SSE-DP-2024-5

Forward and Backward Smoothing for Noisy Nonstationary Time Series

with an Application of Detecting Recent Change Points

Seisho Sato

(University of Tokyo)

and

Naoto Kunitomo

(Institute of Statistical Mathematics)

December 2024

Forward and Backward Smoothing for Noisy Nonstationary Time Series with an Application of Detecting Recent Change Points *

Seisho Sato †
and
Naoto Kunitomo ‡

December 13, 2024

Abstract

We propose a novel smoothing (or filtering) approach for time series analysis to estimate the hidden states of random variables and handle noisy, nonstationary time series data. The method is applicable even when the sample size is small, as is often the case with major macroeconomic time series data. Our approach is based on the frequency decomposition of nonstationary time series, and we address the smoothing and filtering challenges specific to such data. In particular, we introduce two methods: forward and backward SIML smoothing, designed to resolve the initial value problem in nonstationary time series analysis. The proposed smoothing methods offer interpretations in both the time and frequency domains. To demonstrate the effectiveness of our approach, we provide an illustrative empirical example using U.S. manufacturers' new order data and apply the filtering method to the problem of detecting recent breaks in macroeconomic consumption trends.

Key Words

Nonstationary economic time series, Errors-in-variables models, Frequency decomposition, Forward and backward SIML smoothing, Band smoothing, Multi-step smoothing, Detecting Recent Change Points.

*2024-12-13. This is a revised version of the paper (SDS-15, MIMS-RSP Statistics and Data Science Series, MIMS, Meiji University). (http://www.mims.meiji.ac.jp/publications/2020-ds/SDS-15.pdf) The research has been supported by JSPS-Grant JP17H02513.

†Graduate School of Economics, University of Tokyo, Bunkyo-ku,Hongo7-3-1, Tokyo 113-0033, seisho@e.u-tokyo.ac.jp

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

1. Introduction

Time series analysis is a cornerstone of macroeconomic research, providing critical insights into trends, cycles, and other dynamic behaviors in economic data. However, macroeconomic time series often exhibit significant challenges: they are typically nonstationary, noisy, and available in small sample sizes. For example, 30 years of quarterly data comprise only 120 observations, while 20 years of monthly data include just 240 observations. These characteristics limit the applicability of traditional smoothing and filtering methods, which often rely on large sample sizes or strong distributional assumptions.

Existing methods, such as Kalman filtering and the Hodrick-Prescott filter, have been widely employed in econometrics. While these methods are powerful, they often face limitations in multivariate settings or when addressing noisy, nonstationary data. Additionally, many traditional approaches struggle with the initial value problem, where the starting estimate for state variables significantly influences the smoothing outcomes. Another critical challenge lies in detecting structural breaks near the endpoints of observations. Recent macroeconomic shocks, such as the 2008 financial crisis and the COVID-19 pandemic, underscore the importance of timely structural break detection for effective policymaking.

To address these issues, this study proposes a new smoothing and filtering framework tailored to the needs of macroeconomic time series analysis. Our approach, referred to as the SIML smoothing methods, leverages frequency-domain filtering to provide robust and interpretable solutions for nonstationary time series. The proposed methods include forward, backward, and multi-step smoothing procedures, which are designed to resolve the initial value problem systematically. Moreover, this study extends these methods to multivariate settings and applies them to the detection of structural changes near observation endpoints, a task where existing methods are often inadequate.

The contributions of this study are as follows: (i) Development of Novel Smoothing Methods: We introduce forward, backward, and multi-step SIML smoothing techniques that address key challenges in nonstationary time series analysis, including small sample sizes and initial value sensitivity. (ii) Structural Break Detection Framework: A new method for detecting structural changes near observation endpoints is developed, focusing on recent changes in trend-cycle components of noisy data. (iii) Applicability to Multivariate Time Series: Our methods are designed to handle multivariate settings without relying on strong distributional assumptions, making them suitable for multiple macroeconomic data. (iv) Empirical Applications: The proposed methods are demonstrated through real-world examples, including U.S. manufacturers' new order data and the detection of structural breaks in macroeconomic consumption trends in Japan.

By addressing the unique challenges of nonstationary, noisy, and small-sample time series, this study provides a systematic, interpretable, and robust framework for macroeconomic time series analysis. Our methods not only extend the capabilities of existing smoothing and filtering techniques, but also offer new tools for structural break detection, with significant implications for econometric research and policy formulation.

The remainder of this paper is structured as follows. In Section 2, we introduce the general framework of the nonstationary errors-in-variables model and the SIML method. Section 3 develops the SIML smoothing techniques, including forward, backward, and multi-step procedures, and addresses theoretical issues such as convergence. Section 4 generalizes the methods to multivariate settings and provides a frequency-domain interpretation. Section 5 presents numerical examples to illustrate the practical applications of our methods, while Section 6 introduces a novel approach to detecting recent structural breaks. Finally, Section 7 concludes the study and outlines future research directions. Mathematical derivations and additional figures are provided in the Appendix.

2. Nonstationary Errors-in-variables models

Let y_{ji} be the $i-$th observation of the $j-$th time series at i for $i = 1, \cdots, n; j = 1, \cdots, p$. We set $\mathbf{y}_i = (y_{1i}, \cdots, y_{pi})^{'}$ as a $p \times 1$ vector and $\mathbf{Y}_n = (\mathbf{y}_i^{'})$ $(= (y_{ij}))$ as an $n \times p$ matrix of observations, and we denote $\mathbf{y}_0$(or $\mathbf{y}_n$) as the initial $p \times 1$ vector, which is assumed to be observable. Furthermore, we attempt to estimate the underlying nonstationary trend and stationary cycle when the nonstationary state vector $\mathbf{x}_i$ $(= (x_{ji}))$ $(i = 0, 1, \cdots, n)$, and the vector of noise component $\mathbf{v}_i^{'} = (v_{1i}, \cdots, v_{pi})$ are mutually independent. Then, we use the nonstationary errors-in-variables representation

$$\mathbf{y}_i = \mathbf{x}_i + \mathbf{v}_i \quad (i = 0, 1, \cdots, n), \tag{2.1}$$

where the state vector $\mathbf{x}_i$ $(i = 0, 1, \cdots, n)$ is a sequence of the nonstationary I(1) vector process, which satisfies

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

($\mathbf{x}_0$ or $\mathbf{x}_n$ as the initial vector), and the measurement error (or noise) vector $\mathbf{v}_i^{(x)}$ is a sequence of the i.i.d. random vectors with $\mathbf{E}(\mathbf{v}_i^{(x)}) = \mathbf{0}$ and $\mathbf{E}(\mathbf{v}_i^{(x)}\mathbf{v}_i^{(x)'}) = \mathbf{\Sigma}_x$. The random vector $\mathbf{v}_i$ $(i = 0, 1, \cdots, n)$ is a sequence of i.i.d. random variables with $\mathbf{E}(\mathbf{v}_i) = \mathbf{0}$ and $\mathbf{E}(\mathbf{v}_i\mathbf{v}_i^{'}) = \mathbf{\Sigma}_v$.

When we assume 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, and we have the observations of an $n \times p$ matrix $\mathbf{Y}_n = (\mathbf{y}_i^{'})$, then, given the initial condition $\mathbf{y}_0$ as times goes from 0 to n, the $np \times 1$ random vector $(\mathbf{y}_1^{'}, \cdots, \mathbf{y}_n^{'})^{'}$ follows

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

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

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

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

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

where

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

and the $(k, j)-$th element of $\mathbf{P}_n = (p_{kj}^{(n)})$ is defined by

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

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

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

(See Chapter 3 of Kunitomo and Sato (2024) for the details.)

In the general case of (2.1) and (2.2), $\mathbf{y}_i = \mathbf{x}_i + \mathbf{v}_i$, $\Delta \mathbf{x}_i = \mathbf{v}_i^{(x)}$, and the noise component $\mathbf{v}_i (i = 0, 1, \cdots, n)$ and the state variables component $\mathbf{v}_i^{(x)}$ $(= \Delta \mathbf{x}_i)$ are sequences of the stationary processes satisfying

$$\mathbf{v}_i = \sum_{j=-\infty}^{\infty} \mathbf{C}_j^{(v)} \mathbf{e}_{i-j}^{(v)} \tag{2.9}$$

and

$$\mathbf{v}_i^{(x)} = \sum_{j=-\infty}^{\infty} \mathbf{C}_j^{(x)} \mathbf{e}_{i-j}^{(x)} , \tag{2.10}$$

where $\mathbf{e}_i^{(v)}$ and $\mathbf{e}_i^{(x)}$ are sequences of i.i.d. random vectors with $\mathbf{E}(\mathbf{e}_i^{(v)}) = \mathbf{E}(\mathbf{e}_i^{(x)}) = \mathbf{0}$, $\mathbf{E}(\mathbf{e}_i^{(v)}\mathbf{e}_i^{(v)\prime}) = \mathbf{\Sigma}_e^{(v)}$ (a non-negative definite matrix) and $\mathbf{E}(\mathbf{e}_i^{(x)}\mathbf{e}_i^{(x)\prime}) = \mathbf{\Sigma}_e^{(x)}$ (a non-negative definite matrix). For normalization we use $\mathbf{C}_0^{(v)} = \mathbf{C}_0^{(x)} = \mathbf{I}_p$.
The $p \times p$ coefficient matrices $\mathbf{C}_j^{(v)}$ and $\mathbf{C}_j^{(x)} (j = \cdots, -1, 0, 1, \cdots)$ in (2.9) and (2.10) are absolutely summable such that $\sum_{j=-\infty}^{\infty} \|\mathbf{C}_j^{(v)}\| < \infty$ and $\sum_{j=-\infty}^{\infty} \|\mathbf{C}_j^{(x)}\| < \infty$, where $\|\mathbf{C}_j^{(v)}\| = \max_{k,l=1,\cdots,p} |c_{k,l}^{(v)}(j)|$ for $\mathbf{C}_j^{(v)} = (c_{k,l}^{(v)}(j))$ and $\|\mathbf{C}_j^{(x)}\| = \max_{k,l=1,\cdots,p} |c_{k,l}^{(x)}(j)|$ for $\mathbf{C}_j^{(x)} = (c_{k,l}^{(x)}(j))$, respectively.
The measurement error vector $\mathbf{v}_i$ may include the (stationary) seasonal component $\mathbf{s}_i$ when the main interest is to understand the trend factors as the state vector in the low frequency part, which is less than a year. Alternatively, the state vector $\Delta\mathbf{x}_i$ $(= \mathbf{v}_i^{(x)})$ may include the seasonal components satisfying $\mathbf{s}_i = \sum_{j=-\infty}^{\infty} \mathbf{C}_{sj}^{(s)} \mathbf{e}_{i-sj}^{(s)}$ $(s \geq 2)$ and $\mathbf{e}_i^{(s)}$ is a sequence of i.i.d. random vectors with $\mathbf{E}(\mathbf{e}_i^{(s)}) = \mathbf{0}$ and $\mathbf{E}(\mathbf{e}_i^{(s)}\mathbf{e}_i^{(s)\prime}) = \mathbf{\Sigma}_e^{(s)}$ (the coefficient matrices $\mathbf{C}_{sj}^{(s)}$ are absolutely summable $\sum_{j=-\infty}^{\infty} \|\mathbf{C}_j^{(s)}\| < \infty$).

In the time series analysis, the causal MA representation of stationary process has the form $\mathbf{v}_i = \sum_{j=0}^{\infty} \mathbf{C}_j^{(v)} \mathbf{e}_{i-j}^{(v)}$, and $\mathbf{v}_i^{(x)} = \sum_{j=0}^{\infty} \mathbf{C}_j^{(x)} \mathbf{e}_{i-j}^{(x)}$ in (2.9) and (2.10). The non-causal representation has similar form with the reversed time direction $\mathbf{v}_i = \sum_{j=-\infty}^{0} \mathbf{C}_j^{(v)} \mathbf{e}_{i-j}^{(v)}$, and $\mathbf{v}_i^{(x)} = \sum_{j=-\infty}^{0} \mathbf{C}_j^{(x)} \mathbf{e}_{i-j}^{(x)}$. (For a discussion of causal and non-causal representations of stationary processes, see Chapter 3 of Brockwell and Davis (1990).) In the following sections, we interpret the forward SIML filtering as a causal filtering based on the causal MA representation of stationary time series while the backward SIML filtering as a non-causal filtering based on the non-causal representation.

3. The SIML Smoothing Methods

3.1 Forward SIML Smoothing

For the stationary process, we utilize the representation of (2.9) and (2.10). We first consider the nonstationary errors-in-variables model of (2.1), (2.9) and (2.10) with the causal representation $\mathbf{v}_i = \sum_{j=0}^{\infty} \mathbf{C}_j^{(v)} \mathbf{e}_{i-j}^{(v)}$ and $\mathbf{v}_i^{(x)} = \sum_{j=0}^{\infty} \mathbf{C}_j^{(x)} \mathbf{e}_{i-j}^{(x)}$. To disentangle the non-stationarity of time series data, we first use $\mathbf{K}_n-$transformation in (2.5) because the elements of the resulting $n \times p$ random matrix $\mathbf{Z}_n$ take real values in the frequency domain, and their roles are easy to be understood. Unlike the standard time series analysis, however, we use the real-valued Fourier transformation and we have an intuition that they are the orthogonal data process and are nearly distributed as the Gaussian process. As $\mathbf{P}_n$ is a real-valued discrete Fourier transformation, vectors $\mathbf{z}_k$ $(k = 1, \cdots, n)$ in $\mathbf{Z}_n$ are asymptotically uncorrelated. (See discussions on spectral decomposition of stationary time series in Section 4.)

The main idea of the forward filtering is that we consider the partial inversion of the transformed (real-valued causal) orthogonal processes. Let an $n \times p$ matrix be

$$\hat{\mathbf{X}}_n(\mathrm{Q}) = \mathbf{C}_n \mathbf{P}_n \mathbf{Q}_n \mathbf{P}_n \mathbf{C}_n^{-1} (\mathbf{Y}_n - \bar{\mathbf{Y}}_0) \tag{3.1}$$

and

$$\mathbf{Z}_n = \mathbf{P}_n \mathbf{C}_n^{-1} (\mathbf{Y}_n - \bar{\mathbf{Y}}_0) \ , \mathbf{Y}_n = \bar{\mathbf{Y}}_0 + \mathbf{X}_n^{(0)} + \mathbf{V}_n \ , \tag{3.2}$$

where $\mathbf{X}_n^{(0)} = (\mathbf{x}_i^{(0)'})$ and $\mathbf{V}_n = (\mathbf{v}_i^{'})$ are $n \times p$ matrices, and $\mathbf{x}_i^{(0)} = \mathbf{x}_i - \mathbf{x}_0$ $(i = 1, \cdots, n)$. We set the initial vector as $\mathbf{y}_0 = \mathbf{x}_0$.
The stochastic process $\mathbf{Z}_n$ is the orthogonal decomposition of the original time series $\mathbf{Y}_n$ in the frequency domain, and $\mathbf{Q}_n$ is an $n \times n$ filtering matrix. Because $\mathbf{Y}_n$ consists of non-stationary time series, we need a special form of transformation $\mathbf{K}_n$ in (2.5). Let an $m \times n$ choice matrix be $\mathbf{J}_m = (\mathbf{I}_m, \mathbf{O})$, and let an $n \times p$ matrix be

$$\hat{\mathbf{X}}_n(m) = \mathbf{C}_n \mathbf{P}_n \mathbf{Q}_n^{(m)} \mathbf{P}_n \mathbf{C}_n^{-1} (\mathbf{Y}_n - \bar{\mathbf{Y}}_0) \ , \tag{3.3}$$

and we denote an $n \times n$ matrix $\mathbf{Q}_n^{(m)} = \mathbf{J}_m^{'} \mathbf{J}_m$.
For the $n \times p$ hidden state matrix $\mathbf{X}_n$, we construct an estimator only by using the lower-frequency parts in the inverse transformation of $\mathbf{Z}_n$ and by deleting the estimated noise parts. We denote the hidden trend state as

$$\mathbf{X}_n(m) = \mathbf{C}_n \mathbf{P}_n \mathbf{Q}_n^{(m)} \mathbf{P}_n \mathbf{C}_n^{-1} \mathbf{X}_n^{(0)} \ . \tag{3.4}$$

This quantity is different from $\mathbf{X}_n$ because $\mathbf{x}_i$ $(i = 1, \cdots, n)$ in (3.1) and (3.2) contains not only the trend component of $\mathbf{y}_i$ $(i = 1, \cdots, n)$ but also the noise component in the frequency domain. We attempt to estimate the trend component of $\mathbf{x}_i$ by using (3.3) and recover the trend component of $\mathbf{X}_n$, which are close to zero frequency, because the effects of differenced measurement error noises $(\mathbf{v}_i - \mathbf{v}_{i-1})$ are negligible at zero frequency. This method differs from some existing procedures that consider the decomposition of stationary time series only in the time domain. Our arguments can be justified by using the frequency decomposition of $\mathbf{y}_i$ and $\mathbf{r}_i^{(n)} = \Delta \mathbf{y}_i$ $(= \mathbf{y}_i - \mathbf{y}_{i-1}$ and $\mathbf{y}_0$ being fixed). Because the issue has importance consequences, we will discuss this in the detail in Section 4.

3.2 Backward SIML Smoothing

We investigate the role of the initial condition in the nonstationary process and consider the situation when the time is reversed, that is, from n to 0, rather than from 0 to n. We consider the nonstationary errors-in-variables model of (2.1), (2.9) and (2.10) with the non-causal representation $\mathbf{v}_i = \sum_{j=-\infty}^{0} \mathbf{C}_j^{(v)} \mathbf{e}_{i-j}^{(v)}$ and $\mathbf{v}_i^{(x)} = \sum_{j=-\infty}^{0} \mathbf{C}_j^{(x)} \mathbf{e}_{i-j}^{(x)}$.

