to construct

$$\hat{\mathbf{X}}_n = \mathbf{C}_n \mathbf{P}_n \mathbf{Q}_W \mathbf{P}_n \mathbf{C}_n^{-1} (\mathbf{Y}_n - \bar{\mathbf{Y}}_0) \ . \tag{4.8}$$

There are several possibilities to how we incorporate the extraneous information in the smoothing procedure. It is reasonable to consider the case when $\mathbf{Q}_W$ is an idempotent matrix such as $\mathbf{Q}_W^2 = \mathbf{Q}_W$. In our study, we use two alternative smoothing procedures : *Type-I* and *Type II.* Type-I smoothing may be appropriate for change-point smoothing in the trend component and Type-II smoothing may be appropriate for outlier detection in the noise component.

**Type-I Smoothing** :

Type-I is based on Example 1 presented in Section 2. The (trend-cycle) regression part of $\mathbf{Y}_n$ is (4.1) when we take $\mathbf{J}_W = (\mathbf{I}_m, \mathbf{O})$ ($\hat{\mathbf{W}}_n$ represents an $m \times r$ matrix and $\mathbf{J}_m^{'} = (\mathbf{I}_m, \mathbf{O})^{'}$ represents an $n \times m$ matrix) and an $n \times n$ matrix

$$\mathbf{Q}_n^{(0)} = \mathbf{J}_m^{'} \hat{\mathbf{W}}_n (\hat{\mathbf{W}}_n^{'} \hat{\mathbf{W}}_n)^{-1} \hat{\mathbf{W}}_n^{'} \mathbf{J}_m \ . \tag{4.9}$$

If we want to remove the regression effects and use only the trend-cycle part, we need to take $\mathbf{J}_W = \mathbf{J}_m$, $\mathbf{J}_m = (\mathbf{I}_m, \mathbf{O})$ ($m \times n$ choice matrix, $m \leq n$) and

$$\mathbf{Q}_n^{(1)} = \mathbf{J}_m^{'} \mathbf{J}_m - \mathbf{Q}_n^{(0)} = \mathbf{J}_m^{'} [\mathbf{I}_m - \hat{\mathbf{W}}_n (\hat{\mathbf{W}}_n^{'} \hat{\mathbf{W}}_n)^{-1} \hat{\mathbf{W}}_n^{'}] \mathbf{J}_m \ . \tag{4.10}$$

Then we have the decomposition

$$\begin{aligned} \hat{\mathbf{X}}_n &= \mathbf{C}_n \mathbf{P}_n \mathbf{J}_m^{'} \mathbf{J}_m \mathbf{P}_n \mathbf{C}_n^{-1} (\mathbf{Y}_n - \bar{\mathbf{Y}}_0) \\ &= \mathbf{C}_n \mathbf{P}_n \mathbf{J}_m^{'} [\mathbf{Q}_n^{(0)} + \mathbf{Q}_n^{(1)}] \mathbf{J}_m \mathbf{P}_n \mathbf{C}_n^{-1} (\mathbf{Y}_n - \bar{\mathbf{Y}}_0) \ . \end{aligned} \tag{4.11}$$

In this case, we have the property $\mathbf{Q}_n^2 = \mathbf{Q}_n = \mathbf{Q}_n^{(0)} + \mathbf{Q}_n^{(1)} = \mathbf{J}_m^{'} \mathbf{J}_m$ , and we have the decomposition of the trend-cycle part and the regression part. There is a simple interpretation of this smoothing because we use only the regression part at $m$ low frequencies. There are two steps in the smoothing procedure. First, we remove the regression part from $\mathbf{Y}_n$ by taking

$$\mathbf{X}_n^{(1)} = \mathbf{C}_n \mathbf{P}_n [\mathbf{I}_n - \mathbf{Q}_n^{(0)}] \mathbf{P}_n \mathbf{C}_n^{-1} (\mathbf{Y}_n - \bar{\mathbf{Y}}_0) \ . \tag{4.12}$$

Then, as the second step, we apply the 2nd smoothing to $\mathbf{Y}_n^{(1)}$ as

$$\mathbf{X}_n^{(2)} = \mathbf{C}_n \mathbf{P}_n \mathbf{J}_m^{'} \mathbf{J}_m \mathbf{P}_n \mathbf{C}_n^{-1} \mathbf{X}_n^{(1)} \tag{4.13}$$

by taking another transformation.
Then, the resulting transformation is (4.8) with $\mathbf{Q}_W = \mathbf{Q}_n^{(1)}$ after an iteration. There should be some mechanism for performing further iterations.

**Type-II Smoothing** :

Type-II smoothing is based on Example 2 presented in Section 2. When we need to estimate not only the trend component, but also the noise component, it is important to estimate structural changes and outlier components consistently.

For the seasonal adjustment of time series, we need to estimate the seasonal component for obtaining the seasonally adjusted series, and it is related to Example 2. Thus, we construct an $q \times n$ choice matrix $\mathbf{F}_q$ such that the seasonal components can be removed in their frequencies.
When $s = 4$, we want to remove the data with frequencies around $\lambda_s = 1/4, 1/2$ (1/2 corresponds to the cycle of 2 quarters and 1/4 corresponds to the cycle of 4 quarters). However, we cannot distinguish the 4 quarters cycle from the 2 quarters cycle by using quarterly observations. We set $m_1 = [2n/s]$, and an $(n-2h-1) \times n$ choice matrix and an $(n-3h-2) \times (n-2h-1)$ choice matrix as

$$\mathbf{J}_1^Q = \begin{bmatrix} \mathbf{I}_{m_1-(h+1)} & \mathbf{O} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & \mathbf{I}_{n-m_1-h} \end{bmatrix} \ , \ \mathbf{J}_2^Q = [\mathbf{I}_{n-3h-2} \ , \ \mathbf{O}] \ . \tag{4.14}$$

Then we take a $q \times n$ matrix

$$\mathbf{F}_q^Q = \mathbf{J}_2^Q \mathbf{J_1}^Q \tag{4.15}$$

with a small positive integer $h > 0$.
When $s = 12$, we need a more complicated transformation to remove seasonality because we cannot distinguish the 12 month cycle from the 6, 4, 3, 2.4, and the 2 month cycles using monthly observations with frequencies around $\lambda_s = 1/12, 2/12, 3/12, 4/12, 5/12, 6/12$. We set $m_i = i[2n/s]$ and take $(n - i(2h+1)) \times (n - (i-1)(2h+1))$ choice matrices $(i = 1, \cdots, 5)$ and an $(n - 5(2h+1) - (h+1)) \times (n - 5(2h+1))$ choice matrix such that

$$\mathbf{J}_i^M = \begin{bmatrix} \mathbf{I}_{m_i-(i-1)(2h+1)-(h+1)} & \mathbf{O} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & \mathbf{I}_{n-m_i-h} \end{bmatrix} \ , \ \mathbf{J}_6^M = [\mathbf{I}_{n-11h-6} \ , \ \mathbf{O}] \ , \tag{4.16}$$

with a small positive integer $h > 0$. To remove the data with seasonal frequencies around $\lambda_{js}$ $(j = 2, 3, 4, 5)$ using $\mathbf{J}_j^M$ $(j = 1, \cdots, 6)$, we set a $q \times n$ matrix

$$\mathbf{F}_q^M = \prod_{j=1}^{6} \mathbf{J}_{7-j}^M \ . \tag{4.17}$$

More generally, when we have information of the instrumental variables $\mathbf{W}_n$, we can incorporate the estimated coefficient by regressing

$$\mathbf{Y}_m^* = \mathbf{F}_q \mathbf{P}_n \mathbf{C}_n^{-1} (\mathbf{Y}_n - \bar{\mathbf{Y}}_0) \tag{4.18}$$

to

$$\mathbf{W}_m^* = \mathbf{F}_q \mathbf{P}_n \mathbf{C}_n^{-1} (\mathbf{W}_n - \bar{\mathbf{W}}_0) \ , \tag{4.19}$$

where $\mathbf{F}_q$ is either $\mathbf{F}_q^Q$ or $\mathbf{F}_q^M$.
Type-II smoothing is defined by

$$\mathbf{Q}_n^{(2)} = \mathbf{W}_n^* (\mathbf{W}_n^{*'} \mathbf{W}_n^*)^{-1} \mathbf{W}_n^{*'} \ . \tag{4.20}$$

and

$$\mathbf{Q}_n^{(3)} = \mathbf{F}_q^{'} \mathbf{F}_q - \mathbf{Q}_n^{(2)} \ . \tag{4.21}$$

Then, we have the decomposition

$$\begin{aligned} \hat{\mathbf{X}}_n^* &= \mathbf{C}_n \mathbf{P}_n \mathbf{F}_q^{'} \mathbf{F}_q \mathbf{P}_n \mathbf{C}_n^{-1} (\mathbf{Y}_n - \bar{\mathbf{Y}}_0) \\ &= \mathbf{C}_n \mathbf{P}_n [\mathbf{Q}_n^{(2)} + \mathbf{Q}_n^{(3)}] \mathbf{P}_n \mathbf{C}_n^{-1} (\mathbf{Y}_n - \bar{\mathbf{Y}}_0) \ . \end{aligned} \tag{4.22}$$

In this case, we have the decomposition $\mathbf{Q}_n^{(2)} + \mathbf{Q}_n^{(3)} = \mathbf{F}_q^{'} \mathbf{F}_q$ and the corresponding decomposition of the trend-cycle and regression parts.

**Examples of Dummy Variables** :
There are some examples of outlier and trend dummies. For nonstationary time series, we should be careful about normalization because there can be significant effects on smoothing. Although there are many other possible dummy variables, we provide some examples that have been used in official data handling such as official seasonal adjustment. In the X-12-ARIMA and X-13ARIMA-SEATS programs, for instance, the Reg-ARIMA modeling uses the following dummy variables. See Census Bureau (2020) for the details of the X-12ARIMA-SEATS program and the list of variables in the Reg-ARIMA modeling.)

We give the next list because of an illustration, which will be used in an empirical analysis in Section 6 below. Let $w_s$ $(s = 1, \cdots, n)$ be the dummy variable.

**Example 1** :
The level shift (LS) variable can be defined as $w_s = 0$ if $s < t$ and $w_t = 1$ if $s \geq t$ for $s = 1, \cdots, n$. This can be handled by Type-I smoothing.

**Example 2** :
The outlier variable can be defined as $w_s = 1$ if $s = t$ and $w_t = 0$ if $s \neq t$ for $s = 1, \cdots, n$. This variable is often called additive outlier (AO).

**Example 3** :
The ramp variable can be defined by $w_s = 1$ if $s < t_0$, $w_s = 1 - (t - t_0)/(t_1 - t_0)$ if

$t_0 \leq t \leq t_1$, and $w_t = 0$ if $s \geq t_1$.

**Example 4** :
The double ramp variable can be defined by $w_s = 1$ if $s < t_0$, $w_s = 1-(t-t_0)/(t_1-t_0)$ if $t_0 \leq t \leq t_1$, $w_s = (t-t_1)/(t_2-t_1)$ if $t_1 \leq t \leq t_2$, and $w_t = c$ if $s \geq t_2$.