We take the $n \times p$ matrix $\mathbf{Y}_n^* = (\mathbf{y}_{i-1}^{'})$ and set the $np \times 1$ random vector $(\mathbf{y}_0^{'}, \cdots, \mathbf{y}_{n-1}^{'})^{'}$ [1]. Given the initial condition $\mathbf{y}_n$, we rewrite

$$\text{vec}(\mathbf{Y}_n^*) \sim N_{n\times p}\left(\mathbf{1}_n \cdot \mathbf{y}_n^{'}, \mathbf{I}_n \otimes \mathbf{\Sigma}_v + \mathbf{C}_n^{'}\mathbf{C}_n \otimes \mathbf{\Sigma}_x\right) \ , \tag{3.5}$$

where $\mathbf{1}_n^{'} = (1, \cdots, 1)$ and $\mathbf{C}_n$ is given by (2.4). (See (2.3) for $\mathbf{Y}_n$.)
We use $\mathbf{K}_n^*-$transformation that from $\mathbf{Y}_n^*$ to $\mathbf{Z}_n^*$ $(= (\mathbf{z}_k^{*'}))$ by

$$\mathbf{Z}_n^* = \mathbf{K}_n^*\left(\mathbf{Y}_n^* - \bar{\mathbf{Y}}_n^*\right) , \mathbf{K}_n^* = \mathbf{P}_n^*\mathbf{C}_n^{'-1} \ , \tag{3.6}$$

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

$$\mathbf{C}_n^{'-1} = \begin{pmatrix} 1 & -1 & \cdots & 0 & 0 \\ 0 & 1 & -1 & \cdots & 0 \\ 0 & 0 & 1 & -1 & \cdots \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}_{n\times n} , \tag{3.7}$$

and the $(k, j)-$th element of $\mathbf{P}_n^* = (p_{kj}^{*(n)})$ is defined by

$$p_{kj}^{*(n)} = \sqrt{\frac{2}{n+\frac{1}{2}}}\sin\left[\frac{2\pi}{2n+1}(k-\frac{1}{2})j\right] . \tag{3.8}$$

By using the spectral decomposition, $\mathbf{C}_n^{'-1}\mathbf{C}_n^{-1} = \mathbf{P}_n^{*'}\mathbf{D}_n\mathbf{P}_n^*$, and $\mathbf{D}_n$ is a diagonal matrix with the k-th element $d_k = 2[1 - \cos(\pi(\frac{2k-1}{2n+1}))]$ $(k = 1, \cdots, n)$ (see Lemma A-1 in the Appendix). In our formulation of two transformations of (3.2) and (3.6), we have the common latent roots both in the forward and backward smoothing procedures as $a_{kn}^* \ (= d_k) \ = 4\sin^2\left[\frac{\pi}{2}\left(\frac{2k-1}{2n+1}\right)\right]$ $(k = 1, \cdots, n)$.

We consider the partial inversion of the transformed orthogonal processes. Let an $n \times p$ matrix be

$$\hat{\mathbf{X}}_n^*(\mathbf{Q}_n) = \mathbf{C}_n^{'}\mathbf{P}_n^{*'}\mathbf{Q}_n\mathbf{P}_n^*\mathbf{C}_n^{'-1}(\mathbf{Y}_n^* - \bar{\mathbf{Y}}_n^*) \tag{3.9}$$

and

$$\mathbf{Z}_n^* = \mathbf{P}_n^*\mathbf{C}_n^{'-1}(\mathbf{Y}_n^* - \bar{\mathbf{Y}}_n^*) , \mathbf{Y}_n^* = \bar{\mathbf{Y}}_n^* + \mathbf{X}_n^* + \mathbf{V}_n^* \ , \tag{3.10}$$

where $\mathbf{X}_n^* = (\mathbf{x}_{i-1}^{*(n)'})$ and $\mathbf{V}_n^* = (\mathbf{v}_{i-1}^{'})$ are the $n \times p$ matrices, and $\mathbf{x}_{i-1}^{*(n)} = \mathbf{x}_{i-1} - \mathbf{x}_n$ $(i = 1, \cdots, n)$.
The stochastic process $\mathbf{Z}_n^*$ is the orthogonal decomposition of the original time series $\mathbf{Y}_n^*$ in the frequency domain, and $\mathbf{Q}_n$ is an $n \times n$ filtering matrix. Because $\mathbf{Y}_n^*$ consists of non-stationary time series, we need a special form of transformation $\mathbf{K}_n^*$.

[1] Given the initial condition $\mathbf{y}_n$, we consider the joint distribution of $(\mathbf{y}_{n-1}^{'}, \cdots, \mathbf{y}_0^{'})^{'}$, while we use $\mathbf{y}_i$ $(i = 0, 1, \cdots, n)$ as the notation without making confusion..

We provide an explicit form for the trend filtering procedure. Then, let the $n \times p$ matrix be

$$\hat{\mathbf{X}}_n^*(m) = \mathbf{C}_n^{'}\mathbf{P}_n^{*'}\mathbf{Q}_n^{(m)}\mathbf{P}_n^{*}\mathbf{C}_n^{'-1}(\mathbf{Y}_n^* - \bar{\mathbf{Y}}_n^*) \tag{3.11}$$

and $\mathbf{Q}_n^{(m)} = \mathbf{J}_m^{'}\mathbf{J}_m$.

We construct an estimator of $n \times p$ hidden state matrix $\mathbf{X}_n^*$ only in the lower-frequency parts by using the inverse transformation of $\mathbf{Z}_n^*$ and by deleting the estimated noise parts. In this notation, the corresponding hidden trend state is given by

$$\mathbf{X}_n^*(m) = \mathbf{C}_n^{'}\mathbf{P}_n^{*'}\mathbf{Q}_n^{(m)}\mathbf{P}_n^{*}\mathbf{C}_n^{'-1}\mathbf{X}_n^{*(n)} , \tag{3.12}$$

where $\mathbf{X}_n^{*(0)} = (\mathbf{x}_i^{*(0)'})$ is the hidden state matrix and $\mathbf{x}_i^{*(n)} = \mathbf{x}_i - \mathbf{x}_n$ $(i = 0, \cdots, n-1)$. It corresponds to the inversion of the m low frequency parts of the hidden state variables.

3.3 Initial Value Problem and Convergence

When dealing with non-stationary time series observations modeled as a random walk, the initial value plays a crucial role due to the nature of non-stationarity. This contrasts with stationary time series models, where the effect of the initial value becomes asymptotically negligible as the sample size increases. Therefore, it is important to develop a smoothing or filtering procedure for non-stationary time series that minimizes dependence on the initial value. In the context of initial values, two possibilities arise: $\mathbf{y}_0$ and $\mathbf{y}_n$ when we have $n+1$ vector observations $\mathbf{y}_i$ $(i = 0, 1, \cdots, n)$. In this problem, we have an interesting result.

We consider two operators $T_{2k}^{(m)}$ and $T_{2k-1}^{(m)}(k \geq 1)$ to an $n \times 1$ vector. Let $T_0 = I_n$ and define $T_{2k-1}^{(m)}$ and $T_{2k}^{(m)}$ recursively for $k = 1, \cdots, M$ by

$$T_{2k+1}^{(m)}(\mathbf{y}) = \mathbf{C}_n\mathbf{P}_n\mathbf{Q}_n^{(m)}\mathbf{P}_n\mathbf{C}_n^{-1}[\mathbf{y} - \mathbf{1}_n(\mathbf{e}_1^{'}T_{2k}^{(m)}(\mathbf{y}))] + \mathbf{1}_n(\mathbf{e}_1^{'}T_{2k}^{(m,n)}(\mathbf{y})) , \tag{3.13}$$

and

$$T_{2k}^{(m)}(\mathbf{y}) = \mathbf{C}_n^{'}\mathbf{P}_n^{*}{}^{'}\mathbf{Q}_n^{(m)}\mathbf{P}_n^{*}\mathbf{C}_n^{'-1}[\mathbf{y} - \mathbf{1}_n(\mathbf{e}_n^{'}T_{2k-1}^{(m)}(\mathbf{y}))] + \mathbf{1}_n(\mathbf{e}_n^{'}T_{2k-1}^{(m)}(\mathbf{y})) , \tag{3.14}$$

where $\mathbf{Q}_n^{(m)} = \mathbf{J}_m^{'}\mathbf{J}_m$, $\mathbf{1}_n^{'} = (1, \cdots, 1)$, and $\mathbf{e}_1^{'} = (1, 0, \cdots, 0)$ and $\mathbf{e}_n^{'} = (0, \cdots, 0, 1)$ are $n \times 1$ unit vectors.

The operator $T_{2k-1}^{(m)}$ $(k \geq 1)$ is the forward SIML filering with the initial value $\mathbf{y}_0$ at $i = 0$ and $T_{2k}^{(m)}$ $(k \geq 1)$ is the backward filtering with the initial value at $i = n$.

For non-stationary time series, two operators have different roles in smoothing procedure.

Let $\tilde{\mathbf{y}} = (y_0, y_1, \cdots, y_{n-1}, y_n)$ be an $(n+1) \times 1$ vector, and $n \times (n+1)$ choice matrices

$\mathbf{J}_f = (\mathbf{0}, \mathbf{I}_n)$ and $\mathbf{J}_b = (\mathbf{I}_n, \mathbf{0})$. Then, we repeat the forward smoothing procedure such that for $\mathbf{y} = \mathbf{J}_f \tilde{\mathbf{y}}$ and $k \geq 1$,

$$
\begin{aligned}
T^{(m)}_{2k+1}(\mathbf{y}) &= \mathbf{C}_n^{'} \mathbf{P}_n {*}' \mathbf{Q}_n^{(m)} \mathbf{P}_n^{*} \mathbf{C}_n^{'-1} \mathbf{J}_f \tilde{\mathbf{y}} + [\mathbf{I}_n - \mathbf{C}_n \mathbf{P}_n \mathbf{Q}_n^{(m)} \mathbf{P}_n \mathbf{C}_n^{-1}] \\
&\quad \times \mathbf{1}_n \mathbf{e}_1^{'} \left[\mathbf{C}_n^{'} \mathbf{P}_n^{*}{}' \mathbf{Q}_n^{(m)} \mathbf{P}_n^{*} \mathbf{C}^{'-1} \left(\mathbf{J}_b \tilde{\mathbf{y}} - \mathbf{1}_n (\mathbf{e}_n^{'} T^{(m)}_{2k-1}(\mathbf{y})) \right) + \mathbf{1}_n (\mathbf{e}_n^{'} T^{(m)}_{2k-1}(\mathbf{y})) \right] .
\end{aligned}
$$

We also repeat the backward smoothing procedure such that for $\mathbf{y} = \mathbf{J}_b \tilde{\mathbf{y}}$ and $k \geq 1$,

$$
\begin{aligned}
T^{(m)}_{2(k+1)}(\mathbf{y}) &= \mathbf{C}_n \mathbf{P}_n \mathbf{Q}_n^{(m)} \mathbf{P}_n \mathbf{C}_n^{-1} \mathbf{J}_b \tilde{\mathbf{y}} + [\mathbf{I}_n - \mathbf{C}_n^{'} \mathbf{P}_n^{*'} \mathbf{Q}_n^{(m)} \mathbf{P}_n^{*} \mathbf{C}_n^{-1'}] \\
&\quad \times \mathbf{1}_n \mathbf{e}_n^{'} \left[\mathbf{C}_n \mathbf{P}_n \mathbf{Q}_n^{(m)} \mathbf{P}_n \mathbf{C}^{-1} \left(\mathbf{J}_f \tilde{\mathbf{y}} - \mathbf{1}_n (\mathbf{e}_1^{'} T^{(m)}_{2k}(\mathbf{y})) \right) + \mathbf{1}_n (\mathbf{e}_1^{'} T^{(m)}_{2k}(\mathbf{y})) \right] .
\end{aligned}
$$

Then, we have the next proposition on the convergence of the smoothing procedure and the proof is given in the Appendix.

Theorem 3.1 : As $k \to \infty$, for any even number m for $1 < m < n$, we have

$$
T^{(m)}_{2k+1} \to T^{(m)}_{1*} = \sum_{s=0}^{\infty} (\mathbf{A}_2^{(m)})^s \mathbf{A}_1^{(m)} , \tag{3.15}
$$

and

$$
T^{(m)}_{2k} \to T^{(m)}_{2*} = \sum_{s=0}^{\infty} (\mathbf{A}_{2*}^{(m)})^s \mathbf{A}_{1*}^{(m)} , \tag{3.16}
$$

where

$$
\begin{aligned}
\mathbf{A}_1^{(m)} &= \mathbf{C}_n \mathbf{P}_n \mathbf{Q}_n^{(m)} \mathbf{P}_n \mathbf{C}_n^{-1} \mathbf{J}_f + [\mathbf{I}_n - \mathbf{C}_n \mathbf{P}_n \mathbf{Q}_n^{(m)} \mathbf{P}_n \mathbf{C}_n^{-1}] \mathbf{1}_n \mathbf{e}_1^{'} \mathbf{C}_n^{'} \mathbf{P}_n^{*'} \mathbf{Q}_n^{(m)} \mathbf{P}_n^{*} \mathbf{C}_n^{'-1} \mathbf{J}_b, \\
\mathbf{A}_2^{(m)} &= [\mathbf{I}_n - \mathbf{C}_n \mathbf{P}_n \mathbf{Q}_n^{(m)} \mathbf{P}_n \mathbf{C}_n^{-1}] \mathbf{1}_n \times [1 - \mathbf{e}_1^{'} \mathbf{C}_n^{'} \mathbf{P}_n^{*'} \mathbf{Q}_n^{(m)} \mathbf{P}_n^{*} \mathbf{C}_n^{'-1} \mathbf{1}_n] \mathbf{e}_n^{'} , \\
\mathbf{A}_{1*}^{(m)} &= \mathbf{C}_n^{'} \mathbf{P}_n^{*'} \mathbf{Q}_n^{(m)} \mathbf{P}_n^{*} \mathbf{C}_n^{'-1} \mathbf{J}_b + [\mathbf{I}_n - \mathbf{C}_n^{'} \mathbf{P}_n^{*'} \mathbf{Q}_n^{(m)} \mathbf{P}_n^{*} \mathbf{C}_n^{'-1}] \mathbf{1}_n \mathbf{e}_n^{'} \mathbf{C}_n \mathbf{P}_n \mathbf{Q}_n^{(m)} \mathbf{P}_n \mathbf{C}_n^{-1} \mathbf{J}_f, \\
\mathbf{A}_{2*}^{(m)} &= [\mathbf{I}_n - \mathbf{C}_n^{'} \mathbf{P}_n^{*'} \mathbf{Q}_n^{(m)} \mathbf{P}_n^{*} \mathbf{C}_n^{'-1}] \mathbf{1}_n \times [1 - \mathbf{e}_n^{'} \mathbf{C}_n \mathbf{P}_n \mathbf{Q}_n^{(m)} \mathbf{P}_n \mathbf{C}_n^{-1} \mathbf{1}_n] \mathbf{e}_n^{'} .
\end{aligned}
$$

The absolute values of all eigenvalues of $\mathbf{A}_2^{(m)}$ and $\mathbf{A}_{2*}^{(m)}$ are less than one. Then, we can expresss

$$
\sum_{s=0}^{\infty} (\mathbf{A}_2^{(m)})^s = (\mathbf{I}_n - \mathbf{A}_2^{(m)})^{-1} , \ \sum_{s=0}^{\infty} (\mathbf{A}_{2*}^{(m)})^s = (\mathbf{I}_n - \mathbf{A}_{2*}^{(m)})^{-1} .
$$

Given that the initial value is the starting point of nonstationary time series, we need to develop a smoothing procedure that does not depend on the initial value. Practically, often we do want to use the procedure that does not depend on the first or latest observation $\mathbf{y}_0$ or $\mathbf{y}_n$. In these cases, it may be reasonable to use the $T^{(m)}_{2*}$ or $T^{(m)}_{1*}$, respectively.

It may be interesting to find the difference between the forward smoothing and backward smoothing. Let the two operators be $\mathbf{H}_n = (h_{ab}^{(n)}) = \mathbf{C}_n \mathbf{P}_n \mathbf{Q}_n^{(m)} \mathbf{P}_n \mathbf{C}_n^{-1}$ and $\mathbf{F}_n = (f_{ab}^{(n)}) = \mathbf{C}_n^{'} \mathbf{P}_n^{*'} \mathbf{Q}_n^{(m)} \mathbf{P}_n^{*} \mathbf{C}_n^{'-1}$. Then, each term of $h_{ab}^{(n)}$ and $f_{ab}^{(n)}$ are complicated sums of trigonometric functions in the forward and backward SIML smoothing. However, they are similar as we summarize in the next result. The proof is given in the Appendix.

Theorem 3.2 : For any $\delta > 0$, we take $m = [m_n]$ such that $0 < m_n < m_n^{1+\delta} < n$. Then, as $m_n/n \to 0$ and $n \to \infty$,

$$(\frac{n}{m_n^{1+\delta}}) \mathrm{Tr}[\mathbf{H}_n - \mathbf{F}_n] \to 0 \ , \tag{3.17}$$

where the trace of an $n \times n$ matrix $\mathbf{A} = (a_{ij})$ is defined by $\mathrm{Tr}(\mathbf{A}) = \sum_{i=1}^{n} a_{ii}$.

The total norm of two operators is $O(m_n^{1+\delta}/n)$, which is small in typical applications. From this result, the backward SIML smoothing is essentially similar to the forward smoothing when n is large. As we shall discuss in Section 4.2, it is a real (finite- and discrete) Fourier transformation if we take that the time is reversed from n to 0, rather than from 0 to n.

3.4 Band Smoothing

We consider a general filtering based on $\mathbf{K}_n$ and $\mathbf{K}_n^*$ transformations and use the inversion of some frequency parts of the random matrices $\mathbf{Z}_n$ and $\mathbf{Z}_n^*$. The leading example is the seasonal frequency in the discrete time series, and we take s (> 1) as a positive integer.
Let an $m_2 \times [m_1 + m_2 + (n - m_1 - m_2)]$ choice matrix be $\mathbf{J}_{m_1,m_2} = (\mathbf{O}, \mathbf{I}_{m_2}, \mathbf{O})$ (we take $m_1 + m_2 < n$), and let also $n \times p$ matrices be

$$\hat{\mathbf{X}}_n(m_1, m_2) = \mathbf{C}_n \mathbf{P}_n \mathbf{J}_{m_1,m_2}^{'} \mathbf{J}_{m_1,m_2} \mathbf{P}_n \mathbf{C}_n^{-1} (\mathbf{Y}_n - \bar{\mathbf{Y}}_0) \tag{3.18}$$

and

$$\hat{\mathbf{X}}_n^*(m_1, m_2) = \mathbf{C}_n^{'} \mathbf{P}_n^{*'} \mathbf{J}_{m_1,m_2}^{'} \mathbf{J}_{m_1,m_2} \mathbf{P}_n^{*} \mathbf{C}_n^{'-1} (\mathbf{Y}_n - \bar{\mathbf{Y}}_n) \tag{3.19}$$

and an $n \times n$ matrix $\mathbf{Q}_n = \mathbf{Q}_n^{(m_1,m_2)} = \mathbf{J}_{m_1,m_2}^{'} \mathbf{J}_{m_1,m_2}$.

As an example in economic data, when we have the seasonal frequency s (> 1), we can take $m_1 = [2n/s] - [m/2]$ and $m_2 = m$. For instance, we take $s = 4$ for the quarterly data and $s = 12$ for the monthly data. (See Kunitomo and Sato (2021) for details.) In the same way to the trend smoothing problem, the SIML forward-filtering is given by

$$\mathbf{X}_n(m_1, m_2) = \mathbf{C}_n \mathbf{P}_n \mathbf{Q}_n^{(m_1,m_2)} \mathbf{P}_n \mathbf{C}_n^{-1} \mathbf{X}_n^{(0)} \ , \tag{3.20}$$

and the SIML backward-filtering is given by

$$\mathbf{X}_n^*(m_1,m_2)=\mathbf{C}_n^{'}\mathbf{P}_n^{*'}\mathbf{Q}_n^{(m_1,m_2)}\mathbf{P}_n^*\mathbf{C}_n^{'-1}\mathbf{X}_n^*\ , \tag{3.21}$$

respectively, which are based on the estimated frequency components of $\mathbf{x}_i^{(0)}$ $(i=1,\cdots,n)$ or $\mathbf{x}_i^{*(n)}$ $(i=0,\cdots,n-1)$.
In this case, we can define two operators $T_{2k-1}^{(m_1,m_2)}$ and $T_{2k}^{(m_1,m_2)}$ for $k=1,\cdots M$ as (3.13) and (3.14) by using $\mathbf{J}_{m_1,m_2}$, rather than $\mathbf{J}_m$. Then, it is straightforward to to find the next proposition on the convergence of smoothing procedure, and the proof is in the Appendix.

Theorem 3.3 : Let m_1 and m_2 are even numbers such that $1<m_1<m_1+m_2<n$. As $k\to\infty$, we have

$$T_{2k+1}^{(m_1,m_2)}\to T_{1*}^{(m_1,m_2)}=\sum_{s=0}^{\infty}(\mathbf{A}_2^{(m_1,m_2)})^s\mathbf{A}_1^{(m_1,m_2)}\ , \tag{3.22}$$

and

$$T_{2k}^{(m_1,m_2)}\to T_{2*}^{(m_1,m_2)}=\sum_{s=0}^{\infty}(\mathbf{A}_{2*}^{(m_1,m_2)})^s\mathbf{A}_{1*}^{(m_1,m_2)}\ , \tag{3.23}$$

where

$$\begin{aligned}
\mathbf{A}_1^{(m_1,m_2)} &= \mathbf{C}_n\mathbf{P}_n\mathbf{Q}_n^{(m_1,m_2)}\mathbf{P}_n\mathbf{C_n}^{-1}\mathbf{J}_f \\
&+ [\mathbf{I}_n-\mathbf{C}_n\mathbf{P}_n\mathbf{Q}_n^{(m_1,m_2)}\mathbf{P}_n\mathbf{C}_n^{-1}]\mathbf{1}_n\mathbf{e}_1^{'}\mathbf{C}_n^{'}\mathbf{P}_n^*\mathbf{Q}_n^{(m_1,m_2)}\mathbf{P}_n^*\mathbf{C}_n^{'-1}\mathbf{J}_b, \\
\mathbf{A}_2^{(m_1,m_2)} &= [\mathbf{I}_n-\mathbf{C}_n\mathbf{P}_n\mathbf{Q}_n^{(m_1,m_2)}\mathbf{P}_n\mathbf{C}_n^{-1}]\mathbf{1}_n\times[1-\mathbf{e}_1^{'}\mathbf{C}_n^{'}\mathbf{P}_n^{*'}\mathbf{Q}_n^{(m_1,m_2)}\mathbf{P}_n^*\mathbf{C}_n^{'-1}\mathbf{1}_n]\mathbf{e}_n^{'}\ ,
\end{aligned}$$

$$\begin{aligned}
\mathbf{A}_{1*}^{(m_1,m_2)} &= \mathbf{C}_n^{'}\mathbf{P}_n^{*'}\mathbf{Q}_n^{(m_1,m_2)}\mathbf{P}_n^*\mathbf{C}_n^{'-1}\mathbf{J}_b \\
&+ [\mathbf{I}_n-\mathbf{C}_n^{'}\mathbf{P}_n^{*'}\mathbf{Q}_n^{(m_1,m_2)}\mathbf{P}_n^*\mathbf{C}_n^{'-1}]\mathbf{1}_n\mathbf{e}_n^{'}\mathbf{C}_n\mathbf{P}_n\mathbf{Q}_n^{(m_1,m_2)}\mathbf{P}_n\mathbf{C}_n^{-1}\mathbf{J}_f \\
\mathbf{A}_{2*}^{(m_1,m_2)} &= [\mathbf{I}_n-\mathbf{C}_n^{'}\mathbf{P}_n^{*'}\mathbf{Q}_n^{(m_1,m_2)}\mathbf{P}_n^*\mathbf{C}_n^{'-1}]\mathbf{1}_n\times[1-\mathbf{e}_n^{'}\mathbf{C}_n\mathbf{P}_n\mathbf{Q}_n^{(m_1,m_2)}\mathbf{P}_n\mathbf{C}_n^{-1}\mathbf{1}_n]\mathbf{e}_n^{'}
\end{aligned}$$

and $\mathbf{J}_f$ and $\mathbf{J}_b$ are used as Theorem 3.1.

The absolute values of all eigenvalues of $\mathbf{A}_2^{(m_1,m_2)}$ and $\mathbf{A}_{2*}^{(m_1,m_2)}$ are less than one; then, we can express

$$\sum_{s=0}^{\infty}(\mathbf{A}_2^{(m_1,m_2)})^s=(\mathbf{I}_n-\mathbf{A}_2^{(m_1,m_2)})^{-1}\ ,\ \sum_{s=0}^{\infty}(\mathbf{A}_{2*}^{(m_1,m_2)})^s=(\mathbf{I}_n-\mathbf{A}_{2*}^{(m_1,m_2)})^{-1}\ .$$

This result is useful for handling seasonality of economic time series, as an example. Theorem 3.1 can be regarded as a special case of Theorem 3.3 when $m_1=0$ and $m_2=m$. See Chapter 5 of Kunitomo and Sato (2024) for complicated use SIML filters such as the seasonal adjustment.

3.5 Multi-step Smoothing

In the forward and backward smoothing procedures, choosing an appropriate m is important. However, this problem may become difficult when seasonal components exist. Then, it may be normal to repeat smoothing because the forward and backward smoothing several times, which may be called *multi-stage smoothing*, can be run.

Let $T_{2k-1}^{(m)}$ be the first stage forward smoothing with a specific choice of m for $k = 1, \cdots, [n/2]$. Then, we define the double-stage forward smoothing by

$$T_{2k-1}^{(m,m_1,m_2)} = T_{2k-1}^{(m_1,m_2)} T_{2k-1}^{(m)} . \tag{3.24}$$

Similarly, we define the double-stage backward smoothing by

$$T_{2k}^{(m,m_1,m_2)} = T_{2k}^{(m_1,m_2)} T_{2k}^{(m)} . \tag{3.25}$$

More complicated smoothing procedures can also exist. Hence, we need some criterion to find an appropriate smoothing procedure for applications. Consequently, handling complicated seasonal patterns in the frequency domain is possible, for instance.

In real applications, finding an appropriate m or m_1 and m_2 at the beginning might not be certain. The meaning of frequencies may leads to a guide on choosing frequencies. In addition, at the first stage, one strategy in the trend estimation would be to choose a relatively large m_1, which should be less than the seasonality frequency. Then, at the second stage, we choose m_2, which is smaller than m_1 and use the following evaluation criterion.

3.6 Prediction Errors and Evaluation Criteria

The problem of choosing an appropriate filtering, including the choice of m (or m_1 and m_2 in a more general case) in smoothing, is an important question for applications. Given that our procedure does not assume a particular distribution such as Gaussianity and semi-parametric, it is a challenging problem. As we shall discuss in the next section, our filtering method is based on the frequency interpretation.

In the present non-parametric setting, we consider the prediction error based on the orthogonal process in the frequency domain and the resulting predictive criterion. Let $\mathbf{r}_j^{(n)} = \mathbf{y}_j^{(n)} - \mathbf{y}_{j-1}^{(n)}$ $(j = 1, \cdots, n)$; hence we have the expression

$$\mathbf{r}_j^{(n)} = \sum_{k=1}^{n} p_{jk} \mathbf{z}_k , \tag{3.26}$$

where $\mathbf{z}_k$ is the orthogonal process at the frequency $\lambda_k^{(n)} = (k-1/2)/(2n+1)$ $(k = 1, \cdots, n)$.
Then, for $1 \le j \le n$, it may be natural to use the information of $m-$low frequencies and construct the estimator

$$\hat{\mathbf{r}}_j^{(n)}(m) = \sum_{k=1}^{m} p_{j,k}\mathbf{z}_k \ , \tag{3.27}$$

which is a linear combination of m orthogonal processes with different frequencies. Then, for $h \ge 1$, it may be reasonable to use the linear predictor

$$\hat{\mathbf{x}}_{n+h}^{(n)}(m) = \sum_{s=h+1}^{n+h} \hat{\mathbf{r}}_s^{(n)}(m) = \sum_{s=h+1}^{n+h}\sum_{k=1}^{m} p_{sk}\mathbf{z}_k \ . \tag{3.28}$$

Since we have ignored the information of the remaining (higher) frequencies of the process $(k = m+1, \cdots, n)$, by using (3.1) and (3.2), the prediction error can be written as

$$\hat{\mathbf{x}}_{n+h}^{(n)}(m) - \mathbf{x}_{n+h}^{(n)} = \sum_{k=1}^{m}\sum_{s=h+1}^{n+h}\sum_{j=1}^{n} p_{sj}(\mathbf{C}_n^{-1}\mathbf{V}_n)_{kj} \ . + \sum_{k=m+1}^{n}\sum_{s=h+1}^{n+h}\sum_{j=1}^{n} p_{sj}(\mathbf{C}_n^{-1}\mathbf{X}_n)_{kj} \ ,$$

which can be simplified by an elementary evaluation such that

$$\begin{aligned}\sum_{s=h+1}^{n+h} p_{sk} &= \frac{1}{\sqrt{2n+1}}\sum_{s=h+1}^{n+h}[e^{i\frac{2\pi}{2n+1}(n+h)(k-\frac{1}{2})(s-\frac{1}{2})} + e^{-i\frac{2\pi}{2n+1}(n+h)(k-\frac{1}{2})(s-\frac{1}{2})}] \\ &= \frac{1}{\sqrt{2n+1}}\frac{\sin\frac{2\pi}{2n+1}(n+h)(k-\frac{1}{2}) - \sin\frac{2\pi}{2n+1}h(k-\frac{1}{2})}{\sin\frac{2\pi}{2n+1}\frac{1}{2}(k-\frac{1}{2})} \ .\end{aligned}$$

For an illustration, we consider the case when the noise terms of $\mathbf{v}_i$ and $\mathbf{v}_i^{(x)}$ $(i = 1, \cdots, n)$ are sequences of i.i.d. random variables and $p = 1$. (We take the variances as σ_v^2 and σ_x^2, respectively.) By using a_{kn}^* $(k = 1, \cdots, m)$ in (2.8), we can derive the prediction MSE as

$$\begin{aligned}\mathrm{MSE}_m &= \frac{4\sigma_v^2}{2n+1}\sum_{k=1}^{m}[\sin\frac{2\pi}{2n+1}(n+h)(k-\frac{1}{2}) - \sin\frac{2\pi}{2n+1}h(k-\frac{1}{2})]^2 \\ &\quad + \frac{\sigma_x^2}{2n+1}\sum_{k=m+1}^{n}[\frac{\sin\frac{2\pi}{2n+1}(n+h)(k-\frac{1}{2}) - \sin\frac{2\pi}{2n+1}h(k-\frac{1}{2})}{\sin\frac{2\pi}{2n+1}\frac{1}{2}(k-\frac{1}{2})}]^2 \ .\end{aligned} \tag{3.29}$$

As a typical example, we set $\sigma_v^2 = 2, \sigma_x^2 = 1, h = 4, n = 100$. The minimum value of MSE is attained when $m^* = 23$.

We have the trade-off of two terms in the predictive MSE. We notice that the first term is an increasing function of m, while the second term is a decreasing function of

m. A point of m^* exists such that $MSE(m)$ is minimized. Several criteria based on the prediction MSE could be developed by using (3.29). Because the prediction MSE depends on the unknown parameters of σ_x^2 and σ_v^2 even when $p = 1$ and the i.i.d. case, we must replace them in a simple manner. When $p \geq 1$, it would be possible to extend the above argument. Form the above consideration on the predictive MSE of state variables in this section, some method based on frequency domain in Section 4 may be practically important.

4. Frequency Domain Interpretation

4.1 On Spectral Representation and Likelihood

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 $\mathbf{v}_i$ $(i = 1, \cdots, n)$. Then

$$\mathbf{f}_v(\lambda) = (\sum_{j=-\infty}^{\infty} \mathbf{C}_j^{(v)} e^{2\pi i \lambda j}) \mathbf{\Sigma}_e^{(v)} (\sum_{j=-\infty}^{\infty} \mathbf{C}_j^{(v)'} e^{-2\pi i \lambda j}) \quad (-\frac{1}{2} \leq \lambda \leq \frac{1}{2}) \tag{4.1}$$

and

$$\mathbf{f}_{\Delta x}(\lambda) = (\sum_{j=-\infty}^{\infty} \mathbf{C}_j^{(x)} e^{2\pi i \lambda j}) \mathbf{\Sigma}_e^{(x)} (\sum_{j=-\infty}^{\infty} \mathbf{C}_j^{(x)'} e^{-2\pi i \lambda j}) \quad (-\frac{1}{2} \leq \lambda \leq \frac{1}{2}) \ , \tag{4.2}$$

where we set $\mathbf{C}_0^{(v)} = \mathbf{C}_0^{(x)} = \mathbf{I}_p$ as normalizations and $i^2 = -1$.
Then, the relation between the $p \times p$ spectral density matrix of the transformed vector process, which are observable, and the spectral density of the observed difference series $\Delta \mathbf{y}_i$ $(= \mathbf{y}_i - \mathbf{y}_{i-1})$ can be represented as

$$\mathbf{f}_{\Delta y}(\lambda) = \mathbf{f}_{\Delta x}(\lambda) + (1 - e^{2\pi i \lambda}) f_v(\lambda)(1 - e^{-2\pi i \lambda}) \ . \tag{4.3}$$

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

$$\mathbf{\Sigma}_x = \mathbf{f}_{\Delta x}(0) \ (= (\sigma_{gh}^{(x)})) \ , \ \mathbf{\Sigma}_v = f_v(0) \ = (\sigma_{gh}^{(v)}) \ . \tag{4.4}$$

Let $f_v^{(SR)}(\lambda_k)$, $f_s^{(SR)}(\lambda_k)$, and $f_{\Delta x}^{(SR)}(\lambda_k)$ be the symmetrized $p \times p$ spectral matrices of $\mathbf{v}_i$, $\mathbf{s}_i$ and $\Delta \mathbf{x}_i$ at λ_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 $f_{\Delta x}^{(SR)}(\lambda_k) = (1/2)[f_{\Delta x}^{(SR)}(\lambda_k) + \bar{f}_{\Delta x}^{(SR)}(\lambda_k)]$. Then, we find the relation $(1 - e^{2\pi i \lambda_k})(1 - e^{-2\pi i \lambda_k}) = 2[1 - \cos 2\pi \lambda_k] = 4 \sin^2 [\pi \frac{k - \frac{1}{2}}{2n+1}]$.
Since the orthogonal processes are approximately distributed as the Gaussian distribution, given the initial conditions ($\mathbf{y}_0$ in the causal representation or $\mathbf{y}_n$ in the

non-causal representation), (-2) times the conditional log-likelihood function in the general model can be approximated as

$$(4.5) \qquad (-2)l_n(\boldsymbol{\theta}) = \sum_{k=1}^{n} \log |a_{kn}^{*} f_v^{(SR)}(\lambda_k) + f_{\Delta x}^{(SR)}(\lambda_k)| + \sum_{k=1}^{n} \mathbf{z}_k^{'}[a_{kn}^{*} f_v^{(SR)}(\lambda_k) + f_{\Delta x}^{(SR)}(\lambda_k)]^{-1}\mathbf{z}_k \ .$$

When $\mathbf{v}_i^{(x)}$ and $\mathbf{v}_i$ are mutually independent random variables, given the initial conditions ($\mathbf{y}_0$ or $\mathbf{y}_n$), (-2) times the conditional log-likelihood function an be approximated as $(-2)l_n(\boldsymbol{\theta}) = \sum_{k=1}^{n} \log |a_{kn}^{*}\boldsymbol{\Sigma}_v + \boldsymbol{\Sigma}_x| + \sum_{k=1}^{n} \mathbf{z}_k^{'}[a_{kn}^{*}\boldsymbol{\Sigma}_v + \boldsymbol{\Sigma}_x]^{-1}\mathbf{z}_k$ and $\boldsymbol{\theta}$ is a vector of parameters.
As an application, by taking a positive integer m $(= [m_n])$ and $m_n = n^{\alpha}$ $(0 < \alpha < 1)$, Kunitomo and Sato (2021) proposed the SIML estimator of $\hat{\boldsymbol{\Sigma}}_x$ by

$$(4.6) \qquad \hat{\boldsymbol{\Sigma}}_{x,SIML} = \frac{1}{m}\sum_{k=1}^{m} \mathbf{z}_k \mathbf{z}_k^{'} \ .$$

It has consistency when $0 < \alpha < 1$ and the asymptotic normality when $0 < \alpha < .8$ as $n \to \infty$.