## 5. Frequency Domain Analysis and Likelihood

We consider the additive decomposition model $\mathbf{y}_i = \mathbf{x}_i + \mathbf{v}_i \ (i = 1, \cdots, n)$ of (2.1)-(2.4) in the time domain and give an interpretation on the consequence of the transformation of observation vectors by $\mathbf{K}_n^*$ in (2.6)-(2.9). The transformed random variables $\mathbf{z}_k^{(n)}$ $(k = 1, \cdots, n)$ have a particular structure in the frequency domain. For the resulting simplicity, we take positive integers $m$ $(= m_n)$.

Let $\mathbf{f}_{\Delta x}(\lambda)$ and $\mathbf{f}_v(\lambda)$ be the spectral density $(p \times p)$ matrices of $\Delta \mathbf{x}_i$ and $\mathbf{v}_i$ $(i = 1, \cdots, n)$, respectively, which are given by

$$\mathbf{f}_{\Delta x}(\lambda) = (\sum_{j=0}^{\infty} \mathbf{C}_j^{(x)} e^{2\pi i \lambda j}) \mathbf{\Sigma}_e^{(x)} (\sum_{j=0}^{\infty} \mathbf{C}_j^{(x)'} e^{-2\pi i \lambda j}) \quad (-\frac{1}{2} \leq \lambda \leq \frac{1}{2}) \ , \tag{5.1}$$

and

$$\mathbf{f}_v(\lambda) = (\sum_{j=0}^{\infty} \mathbf{C}_j^{(v)} e^{2\pi i \lambda j}) \mathbf{\Sigma}_e^{(v)} (\sum_{j=0}^{\infty} \mathbf{C}_j^{(v)'} e^{-2\pi i \lambda j}) \quad (-\frac{1}{2} \leq \lambda \leq \frac{1}{2}) \ , \tag{5.2}$$

where we set $\mathbf{C}_0^{(x)} = \mathbf{C}_0^{(v)} = \mathbf{I}_p$ as normalizations and $i^2 = -1$. (See Chapter 7 of Anderson (1971) with different notations, for instance.)
Then, the $p \times p$ spectral density matrix of the transformed vector process, which is observable, and the spectral density of the difference series $\Delta \mathbf{y}_i$ $(= \mathbf{y}_i - \mathbf{y}_{i-1})$ can be represented as

$$\mathbf{f}_{\Delta y}(\lambda) = \mathbf{f}_{\Delta x}(\lambda) + (1 - e^{2\pi i \lambda}) f_v(\lambda) (1 - e^{-2\pi i \lambda}) \ . \tag{5.3}$$

We denote the long-run variance-covariance matrices of trends and stationary components for $g, h = 1, \cdots, p$ as

$$\mathbf{\Sigma}_e^{(x)} = \mathbf{f}_{\Delta x}(0) \ (= (\sigma_{gh}^{(x)})) \ , \ \mathbf{\Sigma}_e^{(v)} = f_v(0) \ = (\sigma_{gh}^{(v)}) \ . \tag{5.4}$$

From (5.3), we find that $\mathbf{f}_{\Delta y}(0) = \mathbf{f}_{\Delta x}(0)$ at the frequency $\lambda = 0$ and we can ignore the effects of stationary noise terms in (2.1)-(2.4) if we use only the information around the zero-frequency from data. That is, the information of the non-stationary trend parts is separated from the information of the stationary noise parts in the frequency domain.

Since the spectral matrices in (5.1)-(5.3) are complex-valued, we symmetrize the spectral density matrices. Then, it is possible to relate the complex-valued spectral

matrices to the real-valued random vectors and their likelihood function. By reconsidering the relationship among the continuous-valued discrete time series and the spectral densities, it is possible to interpret the filtered parts and smoothing parts. It has been sometimes neglected in the (standard) statistical time series analysis.
Let $f_v^{(SR)}(\lambda_k)$ and $f_{\Delta x}^{(SR)}(\lambda_k)$ be the symmetrized $p\times p$ spectral matrices of $\mathbf{v}_i$ and $\Delta\mathbf{x}_i$ at $\lambda_k$ $(= (k-\frac{1}{2})/(2n+1))$ for $k=1,\cdots,n$, that is, $f_v^{(SR)}(\lambda_k) = (1/2)[f_v^{(SR)}(\lambda_k) + \bar{f}_v^{(SR)}(\lambda_k)]$ and $f_{\Delta x}^{(SR)}(\lambda_k) = (1/2)[f_{\Delta x}^{(SR)}(\lambda_k) + \bar{f}_{\Delta x}^{(SR)}(\lambda_k)]$. We denote the $n\times p$ matrix $\mathbf{Z}_n = (\mathbf{z}_k^{(n)}(\lambda_k^{(n)})')$ and $\lambda_k^{(n)} = (k-1/2)/(2n+1)$ $(k=1,\cdots,n)$, where $\lambda_k^{(n)}$ corresponds to the frequency of $\mathbf{z}_k^{(n)}$. Then, the transformed random variables are asymptotically orthogonal (or uncorrelated) and the orthogonal processes are approximately distributed as Gaussian distributions when $n$ is large.
If we substitute $\lambda_k$ into (5.3), we find that the variance-covariance matrix of $\mathbf{z}_k^{(n)}$ at $\lambda_k$ is approximately given by $\mathbf{f}_{\Delta y}(\lambda_k) = \mathbf{f}_{\Delta x}(\lambda_k) + a_{kn}^* f_v(\lambda_k)$ because $\|1-e_k^{2\pi i\lambda}\|^2 = 2-2\cos(2\pi\lambda_k) = a_{kn}^*$ for $k=1,\cdots,n$.

Given the initial condition $\mathbf{y}_0$, the (-2) times the conditional log-likelihood function in (2.1)-(2.4) can be approximated except a constant term by

$$
\begin{aligned}
l_n &= \sum_{k=1}^{n} \log|a_{kn}^* f_v^{(SR)}(\lambda_k) + f_{\Delta x}^{(SR)}(\lambda_k))| \\
&\quad + \sum_{k=1}^{n} \mathbf{z}_k'[a_{kn}^* f_v^{(SR)}(\lambda_k) + f_{\Delta x}^{(SR)}(\lambda_k)]^{-1}\mathbf{z}_k
\end{aligned} \tag{5.5}
$$

provided that $a_{kn}^* f_v^{(SR)}(\lambda_k) + f_{\Delta x}^{(SR)}(\lambda_k)$ are positive definite (a.e.).
In particular, we consider the case when $\Delta\mathbf{x}_i$ and $\mathbf{v}_i$ are a sequence of independent random vectors, then we have $\mathbf{\Sigma}_e^{(x)} = f_{\Delta x}^{(SR)}(\lambda_k)$ and $\mathbf{\Sigma}_e^{(v)} = f_v^{(SR)}(\lambda_k)$ for $k = 1,\cdots,n$. This corresponds to the case when

$$
l_n^* = \sum_{k=1}^{n} \log|a_{kn}^*\mathbf{\Sigma}_e^{(v)} + \mathbf{\Sigma}_e^{(x)}| + \sum_{k=1}^{n} \mathbf{z}_k'[a_{kn}^*\mathbf{\Sigma}_e^{(v)} + \mathbf{\Sigma}_e^{(x)}]^{-1}\mathbf{z}_k\ , \tag{5.6}
$$

provided that $a_{kn}^*\mathbf{\Sigma}_e^{(v)} + \mathbf{\Sigma}_e^{(x)}$ are positive definite (a.e.).
Furthermore, if we we take $k=1,\cdots,m_n$ such that $m_n/n \to 0$ and $m_n\to\infty$ as $n\to\infty$, we have $a_{kn}^*\to 0$ $(k=1,\cdots,m_n)$. In this situation, the first $m_n$ terms of (5.5) can be regarded as the log-likelihood function in the low frequency parts of time series, which is given as

$$
l_{1n}^{**} = \sum_{k=1}^{m} \log|\mathbf{\Sigma}_e^{(x)}| + \sum_{k=1}^{m} \mathbf{z}_k'[\mathbf{\Sigma}_e^{(x)}]^{-1}\mathbf{z}_k\ , \tag{5.7}
$$

provided that $\mathbf{\Sigma}_e^{(x)}$ is positive definite.
The last representation corresponds to the Gaussian log-likelihood function based on $m$ mutually independent observations in the statistical multivariate analysis.

(Anderson (2003), for instance.)
In this way, we can separate the likelihood information into different parts of time series components in the SIML (separate information maximum likelihood) method. The SIML estimator of $\mathbf{\Sigma}_e^{(x)}$ is constructed by using $\mathbf{z}_k$ $(k = 1, \cdots, m_n)$, and we have some desirable asymptotic properties for $\mathbf{\Sigma}_e^{(x)}$.

When we have some dummy variables $\mathbf{W}_n$, we need to assume that they are independent of other noise, cycle, seasonal, and trend components. There can be several ways to handle explanatory variables as explained as Type-I and Type II in Section 4, but we explain a typical case. Given the initial condition and the information set of explanatory variables $\mathbf{W}_n$, $(-2)$ times the conditional log-likelihood can be approximated as (5.5) by $\mathbf{z}_k^*(w) = \mathbf{y}_k^* - \mathbf{B}'\mathbf{w}_k^*$, where $\boldsymbol{B}$ is the parameter matrix, $\mathbf{y}_k^*$ and $\mathbf{w}_k^*$ are the transformed explained variables and explanatory variables (using $\mathbf{K}_n^*$−transformation from the observed $\mathbf{y}_i$ and $\mathbf{w}_i$ $(i = 1, \cdots, n)$, respectively.
When we use the explanatory variables $\mathbf{W}_n$, we can estimate the unknown matrix $\mathbf{B}$ by Theorem 4.1 consistently. Let $\hat{\mathbf{B}}$ be the SIML estimator and $\mathbf{z}_k^* = \mathbf{y}_k^* - \hat{\mathbf{B}}'\mathbf{w}_k^*$ $(k = 1, \cdots, n)$, which depend son $\mathbf{W}_n$ and denote $\mathbf{z}_k^*(w)$ $(k = 1, \cdots, n)$. To estimate $\mathbf{\Sigma}_x$ when there are explanatory variables, for instance, it is reasonable to use

$$\mathbf{G}_m^*(w) = \frac{1}{m}\sum_{k=1}^{m} \mathbf{z}_k^*(w)\mathbf{z}_k^{*'}(w) \tag{5.8}$$

because it is consistent and has the asymptotic normality if we take $m = [m_n]$ such that $m_n/n \to 0$, $m_n \to \infty$ and $n \to \infty$.

There are two remarks on the likelihood function of time series data in the above discussion.
First, the ML estimation of unknown parameters in the nonstationary errors-in-variables models may have some difficulty when $p > 1$ without some restrictions of the parameter space. It is because the exact likelihood function can have peculiar shape even when the observations are the sequence of independent random variables. which is known in the statistical multivariate analysis. There could be more complications when we have multivariate time series data with random walks, seasonal components, autocorrelations and measurement errors.
Second, the likelihood functions in this section can be related to the classical topic on the *Wittle-type* likelihood function for stationary time series in the literature, which does not depend on the Gaussian distributions for underlying noise distributions in multivariate stationary processes. The maximum of Wittle-type likelihood has been called the quasi-maximum likelihood (QML) estimation, which is discussed and applied by Hosoya (1997), for instance.