In the forward smoothing, we may use the causal MA representation of the stationary process and discussed its interpretation in their Section 5. In the backward smoothing, we need the non-causal MA representation of the stationary process. For causal and non-causal MA models, we refer to Chapter 4 of Brockwell and Davis (1990).

4.2 Frequency Decomposition

At first glance, the SIML smoothing method might appear to be an *ad-hoc* statistical procedure lacking a solid mathematical foundation. However, upon closer examination, we can establish a robust statistical basis for this method. The justification for SIML smoothing deviates from the standard approaches commonly used in traditional time series analysis in the frequency domain. (For related discussions, see Doob (1953), Brockwell and Davis (1990), and their extensions to nonstationary processes, including Brillinger and Hatanaka (1969) and Brillinger (1980).)

For $\lambda_k^{(n)} = (k - 1/2)/(2n+1)$ $(k = 1, \cdots, n)$, from (2.5) and (3.6), we rewrite

$$(4.7) \qquad \mathbf{z}_k(\lambda_k^{(n)}) = \sum_{j=1}^{n} \mathbf{r}_j^{(n)}[\frac{2}{\sqrt{2n+1}} \cos[2\pi\lambda_k^{(n)}(j - \frac{1}{2})]$$

where $\mathbf{r}_j^{(n)} = \mathbf{y}_j^{(n)} - \mathbf{y}_{j-1}^{(n)}$ $(j = 1, \cdots, n)$ with the initial condition $\mathbf{y}_0$, and

$$(4.8) \qquad \mathbf{z}_k^{*}(\lambda_k^{(n)}) = \sum_{j=1}^{n} \mathbf{r}_{j-1}^{(n)*}[\frac{2}{\sqrt{2n+1}} \sin[2\pi\lambda_k^{(n)} j]$$

where $\mathbf{r}_{j-1}^{(n)*} = \mathbf{y}_{j-1}^{(n)} - \mathbf{y}_j^{(n)}$ $(j = 1, \cdots, n)$ with the initial condition $\mathbf{y}_n$.
The situation is that under the assumption of (4.1)-(4.3), both $\mathbf{r}_j^{(n)}$ and $\mathbf{r}_j^{(n)*}$ are stationary processes with the spectral density, which are consistent with (4.7)-(4.8). Then, by using the inversion transformations with $\mathbf{P}_n$ and $\mathbf{P}_n^*$, we find that

$$\mathbf{r}_s^{(n)} = \sum_{k=1}^{n} p_{sk}\mathbf{z}_k(\lambda_k^{(n)}) \ , \ \mathbf{r}_{s-1}^{(n)*} = \sum_{k=1}^{n} p_{sk}^*\mathbf{z}_k^*(\lambda_k^{(n)}) \tag{4.9}$$

They correspond to the representation of $\mathbf{R}_n = (\mathbf{r}_i^{(n)'}) = \mathbf{C}_n^{-1}\hat{\mathbf{X}}_n(\mathrm{Q})$ with $\mathbf{Q}_n = \mathbf{I}_n$, and $\mathbf{R}_n^* = (\mathbf{r}_{i-1}^{*(n)'}) = \mathbf{C}_n^{'-1}\hat{\mathbf{X}}_n^*(\mathrm{Q}^*)$ with $\mathbf{Q}_n^* = \mathbf{I}_n$, respectively. Then, by using $\mathbf{Y}_n = \mathbf{C}_n\mathbf{R}_n$ and $\mathbf{Y}_n^* = \mathbf{C}_n^{'}\mathbf{R}_n^{'}$, we recover the nonstationary processes given the initial condition as $\mathbf{y}_t^{(n)} = \mathbf{y}_0 + \sum_{s=1}^{t}\mathbf{r}_s^{(n)}$ and $\mathbf{y}_t^{(n)} = \mathbf{y}_n + \sum_{s=1}^{n-(t+1)}\mathbf{r}_{n-s}^{*(n)}$.

In the statistical time series analysis, a set of observation is regarded as a realization of stochastic proces with discrete time as $j = 1, \cdots, n$ (the time interval is fixed) while the spectral density matrix is represented in the continuous frequency variable $\lambda \in [0, \frac{1}{2}]$, some care should be necessary to interpret our smoothing (or filtering) method.
Define the coefficients as

$$\alpha_n(\lambda_m^{(n*)}, j - \frac{1}{2}) = \frac{1}{n}\sum_{k=1}^{m}[2\cos 2\pi\lambda_k^{(n)}(j - \frac{1}{2})] \tag{4.10}$$

and

$$\beta_n(\lambda_m^{(n*)}, j) = \frac{1}{n}\sum_{k=1}^{m}[2\sin 2\pi\lambda_k^{(n)}j] \ , \tag{4.11}$$

where $\lambda_k^{(n)} = (k - \frac{1}{2})/(2n+1)$ and $\lambda_m^{(n*)} = m/(2n+1)$ $(k, m = 1, \cdots, n)$. By utilizing some trigonometric relations, we obtain $\alpha_n(\lambda_m^{(n*)}, j) = 2\sin 2\pi[m/(2n+1)](j - \frac{1}{2})/[2\sin\pi\ (j - \frac{1}{2})/(2n+1)]$ and $\beta_n(\lambda_m^{(n*)}, j) = 2[1 - \cos(2\pi mj/(2n+1))]/[2\sin\pi j/(2n+1)]$. When $\lambda_m^{*(n)} \to \lambda$ as $n \to \infty$ $(0 < \lambda < \frac{1}{2})$, we find

$$\alpha_n(\lambda_m^{(n*)}, j - \frac{1}{2}) \to \alpha(\lambda, j) = \frac{2\sin 2\pi\lambda\ (j - \frac{1}{2})]}{\pi\ (j - \frac{1}{2})}$$

and

$$\beta_n(\lambda_m^{(n*)}, j) \to \beta(\lambda, j) = \frac{2[1 - \cos 2\pi\lambda\ j]}{\pi\ j} \ ,$$

respectively.
If we set the stochastic processes of uncorrelated increments with continuous parameter λ $(0 \le \lambda \le \frac{1}{2})$ as $A_n(\lambda) = \sum_{j=1}^{n}\alpha(\lambda, j - \frac{1}{2})\mathbf{r}_j^{(n)}$ and $B_n(\lambda) = \sum_{j=1}^{n}\beta(\lambda, j)\mathbf{r}_{j-1}^{*(n)}$, then $dA_n(\lambda) = 4[\sum_{j=1}^{n}(\cos 2\pi\lambda(j - \frac{1}{2}))\mathbf{r}_j^{(n)}]d\lambda$ and $dB_n(\lambda) = 4[\sum_{j=1}^{n}(\sin 2\pi\lambda j)\mathbf{r}_{j-1}^{*(n)}]d\lambda$. Hence, we find

$$\int_0^{\frac{1}{2}}\cos[2\pi\lambda(s - \frac{1}{2})]dA_n(\lambda) = \mathbf{r}_s^{(n)} \ \ (s = 1, \cdots, n) \ . \tag{4.12}$$

and

$$\int_0^{\frac{1}{2}} \sin[2\pi\lambda s]dB_n(\lambda) = \mathbf{r}_{s-1}^{*(n)} \quad (s = 1, \cdots, n) \ , \tag{4.13}$$

respectively.
They correspond to the continuous representation of a discrete (real-valued) stationary time series in the frequency domain (refer to Chapter 7.4 of Anderson (1971)). If we write the limits of $\mathbf{A} = \lim_{n\to\infty} \mathbf{A}_n(\lambda)$ and $\mathbf{B} = \lim_{n\to\infty} \mathbf{B}_n(\lambda)$ (assuming that they exist), then the (real-valued) spectral distribution matrix F_{RS} for any $0 \leq \lambda_1 < \lambda_2 \leq 1/2$ can be defined as

$$\begin{aligned} F_{RS}(\lambda_2 - \lambda_1) &= \mathbf{E}[(\mathbf{A}(\lambda_2 - \lambda_1)\mathbf{A}(\lambda_2 - \lambda_1)^{'}] = \mathbf{E}[(\mathbf{B}(\lambda_2 - \lambda_1)\mathbf{B}(\lambda_2 - \lambda_1)^{'}] \\ &= \int_{\lambda_1}^{\lambda_2} f_{RS}(\lambda)d\lambda \end{aligned}$$

if F_{RS} is absolutely continuous and the matrix-valued density process $f_{RS}(\lambda)$ $(0 \leq \lambda_1 < \lambda_2 \leq 1/2)$ exists.

We set $\hat{\mathbf{R}}_n(m) = (\hat{\mathbf{r}}_i^{(m,n)'}) = \mathbf{C}_n^{-1}\hat{\mathbf{X}}_n(m)$ and $\hat{\mathbf{r}}_i^{(m,n)}$ are $p \times 1$ vectors for $i = 1, \cdots, n$ in the forward filter. Also we set $\hat{\mathbf{R}}_n^*(m) = (\hat{\mathbf{r}}_i^{*(m,n)'}) = \mathbf{C}_n^{'-1}\hat{\mathbf{X}}_n^*(m)$ and $\hat{\mathbf{r}}_i^{*(m,n)}$ are $p \times 1$ vectors for $i = 1, \cdots, n$ in the backward filter. Because $\hat{\mathbf{R}}_n(m) = \mathbf{C}_n^{-1}\mathbf{P}_n\mathbf{Q}_m^{(n)}\mathbf{Z}_n$ and $\hat{\mathbf{R}}_n^*(m) = \mathbf{C}_n^{'-1}\mathbf{P}_n^*\mathbf{Q}_m^{(n)}\mathbf{Z}_n^*$, we may write

$$\hat{\mathbf{r}}_s^{(m,n)} = \sum_{k=1}^{m} p_{sk}\mathbf{z}_k(\lambda_k^{(n)}) \quad (s = 1, \cdots, m; 0 < m < n), \tag{4.14}$$

and

$$\hat{\mathbf{r}}_{s-1}^{*(m,n)} = \sum_{k=1}^{m} p_{sk}^*\mathbf{z}_k^*(\lambda_k^{(n)}) \quad (s = 1, \cdots, m; 0 < m < n). \tag{4.15}$$

For the trend-cycle smothing with $[1, m]$ $(1 < m \leq n)$ and $\lambda_m^{(n*)} = m/(2n+1)$ $(\sim (1/2)(m/n)$ in the low frequency, which corresponds to the maximum frequency of the trend-cycle part of the time series. Then (4.14) and (4.15) correspond to

$$\mathbf{r}_s^{(n)}(0, \lambda_m^{(n*)}) = \int_0^{\lambda_m^{(n*)}} \cos[2\pi\lambda\ (s - \frac{1}{2})]dA_n(\lambda) \tag{4.16}$$

and

$$\mathbf{r}_{s-1}^{(n)}(0, \lambda_m^{(n*)}) = \int_0^{\lambda_m^{(n*)}} \sin[2\pi\lambda\ s]dB_n(\lambda) \ , \tag{4.17}$$

respectively, where $A_n(\lambda) = \sum_{j=1}^n \alpha(\lambda, j - \frac{1}{2})\mathbf{r}_j^{(n)}$ and $B_n(\lambda) = \sum_{j=1}^n \beta(\lambda, j)\mathbf{r}_{j-1}^{*(n)}$. Thus, we have an analogous interpretation of each term as the trend-cycle filtering representation of orthogonal processes.

Similarly, for the band smoothing with $[m_1+1, m_1+m_2]$ $(1 < m_1 < m_1+m_2 < n)$, we set

$$\hat{\mathbf{r}}_s^{(m_1,m_2,n)} = \sum_{k=m_1+1}^{m_1+m_2} p_{sk}\mathbf{z}_k^{(n)}(\lambda_k^{(n)}) = \hat{\mathbf{r}}_s^{(m_2,n)} - \hat{\mathbf{r}}_{s-1}^{(m_1,n)}$$

and

$$\hat{\mathbf{r}}_{s-1}^{*(m_1,m_2,n)} = \sum_{k=m_1+1}^{m_1+m_2} p_{sk}\mathbf{z}_k^{*(n)}(\lambda_k^{(n)}) = \hat{\mathbf{r}}_{s-1}^{*(m_2,n)} - \hat{\mathbf{r}}_{s-1}^{*(m_1,n)}$$

$(s = 1, \cdots, m; 0 < m_1 < m_2 < n)$. Then, we have an analogous interpretation of each term as the band filtering representation of orthogonal processes. They correspond to

$$\mathbf{r}_s^{(n)}(\lambda_{m_1}^{(n*)}, \lambda_{m_2}^{(n*)}) = \int_{\lambda_{m_1}^{(n*)}}^{\lambda_{m_2}^{(n*)}} \cos[2\pi\lambda\,(s-\frac{1}{2})]dA_n(\lambda) \tag{4.18}$$

and

$$\mathbf{r}_{s-1}^{*(n)}(\lambda_{m_1}^{(n*)}, \lambda_{m_2}^{(n*)}) = \int_{\lambda_{m_1}^{(n*)}}^{\lambda_{m_2}^{(n*)}} \sin[2\pi\lambda\, s]dB_n(\lambda)\ . \tag{4.19}$$

In this way, it is possible to interpret the smoothing method from the frequency domain of the multivariate time series.

5. A Numerical Example

We applied the SIML-forward and SIML-backward smoothing methods to the monthly U.S. Manufacturers' new orders data from 2010 to 2020 to illustrate their effectiveness. This dataset is well-suited for demonstrating our methodology due to its nonstationary nature, combining trends, strong seasonal fluctuations, noise, and occasional abrupt changes, including potential outliers. These characteristics make it critical to investigate the effects of initial conditions on the smoothing results, particularly when dealing with economic time series. Figures 1 and 2 display the original data, serving as a reference for the smoothing outcomes.

In the forward smoothing results, the red curve in Figure 1 represents the smoothed series using the first observation as the initial condition ($m = 5$). The green curve shows T_1^*, the limit obtained through forward-backward iterations, while the violet curve corresponds to a two-step forward filtering process where the first smoothing used $m_1 = 15$ and the second smoothing used $m = 5$. Similarly, in the backward smoothing results shown in Figure 2, the blue curve represents the smoothed series using the last observation as the initial condition $m = 5$), the sky-blue curve shows T_2^*, the limit obtained through backward-backward iterations, and the violet curve corresponds to a two-step backward filtering process with $m_1 = 15$ for the first smoothing and $m_2 = 5$ for the second smoothing.

The analysis reveals that the initial conditions for both forward and backward smoothing have a significant impact on the smoothed time series near the starting

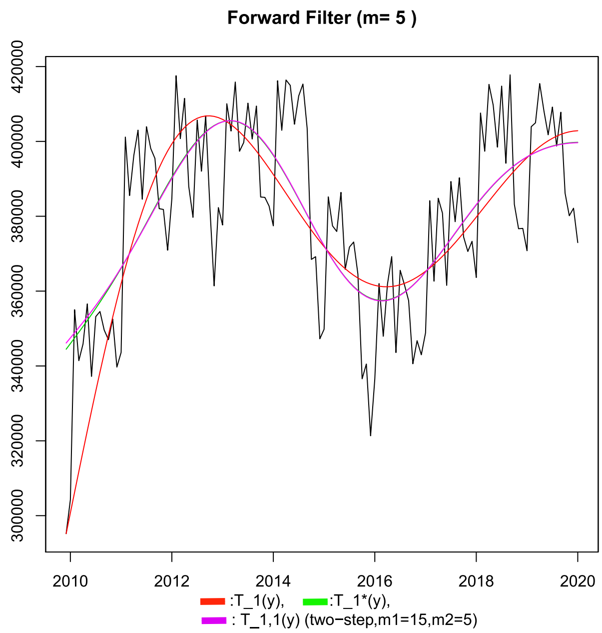

Figure 1: Forward filtering results for monthly US Manufacturers' New Orders from 2010 to 2020. (https://www.census.gov/manufacturing/m3/index.html)

points. However, as iterations are repeated or multi-step smoothing is applied, the influence of the initial values diminishes. After a few steps, the differences between the forward and backward smoothing results become negligible for practical purposes. This observation aligns with the theoretical findings discussed in Section 3 regarding the impact of initial conditions on SIML smoothing.

The U.S. Manufacturers' new orders data provided an extreme case where the effects of initial conditions were particularly pronounced. However, in many empirical applications, the forward and backward smoothing procedures yield similar results without requiring additional iterations. This consistency underscores the robustness of the SIML smoothing methods for handling complex time series with trend, cycle, seasonal component, and noise.

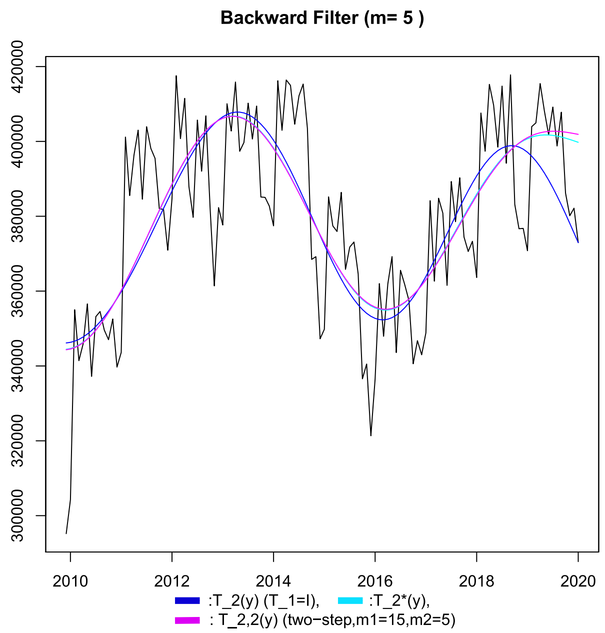

Figure 2: Backward filtering results for monthly US Manufacturers' New Orders from 2010 to 2020. (https://www.census.gov/manufacturing/m3/index.html)

6 Detecting Recent Hidden Change Point

6.1 A Simple Case

We consider the problem of detecting recent breaks in time series. In the standard time analysis, it is often difficult to distinguish the structural breaks and temporal irregular noise components based on the observed noisy non-stationary time series. We propose to use the state estimation and filtering based on the SIML method. For the simplicity, we consider the case when $p = 1$ and the basic non-stationary errors-in-variables model in this subsection.

For an $n \times 1$ vector $\mathbf{a}_n(1) = \mathbf{e}_n - \mathbf{e}_{n-1}$, we denote the difference of the state variable followed by I(1) process, $\mathbf{a}_n(1)'\mathbf{X}_n = x_n - x_{n-1}$, where $n \times 1$ state vector is given as $\mathbf{X}_n = (x_i), i = 1, \cdots, n$.

To estimate the recent change of the true state variable, it is useful to investigate whether the estimated difference is larger than the usual quantity. For this purpose, first we use the estimator based on the forward filtering of $\hat{\mathbf{X}}_n$ as

$$\hat{\Delta}^{(f)} x_{n-h} = \mathbf{a}_{n-h}(1)' \hat{\mathbf{X}}^{(f)}_{n-h} = \hat{x}^{(f)}_{n-h} - \hat{x}^{(f)}_{n-h-1} , \tag{6.1}$$

where h is a fixed nonnegative integer.
By using $\mathbf{a}_n^{'}(1)\mathbf{C}_n = \mathbf{e}_n^{'}$, we investigate the asymptotic properties of the statistic $\mathbf{\Delta}^{(f)} x_n$, it is asymptotically normal. We summarize the result.

Proposition 6.1 : In the basic non-stationary errors-in variables model (p=1) where $\Delta x_i = x_i - x_{i-1} = v_i^{(x)}$ and v_i $(i = 1, \cdots, n$ are i.i.d. random variables with $\mathbf{E}[v_i^{(x)}] = \mathbf{E}[v_i]0$, $\mathbf{E}[(v_i^{(x)})^2] = \sigma_x^2$ and $\mathbf{E}[v_i^2] = \sigma^2$. As $m, n \to \infty$ and $m/n \to 0$

$$W_{f,n-h} = \sqrt{[\frac{3}{2\pi^2(h+1)^2}(\frac{n^3}{m^3})\hat{\Delta}^{(f)}x_{n-h}} \xrightarrow{d} N(0, f_{\Delta y}(0)) \ , \tag{6.2}$$

where h is a (fixed) non-negative integer and $f_{\Delta y}(0)$ is the spectral density of Δy_n at $\lambda = 0$.

Given the initial condition y_0, it is possible to estimate the value of the spectral density at $\lambda = 0$ by using the data Δy_i $(i = 1, \cdots, n)$.

Then, the t-type statistic can be defined by

$$T_n = \frac{W_{f,n}}{\sqrt{\hat{f}_{\Delta y}(0)}} \ . \tag{6.3}$$

When n is large, the distribution of this statistic follows approximately as $N(0, 1)$. It is important notice that we do not assume a particular underlying distribution and it is a non-parametric method of detecting recent structural changes.

WhenΔx_n is a sequence of i.i.d. random variables, we have $\mathbf{Var}[\Delta y] = \sigma_x^2 + 2\sigma_v^2$. Then, the spectral density at zero frequency becomes the volatility of $\mathbf{y}$ as $f_{\Delta y}(0) = \sigma_{\Delta x}^2$.

In many macroeconomic variables, we occasionally observe jumps in the data. However, it is often challenging to determine whether these jumps are merely irregular temporal noise or the onset of a significant break or shift in the trend-cycle component of the time series. If the observed jump represents a change point marking the beginning of a new trend, the series will not revert to its original trend-cycle components. In such cases, the following result is useful for detecting breaks in time series data. Then, the next result is useful to detect breaks of time series.

Proposition 6.2 : In the basic non-stationary errors-in variables model, as $m, n \to \infty$ and $m/n \to 0$,

$$\frac{\sum_{k=0}^{h} W_{f,n-k}^2}{c(h)\hat{f}_{\Delta y}(0)} \xrightarrow{d} \chi^2(1) \tag{6.4}$$

where h is a non-gegative (fixed) integer, $c(h) = \sum_{k=0}^{h} k^2$, and $\hat{f}_{\Delta y}(0)$ is a consistent estimate of the spectral density of Δy_n at zero frequency.

If we take $h = 0$ and $c(0) = 1$, it corresponds to detect breaks at one time n. When there is a structural break in the trend-cycle components, the effects continue. In this sense, it would be desirable take several periods as $h > 1$. The approximation reported in Proposition 6.2 is reasonable when h is not large $(c(0) = 1, c(1) = 5, c(1) = 14, \cdots$. Whever, we expect that when h is large, the power of detecting structural change becomes low.

Next, we consider the estimation of the hidden state variable based on the backward-filtering. Let the forward-filtering estimate of the state at the period $n-h$ ($h \geq 0$) be $\hat{\Delta}^{(f)}x_{n-h}$, and the backward-filtering estimate of the state at the period $n-h$ ($h \geq 0$) be $\hat{\Delta}^{(b)}x_{n-h}$. Then, it is possible to use the statistic

$$d_{n-h} = \hat{\Delta}^{(b)}x_{n-h} - \hat{\Delta}^{(f)}x_{n-h} \ .$$

Because the difference of the state variable, which is an $I(1)$-process, can be written as $\mathbf{a}_{n-h}(1)'\mathbf{X}_n = x_{n-h} - x_{n-h-1}$, we define the backward-filtering estimate as $\hat{\Delta}^{(b)}x_{n-h} = -\mathbf{a}_{n+1-h}(1)'\hat{\mathbf{X}}_n^* = \hat{x}_{n-h}^{(b)} - \hat{x}_{n-h-1}^{(b)}$, we have the following result. [2].

Proposition 6.3 : Let h be a positive integer. In the basic non-stationary errors-in variables model (p=1) as Proposition 6.1. As $m,n \to \infty$ and $m/n \to 0$,

$$W_{b,n-h} = \sqrt{(\frac{1}{2})(\frac{n}{m})}\hat{\Delta}^{(b)}x_{n-h} \xrightarrow{d} N(0, f_{\Delta y}(0)) \ , \tag{6.5}$$

where $f_{\Delta y}(0)$ is the spectral density of Δy_n at $\lambda = 0$.
(ii) As $n \to \infty$ and $m/n \to 0$,

$$W_{b,f,n} = \sqrt{(\frac{1}{2})(\frac{n}{m})}d_{n-h} \xrightarrow{d} N(0, f_{\Delta y}(0)) \ . \tag{6.6}$$

(iii) As $m, n \to \infty$, $m/n \to 0$,

$$\frac{\sum_{k=0}^{h} W_{b,f,n-k}^2}{h\hat{f}_{\Delta y}(0)} \xrightarrow{d} \chi^2(1) \ . \tag{6.7}$$

Although the asymptotic distributions of two statistics have the same form, they can be considerably different because the second statistic incorporates both forward-backward information. Then, the power of detecting change points may be larger than the first one.

There can be many ways to combine the forward-filtering and the backward-filtering such as the multi-step filtering. Let the number of frequencies in the forward-filtering

[2] In the backward-filtering method, the initial period is n and the asymptotic behavior is different from $\hat{\Delta}^{(b)}x_n = \hat{x}_n^{(f)} - \hat{x}_{n-1}^{(b)}$. The exact distribution of $y_n^{(f)} - y_{n-1}^{(b)}$ depends on the initial conditions and noise at n.

be m_1 and the number of frequencies in the backward-filtering be m_2, and we denote the resulting statistic as $\hat{\Delta}^{(s)} x_{n-h}$. We assume the condition

$$\text{(Condition A)} \qquad \frac{m_1}{n} \longrightarrow c \ \ (0 \leq c \leq 1).$$

Let h be a positive integer. In the basic non-stationary errors-in variables model, assume (Condition A) and $m_2, n \to \infty$. Then, the asymptotic distribution of $W_{s,n-h} = \sqrt{(\frac{1}{2})(\frac{n}{m_2})}\hat{\Delta}^{(s)} x_{n-h}$ can be written as $N(0, v_n f_{\Delta y}(0))$, where $f_{\Delta y}(0)$ is the spectral density of Δy_n at $\lambda = 0$,

$$v_n = 1 + 2a(\frac{m_2}{n}) \ \ , \ a = \frac{1}{2}\int_c^1 [\frac{\cos \frac{\pi}{2}x}{\sin \frac{\pi}{2}x}]^2 dx \ , \tag{6.8}$$

See the Appendix for the derivation.

The constant a depends on c such that

$$a = \frac{1}{2}[\frac{2}{\pi}(\frac{\cos \frac{\pi}{2}c}{\sin \frac{\pi}{2}c}) - (1-c)] \ .$$

When $c = 0$ or $c = 1$, $v_n = 1$ because $a = 0$.

6.2 The General Case

In the general case of (2.1) and (2.2) with $p = 1$, let $\mathbf{y}_i = \mathbf{x}_i + \mathbf{v}_i$, $\Delta \mathbf{x}_i = \mathbf{v}_i^{(x)}$, and $\mathbf{v}_i (i = 0, 1, \cdots, n)$. The noise component and the state variables component $\mathbf{v}_i^{(x)}$ $(= \Delta \mathbf{x}_i)$ are sequences of the stationary processes satisfying

$$\mathbf{v}_i = \sum_{j=-\infty}^{\infty} \mathbf{C}_j^{(v)} \mathbf{e}_{i-j}^{(v)} \tag{6.9}$$

and

$$\mathbf{v}_i^{(x)} = \sum_{j=-\infty}^{\infty} \mathbf{C}_j^{(x)} \mathbf{e}_{i-j}^{(x)} \ , \tag{6.10}$$

where $\mathbf{e}_i^{(v)}$ and $\mathbf{e}_i^{(x)}$ are sequences of i.i.d. random variables with $\mathbf{E}(\mathbf{e}_i^{(v)}) = \mathbf{E}(\mathbf{e}_i^{(x)}) = \mathbf{0}$, $\mathbf{E}(\mathbf{e}_i^{(v)} \mathbf{e}_i^{(v)\prime}) = \mathbf{\Sigma}_e^{(v)}$ and $\mathbf{E}(\mathbf{e}_i^{(x)} \mathbf{e}_i^{(x)\prime}) = \mathbf{\Sigma}_e^{(x)}$.
For normalization we use $\mathbf{C}_0^{(v)} = \mathbf{C}_0^{(x)} = \mathbf{I}_p$. The $p \times p$ coefficient matrices $\mathbf{C}_j^{(v)}$ and $\mathbf{C}_j^{(x)} (j = \cdots, -1, 0, 1, \cdots)$ in (6.9) and (6.10) are absolutely summable such that $\sum_{j=-\infty}^{\infty} \|\mathbf{C}_j^{(v)}\| < \infty$ and $\sum_{j=-\infty}^{\infty} \|\mathbf{C}_j^{(x)}\| < \infty$, where $\|\mathbf{C}_j^{(v)}\| = \max_{k,l=1,\cdots,p} |c_{k,l}^{(v)}(j)|$ for $\mathbf{C}_j^{(v)} = (c_{k,l}^{(v)}(j))$ and $\|\mathbf{C}_j^{(x)}\| = \max_{k,l=1,\cdots,p} |c_{k,l}^{(x)}(j)|$ for $\mathbf{C}_j^{(x)} = (c_{k,l}^{(x)}(j))$, respectively. The measurement error vector $\mathbf{v}_i$ may include the seasonal component $\mathbf{s}_i$

when the main interest is to understand the trend factors as the state vector in the low frequency part, which is less than a year.

Alternatively, the state variables $\Delta \mathbf{x}_i$ $(= \mathbf{v}_i^{(x)})$ may include the seasonal components satisfying $\mathbf{s}_i = \sum_{j=-\infty}^{\infty} \mathbf{C}_{sj}^{(s)} \mathbf{e}_{i-sj}^{(s)}$ $(s \geq 2)$ and $\mathbf{e}_i^{(s)}$ is a sequence of i.i.d. random vectors with $\mathbf{E}(\mathbf{e}_i^{(s)}) = \mathbf{0}$ and $\mathbf{E}(\mathbf{e}_i^{(s)} \mathbf{e}_i^{(s)'}) = \mathbf{\Sigma}_e^{(s)}$ (the coefficients $\mathbf{C}_{sj}^{(s)}$ are absolutely summable $\sum_{j=-\infty}^{\infty} \|\mathbf{C}_j^{(s)}\| < \infty$).

The above formulation of (4.1) and (4.2) includes the non-stationary errors-in-variables model with the causal representation $\mathbf{v}_i = \sum_{j=0}^{\infty} \mathbf{C}_j^{(v)} \mathbf{e}_{i-j}^{(v)}$, $\mathbf{v}_i^{(x)} = \sum_{j=0}^{\infty} \mathbf{C}_j^{(x)} \mathbf{e}_{i-j}^{(x)}$, and the non-causal representation $\mathbf{v}_i = \sum_{j=-\infty}^{0} \mathbf{C}_j^{(v)} \mathbf{e}_{i-j}^{(v)}$, $\mathbf{v}_i^{(x)} = \sum_{j=-\infty}^{0} \mathbf{C}_j^{(x)} \mathbf{e}_{i-j}^{(x)}$, respectively, in (6.9) and (6.10).

In the general non-stationary errors-in-variables model with $p = 1$, Proposition 6.1 for the forward filtering method can be extended as follows.

Theorem 6.4 : In the general non-stationary errors-in-variables model given by (2.1) and (2.2), as $m, n \to \infty$ and $m/n \to 0$,

$$W_{f,n} = \sqrt{[\frac{3}{2\pi^2 (h+1)^2}](\frac{n^3}{m^3})} \hat{\Delta} x_{n-h} \xrightarrow{d} N(0, f_{\Delta y}(0)) \ , \tag{6.11}$$

where $f_{\Delta y}(0)$ is the spectral density of Δy_n at $\lambda = 0$.

Propositions 6.3 for the backward-forward filtering methods can be extended as follows.

Theorem 6.5 : In the general non-stationary errors-in variables model with $p = 1$ given by (2.1) and (2.2), as $m, n \to \infty$ and $m/n \to 0$,

$$W_{b,n-h} = \sqrt{(\frac{1}{2})(\frac{n}{m})} \hat{\Delta}^{(b)} x_{n-h} \xrightarrow{d} N(0, f_{\Delta y}(0)) \ , \tag{6.12}$$

where $f_{\Delta y}(0)$ is the spectral density of Δy_n at $\lambda = 0$.
(ii) As $n \to \infty$ and $m/n \to 0$,

$$W_{b,f,n} = \sqrt{(\frac{1}{2})(\frac{n}{m})} d_{n-h} \xrightarrow{d} N(0, f_{\Delta y}(0)) \ . \tag{6.13}$$

(iii) As $m, n \to \infty$, $m/n \to 0$,

$$\frac{\sum_{k=0}^{h} W_{b,f,n-k}^2}{h \hat{f}_{\Delta y}(0)} \xrightarrow{d} \chi^2(1) \ . \tag{6.14}$$

6.3 An Application of Macroconsumption

As an example, we applied the SIML-forward and SIML-backward filtering methods to analyze real final consumption (shouhi) data published by the Cabinet Office of Japan, covering the period from 1994Q1 to 2020Q4. The analysis focuses on the period from 2020Q1 to 2020Q4, as the COVID-19 pandemic caused significant disruptions. In particular, macroeconomic consumption exhibited a sharp decline in 2020Q2, followed by a gradual recovery. However, it took considerable time to return to pre-pandemic levels.

For this analysis, we set $m = [n/6]$, corresponding to trend-cycle components with cycles longer than three years. This choice ensures that the hidden trend variables are defined in the low-frequency range. The red curves in Figures 6.1 and 6.2 show the estimated results of the forward-filtering procedure.

Table 6.1 : t-statistic in the forward-filtering

h/Period	2019Q3	2023Q4	2020Q1	2020Q2	2020Q3	2020Q4	2021Q1	2021Q2
h=0	0.295	0.170	-0.275	-2.217	-2.556	-2.435	-2.119	-1.702
h=1	0.247	0.131	-0.277	-2.054	-2.232	-1.963	-1.602	-1.098
h=2	0.179	0.075	-0.276	-1.802	-2.082	-1.923	-1.725	-1.447
h=3	0.105	0.015	-0.267	-1.486	-1.722	-1.545	-1.432	-1.253

Table 6.2 : t-statistic in the backward-filtering

h/Period	2019Q3	2023Q4	2020Q1	2020Q2	2020Q3	2020Q4	2021Q1	2021Q2
h=0	0.140	-0.105	-0.712	-3.153	-2.937	-1.542	-0.925	-0.464
h=1	0.0996	-0.074	-0.579	-2.734	-2.720	-1.846	-1.450	-1.097
h=2	0.053	-0.041	-0.419	-2.193	-2.430	-2.189	-2.074	-1.867
h=3	0.016	-0.014	-0.268	-1.612	-2.097	-2.468	-2.644	-2.602

Tables 6.1 and 6.2 summarize the t-statistics obtained from the forward and backward filtering methods, respectively. These statistics were calculated using data sequentially for the periods 1994Q12020Q1, 1994Q12020Q2, 1994Q12020Q3, and 1994Q12020Q4. The standard errors were derived based on the asymptotic distributions described in Propositions 6.1 and 6.3. An estimate of $f(\Delta y_0)$ was obtained as $\hat{\sigma}_x^2 = \frac{1}{m}\sum_{i=1}^{m} z_i^2$. (See Chapter 2 of Kunitomo and Sato (2024).)