## 6. An Example of macro-consumption of durable goods

We use the official macro-consumption data of durable goods in Japan from 1994Q1 to 2019Q4 to illustrate the regression smoothing method [2]. To construct the seasonal-adjusted data, we need to estimate the seasonal factor to construct the seasonal adjustment data. But then we need to estimate the trend and noise components to estimate the seasonal component from the original quarterly time series at the same time. The traditional X-13ARIMA-SEATS program of the U.S. Census Bureau has a rather complicated procedure to do this by using moving average filtering repeatedly.

In our analysis, we applied the SIML smoothing procedure with $m = 29$ (Example 1 in Section 2). and $h = 2$ (Type-2 smoothing in Section 4) , which yield the minimum numbers of AIC. Hence at each seasonal frequencies 1/4 and 1/2, we have chosen 5 and 3 frequency data points in (4.14) and (4.15) to estimate the seasonal state variable by inverting the frequency bounds. All corresponding figures are presented in Appendix B.

Figure 1 shows a summary of SIML smoothing for log-transformed data. This was done because the original series has a significant heteroscedastic seasonality. In Figures 1-4, "org " stands for the original series, "trend ", "seasonal ", "noise " mean the estimated trend, seasonal, and noise components, while "adj " means the estimated seasonally adjusted series, i.e., the observed series minus the estimated seasonal component. "Z " means the transformed orthogonal series. In Figures 2-4, "reg " stands for the dummy variable.

The original time series has typical characteristics of major macroeconomic time series in Japan, i.e., it is a realization of nonstationary time series and exhibits rather clear trend, cycle, seasonal and irregular components. We applied the SIML filtering with $m = 29$; red curve indicates the estimated trend-cycle component. Since $\lambda_m = 29/[2n] \sim 0.14$, which means 1.8 year, that is, we have estimated the trend-cycle components over about 2 years cycle. It may be practically reasonable choice for the trend-cycle component from our macro data. The noise component is constructed as the observed data minus the estimated trend-cycle and seasonal components. By using $\mathbf{Z}_n-$transformed data, we capture the significant effects at the seasonal frequencies. Because we use the quarterly data, we have sharp peaks and troughs at frequency 1/4 and 1/2. The estimated seasonal component moves regularly, which may change over time rather smoothly. We have found that the estimated seasonal component by using the X-13ARIMA-SEATS tend to exhibit

[2] We have taken data from https://www.esri.cao.go.jp/jp/sna/menu.html (Economic and Social Research Institute (ESRI), Cabinet Office, Japan). They are original series in real terms and ESRI uses the X-12-ARIMA smoothing program for constructing seasonal-adjusted official data. We use the consumption series of durable goods as a typical component of GDP.

more rigid seasonality. Although the estimated seasonal component gives regular seasonal pattern, the estimated trend-cycle and noise components suggest there were some abrupt changes around the year of 2008-2009, 2011, and 2014, which are different from the usual noise component. (One way to deal with these effects it to use outlier detections and Reg-ARIMA model in X-13ARIMA-SEATS program.) It may be appropriate to consider the possibility that there are major breaks and institutional changes during the sample period.

First, there was a rapid downward effect attributed by the 2008 financial crisis, and we may consider this event for being appropriate to use the ramp-dummy at 2008Q3-2009Q1. Figures 2 and 3 summarize SIML smoothing and frequency regression results for the cases.

Second, we applied two AO-dummy variables at 2011Q1 and 2014Q1. (See Example 2 in Section 4.) In these periods, there were large effects caused by the 2011 earthquake in Japan and an increase of consumption tax in April 2014. There was a temporary increase of durable consumption in the 2014Q1 period. Both events had significant effects on the macroeconomy and consumption in Japan.

Finally, Figure 4 represents a summary of SIML smoothing and frequency regression with three dummy variables considered simultaneously. (See Example 2 and 4 in Section 4.) Based on the criteria of AIC, we selected the last case for the best modeling for the macro-consumption of durable goods; these effects are captured by our method. By using the transformed data of (4.14) and (4.15) and the dummy variables, the AIC(w) was calculated based on the regression equation by

$$\mathrm{AIC}(w) = n \log \hat{\sigma}_w^2 + 2r \tag{6.1}$$

where we use $\hat{\sigma}_w^2$ calculated from the residuals of the dummy regression ((4.2) with $p = 1$) and $r$ denotes the number of dummy variables [3].
In our example, the main purpose of data analysis was to evaluate the appropriateness of the published data. This type of task was not easy because the published data used the X-12-ARIMA program of the Cnensus Bureau and it is a complicated procedure in practice [4]. For each models, we have calculated two AICs were calculated: the first AIC in figures was calculated using all frequency data while the AIC in the parenthesis was calculated using all frequency data except data around the seasonal frequency.

By using the model selection criteria for minimizing these AICs, we find that SIML smoothing with three dummy variables (i.e., two AOs and a double ramp)

[3] This AIC(w) is based on (5.6) and (5.7) with dummy variables, which can be implemented easily. However, we have taken the case as if $a_{kn}$ were constant with respect to $k$ because we use the procedure, that is free from the maximum likelihood (ML) estimation of unknown parameters needed. In this sense, our AIC(w) is an approximate one.

[4] The details of their estimation procedure from original series are explained at the web-cite of ESR (Economic and Social Research Institute, Cabinet office of Japan), https://www.esri.cao.go.jp/en/sna/sokuhou/sokuhou_top.html.

is the best model. We have reasonable result on the decomposition of original time series into trend-cycle, seasonal, and noise components. In the selected model, the trend-cycle component includes one structural change and the noise component includes two outliers.

This empirical analysis illustrates that we need to consider the important role of incorporating the effects of the change point problem and abrupt changes in the seasonal adjustment procedure. In this respect, we have illustrated our methodology based on the frequency regression and smoothing.

## 7. Concluding Remarks

In many original macro-economic time series, it is common to observe nonstationary trend, cycles, seasonal, and measurement errors simultaneously. In addition to these components, we sometimes observe abrupt changes, trading-day effects, and other irregular components. Thus, it seems difficult to remove the seasonal component from the original time series in the seasonal adjustment and construct macro-index, which involve multiple nonstationary time series.

This paper presents a new approach to handle nonstationary time series using frequency regression based on the SIML modelling in a systematic manner. We use the SIML method because we can separate the likelihood information of time series data into different frequency parts of their components. Our method sheds a new light on some practical approache to handle economic time series, which have been practically used in official seasonal adjustments without formal justifications. There shall be many empirical examples.

There are further problems to be investigated. The present study is based on the time series decomposition in (2.1)-(2.4). There can be more complicated decomposition models including trend, cycle and seasonal components in a different way. Then we need to investigate the relationships among trends, cycles, seasonals and irregular noises components of nonstationary and stationary time series both in the time and frequency domains. It seems that some extensions of Theorems 3.1 and 4.1 in this paper can be developed.

Another issue would be the computation of the procedure we explained in this paper. We have been developping an R-program of the SIML method, which will be available hopefully in the near future.

## References


[1] Anderson, T. W. (1971), *The Statistical Analysis of Time Series*, Wiley.

[2] Anderson, T.W. (1984), "Estimating Linear Statistical Relationships," *Annals of Statistics*, 12, 1-45.

[3] Anderson, T. W. (2003), *An Introduction to Multivariate Statistical Analysis*, 3rd Edition, Wiley.

[4] Baxter, M. and King, R.G. (1999), "Approximate Band-Pass Filters for Economic Time Series," *The Review of Economics and Statistics*, 81-4, 575-593.

[5] Brillinger, D. and Hatanaka, M. (1969), "An Harmonic Analysis of Nonstationary Multivariate Economic Processes," *Econometrica*,35, 131-141.

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

[7] Census Bureau (2020), "X-13ARIMA-SEATS Reference Manual" (Accessible HTML Outout Version), U.S. Census Bureau.
https://www.census.gov/data/software/x13as.html

[8] Engle, R. (1974), "Band spectrum regression," *Econometrica*, Vol. 15-1, 1-11.

[9] Fuller, W. (1987), *Measurement Error Models*, John-Wiley.

[10] Granger, C. and M. Hatanaka (1964), *Spectral Analysis of Economic Time Series*, Princeton University Press.

[11] Hosoya, Y. (1997), "A limit theory for long range dependence and statistical inference on related models," *The Annals of Statistics*, vol.25, no.1, pp.105-137.

[12] Kitagawa, G. (2020), *Introduction to Time Series Modeling with Applications in R*, 2nd Edition, CRC Press.

[13] Kunitomo, N. and S. Sato (2021), "A robust filtering method for noisy nonstationary multivariate time series with an application to macroconsumption," Open Access, *Japanese Journal of Statistics and Data Science*, 4, 373-410, Springer.

[14] Müller, U. and M. Watson (2018), "Long-run covariability," *Econometrica*, Vol. 86-3, 775-804.

## APPENDIX A : Mathematical Derivations

We present the derivations of Theorem 3.1 and Theorem 4.1 as an application of Theorem A.1 below. Theorem A.1 gives the asymptotic properties of the estimation of long-run variance-covariance matrix for the nonstationary errors-in-variables models. We first provide the intuition for our result and then give some details of its derivation. Some derivations are omitted because they are direct extensions of the results reported by Kunitomo and Sato (2021).

**A-I A Heuristic Derivation** : We consider the nonstationary process $\mathbf{y}_i$ $(i = 1, \cdots, n)$ defined by (2.1)-(2.4) given the initial conditions $\mathbf{y}_i, \mathbf{x}_i$ $(i \leq 0)$. (Without loss of generality, we often ignore the effects of initial conditions whenever they are asymptotically negligible.)
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 for $\mathbf{Y}_n = (\mathbf{y}_i^{'})$ we write $\mathbf{z}_k$ $(k = 1, \cdots, n)$ as

(A.1) $$\mathbf{z}_k(\lambda_k^{(n)}) = \sum_{j=1}^{n} p_{jk}^{(n)} \mathbf{r}_j \ , \ \mathbf{r}_j = \mathbf{y}_j - \mathbf{y}_{j-1} \ ,$$

which is a (real-valued) Fourier type transformation and the initial conditions are fixed.
Then, we find that $\mathbf{z}_k(\lambda_k^{(n)}) = (\mathbf{z}_{in}(\lambda_k^{(n)}))$ $(k = 1, \cdots, n)$ are the (real-valued) Fourier-transformation of data at the frequency $\lambda_k^{(n)}$ $(= (k-1/2)/(2n+1))$, which is a (real-part of) estimate of the orthogonal incremental vector process $(p \times 1)$ $\mathbf{z}(\lambda) = (\mathbf{z}_i(\lambda))$ $(0 \leq \lambda \leq 1/2)$ and $\mathbf{z}(\lambda)$ is a continuous process in the frequency domain.
By evaluating

$$\mathbf{E}\left[\mathbf{z}_k(\lambda_k^{(n)})\mathbf{z}_{k'}(\lambda_{k'}^{(n)})^{'}\right] = \left[\frac{1}{2n+1}\right] \sum_{j,j'=1}^{n} (e^{i\theta_{jk}} + e^{-i\theta_{jk}})(e^{i\theta_{j'k'}} + e^{-i\theta_{j'k'}})\mathbf{E}[\mathbf{r}_j \mathbf{r}_{j'}^{'}] \ ,$$

we find that the effects of each term with $k \neq k^{'}$ are asymptotically negligible, and the dominant sum with $k = k^{'}$ is asymptotically equivalent to

$$\left[\frac{n}{2n+1}\right] \sum_{h=-(n-1)}^{n-1} \left[\cos 2\pi \frac{k-1/2}{2n+1} h\right] [\mathbf{\Gamma}(h) + \mathbf{\Gamma}(-h)] \ ,$$

where we use the notation $\mathbf{\Gamma}(h) = \mathbf{E}(\mathbf{r}_j \mathbf{r}_{j-h}^{'})$ (we have ignored the constant means with $\mathbf{E}[\mathbf{r}_j] = \mathbf{0}$).
Then, by using the conditions in (2.3)-(2.4), it converges to

(A.2) $$\mathbf{f}_{SR}(\lambda) = \sum_{h=-\infty}^{\infty} \cos(2\pi h \lambda) \mathbf{\Gamma}(h) \ , \ 0 \leq \lambda \leq \frac{1}{2} \ ,$$

where the symmetrized spectral density matrix for $\mathbf{r}_j$ is given by $\mathbf{f}_{SR}(\lambda) = (1/2)[\mathbf{f}_{\Delta y}(\lambda) + \bar{\mathbf{f}}_{\Delta y}(\lambda)$.