The following observations are noteworthy: (i) The forward filtering method (6.2) detected a significant structural break in hidden consumption during 2020Q2, as evidenced by the t–statistics. Subsequent t-values remained significant, reflecting

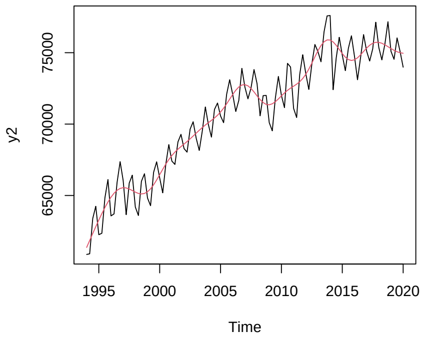

Figure 3: Original Series and Estimated Trend (1994Q1-2020Q1)

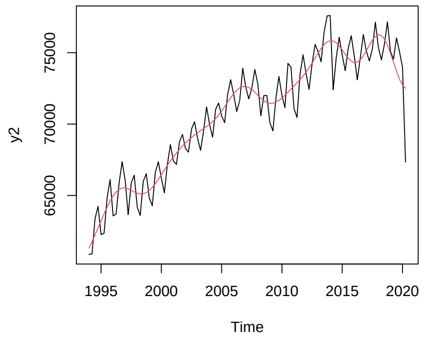

Figure 4: Original Series and Estimated Trend (1994Q1-2020Q2)

the lasting effects of state variable estimation. (ii) The backward filtering method (6.5) also identified a significant structural break in 2020Q2. However, the influence of state variable estimation diminished in later periods, suggesting that the effects of the initial conditions do not persist for long durations. Figures 3-6 show the original series and estimated trends for each analyzed period. These results underscore the practical applicability of the SIML method in identifying structural breaks and understanding long-term trends in economic time series.

These results confirm the ability of both forward and backward filterings to detect recent change points in macroeconomic data

7. Conclusions

When observed nonstationary time series contain noise, disentangling the effects of trends, cycles, and noise from the original data can be challenging, particularly when the sample size is small, as is often the case with major macroeconomic time series. This study introduces a new statistical smoothing procedure designed to decompose time series into nonstationary trend, seasonal, and stationary noise (or measurement error) components.

The proposed smoothing or filtering method for nonstationary series is simple and does not rely on the underlying distributions of the noise or the state vector. As a result, it is robust against potential misspecifications in the nonstationary time series models, making it a reliable tool for practical applications.

It is worth noting that, while there has been extensive research on structural breaks in statistical time series, relatively few studies have focused on detecting recent structural breaks in noisy nonstationary time series. This gap highlights the importance of methods capable of identifying such breaks effectively. To demon-

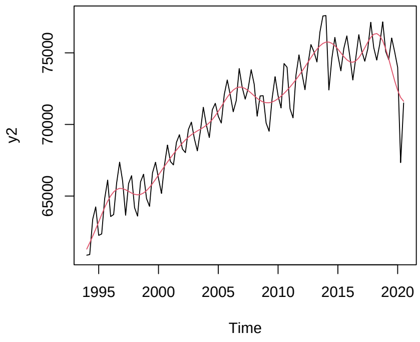

Figure 5: Original Series and Estimated Trend (1994Q1-2020Q3)

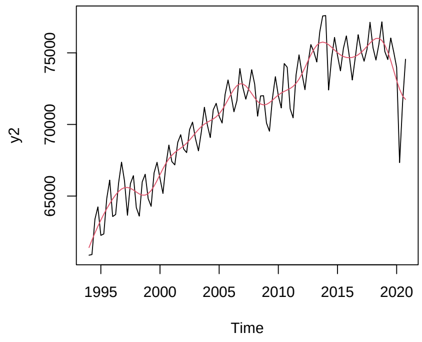

Figure 6: Original Series and Estimated Trend (1994Q1-2020Q4)

strate the practical utility of the proposed method, we provide an empirical examples using macroeconomic data in Japan. The recent example suggests that our approach is valuable for interpreting macroeconomic data and improving seasonal adjustment procedures.

References

[1] Akaike, H. (1980), "Seasonal Adjustment by a Bayesian Modeling," *Journal of Time Series Analysis*, 1, 1-13.

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

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

[4] Brillinger, D. (1980) *Time Series : Data Analysis and Theory*, Expanded Edition, Holden-Day.

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

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

[7] Harvey, A. and T. Trimbur (2008), "Trend Estimation and The Hodorick-Prescott Filter," Journal of Japan Statistical Society, 38-1, 41-49.

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

[9] Kitagawa, G. and Gersch, W. (1996), *smoothing Priors Analysis of Time Series*, Lecture Notes in Statistics, No. 116, Springer.

[10] Kunitomo, N. and S. Sato (2021), "A Robust-filtering Method for Noisy Non-Stationary Multivariate Time Series with Econometric Applications," Japanese Journal of Statistics and Data Science, 4, 373-410. (JJSD, Open Access), https://doi.org/10.1007/s42081-020-00102-y.

[11] Kunitomo, N. and S. Sato (2024), *The SIML Filtering Method for Noisy Non-stationary Economic Time Series*, JSS-Springer Series, Springer, forthcoming.

APPENDIX : Mathematical Derivations

We present here some details of derivations that we have omitted in the previous sections. Most of our derivations is to apply trigonometric relations, which are mathematically elementary and straightforward. Hence, we show only the essential parts of derivations, and first prepare a key lemma on the characteristic roots and eigen vectors of a patterned matrix, Then, we show the proof of theorems.

Lemma A.1 : (i) Define $n \times n$ matrices $\mathbf{A}_n$ and $\mathbf{A}_n^*$ by

(A.1)
$$\mathbf{A}_n = \frac{1}{2}\begin{pmatrix} 1 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 1 & \cdots & 0 \\ 0 & 1 & 0 & 1 & \cdots \\ 0 & 0 & \cdots & 0 & 1 \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix}$$

and

(A.2)
$$\mathbf{A}_n^* = \frac{1}{2}\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 1 & \cdots & 0 \\ 0 & 1 & 0 & 1 & \cdots \\ 0 & 0 & \cdots & 0 & 1 \\ 0 & \cdots & 0 & 1 & 1 \end{pmatrix} .$$

Then, $\cos\pi(\frac{2k-1}{2n+1})$ $(k = 1, \cdots, n)$ are eigen-values of $\mathbf{A}_n$ and $\mathbf{A}_n^*$, and their eigen-vectors are

(A.3)
$$\begin{bmatrix} \cos[\pi(\frac{2k-1}{2n+1})(1-\frac{1}{2})] \\ \cos[\pi(\frac{2k-1}{2n+1})(2-\frac{1}{2})] \\ \vdots \\ \cos[\pi(\frac{2k-1}{2n+1})(n-\frac{1}{2})] \end{bmatrix} , \begin{bmatrix} \sin[\pi(\frac{2k-1}{2n+1})1] \\ \sin[\pi(\frac{2k-1}{2n+1})2] \\ \vdots \\ \sin[\pi(\frac{2k-1}{2n+1})n] \end{bmatrix} \ (k = 1, \cdots, n),$$

respectively.
(ii) We have the spectral decompositions

$$\mathbf{C}_n^{-1}\mathbf{C}_n^{'-1} = \mathbf{P}_n^{'}\mathbf{D}_n\mathbf{P}_n = 2\mathbf{I}_n - 2\mathbf{A}_n \ , \tag{A.4}$$

and

$$\mathbf{C}_n^{'-1}\mathbf{C}_n^{-1} = \mathbf{P}_n^{*'}\mathbf{D}_n\mathbf{P}_n^{*} = 2\mathbf{I}_n - 2\mathbf{A}_n^{*} \ , \tag{A.5}$$

where $\mathbf{P}^{*'}$ is the matrix consisting of eigen-vectors in (A.2), $\mathbf{D}_n$ is a diagonal matrix with the $k-$th element

$$d_k = 2\left[1-\cos(\pi(\frac{2k-1}{2n+1}))\right] \ \ (k=1,\cdots,n) \ , \tag{A.6}$$

$$\mathbf{C}_n^{'-1} = \begin{pmatrix} 1 & -1 & \cdots & 0 & 0 \\ 0 & 1 & -1 & \cdots & 0 \\ 0 & 0 & 1 & -1 & \cdots \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{A.7}$$

and the $(k,j)-$the element of $\mathbf{P}_n = (p_{kj})$ and $\mathbf{P}_n^{*} = (p_{kj}^{*})$ are given by

$$p_{kj} = \sqrt{\frac{2}{n+\frac{1}{2}}}\cos\left[\frac{2\pi}{2n+1}(k-\frac{1}{2})\ (j-\frac{1}{2})\right] \ , \tag{A.8}$$

and

$$p_{kj}^{*} = \sqrt{\frac{2}{n+\frac{1}{2}}}\sin\left[\frac{2\pi}{2n+1}(k-\frac{1}{2})\ j\right] \ . \tag{A.9}$$

Proof of Lemma A.1 : The proof of $\mathbf{A}_n^{*}$ is a direct calculation and we omit the proof for $\mathbf{A}_n$ because it is essentially the same.
(i) Let $\mathbf{A}_n^{*} = (a_{ij}^{*})$ $(i,j=1,\cdots,n)$ and an $n\times 1$ vector $\mathbf{x}=(x_t)$ $(t=1,\cdots,n)$ satisfying $\mathbf{A}_n^{*}\mathbf{x}=\lambda\mathbf{x}$. Then,

$$\begin{aligned} &\frac{x_2}{2} = \lambda x_1 \ , \\ &\frac{x_{t-1}+x_{t+1}}{2} = \lambda x_t \ (t=2,\cdots,n-1) \ , \\ &\frac{1}{2}[x_{n-1}+x_n] = \lambda x_n \ . \end{aligned} \tag{A.10}$$

Let ξ_i $(i=1,2)$ be the solutions of $\xi^2-2\lambda\xi+1=0$. Because $2\lambda=\xi_1+\xi_2$ and $\xi_1\xi_2=1$, we have the solution as $x_t = c_1\xi_1^t + c_2\xi_1^{-t}$ $(t=1,\cdots,n)$ and c_i $(i=1)$ are real constants. The first equation implies $0=c_1\xi_1^2+c_2\xi_1^{-2}-(\xi_1+\xi_1^{-1})(c_1\xi_1+c_2\xi_1^{-1})$, and $c_1+c_2=0$. Then, we find that $x_t=c_1[\xi_1^t-\xi_1^{-t}]$, and the third equation implies $(\xi_1^{2n+1}+1)(1-\xi_1)=0$. Because $\xi_1\neq 1$, we find that $\xi_1^{2n+1}=-1=e^{\pi i(2k-1)}$ for any

positive integer k.
Then,

(A.11) $$\lambda_k = \cos[\pi \frac{2k-1}{2n+1}] \ (k=1,\cdots,n) \ .$$

By taking $c_1 = (1/2i)$, the elements of the characteristic vectors of $\mathbf{A}_n^*$ with $\cos[\pi(2k-1)/(2n+1)]$ are

(A.12) $$x_t = \frac{1}{2i}\left[\xi_1^t - \xi_1^{-t}\right] = \sin\left[\pi \frac{2k-1}{2n+1} t\right] \ .$$

(ii) The rest of the proof involves the standard arguments of spectral decomposition in linear algebra. **Q.E.D.**

Proof of Theorem 3.1 :

(**Step I**)
We consider the case of $T_{2k-1}^{(m)}$ $(k \geq 1)$. By using the recursive relations, for $k \geq 1$ we represent

(A.13) $$T_{2k+1}^{(m)} = A_1^{(m)} + A_2^{(m)} T_{2k-1}^{(m)} \ ,$$

where an $n \times n$ matrix $A_2^{(m)}$ is defined by

(A.14) $$A_2^{(m)} = (\mathbf{I}_n - \mathbf{C}_n \mathbf{P}_n \mathbf{J}_m^{'} \mathbf{J}_m \mathbf{P}_n \mathbf{C}_n^{-1}) \mathbf{1}_n \mathbf{e}_1^{'} (\mathbf{I}_n - \mathbf{C}_n^{'} \mathbf{P}_n^{*'} \mathbf{J}_m^{'} \mathbf{J}_m \mathbf{P}_n^{*} \mathbf{C}_n^{'-1}) \mathbf{1}_n \mathbf{e}_n^{'} \ .$$

Then, we consider the characteristic roots of the coefficient matrix $\mathbf{A}_2^{(m)}$. Because the rank of $\mathbf{A}_2^{(m)}$ is one, there are $n-1$ zero roots and one non-zero root, which is

(A.15) $$\begin{aligned} a_{2n} &= \mathbf{e}_n^{'} (\mathbf{I}_n - \mathbf{C}_n \mathbf{P}_n \mathbf{J}_m^{'} \mathbf{J}_m \mathbf{P}_n \mathbf{C}_n^{-1}) \mathbf{1}_n \mathbf{e}_1^{'} (\mathbf{I}_n - \mathbf{C}_n^{'} \mathbf{P}_n^{*'} \mathbf{J}_m^{'} \mathbf{J}_m \mathbf{P}_n^{*} \mathbf{C}_n^{'-1}) \mathbf{1}_n \\ &= [1 - \mathbf{1}_n^{`} \mathbf{P}_n \mathbf{J}_m^{'} \mathbf{J}_m \mathbf{P}_n \mathbf{e}_1][1 - \mathbf{1}_n^{`} \mathbf{P}_n^{*'} \mathbf{J}_m^{'} \mathbf{J}_m \mathbf{P}_n^{*} \mathbf{e}_n] \ . \end{aligned}$$

(We have used the relations as $\mathbf{e}_n^{'} \mathbf{C}_n = \mathbf{1}_n^{`}$ and $\mathbf{e}_1^{'} \mathbf{C}_n^{'} = \mathbf{1}_n^{`}$.)
By using the relation $1 - \mathbf{1}_n^{'} \mathbf{P}_n \mathbf{J}_m^{'} \mathbf{J}_m \mathbf{P}_n \mathbf{e}_1 = \mathbf{1}_n^{'} \mathbf{P}_n \mathbf{J}_{n-m}^{'} \mathbf{J}_{n-m} \mathbf{P}_n \mathbf{e}_1$ for $\mathbf{J}_{n-m} = (\mathbf{O}, \mathbf{I}_{n-m})$ ($(n-m) \times [m + (n-m)]$ matrix), we evaluate two terms in (A.15). The first term of (A.15) becomes

$$[\sqrt{\frac{2}{n+\frac{1}{2}}}]^2 \sum_{k=m+1}^{n} [\sum_{j=1}^{n} \cos \frac{2\pi}{2n+1}(j-\frac{1}{2})(k-\frac{1}{2})] \times \cos \frac{2\pi}{2n+1}(k-\frac{1}{2})(1-\frac{1}{2})] \ ,$$

which is less than 1. We show this key fact in the following. We take the summation with respect to n as

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

and the relation

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

Then the first term of (A.15) becomes

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

By using the facts that (i) $\sin\pi(k-\frac{1}{2})$ takes $+1$ and -1 alternatively, (ii) the absolute value of $\frac{[\cos\frac{\pi}{2n+1}(k-\frac{1}{2})]^2}{\sin\frac{\pi}{2n+1}(k-\frac{1}{2})}$ is a monotone decreasing for $k=1,\cdots,n$ and (iii)

$$[\sqrt{\frac{2}{n+\frac{1}{2}}}]^2\sum_{k=1}^{n}[\sum_{j=1}^{n}\cos\frac{2\pi}{2n+1}(j-\frac{1}{2})(k-\frac{1}{2})]\times\cos\frac{2\pi}{2n+1}(k-\frac{1}{2})(1-\frac{1}{2})]=1$$

(See Lemma A.2 below), then, $\mathbf{1}_n^{'}\mathbf{P}_n\mathbf{J}_{n-m}^{'}\mathbf{J}_{n-m}\mathbf{P}_n\mathbf{e}_1$ is less than one.
Similarly, for the second term of (A.14), we use

$$[1-\mathbf{1}_n^{'}\mathbf{P}_n^{*'}\mathbf{J}_m^{'}\mathbf{J}_m\mathbf{P}_n^{*}\mathbf{e}_n]=\mathbf{1}_n^{'}\mathbf{P}_n^{*'}\mathbf{J}_{n-m}^{'}\mathbf{J}_{n-m}\mathbf{P}_n^{*}\mathbf{e}_n \; .$$

Then, the second term of (A.15) becomes

$$[\sqrt{\frac{2}{n+\frac{1}{2}}}]^2\sum_{k=1}^{n}\sum_{j=m+1}^{n}[\sin\frac{2\pi}{2n+1}(k-\frac{1}{2})j\sin\frac{2\pi}{2n+1}(n-\frac{1}{2})j \; ,$$

which is less than 1. In this evaluation, we utilized the relations (i)

$$\text{(A.16)}\qquad \begin{aligned}\sum_{k=1}^{n}\sin\frac{2\pi}{2n+1}(k-\frac{1}{2})j &= \frac{1}{2i}\frac{e^{i\frac{2\pi}{2n+1}jn}+e^{-i\frac{2\pi}{2n+1}jn}-2}{e^{i\frac{2\pi}{2n+1}j\frac{1}{2}}-e^{-i\frac{2\pi}{2n+1}j\frac{1}{2}}} \\ &= \frac{1}{2}\frac{1-\cos\frac{2\pi}{2n+1}jn}{\sin\frac{2\pi}{2n+1}j\frac{1}{2}} \; ,\end{aligned}$$

(ii) $1-\cos\frac{2\pi}{2n+1}jn=1-\cos\pi j\cos\frac{2\pi}{2n+1}j=1-(-1)^j\cos\frac{2\pi}{2n+1}j$ and (iii) for $2(n-1/2)=(2n+1)-2$, $\sin\frac{2\pi}{2n+1}(n-\frac{1}{2})j=[-\cos\pi j]\sin\frac{2\pi}{2n+1}j=2[-\cos\pi j][\sin\frac{\pi}{2n+1}j][\cos\frac{\pi}{2n+1}j]$.
Then, the second term becomes

$$[\frac{4}{2n+1}]\sum_{k=1}^{n}\sum_{j=m+1}^{n}(-1)^{j+1}[1-(-1)^j\cos\frac{\pi}{2n+1}j]\cos\frac{\pi}{2n+1}j \; .$$

Since for any odd number j and $1\leq j\leq n$, $[1+\cos\frac{\pi}{2n+1}j]\cos\frac{\pi}{2n+1}j>[1-\cos\frac{\pi}{2n+1}(j+1)]\cos\frac{\pi}{2n+1}(j+1)$, for any even muber m, we have $\sum_{j=1}^{m}(-1)^{j+1}[1-(-1)^j\cos\frac{\pi}{2n+1}j]\cos\frac{\pi}{2n+1}j>0$. Then, by using the fact

$$[\sqrt{\frac{2}{n+\frac{1}{2}}}]^2\sum_{k=1}^{n}\sum_{j=+1}^{n}[\sin\frac{2\pi}{2n+1}(k-\frac{1}{2})j\sin\frac{2\pi}{2n+1}(n-\frac{1}{2})j=1$$

(see Lemma A.2 below), $\mathbf{1}_n^{'}\mathbf{P}_n^{*}\mathbf{J}_{n-m}^{'}\mathbf{J}_{n-m}\mathbf{P}_n^{*'}\mathbf{e}_1$ is less than one.
Because each term of (A.15) is less than one, we have $|a_{2n}| < 1$. Then, by using (A.13), we have convergence of T_{2k+1} as $k \to \infty$.