Under the assumption of stationarity of $\mathbf{r}_j$ it has been known that $\mathbf{z}_k(\lambda_k^{(n)})$ are asymptotically uncorrelated random variables. (See Chapters 8-9 of Anderson (1971), for instance.) By using straightforward (but lengthy) evaluations, we find that for $k \neq k^{'}$

$$\text{(A.3)} \qquad \mathbf{E}[\mathbf{z}_{ik}(\lambda_k^{(n)})\mathbf{z}_{jk}(\lambda_k^{(n)})\mathbf{z}_{hk}(\lambda_{k'}^{(n)})\mathbf{z}_{lk}(\lambda_{k'}^{(n)})] = \sigma_{ij}(\lambda_k^{(n)})\sigma_{hl}(\lambda_{k'}^{(n)}) + o(1)$$

and for $k = k^{'}$

$$\begin{aligned} \text{(A.4)} \qquad & \mathbf{E}[\mathbf{z}_{ik}(\lambda_k^{(n)})\mathbf{z}_{jk}(\lambda_k^{(n)})\mathbf{z}_{hk}(\lambda_k^{(n)})\mathbf{z}_{lk}(\lambda_k^{(n)})] \\ = \; & \sigma_{ij}(\lambda_k^{(n)})\sigma_{hl}(\lambda_k^{(n)}) + \sigma_{ih}(\lambda_k^{(n)})\sigma_{jl}(\lambda_k^{(n)}) + \sigma_{il}(\lambda_k^{(n)})\sigma_{jh}(\lambda_k^{(n)}) + o(1) \;, \end{aligned}$$

where $\mathbf{\Gamma}(h) = (\mathbf{\Gamma}_{ij}(h))$ and

$$\text{(A.5)} \qquad \sigma_{ij}(\lambda_k^{(n)}) = \sum_{h=-(n-1)}^{n-1} \left[\cos 2\pi\lambda_k^{(n)} h\right] \mathbf{\Gamma}_{ij}(h) \;.$$

From (A.3)-(A.4), we notice that $\mathbf{z}_k(\lambda_k^{(n)})$ are (as if) normally distributed random variables with (A.2).

As $n \to \infty$ and $m/n \to 0$, we have $\lambda_k^{(n)} \to 0$ for $1 \leq k \leq m$. We write for $k = 1, \cdots, m$ and as $m/n \to 0$,

$$\text{(A.6)} \qquad \lim_{n\to\infty} \sigma_{ij}(\lambda_k^{(n)}) = \sigma_{ij}^{(x)} \; (i, j = 1, \cdots, p)$$

and $\mathbf{\Sigma}_x = (\sigma_{ij}^{(x)})$. Then in this situation

$$\text{(A.7)} \qquad \mathbf{Var}[\frac{1}{\sqrt{m}} \sum_{k=1}^{m} \mathbf{z}_{ik}(\lambda_k^{(n)})\mathbf{z}_{jk}(\lambda_k^{(n)})] \longrightarrow \sigma_{ii}^{(x)}\sigma_{jj}^{(x)} + \sigma_{ij}^{(x)2} \;.$$

We construct a sequence of random variables, which are approximately uncorrelated and for $i, j = 1, \cdots, p$

$$s_{ij}(t) = \mathbf{z}_{ik}(\lambda_t^{(n)})\mathbf{z}_{jk}(\lambda_t^{(n)}) - \mathbf{E}[\mathbf{z}_{ik}(\lambda_t^{(n)})\mathbf{z}_{jk}(\lambda_t^{(n)})]$$

and

$$M_{ij}(n, k) = \sum_{t=1}^{k} s_{ij}(t) \;.$$

Then, heuristically, we can apply the central limit theorem (CLT) for the multivariate Gaussian stationary process to obtain the asymptotic normality of the normalized quadratic quantities. However, to show this argument in a rigorous way, we need further developments.

**A-II Proof of Main Results** : We first prepare a general result on the consistency and asymptotic normality of the SIML estimation in nonstationary time series; the result may have some new aspect.

**Theorem A.1**: Assume that the fourth order moments of each element of $\mathbf{v}_i^{(x)}$ and $\mathbf{v}_i$ in (2.1)-(2.4) are bounded. Let

$$\text{(A.8)} \qquad \hat{\mathbf{\Sigma}}_x (= (\hat{\sigma}_{gh}^{(x)})) = \frac{1}{m} \mathbf{Z}_m^{*'} \mathbf{Z}_m^* ,$$

which is $\mathbf{G}_m^*$ in (3.2). Then
(i) For $m_n = n^\alpha$ ($[m_n] = m$) and $0 < \alpha < 1$, as $n \longrightarrow \infty$

$$\text{(A.9)} \qquad \hat{\mathbf{\Sigma}}_x - \mathbf{\Sigma}_x \xrightarrow{p} \mathbf{O} .$$

(ii) We set $\mathbf{\Sigma}_x = (\sigma_{gh}^{(x)})$. For $m_n = n^\alpha$ ($[m_n] = m$) and $0 < \alpha < 0.8$, as $n \longrightarrow \infty$

$$\text{(A.10)} \qquad \sqrt{m_n} \left[ \hat{\sigma}_{gh}^{(x)} - \sigma_{gh}^{(x)} \right] \xrightarrow{\mathcal{L}} N \left( 0, \sigma_{gg}^{(x)} \sigma_{hh}^{(x)} + \left[ \sigma_{gh}^{(x)} \right]^2 \right) .$$

The covariance of the limiting distributions of $\sqrt{m_n}[\hat{\sigma}_{gh}^{(x)} - \sigma_{gh}^{(x)}]$ and $\sqrt{m_n}[\hat{\sigma}_{kl}^{(x)} - \sigma_{kl}^{(x)}]$ is given by $\sigma_{gk}^{(x)} \sigma_{hl}^{(x)} + \sigma_{gl}^{(x)} \sigma_{hk}^{(x)}$ $(g, h, k, l = 1, \cdots, p)$.

**Proof of Theorem A.1** : The proof consists of several steps.
(**Step 1**) : We set $m = [m_n]$ and $m_n$ depends on $n$. Let $\mathbf{z}_k^{(x)} = (z_{kj}^{(x)})$ and $Z_k^{(v)} = (z_{kj}^{(v)})$ $(k = 1, \cdots, n)$ be the $k$-th row vector elements of $n \times p$ matrices

$$\text{(A.11)} \qquad \mathbf{Z}_n^{(x)} = \mathbf{K}_n^* (\mathbf{X}_n - \bar{\mathbf{X}}_0) \ , \ \mathbf{Z}_n^{(v)} = \mathbf{K}_n^* \mathbf{V}_n \ , \ \mathbf{K}_n^* = \mathbf{P}_n \mathbf{C}_n^{-1} ,$$

respectively, where we denote $\mathbf{X}_n = (\mathbf{x}_k^{'}) = (x_{kg})$, $\mathbf{V}_n = (\mathbf{v}_k^{'}) = (v_{kg})$, $\mathbf{Z}_n = (\mathbf{z}_k^{'})$ $(= (z_{kg}))$ as $n \times p$ matrices with $z_{kg} = z_{kg}^{(x)} + z_{kg}^{(v)}$. Here, we set the initial conditions $\bar{\mathbf{X}}_0 = \bar{\mathbf{Y}}_0$ and we find that the effects of initial condition are stochastically negligible in the frequency regression. We write $z_{kg}, z_{kg}^{(x)}, z_{kg}^{(v)}$ as the $g-$th component of $\mathbf{z}_k, \mathbf{z}_k^{(x)}$, and $\mathbf{z}_k^{(v)}$ $(k = 1, \cdots, n; g = 1, \cdots, p)$.
By decomposing $\hat{\mathbf{\Sigma}}_x - \mathbf{\Sigma}_x (= (\hat{\sigma}_{gh}^{(x)} - \sigma_{gh}^{(x)})$ for $g, h = 1, \cdots, p)$ into the effects of $\mathbf{z}_k^{(x)}$ and $\mathbf{z}_k^{(v)}$, we rewrite

$$\text{(A.12)} \qquad \sqrt{m_n} \left[ \frac{1}{m} \sum_{k=1}^m \mathbf{z}_k \mathbf{z}_k^{'} - \mathbf{\Sigma}_x \right]$$
$$= \sqrt{m_n} \left[ \frac{1}{m} \sum_{k=1}^m \mathbf{z}_k^{(x)} \mathbf{z}_k^{(x)'} - \mathbf{\Sigma}_x \right] + \frac{\sqrt{m_n}}{m} \sum_{k=1}^m \mathbf{E}[\mathbf{z}_k^{(v)} \mathbf{z}_k^{(v)'}]$$
$$+ \frac{\sqrt{m_n}}{m} \sum_{k=1}^m \left[ \mathbf{z}_k^{(v)} \mathbf{z}_k^{(v)'} - \mathbf{E}[\mathbf{z}_k^{(v)} \mathbf{z}_k^{(v)'}] \right] + \frac{\sqrt{m_n}}{m} \sum_{k=1}^m \left[ \mathbf{z}_k^{(x)} \mathbf{z}_k^{(v)'} + \mathbf{z}_k^{(v)} \mathbf{z}_k^{(x)'} \right] .$$

Then we shall show that three terms except the first term of (A.12) in the right-hand side are $o_p(1)$ under the condition of $0 < \alpha < 0.8$. To show (ii) of Theorem A.1, we need to show that the dominant term in (A.12) is

$$\text{(A.13)} \qquad \sqrt{m_n}\left[\frac{1}{m}\sum_{k=1}^{m}\mathbf{z}_k^{(x)}\mathbf{z}_k^{(x)'} - \boldsymbol{\Sigma}_x\right] ,$$

and it is asymptotically normal as $m_n \to \infty$ $(n \to \infty)$. By denoting $\boldsymbol{\Gamma}_x(h) = \mathbf{E}[\Delta\mathbf{x}_t\Delta\mathbf{x}_{t-h}^{'}]$, we express

$$\text{(A.14)} \qquad \boldsymbol{\Sigma}_x = \mathbf{f}_{\Delta x}(0) = \sum_{h=-\infty}^{+\infty}\boldsymbol{\Gamma}_x(h) .$$

From $\theta_{jk} = [2\pi/(2n+1)](j-\frac{1}{2})(k-\frac{1}{2})$, we set $c_{ij} = (2/m)\sum_{k=1}^{m}\cos(\theta_{ik})\cos(\theta_{jk})$ $(i,j = 1,\cdots,n)$. Then for any (non-zero $p \times 1$) constant vector $\boldsymbol{a}$, we can evaluate

$$\begin{aligned}\mathbf{E}\left[\frac{1}{m}\sum_{k=1}^{m}(\boldsymbol{a}^{'}\mathbf{z}_k^{(x)})^2 - \boldsymbol{a}^{'}\Sigma_x\boldsymbol{a}\right]^2 &= (\frac{2}{2n+1})^2\mathbf{E}\left[\sum_{j,j^{'}=1}^{n}c_{jj^{'}}\boldsymbol{a}^{'}(\mathbf{v}_j^{(x)}\mathbf{v}_{j^{'}}^{(x)'} - \mathbf{E}(\mathbf{v}_j^{(x)}\mathbf{v}_{j^{'}}^{(x)'})\boldsymbol{a}\right]^2 \\ &\leq K_1[\frac{2}{2n+1}]^2\sum_{j,j^{'}=1}^{n}c_{j,j^{'}}^2 ,\end{aligned}$$

where $K_1$ is a positive constant and we have used the boundedness of fourth moments. Since $m\ \sum_{j,j^{'}=1}^{n}c_{j,j^{'}}^2 = (n+1/2)^2$, we can show that the above term converges to zero as $m \to \infty, m/n \to 0$.

(**Step 2**) :

Let $\mathbf{b}_k = (b_{kj}) = \boldsymbol{\alpha}_k^{'}\mathbf{P}_n\mathbf{C}_n^{-1} = (b_{kj})$ and $\boldsymbol{\alpha}_k^{(n)'} = (0,\cdots,1,0,\cdots)$ be an $n \times 1$ vector. We write $z_{kg}^{(v)} = \sum_{j=1}^{n}b_{kj}\ v_{jg}$ for the seasonal and noise part and use the relation

$$\text{(A.15)} \quad (\mathbf{P}_n\mathbf{C}_n^{-1}\mathbf{C}_n^{'-1}\mathbf{P}_n^{'})_{k,k^{'}} = \delta(k,k^{'})4\sin^2[\frac{\pi}{2n+1}(k-\frac{1}{2})] = \delta(k,k^{'})a_{kn}^{*} .$$

Then under the conditions that $\|\mathbf{C}_j^{(v)}\| = O(\rho^j)$ $(0 \leq \rho < 1)$, we can find $K_2$ (a positive constant) such that

$$\text{(A.16)} \qquad \mathbf{E}[(z_{kg}^{(v)})]^2 = \mathbf{E}[\sum_{i=1}^{n}b_{ki}v_{ig}\sum_{j=1}^{n}b_{kj}v_{jg}] \leq K_2 \times a_{kn}^{*} .$$

It is because $\mathbf{E}[(z_{kg}^{(v)})]^2 = \sum_{i,j=1}^{n}b_{ki}b_{kj}\sigma_{gg}^{(v)}(i-j)$ , where $\sigma_{gg}^{(v)}(i-j)$ is the $(i-j)-$th autocovariance of $v_{ig}$ and $v_{jg}$. (We denote $b_{ki} = 0$ for $i < 0$ and $i > n$.) Then

$$\mathbf{E}[(z_{kg}^{(v)})]^2 = \sum_{l=-(n-1)}^{n-1}[\sum_{j=1}^{n}b_{kj}b_{k,j+l}\sigma_{gg}^{(v)}(l)] \leq [\sum_{j=1}^{n}b_{kj}^2]\sum_{l=-\infty}^{\infty}|\sigma_{gg}^{(v)}(l)| .$$

Because $\|\mathbf{C}_j^{(v)}\| = O(\rho^j)$, $\sum_{l=-\infty}^{\infty} |\sigma_{gg}^{(v)}(l)|$ is bounded. Also it is straight-forward to find that

$$\text{(A.17)} \qquad \frac{1}{m}\sum_{k=1}^{m} a_{kn}^{*} = \frac{1}{m} 2 \sum_{k=1}^{m} \left[ 1 - \cos(\pi \frac{2k-1}{2n+1}) \right] = O(\frac{m^2}{n^2}) \ ,$$

by using the relation

$$\sum_{k=1}^{m} 2\cos(\pi \frac{2k-1}{2n+1}) = \sum_{k=1}^{m} [e^{i\frac{2\pi}{2n+1}(k-\frac{1}{2})} + e^{-i\frac{2\pi}{2n+1}(k-\frac{1}{2})}] = \frac{\sin(\frac{2\pi}{2n+1}m)}{\sin(\frac{\pi}{2n+1})} \ .$$

Then the second term of (A.12) becomes

$$\text{(A.18)} \qquad \frac{1}{m}\sum_{k=1}^{m} \mathbf{E}[z_{kg}^{(v)}]^2 \leq K_3 \frac{1}{m}\sum_{k=1}^{m} a_{kn}^{*} = O(\frac{m^2}{n^2}) \ ,$$

which is $o(1)$ if we set $\alpha$ such that $0 < \alpha < 1$ and $K_3$ is a positive constant. For the fourth term,

$$\mathbf{E}\left[ \frac{1}{m}\sum_{j=1}^{m} z_{kg}^{(x)} z_{kg}^{(v)} \right]^2 = \frac{1}{m^2} \sum_{k,k'=1}^{m} \mathbf{E}\left[ z_{kg}^{(x)} z_{k',g}^{(x)} z_{kg}^{(v)} z_{k',g}^{(v)} \right] = O(\frac{m}{n^2}) \ .$$

In the above evaluation we have used the evaluation that if we set $s_{jk} = \cos\theta_{jk}$ $(j,k = 1,2,\cdots,n)$, then we have the relation

$$|\sum_{j=1}^{n} s_{jk} s_{j,k'}| \leq [\sum_{j=1}^{n} s_{jk}^2] = \frac{n}{2} + \frac{1}{4} \ \text{ for any } k \geq 1 \ .$$

Finally, for the third term, we need to consider the variance of

$$(z_{kg}^{(v)})^2 - \mathbf{E}[(z_{kg}^{(v)})^2] = \sum_{j,j'=1}^{n} b_{kj} b_{k,j'} \left[ v_{jg} v_{j'g} - \mathbf{E}[v_{jg} v_{j',g}] \right] .$$

Then by using the assumptions, we have after some evaluations, we find a positive constant $K_4$ such that

$$\begin{aligned}
\text{(A.19)} \qquad & \mathbf{E}\left[ \frac{1}{m_n} \sum_{k=1}^{m} ((z_{kg}^{(v)})^2 - \mathbf{E}[(z_{kg}^{(v)})^2]) \right]^2 \\
&= \frac{1}{m^2} \sum_{k_1,k_2=1}^{m} \mathbf{E}\left[ \sum_{j_1,j_2,j_3,j_4=1}^{n} b_{k_1,j_1} b_{k_1,j_2} \left( v_{j_1,g} v_{j_2,g} - \mathbf{E}(v_{j_1,g} v_{j_2,g}) \right) \right. \\
&\qquad \left. \times b_{k_2,j_3} b_{k_2,j_4} \left( v_{j_3,g} v_{j_4,g} - \mathbf{E}(v_{j_3,g} v_{j_4,g}) \right) \right] \\
&\leq K_4 \frac{1}{m^2} [\sum_{k=1}^{m} a_{kn}^{*}]^2 \\
&= O(\frac{1}{m^2} \times (\frac{m^3}{n^2})^2) \ ,
\end{aligned}$$

which is $O(m^4/n^4)$ by straight-forward calculations. Here we just give an illustration of our derivations when $p=1$ and we re-write $v_i = \sum_{j=0}^{\infty} c_j^{(v)} e_{i-j}^{(v)}$ ($c_j^{(v)} = \mathbf{C}_j^{(v)}$) and we evaluate

$$
\begin{aligned}
& \sum_{k_1,k_2=1}^{m} \sum_{j_1,j_2,j_3,j_4} b_{k_1,j_1} b_{k_1,j_2} b_{k_2,j_3} b_{k_2,j_4} \times \mathbf{E}\{[v_{j_1}v_{j_2} - \mathbf{E}(v_{j_1}v_{j_2})][v_{j_3}v_{j_4} - \mathbf{E}(v_{j_3}v_{j_4})]\} \\
= & \sum_{k_1,k_2=1}^{m} \sum_{j_1,j_2,j_3,j_4} b_{k_1,j_1} b_{k_1,j_2} b_{k_2,j_3} b_{k_2,j_4} \sum_{l_1,l_2,l_3,l_4=0}^{\infty} c_{l_1}^{(v)} c_{l_2}^{(v)} c_{l_3}^{(v)} c_{l_4}^{(v)} \\
& \quad \times \mathbf{E}\{[e_{j_1-l_1}e_{j_2-l_2} - \mathbf{E}(e_{j_1-l_1}e_{j_2-l_2})][e_{j_3-l_3}e_{j_4-l_4} - \mathbf{E}(e_{j_3-l_3}e_{j_4-l_4})]\} \; .
\end{aligned}
$$

We need to evaluate the corresponding terms for four cases when (i) $j_1-l_1 = j_2-l_2 = j_3-l_3 = j_4-l_4$, (ii) $j_1-l_1 = j_2-l_2 \neq j_3-l_3 = j_4-l_4$, (iii) $j_1-l_1 = j_3-l_3 \neq j_2-l_2 = j_4-l_4$, (iv) $j_1-l_1 = j_4-l_4 \neq j_2-l_2 = j_4-l_4$. By using the condition in the general case that $\|\mathbf{C}_j^{(v)}\| = O(\rho^j)$ $(j \geq 0, 0 \leq \rho < 1)$, we have $\sum_{j=0}^{\infty} |c_j^{(v)}| < \infty$ in this special case. We also utilize the relation such as $\sum_{j=1}^{n} b_{kj} b_{k'j} = \delta(k,k')a_{kn}^{*}$ in the general case and we have the notation that $b_{k,j} = 0$ for $k=1,\cdots,m$, $j<0, j>n$ and $c_j = 0$ $(j<0)$.
Then in each (i)-(iv) case, we can take a positive constant $K_5$ such that (A.19) is less than

$$
K_5 \sum_{k_1,k_2=1}^{m} [\sum_{j_1=1}^{n} b_{k_1,j_1}^2]^{1/2} [\sum_{j_2=1}^{n} b_{k_1,j_2}^2]^{1/2} [\sum_{j_3=1}^{n} b_{k_2,j_3}^2]^{1/2} [\sum_{j_4=1}^{n} b_{k_2,j_4}^2]^{1/2} \; .
$$