Lemma A.2 : Let $p_{kj}^{(n)} = \sqrt{\frac{2}{n+\frac{1}{2}}}\cos\left[\frac{2\pi}{2n+1}(k-\frac{1}{2})(j-\frac{1}{2})\right]$ and $p_{kj}^{*(n)} = \sqrt{\frac{2}{n+\frac{1}{2}}}\sin\left[\frac{2\pi}{2n+1}(k-\frac{1}{2})j\right]$ for $k, j = 1, \cdots, n$. Then, for any $a, b = 1, \cdots, n$,

$$\text{(A.17)} \qquad \sum_{a=1}^{n}\sum_{k=1}^{n} p_{ka}^{(n)}p_{kb}^{(n)} = \sum_{a=1}^{n}\sum_{k=1}^{n} p_{ka}^{*(n)}p_{kb}^{*(n)} = 1 ,$$

Proof of Lemma A.2 : We use the orthogonal relations such that for any $a, b = 1, \cdots, n$,

$$\text{(A.18)} \qquad \sum_{k=1}^{n} p_{ka}^{(n)}p_{kb}^{(n)} = \sum_{k=1}^{n} p_{ka}^{*(n)}p_{kb}^{*(n)} = \delta(a,b) ,$$

where $\delta(a,a) = 1$ and $\delta(a,b) = 0$ for any $a \neq b$. Then we have the result.
Q.E.D.

(**Step II**)
We can apply the similar arguments to $T_{2k}^{(m)}$ $(k \geq 1)$. By using the recursive relations, for $k \geq 1$ we can represent

$$\text{(A.19)} \qquad T_{2k}^{(m)} = A_{1*}^{(m)} + A_{2*}^{(m)}T_{2(k-1)}^{(m)} ,$$

where $A_{1*}^{(m)}$ and $A_{2*}^{(m)}$ are $n \times n$ matrices as defined in Theorem 3.1. By evaluating the eigenvalues of $A_{2*}^{(m)}$, we find that the absolute value of eigenvalues are less than one, and we have convergence of $T_{2k}^{(m)}$ as $k \to \infty$.
(**Q.E.D.**)

Proof of Theorem 3.2 :

Our proof is the direct evaluation of trace of $\mathbf{H}_n$ and $\mathbf{F}_n$ by using the trigonometric relations, which are elementary. We utilize the fact that $\mathrm{Tr}(\mathbf{H}_n) = \mathrm{Tr}(\mathbf{P}n\mathbf{Q}_n^{(m)}\mathbf{P}_n)$ and $\mathrm{Tr}(\mathbf{F}_n) = \mathrm{Tr}(\mathbf{P}_n^{*'}\mathbf{Q}_n^{(m)}\mathbf{P}_n^{*})$. We set $\mathbf{H}_n^{*} = (h_{ab}^{*}) = \mathbf{P}_n\mathbf{Q}_n^{(m)}\mathbf{P}_n$ and $\mathbf{F}_n^{*} = (f_{ab}^{*}) = \mathbf{P}_n^{*'}\mathbf{Q}_n^{(m)}\mathbf{P}_n^{*}$. From (2.7) and (3.8),

$$\text{(A.20)} \qquad \sum_{a=1}^{n} h_{aa}^{*} = \frac{4}{2n+1}\sum_{a=1}^{n}\sum_{j=1}^{m}[\cos\frac{2\pi}{2n+1}(a-\frac{1}{2})(j-\frac{1}{2})]^2$$
$$= .\frac{4}{2n+1}\sum_{a=1}^{n}\sum_{j=1}^{m}\frac{1}{2}[1+\cos\frac{4\pi}{2n+1}(a-\frac{1}{2})(j-\frac{1}{2})]$$

By taking the summation w.r.t. a first, we find

$$\frac{2n+1}{4}h_{aa}^{*}$$

$$
\begin{aligned}
&= \frac{mn}{2} + \frac{1}{2}\sum_{j=1}^{m}\sum_{a=1}^{n}\frac{1}{2}[e^{i\frac{4\pi}{2n+1}(a-\frac{1}{2})(j-\frac{1}{2})} + e^{-i\frac{4\pi}{2n+1}(a-\frac{1}{2})(j-\frac{1}{2})}] \\
&= \frac{mn}{2} + \frac{1}{4}\sum_{j=1}^{m}[\frac{e^{i\frac{4\pi}{2n+1}(j-\frac{1}{2})\frac{1}{2}}(1-e^{i\frac{4\pi}{2n+1}(j-\frac{1}{2})n})}{1-e^{i\frac{4\pi}{2n+1}(j-\frac{1}{2})}} + \frac{e^{-i\frac{4\pi}{2n+1}(j-\frac{1}{2})\frac{1}{2}}(1-e^{-i\frac{4\pi}{2n+1}(j-\frac{1}{2})n})}{1-e^{-i\frac{4\pi}{2n+1}(j-\frac{1}{2})}}] \\
&= \frac{mn}{2} + \frac{1}{4}\sum_{j=1}^{m}[\frac{1-e^{i\frac{4\pi}{2n+1}(j-\frac{1}{2})n}}{e^{-i\frac{2\pi}{2n+1}(j-\frac{1}{2})}-e^{i\frac{2\pi}{2n+1}(j-\frac{1}{2})}} + \frac{1-e^{-i\frac{4\pi}{2n+1}(j-\frac{1}{2})n}}{e^{i\frac{2\pi}{2n+1}(j-\frac{1}{2})}-e^{-i\frac{2\pi}{2n+1}(j-\frac{1}{2})}}] \\
&= \frac{mn}{2} + \frac{1}{4}\sum_{j=1}^{m}\frac{\sin\frac{4\pi}{2n+1}(j-\frac{1}{2})n}{\sin\frac{2\pi}{2n+1}(j-\frac{1}{2})} .
\end{aligned}
$$

Because $\sin\frac{4\pi}{2n+1}(j-\frac{1}{2})n = \sin\pi[(2j-1-\frac{2j-1}{2n+1})] = [-\cos\pi(2j-1)][\sin\pi(2j-1)/(2n+1)]$ and $-\cos\pi(2j-1) = 1$ for any positive integer j, we have

$$
\text{(A.21)} \qquad \frac{2n+1}{4}\sum_{a=1}^{n} h_{aa}^{*} = \frac{mn}{2} + \frac{m}{4} .
$$

Similarly,

$$
\begin{aligned}
\sum_{a=1}^{n} f_{aa}^{*} &= \frac{4}{2n+1}\sum_{a=1}^{n}\sum_{j=1}^{m}[\sin\frac{2\pi}{2n+1}(a-\frac{1}{2})j]^2 \\
&\quad . \frac{4}{2n+1}\sum_{a=1}^{n}\sum_{j=1}^{m}\frac{1}{2}[1-\cos\frac{4\pi}{2n+1}(a-\frac{1}{2})j]
\end{aligned}
$$

By taking the summation w.r.t. a and j, we find that ,

$$
\text{(A.22)} \qquad \frac{2n+1}{4}\sum_{a=1}^{n} f_{aa}^{*} = \frac{mn}{2} - \frac{m}{4} .
$$

Because $\mathrm{Tr}[\mathbf{H}_n - \mathbf{F}_n] = \sum_{a=1}^{n}[h_{aa}^{*} - f_{aa}^{*}]$, we have the result.
(**Q.E.D.**)

Proof of Theorem 3.3 :

The proof is basically the same as Theorem 3.1. We replace $\mathbf{Q}_n^{(m)} = \mathbf{J}_m^{'}\mathbf{J}_m$ by $\mathbf{Q}_n^{(m_1,m_2)} = \mathbf{J}_{m_1,m_2}^{'}\mathbf{J}_{m_1,m_2}$.
Consider $T_{2k+1}^{(m_1,m_2)} = A_1^{(m_1,m_2)} + A_2^{(m_1,m_2)}T_{2k-1}^{(m_1,m_2)}$. Then, the non-zero eigenvalue of $A_2^{(m_1,m_2)}$ is

$$
\begin{aligned}
a_{2n} &= \mathbf{e}_n^{'}(\mathbf{I}_n - \mathbf{C}_n\mathbf{P}_n\mathbf{J}_{m_1,m_2}^{'}\mathbf{J}_{m_1,m_2}\mathbf{P}_n\mathbf{C}_n^{-1})\mathbf{1}_n\mathbf{e}_1^{'}(\mathbf{I}_n - \mathbf{C}_n^{'}\mathbf{P}_n^{*'}\mathbf{J}_{m_1,m_2}^{'}\mathbf{J}_{m_1,m_2}\mathbf{P}_n^{*}\mathbf{C}_n^{'-1})\mathbf{1}_n \\
&= [1 - \mathbf{1}_n^{'}\mathbf{P}_n\mathbf{J}_{m_1,m_2}^{'}\mathbf{J}_{m_1,m_2}\mathbf{P}_n\mathbf{e}_1][1 - \mathbf{1}_n^{'}\mathbf{P}_n^{*}\mathbf{J}_{m_1,m_2}^{'}\mathbf{J}_{m_1,m_2}\mathbf{P}_n^{*}\mathbf{e}_n] .
\end{aligned}
$$

In this case, we use the relation

$$
\begin{aligned}
[1 - \mathbf{1}_n^{'}\mathbf{P}_n\mathbf{J}_{m_1,m_2}^{'}\mathbf{J}_{m_1,m_2}\mathbf{P}_n\mathbf{e}_1] &= \mathbf{1}_n^{'}\mathbf{P}_n\mathbf{J}_{m_1}^{'}\mathbf{J}_{m_1}\mathbf{P}_n\mathbf{e}_1 \\
&+ \mathbf{1}_n^{'}\mathbf{P}_n\mathbf{J}_{m_1+m_2+1,n}^{'}\mathbf{J}_{m_1+m_2+1,n}\mathbf{P}_n\mathbf{e}_1
\end{aligned}
$$

and

$$[1-\mathbf{1}_n^{'}\mathbf{P}_n^{*'}\mathbf{J}_{m_1,m_2}^{'}\mathbf{J}_{m_1,m_2}\mathbf{P}_n^{*}\mathbf{e}_n] = \mathbf{1}_n^{'}\mathbf{P}_n^{*'}\mathbf{J}_{m_1}^{'}\mathbf{J}_{m_1}\mathbf{P}_n^{*}\mathbf{e}_n + \mathbf{1}_n^{'}\mathbf{P}_n^{*'}\mathbf{J}_{m_1+m_2+1,n}^{'}\mathbf{J}_{m_1+m_2+1,n}\mathbf{P}_n^{*}\mathbf{e}_n\ .$$

Then, by using the same argument as the proof of Theorem 3.1 and Lemma A.2, we find that the absolute values of the eigenvalues of $\mathbf{A}_2^{(m_1,m_2)}$ and $\mathbf{A}_{2*}^{(m_1,m_2)}$ are less than one when m_1 and m_2 are even numbers. Then, we have the convergence of the repeated smoothing procedures.
(**Q.E.D.**)

Proof of Proposition 6.1 :
We consider the case when $\Delta x_i = x_i - x_{i-1} = v_i^{(v)}$ and v_i $(t = i,\cdots,n)$ are i.i.d. sequences. We set the initial conditions $x_0 = v_0 = 0$ for the convenience and take h as a fixed non-negative integer.
Let $X_{n-h} = X_{1,n-h} + X_{2,n-h}$ $(h = 0,1,\cdots,n)$, where

$$X_{1,n-h} = \sum_{i=1}^{n}[\sum_{j=1}^{m}\cos\frac{2\pi}{2n+1}(n-h-\frac{1}{2})(j-\frac{1}{2})\cos\frac{2\pi}{2n+1}(j-\frac{1}{2})(i-\frac{1}{2})]v_i^{(x)}\ ,$$

and

$$X_{2,n-h} = \sum_{i=1}^{n}[\sum_{j=1}^{m}\cos\frac{2\pi}{2n+1}(n-h-\frac{1}{2})(j-\frac{1}{2})\cos\frac{2\pi}{2n+1}(j-\frac{1}{2})(i-\frac{1}{2})]\Delta v_i\ .$$

We first evaluate the variance of the first term

$$\begin{aligned}\mathbf{Var}[X_{1,n-h}] &= \sigma^{(x)2}\sum_{i=1}^{n}[\sum_{j=1}^{m}\cos\frac{2\pi}{2n+1}(n-h-\frac{1}{2})(j-\frac{1}{2}\cos\frac{2\pi}{2n+1}(j-\frac{1}{2})(i-\frac{1}{2})]^2\\ &= \sigma^{(x)2}\sum_{j,j^{'}=1}^{m}\cos\frac{2\pi}{2n+1}(n-h-\frac{1}{2})(j-\frac{1}{2})\cos\frac{2\pi}{2n+1}(n-h-\frac{1}{2})(j^{'}-\frac{1}{2})\\ &\quad\times[\sum_{t=1}^{n}\cos\frac{2\pi}{2n+1}(j-\frac{1}{2})(t-\frac{1}{2})\cos\frac{2\pi}{2n+1}(j^{'}-\frac{1}{2})(i-\frac{1}{2})]\end{aligned}$$

We use the orthogonal relation

$$\sum_{i=1}^{n}\cos[\frac{2\pi}{2n+1}(j-\frac{1}{2})(i-\frac{1}{2})]\times\cos[\frac{2\pi}{2n+1}(j^{'}-\frac{1}{2})(i-\frac{1}{2})] = \delta(j,j^{'})[\frac{n}{2}+\frac{1}{4}]\ .$$

Then, we need to evaluate

$$\mathbf{Var}[X_{1,n-h}] = \sigma^{(x)2}[\frac{n}{2}+\frac{1}{4}]\sum_{j=1}^{m}[\cos\frac{2\pi}{2n+1}(n-h-\frac{1}{2})(j-\frac{1}{2})]^2\ .$$

Because $2(n-h-1/2) = 2n+1-2(h+1)$, we use the trigonometric relation $\cos\frac{2\pi}{2n+1}(n-h-\frac{1}{2})(j-\frac{1}{2}) = \cos\pi[1-\frac{2(h+1)}{2n+1}(j-\frac{1}{2})] = \sin\pi(j-\frac{1}{2})\sin\pi(h+1)\frac{2j-1}{2n+1}$. When $m/n \to 0$ as $n\to\infty$,

$$\begin{aligned}\sum_{j=1}^{m}[\cos\frac{2\pi}{2n+1}(n-h-\frac{1}{2})(j-\frac{1}{2})]^2 &= \sum_{j=1}^{m}[\sin\pi(h+1)\frac{2j-1}{2n+1}]^2 \\ &\sim [\frac{\pi}{n}(h+1)]^2\sum_{j=1}^{m}(\frac{j}{n})^2 \\ &\sim [\frac{\pi^2(h+1)^2}{3}]\frac{m^3}{n^2}\end{aligned}$$

and, then

$$\mathbf{Var}[X_{1,n-h}] \sim \sigma^{(x)2}(\frac{\pi^2(h+1)^2}{6})\frac{m^3}{n} \ .$$

Because the spectral decomposition of $\mathbf{C}_n^{-1}\mathbf{C}_n^{'-1} = \mathbf{P}_n\mathbf{D}_n\mathbf{P}_n$, we have

$$\mathbf{P}_n\mathbf{Q}_n\mathbf{P}_n\mathbf{C}_n^{-1}\mathbf{C}_n^{'-1}\mathbf{P}_n\mathbf{Q}_n\mathbf{P}_n = \mathbf{P}_n\mathbf{Q}_n\mathbf{D}_n\mathbf{Q}_n\mathbf{P}_n \ .$$

Then, the variance of the second term is

$$\begin{aligned}\mathbf{Var}[X_{2,n-h}] &= \sigma_v^2\sum_{j=1}^{m}[\cos\frac{2\pi}{2n+1}(n-h-\frac{1}{2})(j-\frac{1}{2})]^2\times 4\sin^2[\frac{\pi}{2}\frac{2j-1}{2n+1}] \\ &= 4\times\sum_{j=1}^{m}[\sin\frac{\pi(h+1)}{2n+1}(2j-1))]^2\times\sin^2[\frac{\pi}{2}\frac{2j-1}{2n+1}] \\ &= O(\sum_{j=1}^{m}(\frac{j}{n})^4) = O(\frac{m^5}{n^4}) \ .\end{aligned}$$

Hence, for their variance of X_{n-h} the first term is dominant and the order of the variance becomes

$$\mathbf{Var}[W_{f,n}] \sim (\frac{2}{n})^2[\mathbf{Var}[X_{1,n-h}]\times[\frac{3}{2\pi^2(h+1)^2}(\frac{n}{m})^3] \sim \sigma_{\Delta x}^2$$

when $n, m\ (= m_n) \to \infty$ and $m/n \to 0$.
By using CLT to $\sqrt{\frac{n^3}{m^3}}X_{1,n-h} = \sum_{i=1}^{n}a_i^{(m,n)}v_i^{(x)}$ since the coefficients $a_t^{(m,n)}$ $t = 1,\cdots,n$ are small as $n, m\to\infty$ while $m_n \to 0$,, we have the asymptotic normality. (**Q.E.D.**)