The above evaluation methd works in the general case with a complication of notations. Therefore, by using (A.17), the third term of (A.12) is negligible if we set $\alpha$ such that $0 \leq \alpha < 1$. (The derivations are similar to the ones in Kunitomo and Sato (2021).)
(**Step 3**) : When we have the condition $0 < \alpha < 0.8$, we need to evaluate the limiting distribution of the first term of (A.12) because of (A.15). Instead of (A.13), we consider the asymptotic distribution of

$$
\text{(A.20)} \qquad s_{ij}^{(m)*} = \frac{1}{\sqrt{m}} [g_{ij}^{(m*)} - \mathbf{E}(g_{ij}^{(m*)})]
$$

and

$$
\text{(A.21)} \qquad g_{ij}^{(m*)} = (\frac{1}{m} \sum_{k=1}^{m} \mathbf{z}_k^{(x)} \mathbf{z}_k^{(x)'})_{ij} \; (i,j=1,\cdots,p) \; .
$$

By using (A.11) and $\mathbf{P}_n = (p_{jk}^{(n)})$, we decompose

$$
\begin{aligned}
\text{(A.22)} \qquad s_{ij}^{(m)*} = & \frac{1}{\sqrt{m}} \sum_{k=1}^{m} [\sum_{s=t=1}^{n} p_{ks}^{(n)2} (v_{is}^{(x)} v_{js}^{(x)} - \mathbf{E}(v_{is}^{(x)} v_{js}^{(x)}))] \\
& + \frac{1}{\sqrt{m}} \sum_{k=1}^{m} [\sum_{s \neq t=1}^{n} p_{ks}^{(n)} p_{kt}^{(n)} (v_{is}^{(x)} v_{jt}^{(x)} - \mathbf{E}(v_{is}^{(x)} v_{jt}^{(x)}))]
\end{aligned}
$$

and $\mathbf{v}_t^{(x)}$ $(= (v_{it}^{(x)})) = \Delta \mathbf{x}_t = \sum_{s=0}^{\infty} \boldsymbol{C}_s^{(x)} \mathbf{e}_{t-s}^{(x)}$ , where $\boldsymbol{C}_s^{(x)}$ $(= (C_{is}))$ are $p \times p$ matrices with $C_{is} = O(\rho^{|h|})$ $(0 \leq \rho < 1)$, and $\mathbf{e}_s^{(x)}$ are a sequence of mutually independent random vectors with $\mathbf{E}[\mathbf{e}_s^{(x)}] = 0$, $\mathbf{E}[\mathbf{e}_s^{(x)} \mathbf{e}_s^{(x)\prime}] = \boldsymbol{\Sigma}_v^{(x)}$ $(> 0)$.
When we have the condition $m_n/n \to 0$ as $n \to \infty$, we have $\frac{1}{\sqrt{m_n}}[\sigma_{ij}^{(x)} - \mathbf{E}(g_{ij}^{(m*)})] = o(1)$ . The evaluation of the limiting distribution of (A.20) or (A.22) is considerably simpler than that for (A.13).

We use the relations $p_{ks}^{(n)2} = [4/(2n+1)][\cos \theta_{ks}]^2$, $\sum_{k=1}^{m} p_{ks}^{(n)2} = [2m/(2n+1)]c_{ss}$, and $\sum_{s=1}^{n} c_{ss}^2 = O(n)$, $c_{st} = [2/m] \sum_{k=1}^{m} \cos \theta_{sk} \cos \theta_{tk}$, and $\theta_{jk} = \frac{2\pi}{2n+1}(j - \frac{1}{2})(k - \frac{1}{2})$. Because $c_{ss}$ in bounded, and $\mathbf{v}_s^{(x)}$ has a MA representaion with conditions on its coefficients in (2.3), it is possible to evaluate the variances of $[2\sqrt{m}/(2n+1)] \sum_{s=1}^{n} c_{ss}[v_{is}^{(x)} v_{jt}^{(v)} - \mathbf{E}(v_{is}^{(v)} v_{jt}^{(x)})]$, which converge to zeros in probability when $m_n/n \to 0$ as $n \to \infty$. Hence, the first term of (A.22) is asymptotically negligible because of the condition $m_n/n \to 0$ as $n \to \infty$.
Then, we only need to show the asymptotic normality of the leading term of (A.20), which is

$$\text{(A.23)} \qquad s_{ij}^{(m)**} = \frac{2\sqrt{m}}{2n+1} \sum_{s \neq t=1}^{n} c_{st}[v_{is}^{(x)} v_{jt}^{(v)} - \mathbf{E}(v_{is}^{(v)} v_{jt}^{(x)})] \ .$$

Under the stationarity condition of $\mathbf{v}_s^{(x)}$, the difference between (A.23) and the second term of (A.22) is asymptotically negligible. Also under the stationarity and the conditions on coefficients in (2.3), it has been known in time series analysis that the effects of initial conditions on $\mathbf{v}_s^{(x)}$ $(s \leq 0)$ are asymptotically negligible. (We omit the detail of this arguments because it may be straightforward.)
(**Step 4**) : Our proof of the asymptotic normality requires a further derivation, which is a modification of the method for the spectral density estimation used in the proof of Theorem 9.4.1 in Anderson (1971). Because some of our arguments are similar, we only repeat the essential arguments and some differences. We provide the proof for the case when $p = 1$ and use the notation $\boldsymbol{C}_s^{(x)} = c_s$ $(s = 0, 1, \cdots)$, $\mathbf{v}_s^{(x)} = v_s^{(x)}$, $\mathbf{e}_s^{(x)} = e_s$ and $s^{(m)**} = s_{ij}^{(m)**}$ $(i = j = 1)$ because the proof of the general case when $p \geq 1$ can be obtained by using the standard device of $v_j^* = \mathbf{a}' \mathbf{v}_j^{(x)}$ $(j = 1, \cdots, n)$ with an arbitrary ($p \times 1$ non-zero constant) vector $\mathbf{a}$.

We take $K_n = [n/m_n]$ be a sequence of positive integers and $K_n \to \infty$ $(n \to \infty)$. Then, given $s$, $c_{st} \to 0$ for $t - s > K_n$ as $m_n, n \to \infty$ and $m_n/n \to 0$. Then, by taking $t = s + k$ $(k = 1, \cdots, n - s)$ we rewrite

$$\text{(A.24)} \qquad \begin{aligned} s^{(m)**} &= \frac{4\sqrt{m}}{2n+1} \sum_{t>s=1}^{n} c_{st}[v_s^{(x)} v_t^{(x)} - \mathbf{E}(v_s^{(x)} v_t^{(x)})] \ , \\ &= \frac{4\sqrt{m}}{2n+1} \sum_{l,l'=1}^{\infty} c_l c_{l'} \sum_{s=1}^{n} \sum_{t=s+1, s-l \neq t-l'}^{n} c_{st} e_{s-l} e_{t-l'} \ . \end{aligned}$$

$$= \frac{4\sqrt{m}}{2n+1}\sum_{l,l'=1}^{\infty} c_l c_{l'} \sum_{s=1}^{n}\sum_{k=1}^{n-s} c_{s,s+k} e_{s-l} e_{s+k-l'} \ .$$

We truncate the sum $\sum_{l,l'=1}^{\infty}[\ \cdot\ ]$ by a sub-sequence $r_n$ ($r_n \to \infty$ as $n \to \infty$), and decompose the sum as $(\sum_{l=1}^{r_n} + \sum_{l=r_n+1}^{\infty})\ (\sum_{l'=1}^{r_n} + \sum_{l'=r_n+1}^{\infty})[\ \cdot\ ]$. We can take a sequence of sums $\sum_{l,l'=1}^{r_n}[\ \cdot\ ]$ such that $r_n \to \infty$ and $\sum_{l=r_n+1}^{\infty}|\gamma_l| \to 0$. Then we approximate the infinite sum by a finite sum because the remaing terms are of smaller order asymptotically. The main term is

$$\text{(A.25)} \qquad s_1^{(m)**} = \frac{4\sqrt{m}}{2n+1}\sum_{l,l'=1}^{r_n} c_l c_{l'} \sum_{s=1}^{n}\sum_{k=1}^{n-s} c_{s,s+k} e_{s-l} e_{s+k-l'}$$

$$= \frac{4\sqrt{m}}{2n+1}\sum_{l,l'=1}^{r_n} c_l c_{l'} \sum_{h=l-l'+1,h\neq 0}^{n-q-l'} \ \sum_{q=1-l}^{n-l} c_{q+l,q+h+l'} e_q e_{q+h} \ .$$

Since some parts of the above summation (i.e., the terms in $\sum_{l-l'\leq h<1}[\ \cdot\ ]$, $\sum_{n-q-l'\leq h<n-q}[\ \cdot\ ]$, $\sum_{1-l\leq q<0}[\ \cdot\ ]$ and $\sum_{n-l\leq q\leq n-1}[\ \cdot\ ]$ ) can be of negligible order asymptotically, we can approximate the summation as

$$\text{(A.26)} \qquad s_{11}^{(m)**} = \frac{4\sqrt{m}}{2n+1}[\sum_{l=1}^{r_n} c_l \sum_{l'=1}^{r_n} c_{l'}] \sum_{h=l-l'+1}^{n-q-l'} \ \sum_{q=1-l}^{n-l} c_{q+l,q+h+l'} e_q e_{q+h}$$

$$\sim \frac{4\sqrt{m}}{2n+1}[\sum_{l=1}^{r_n} c_l \sum_{l'=1}^{r_n} c_{l'}] \sum_{h=1}^{n-q}\sum_{q=1}^{n} c_{q+l,q+h+l'} e_q e_{q+h} \ ,$$

where we denote $c_{s,t} = 0$ ($s > n$ or $t > n$) for the notational convenience.

(**Step 5**) : As the final step with $p = 1$, we approximate the sequence of weakly dependent random variables by a sum of independent noise random variables, and apply the CLT.

Let $m_n = n^{\alpha}$ ($0 < \alpha < 0.8$), $K_n = [n/m_n]$, $N_n = [n^{\delta/2}]$ ($\delta > 0$), and $M_n = [n^{1-\delta/2}]$ such that $1 - \delta/2 > 0$ and $\alpha + \delta/2 > 1$. Then, $K_n/N_n \to 0$, $N_n/n \to 0$ $\sqrt{m_n}/n \sim [1/\sqrt{n}][1/\sqrt{K_n}]$ and $M_n \sim n/N_n$ as $n \to \infty$. In the following we utilize the relation $c_{q+l,q+h+l'} - c_{q,q+h} = o(1)$ for $l, l' = 1, \cdots, r_n$ if we take $r_n$ such that $r_n \times m_n/n \to 0$ as $n, m_n \to \infty$. This is because

$$\sin 2\pi m[\frac{2q+h+l+l'}{2n+1}] - \sin 2\pi m[\frac{2q+h}{2n+1}]$$

$$= \sin 2\pi m[\frac{2q+h}{2n+1}][\cos 2\pi m(\frac{l+l'}{2n+1}) - 1] + \cos 2\pi m[\frac{2q+h}{2n+1}]\sin 2\pi m[\frac{l+l'}{2n+1}] \to 0$$

as $n \to \infty$.

Furthermore, by using that some parts of (A.26) are of smaller orders as $n \to \infty$ (the terms in $\sum_{h=K_n+1}[\ \cdot\ ]$), we can apply the CLT to

$$\text{(A.27)} \qquad s_{11}^{(m)***} = 2[\sum_{l=1}^{r_n} c_l]^2 \frac{1}{\sqrt{n}}\frac{1}{\sqrt{K_n}}\sum_{q=1}^{n}\sum_{h=1}^{K_n} c_{q,q+h} e_q e_{q+h} \ ,$$

where we denote $c_{q,q+h}=0$ $(q+h>n)$ for notational convenience. We notice that $c_{q,q+h}$ $(q=1,\cdots,n)$ is a sequence of bounded real numbers.

Let

$$V_{qn}=\frac{1}{\sqrt{K_n}}\sum_{h=1}^{K_n}c_{q,q+h}e_qe_{q+h} \tag{A.28}$$

and

$$U_{jn}=\frac{1}{\sqrt{N_n}}[V_{(j-1)N_n+1,n}+\cdots+V_{jN_n-K_n,n}]\ \ (j=1,\cdots,M_n)\ . \tag{A.29}$$

Then, we find that $\mathbf{E}[V_{q,n}]=0$, $\mathbf{E}[V_{q,n}V_{q+h,n}]=0$ ($h$ is any non-zero integer), $\mathbf{E}[V_{q,n}^2]$ are bounded. Further, we have that $U_{1,n},\cdots,U_{M_n,n}$ are mutually independent and $\mathbf{E}[U_{i,n}^4]$ $(i=1,M_n)$ are uniformly bounded using the assumption of the boundedness of the 4-th order moments of $V_q$ $(q=1,\cdots,n)$. Since other terms except the leading term are stochastically of the smaller order, we can ignore them for evaluating the limiting distribution, and we apply the Lyapounov-type CLT. By using the relation that

$$\frac{1}{\sqrt{n}}\sum_{q=1}^{n}V_{qn}-\frac{1}{\sqrt{M_n}}\sum_{j=1}^{M_n}U_{jn}\xrightarrow{p}0 \tag{A.30}$$

as $n\to\infty$. The remaining terms in the above approximation are of smaller order (i.e. $K_n$ terms in each $U_{jn}$ $(j=1,\cdots,M_n)$) when $m_n,n\to\infty$ and $m_n=[n^\alpha],0<\alpha<0.8$ because of the condition $K_n/N_n\to 0$. Then we have the asymptotic normality of (A.26) when $p=1$. By using the relation $m\ \sum_{s,t=1}^{n}c_{st}^2=(n+1/2)^2$ and

$$4[\sum_{j=-\infty}^{\infty}c_j]^2\sum_{g=1}^{n}\sum_{h=1}^{K_n}c_{g,g+h}[\sigma_v^{(x)}]^4\sim 2[\sum_{j=-\infty}^{\infty}c_j]^2\sum_{s,t=1}^{n}c_{st}^2[\sigma_v^{(x)}]^4\ , \tag{A.31}$$

we have the desired result of the asymptotic variance when $p=1$.
(**Step 6**) : When $p\geq 1$, we can evaluate the asymptotic covariance by calculating the covariance of $\sum_{q,h}c_{q,q+h}\boldsymbol{\gamma}_a^{'}\mathbf{v}_q\boldsymbol{\gamma}_b^{'}\mathbf{v}_{q+h}$ and $\sum_{q^{'},h^{'}}c_{q^{'},q^{'}+h^{'}}\boldsymbol{\gamma}_c^{'}\mathbf{v}_{q^{'}}\boldsymbol{\gamma}_d^{'}\mathbf{v}_{q^{'}+h^{'}}$, where $\boldsymbol{\gamma}_a$ represents any constant $p\times 1$ vector. Then, after straightforward evaluations, we finally find the asymptotic covariance in Theorem A.1 as $\sigma_{ac}^{(x)}\sigma_{bd}^{(x)}+\sigma_{ad}^{(x)}\sigma_{bc}^{(x)}$ $(a,b,c,d=1,\cdots,p)$.
(**Q.E.D**)

**Proof of Theorem 3.1** : We use the representation

$$\sqrt{m_n}[\hat{\boldsymbol{\beta}}_m-\boldsymbol{\beta}]=\sqrt{m_n}\mathbf{G}_{22}^{*-1}(\mathbf{0},\mathbf{I}_k)\mathbf{G}_m^{*}\begin{pmatrix}1\\-\boldsymbol{\beta}\end{pmatrix}\ . \tag{A.32}$$

Because $\mathbf{G}_{22}^{*}\xrightarrow{p}\boldsymbol{\Sigma}_{22}^{(x)}$ $(m/n\to 0,n\to\infty)$ and under the assumption that $\boldsymbol{\Sigma}_{22}^{(x)}$ is a positive definite matrix, we investigate the asymptotic distribution of

$$\sqrt{m_n}[\hat{\boldsymbol{\beta}}_m^{*}-\boldsymbol{\beta}]=\boldsymbol{\Sigma}_{22}^{(x)-1}\frac{1}{\sqrt{m}}(\mathbf{0},\mathbf{I}_k)\mathbf{G}_m^{*}\begin{pmatrix}1\\-\boldsymbol{\beta}\end{pmatrix}\ , \tag{A.33}$$

which is asymptotically equaivalent to (A.33). Then, its asymptotic variance-covariance matrix can be written as

$$\text{(A.34)} \qquad \mathbf{AV}[\hat{\boldsymbol{\beta}}_m] = \boldsymbol{\Sigma}_{22}^{(x)-1}\mathbf{Cov}\left[(\mathbf{0},\mathbf{I}_k)\mathbf{S}\mathbf{b}\ ,\ \mathbf{b}^{'}\mathbf{S}\begin{pmatrix}\mathbf{0}^{'}\\ \mathbf{I}_k\end{pmatrix}\right]\boldsymbol{\Sigma}_{22}^{(x)-1}\ ,$$

where $\mathbf{S} = \sqrt{m_n}[\mathbf{G}_m^* - \boldsymbol{\Sigma}_x]\ (= (s_{jk}))$ and $\mathbf{b} = \begin{pmatrix}1\\ -\boldsymbol{\beta}\end{pmatrix}\ (= (b_j))$ .
By using Theorem A.1, we can evaluate the $(l,l^{'})$-th element $(l,l^{'} = 2,\cdots,k+1 = p)$ of $\boldsymbol{\Sigma}_x = (\sigma_{l,l^{'}}^{(x)})$ as

$$\begin{aligned}
\mathbf{Cov}[\sum_{j=1}^{k+1} b_j s_{jl} \sum_{j^{'}=1}^{k+1} b_{j^{'}} s_{j^{'}l^{'}}] &= \sum_{j,j^{'}=1}^{k+1} b_j b_{j^{'}} (\sigma_{j,j^{'}}^{(x)}\sigma_{l,l^{'}}^{(x)} + \sigma_{j,l^{'}}^{(x)}\sigma_{j^{'},l}^{(x)}) \\
&= \sigma_{l,l^{'}}^{(x)}\sum_{j=1}^{k+1} b_j[\sum_{l^{'}=1}^{k+1} b_{j^{'}}\sigma_{j,j^{'}}^{(x)}] + [\sum_{j=1}^{k+1} b_j\sigma_{j,l^{'}}^{(x)}][\sum_{j^{'}=1}^{k+1} b_{j^{'}}\sigma_{l,j^{'}}^{(x)}] \\
&= \sigma_{l,l^{'}}^{(x)}\sigma_{11.2}^{(x)}
\end{aligned}$$

because $[\boldsymbol{\sigma}_{21}^{(x)},\boldsymbol{\Sigma}_{22}^{(x)}]\mathbf{b} = \mathbf{0}$ and

$$\text{(A.35)} \qquad [\boldsymbol{\sigma}_{11}^{(x)},\boldsymbol{\sigma}_{12}^{(x)}]\mathbf{b} = \sigma_{11}^{(x)} - \boldsymbol{\sigma}_{12}^{(x)}\boldsymbol{\Sigma}_{22}^{(x)-1}\boldsymbol{\sigma}_{21}^{(x)}\ .$$

Then we have the result of the asymptotic variance-covariance matrix of (A.30) in Theorem 3.1.
(**Q.E.D**)

**Proof of Theorem 4.1** : We use the representation

$$\text{(A.36)} \qquad \hat{\mathbf{B}}_m - \mathbf{B} = (\mathbf{W}_m^{*'}\mathbf{W}_m^*)^{-1}\mathbf{W}_m^{*'}\mathbf{U}_m^*\ ,$$

where $\mathbf{U}_m^* = \mathbf{J}_m\mathbf{P}_n\mathbf{C}_n^{-1}\mathbf{U}_n$. Rewrite (A.36) as $\sqrt{m_n}[\hat{\mathbf{B}}_m - \mathbf{B}] = (\frac{1}{m_n}\mathbf{W}_m^{*'}\mathbf{W}_m^*)^{-1}\frac{1}{\sqrt{m_n}}\mathbf{W}_m^{*'}\mathbf{U}_m^*$. By using a similar argument as the proof of Theorem 3.1 under the assumption of (4.4), we find that

$$\text{(A.37)} \qquad \mathbf{AV}[\hat{\mathbf{B}}_m] = \boldsymbol{\Sigma}_{w^*}^{-1}\mathbf{Cov}[\frac{1}{\sqrt{m}}\mathbf{W}_m^{*'}\mathbf{U}_m^*, \frac{1}{\sqrt{m}}\mathbf{W}_m^{*'}\mathbf{U}_m^*]\boldsymbol{\Sigma}_{w^*}^{-1}\ .$$

Then, by using Theorems A.1 and 3.1 we have the result.
(**Q.E.D**)

## APPENDIX B : Figures

In this Appendix, we gather some figures cited in Section 6. All computations have been done by the program called x12siml written in R, which will be available in the near future.