Proof of Proposition 6.2 :
We use the simple relations $2(n-h-\frac{1}{2}) = 2n+1-2(h+1)$, $2(n-\frac{1}{2}) = 2n+1-2$ and the trigonometric relations such as $\cos\frac{2\pi}{2n+1}[2n+1-2(h+1)]$ for a fixed $h\ge 0$

and $j = 1, \cdots, m$. Then, we find that

$$
\begin{aligned}
& \cos \frac{2\pi}{2n+1}(n-h-\frac{1}{2})(j-\frac{1}{2}) - \cos \frac{2\pi}{2n+1}(n-\frac{1}{2})(j-\frac{1}{2})] \\
\sim \quad & [(-1)^{j+1}\pi(h+1)\frac{2j-1}{2n+1}] - [(-1)^{j+1}\pi\frac{2j-1}{2n+1}] \\
= \quad & O(\frac{m}{n}) \ .
\end{aligned}
$$

Thus, we have the result of Proposition 3.2 because the limiting random variables $W_{f,n-k} (k = 0, \cdots, h)$ are degenerated.
(**Q.E.D.**)

Proof of Proposition 6.3 :
For the backward filtering, we use a similar evaluation as Proposition 6.1. We consider the case when $\Delta x_i = x_i - x_{i-1} = v_i^{(v)}$ and v_i $(i = 1, \cdots, n)$ are i.i.d. sequences. We set the initial conditions $x_n = v_n = 0$ for the convenience and take h as a fixed positive integer.
Let $X_{n-h} = X_{3,n-h} + X_{4,n-h}$ $(h = 1, \cdots, n)$, where

$$
X_{3,n-h} = \sum_{j=1}^{n} [\sum_{k=1}^{m} \sin \frac{2\pi}{2n+1}(k-\frac{1}{2})(n-h) \sin \frac{2\pi}{2n+1}(k-\frac{1}{2})j] v_j^{(x)}
$$

and

$$
X_{4,n-h} = \sum_{t=1}^{n} \sum_{k=1}^{m} [\sin \frac{2\pi}{2n+1}(n-h)(k-\frac{1}{2}) \sin \frac{2\pi}{2n+1}(k-\frac{1}{2})t] \Delta v_t \ .
$$

Then, by using the orthogonal relation

$$
\sum_{i=1}^{n} \sin[\frac{2\pi}{2n+1}(k-\frac{1}{2})i] \times \sin[\frac{2\pi}{2n+1}(k^{'}-\frac{1}{2})i] = \delta(k,k^{'})[\frac{n}{2}] \ ,
$$

we find that

$$
\mathbf{Var}[X_{3,n-h}] = \sigma^{(x)2}[\frac{n}{2}] \sum_{k=1}^{m} [\sin \frac{2\pi}{2n+1}(k-\frac{1}{2})(n-h)]^2 \ .
$$

We use the trigonometric relation $\sin \frac{\pi}{2n+1}[2n+1-(2h+1)](k-\frac{1}{2}) = \sin \pi(k-\frac{1}{2}) \cos \pi(2h+1)\frac{k-1/2}{2n+1}$. When $m/n \to 0$ as $n \to \infty$,

$$
\sum_{k=1}^{m} [\cos \pi \frac{2h+1}{2n+1}(k-\frac{1}{2})]^2 \sim m
$$

and

$$
\begin{aligned}
\mathbf{Var}[X_{4,n-h}] \quad &= \quad O(\sum_{j=1}^{m} [\cos \frac{\pi}{2n+1}(2h+1)(j-\frac{1}{2}))]^2 \times \sin^2[\frac{\pi}{2}\frac{2j-1}{2n+1}]) \\
&= \quad \sum_{j=1}^{m} (\frac{j}{n})^2 = O(\frac{m^3}{n^2}) \ .
\end{aligned}
$$

Then, for a fixed h ($h \geq 0$)

$$\mathbf{Var}[\hat{\Delta}^{(b)} x_{n-h}] \sim \sigma^{(x)2} \frac{4}{2n+1} m \sim \sigma^{(x)2} \frac{2m}{n} \ .$$

By using CLT to $\sqrt{\frac{n^3}{m^3}} X_{3,n-h}$ since the coefficients $a_t^{(m,n)}$ $t = 1, \cdots, n$ are small as $n, m \to \infty$ while $m_n \to 0$, we have the asymptotic normality.
(**Q.E.D.**)

Proof of Theorem 6.4 :
In the more general case, we assume that

$$r_i = \Delta x_i = \sum_{s=-\infty}^{\infty} \gamma_s v_{i-s}^{(x)} \ ,$$

where (a) $\sum_{s=-\infty}^{\infty} |\gamma_s| < \infty$ and (b) $\sup \int_{|v_t|>c} v^2 dF(v) \longrightarrow 0$ ($c \longrightarrow \infty$).
Let $r_{i,k} = \sum_{s=-k}^{k} \gamma_s v_{i-s}^{(x)}$. Then, by using the method of Anderson (1971, Page 482), we can apply the CLT to the truncated sum to obtain the asymptotic normality because the difference of r_i and $r_{i,k}$ is stochastically negligible as $k \to \infty$.
For a fixed non-negative integer h, let $X_{n-h}^* = X_{1,n-h}^* + X_{2,n-h}^*$ ($h = 1, \cdots, n$), where

$$\begin{aligned} X_{1,n-h}^* &= \sum_{i=1}^{n} \sum_{j=1}^{m} A(n-h,i,j) r_{i,k} \\ &= \sum_{i=1}^{n} \sum_{s=-k}^{k} \gamma_s \sum_{j=1}^{m} A(n-h,i,j) v_{i-s} \\ &= \sum_{i=k+1}^{n-k} \sum_{s=-k}^{k} \gamma_s \sum_{j=1}^{m} A(n-h,i+s,j) v_i^{(x)} + \sum_{i=1-k}^{k} \sum_{s=-i+1}^{k} \gamma_s \sum_{j=1}^{m} A(n-h,i+s,j) v_i^{(x)} \\ &\qquad + \sum_{i=n-k+1}^{n+k} \sum_{s=-k}^{n-i} \gamma_s \sum_{j=1}^{m} A(n-h,i+s,j) v_i^{(x)} \\ &= W_{f,1n} + W_{f,2n} + W_{f,3n} \ \text{(say)} \ , \end{aligned}$$

where $A(n-h,i,j) = \cos \frac{2\pi}{2n+1}(n-h-\frac{1}{2})(j-\frac{1}{2}) \cos \frac{2\pi}{2n+1}(j-\frac{1}{2})(i-\frac{1}{2})$.
Then, the second and third terms are stochastically negligible because of the evaluation of their variances. We use the relation that for $j = 1, \cdots, m$ and $s, s^{'} = -k, \cdots, k$(k is fixed)

$$\cos \frac{2\pi}{2n+1}(j-\frac{1}{2})(i+s^{'}-\frac{1}{2}) = \cos \frac{2\pi}{2n+1}(j-\frac{1}{2})(i+s-\frac{1}{2}) + O(\frac{1}{n})$$

and the variance of the first term becomes approximately as $m/n \to 0$,

$$\begin{aligned} \mathbf{Var}[X_{1,n-h}] &= (\sum_{s=-k}^{k} \gamma_s)^2 \sigma^{(x)2} [\frac{n}{2} + \frac{1}{4}] \sum_{j=1}^{m} [\cos \frac{2\pi}{2n+1}(n-h-\frac{1}{2})(j-\frac{1}{2})]^2 \\ &\sim (\sum_{s=-k}^{k} \gamma_s)^2 \sigma^{(x)2} [\frac{\pi^2 (h+1)^2}{6}] \frac{m^3}{n} \ . \end{aligned}$$

Because the spectral density at zero is given by $f_{\Delta x}(0) = (\sum_{s=-\infty}^{\infty} \gamma_s)^2 \sigma^{(x)2}$, we have the result by taking $k \to \infty$.
(**Q.E.D**)

Derivation of (6.8) : Let $y_n^{(f)}$ be the estimate of y_n by using the forward filter method. It can be given as

$$y_n^{(f)} = \mathbf{e}_n^{'}[\mathbf{C}_n \mathbf{P}_n \mathbf{Q}_n \mathbf{P}_n \mathbf{C}_n^{-1}(\mathbf{y} - \mathbf{1}_n y_0) + \mathbf{1}_n y_n] ,$$

Then

$$\begin{aligned} y_n^{(f)} - y_n &= -\mathbf{1}_n^{'}(\mathbf{I}_n - \mathbf{P}_n \mathbf{Q}_n \mathbf{P}_n) \begin{bmatrix} y_1 - y_0 \\ y_2 - y_1 \\ \vdots \\ y_n - y_{n-1} \end{bmatrix} \\ &= -\mathbf{1}_n^{'} \mathbf{P}_n \bar{\mathbf{Q}}_n \mathbf{P}_n \begin{bmatrix} y_1 - y_0 \\ y_2 - y_1 \\ \vdots \\ y_n - y_{n-1} \end{bmatrix} \end{aligned}$$

where $\bar{\mathbf{Q}}_n = \mathbf{I}_n - \mathbf{Q}_n$.
The forward-backward estimate of x_t at $t = n - h$ is given by

$$\begin{aligned} y_{n-h}^{(s)} &= \mathbf{e}_{n-h}^{'} \mathbf{C}_n^{'} \mathbf{P}_n^{'*} \mathbf{Q}_n \mathbf{P}_n^{*} \mathbf{C}_n^{'-1}(\mathbf{Y}_n^{*} - \mathbf{1}_n y_n^{(f)}) \\ &= \mathbf{e}_{n-h}^{'} \mathbf{C}_n^{'} \mathbf{P}_n^{'*} \mathbf{Q}_n \mathbf{P}_n^{*} \mathbf{C}_n^{'-1}(\mathbf{Y}_n^{*} - \mathbf{1}_n y_n) + \mathbf{e}_{n-h}^{'} \mathbf{C}_n^{'} \mathbf{P}_n^{'*} \mathbf{Q}_n \mathbf{P}_n^{*} \mathbf{C}_n^{'-1} \mathbf{1}_n (y_n - y_n^{(f)}) . \end{aligned}$$

We set $X_{n-h}^{*} = X_{1,n-h}^{*} + X_{2,n-h}$ for

$$X_{1,n-h}^{*} = \mathbf{e}_{n-h}^{'} \mathbf{P}_n^{'*} \mathbf{Q}_n \mathbf{P}_n^{*}(-1)\mathbf{v}^{(x)} - \mathbf{e}_{n-h}^{'} \mathbf{P}_n^{'*} \mathbf{Q}_n \mathbf{P}_n^{*} \mathbf{e}_n \mathbf{1}_n^{'} \mathbf{P}_n \bar{\mathbf{Q}}_n \mathbf{P}_n \mathbf{v}^{(x)}$$

and

$$X_{2,n-h}^{*} = \mathbf{e}_{n-h}^{'} \mathbf{P}_n^{'*} \mathbf{Q}_n \mathbf{P}_n^{*}(-1)\Delta \mathbf{v} - \mathbf{e}_{n-h}^{'} \mathbf{P}_n^{'*} \mathbf{Q}_n \mathbf{P}_n^{*} \mathbf{e}_n \mathbf{1}_n^{'} \mathbf{P}_n \bar{\mathbf{Q}}_n \mathbf{P}_n \Delta \mathbf{v} ,$$

where $\mathbf{v}^{(x)} = (v_t^{(x)})$ and $\mathbf{v} = (v_t)$.
Because $\mathbf{Q}_n$ ane $\bar{\mathbf{Q}}_n$ are idempotent matrices, we can use the relation that for $\mathbf{d}_{n-h}^{'} = \mathbf{e}_{n-h}^{'} \mathbf{P}_n^{'*} \mathbf{Q}_n \mathbf{P}_n^{*} + \mathbf{e}_{n-h}^{'} \mathbf{P}_n^{'*} \mathbf{Q}_n \mathbf{P}_n^{*} \mathbf{e}_n \mathbf{1}_n^{'} \mathbf{P}_n \bar{\mathbf{Q}}_n \mathbf{P}_n$,

$$\mathbf{d}_{n-h}^{'} \mathbf{d}_{n-h} = \mathbf{e}_{n-h}^{'} \mathbf{P}_n^{'*} \mathbf{Q}_n \mathbf{P}_n^{*} \mathbf{e}_{n-h} + \mathbf{1}_n^{'} \mathbf{P}_n \bar{\mathbf{Q}}_n \mathbf{P}_n \mathbf{1}_n [\mathbf{e}_{n-h}^{'} \mathbf{P}_n^{'*} \mathbf{Q}_n \mathbf{P}_n^{*} \mathbf{e}_n]^2 .$$

Then by calculating the term of $\mathbf{e}_{n-h}^{'} \mathbf{P}_n^{'*} \mathbf{Q}_n \mathbf{P}_n^{*} \mathbf{e}_n$ as

$$\begin{aligned} & \sum_{k=1}^{n} \sin \frac{2\pi}{2n+1}(k - \frac{1}{2})(n-h) \sin \frac{2\pi}{2n+1}(k - \frac{1}{2})n \\ = \ & (-\frac{1}{4}) \sum_{k=1}^{n} [e^{i\frac{2\pi}{2n+1}(k-\frac{1}{2})(n-h)} - e^{-i\frac{2\pi}{2n+1}(k-\frac{1}{2})(n-h)}][e^{i\frac{2\pi}{2n+1}(k-\frac{1}{2})n} - e^{-i\frac{2\pi}{2n+1}(k-\frac{1}{2})n}] \end{aligned}$$

$$
\begin{aligned}
&= (-\frac{1}{4})\sum_{k=1}^{n}\Big[e^{i\frac{2\pi}{2n+1}(k-\frac{1}{2})[2n+1-(h+1)]} + e^{-i\frac{2\pi}{2n+1}(k-\frac{1}{2})[2n+1-(h+1)]} \\
&\qquad -e^{i\frac{2\pi}{2n+1}(k-\frac{1}{2})h} - e^{-i\frac{2\pi}{2n+1}(k-\frac{1}{2})h}\Big] \\
&= (-\frac{1}{4})\Bigg[e^{i\frac{2\pi}{2n+1}\frac{1}{2}[2n+1-(h+1)]}\frac{1-e^{i\frac{2\pi}{2n+1}\frac{1}{2}[2n+1-(h+1)]m}}{1-e^{i\frac{2\pi}{2n+1}\frac{1}{2}[2n+1-(h+1)]}} \\
&\qquad +e^{-i\frac{2\pi}{2n+1}\frac{1}{2}[2n+1-(h+1)]}\frac{1-e^{-i\frac{2\pi}{2n+1}\frac{1}{2}[2n+1-(h+1)]m}}{1-e^{-i\frac{2\pi}{2n+1}\frac{1}{2}[2n+1-(h+1)]}} \\
&\qquad -e^{i\frac{2\pi}{2n+1}\frac{1}{2}h}\frac{1-e^{i\frac{2\pi}{2n+1}hm}}{1-e^{i\frac{2\pi}{2n+1}h}} - e^{-i\frac{2\pi}{2n+1}\frac{1}{2}h}\frac{1-e^{-i\frac{2\pi}{2n+1}hm}}{1-e^{-i\frac{2\pi}{2n+1}h}}\Bigg] ,
\end{aligned}
$$

which is asymptotically equivalent to

$$
(-\frac{1}{4})[-\frac{\sin\frac{2\pi}{2n+1}2hm}{\sin\frac{2\pi}{2n+1}h} - \frac{\sin\frac{2\pi}{2n+1}hm}{\sin\frac{2\pi}{2n+1}\frac{h}{2}}] .
$$

When $m.n \to \infty$ and $m/n \to 0$, it is asymptotically equivalent to m. Also for the term of $\mathbf{1}_n^{'}\mathbf{P}_n\bar{\mathbf{Q}}_n\mathbf{P}_n\mathbf{1}_n$, we find that

$$
\sum_{j,,j^{'}=1}^{n}\sum_{k=m+1}^{n}\cos\frac{2\pi}{2n+1}(j-\frac{1}{2})(k-\frac{1}{2})\cos\frac{2\pi}{2n+1}(j^{'}-\frac{1}{2})(k-\frac{1}{2}) \sim \sum_{k=m+1}^{n}\left[\frac{1}{2}\frac{\cos\frac{\pi}{2}\frac{2k-1}{2n+1}}{\sin\frac{\pi}{2}\frac{2k-1}{2n+1}}\right]^2
$$

because

$$
\begin{aligned}
&\sum_{j=1}^{n}\cos\frac{2\pi}{2n+1}(j-\frac{1}{2})(k-\frac{1}{2}) \\
&= \frac{1}{2}\sum_{j=1}^{n}[e^{i\frac{2\pi}{2n+1}(j-\frac{1}{2})(k-\frac{1}{2})} + e^{-i\frac{2\pi}{2n+1}(j-\frac{1}{2})(k-\frac{1}{2})}] \\
&= \frac{1}{2}\left[e^{i\frac{2\pi}{2n+1}\frac{1}{2}(k-\frac{1}{2})}\frac{1-e^{i\frac{2\pi}{2n+1}n(k-\frac{1}{2})}}{1-e^{i\frac{2\pi}{2n+1}(k-\frac{1}{2})}} + e^{-i\frac{2\pi}{2n+1}\frac{1}{2}(k-\frac{1}{2})}\frac{1-e^{-i\frac{2\pi}{2n+1}n(k-\frac{1}{2})}}{1-e^{-i\frac{2\pi}{2n+1}(k-\frac{1}{2})}}\right] \\
&= \frac{1}{2}\frac{\sin\frac{2\pi}{2n+1}n(k-\frac{1}{2})}{\sin\frac{2\pi}{2n+1}\frac{1}{2}(k-\frac{1}{2})} \\
&= (-1)^{k+1}\frac{1}{2}\frac{\cos\frac{\pi}{2}\frac{2k-1}{2n+1}}{\sin\frac{\pi}{2}\frac{2k-1}{2n+1}} .
\end{aligned}
$$

Then, by using the fact that when $n, m_1 \to \infty$ and $m_1/n \to c$ ($0 \le c <\le 1$)

$$
\frac{1}{n}\sum_{k=m_1+1}^{n}\left[\frac{\cos\frac{\pi}{2}\frac{2k-1}{2n+1}}{\sin\frac{\pi}{2}\frac{2k-1}{2n+1}}\right]^2 \sim \int_c^1\left[\frac{\cos\frac{\pi}{2}z}{\sin\frac{\pi}{2}z}\right]^2 dz ,
$$

we have the result.

(**Q.E.D**)