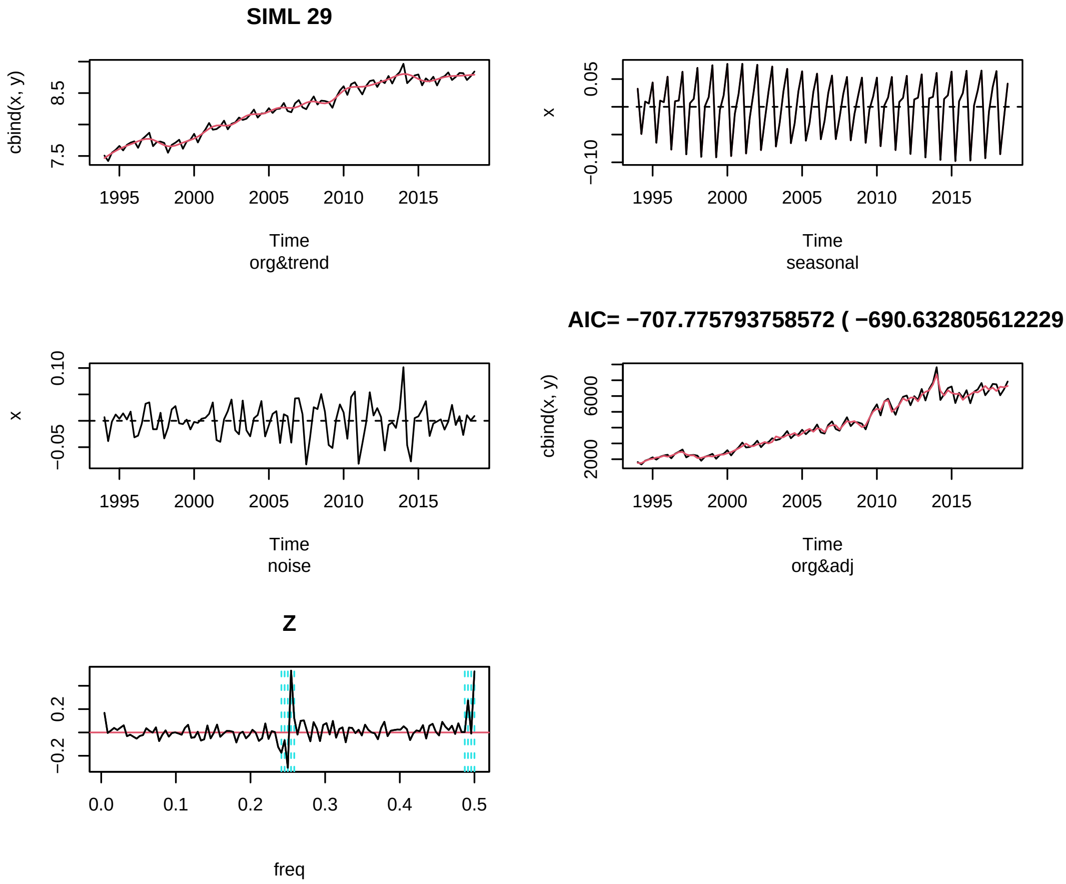


Figure 1: Macro-consumption (Data are the Quarterly real consumption of durable goods (after log-transformation) between 1994Q1-2019Q4 published by the Economic Social Research Institute (ESRI), Cabinet Office, Japan.)

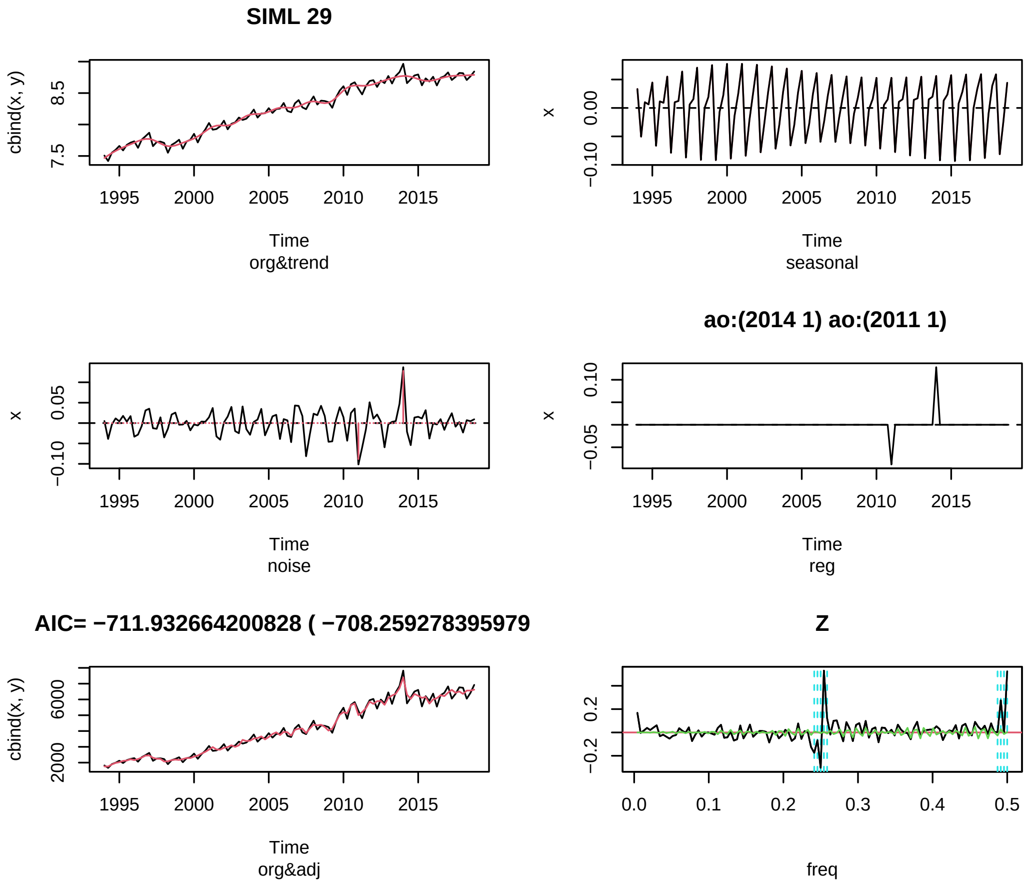


Figure 2: Macro-consumption (Data are the Quarterly real consumption of durable goods (after log-transformation) between 1994Q1-2019Q4 published by the Economic Social Research Institute (ESRI), Cabinet Office, Japan.)

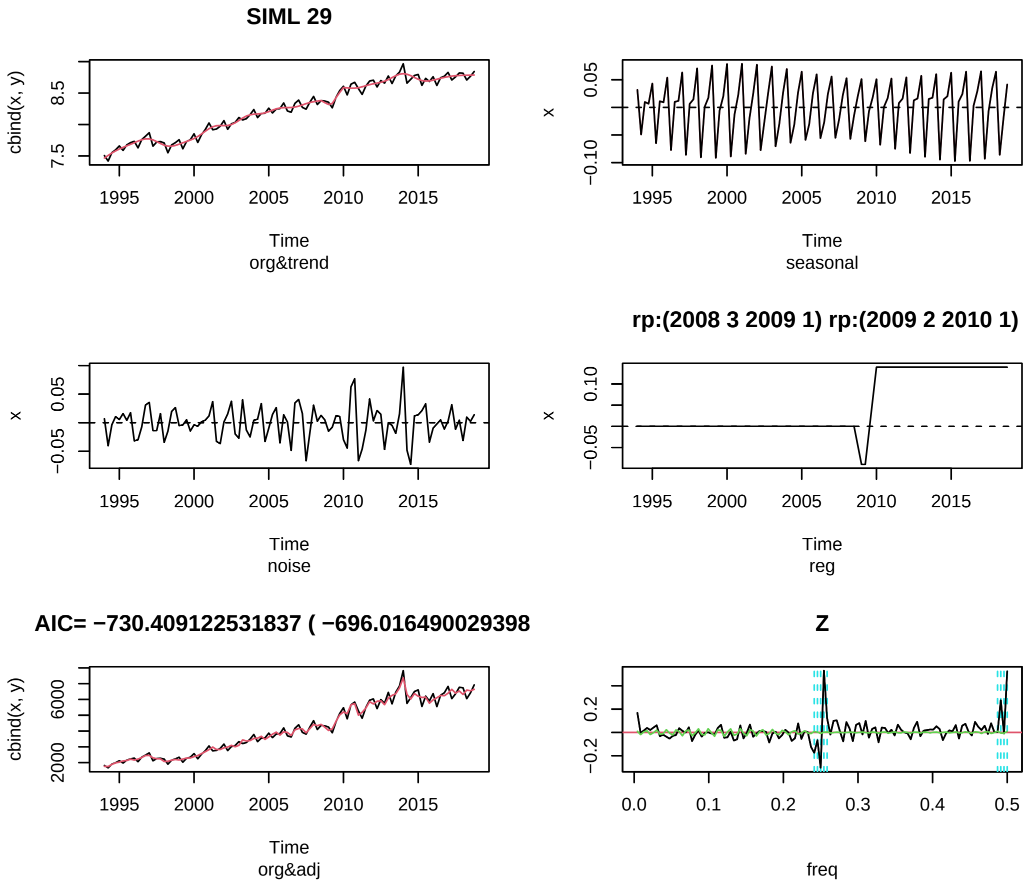


Figure 3: Macro-consumption (Data are the Quarterly real consumption of durable goods (after log-transformation) between 1994Q1-2019Q4 published by the Economic Social Research Institute (ESRI), Cabinet Office, Japan.)

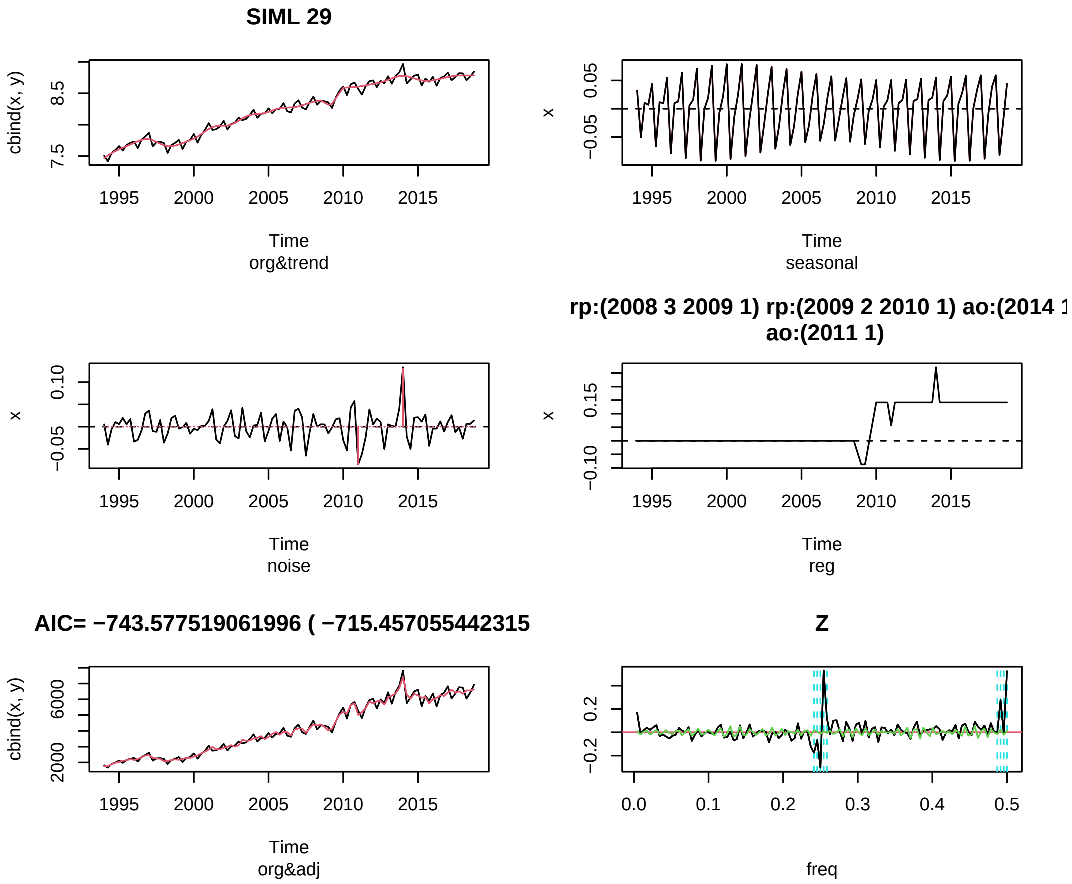


Figure 4: Macro-consumption (Data are the Quarterly real consumption of durable goods (after log-transformation) between 1994Q1-2019Q4 published by the Economic Social Research Institute (ESRI), Cabinet Office, Japan.)