SDS-15

On Backward Smoothing for Noisy Non-stationary Time Series

Seisho Sato and Naoto Kunitomo

June 2020

Statistics & Data Science Series back numbers:
http://www.mims.meiji.ac.jp/publications/datascience.html

On Backward Smoothing for Noisy Non-stationary Time Series *

Seisho Sato †

and

Naoto Kunitomo ‡

June 6, 2020

Abstract

We investigate a new smoothing method to estimate the hidden states of random variables and to handle multiple noisy non-stationary time series data. Kunitomo and Sato (2019) utilized the results of Kunitomo, Sato and Kurisu (2018) to solve the smoothing problem of hidden random variables and the resulting method gives a new way to handle multivariate non-stationary time series. We continue to investigate the filtering problem and in particular, we investigate *the backward SIML smoothing method* to solve the initial value issue. It is possible to interpret the existing filtering methods in the time and frequency domains.

Key Words

Non-stationary mutivariate economic time series, Measurement Error and Errors-variables models, SIML-backward-smoothing, Backward Smoothing, Initial Value Problem

---

*2020-6-5. A preliminary memorandum. The research has been supported by Research Grant JP17H02513.

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

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

1. Introduction

We investigate a new smoothing method to estimate the hidden states of random variables and to handle multiple noisy non-stationary time series data, and particularly to deal with small sample economic time series. Kunitomo and Sato (2019), (see Kunitomo, Sato and Kurisu (2018) also) have developed the separating information maximum likelihood (SIML) method for estimating the non-stationary errors-in-variables models. They have discussed the asymptotic properties and finite sample properties of the estimation of unknown parameters. Kunitomo and Sato (2019) utilized their results to solve the smoothing or filtering problem of hidden random variables, which gives a new estimation method to handle macro-economic time series. In this paper, we continue to investigate the smoothing or filtering problem and in particular, we develop *the backward SIML smoothing method* to solve the initial value issue in the procedure. From our analysis, it is possible to interpret the existing smoothing and filtering methods in the time domain and frequency domain. Although some econometrician may not distinguish smoothing from filtering and the latter terminology has been sometimes used, we shall use *smoothing* mainly in stead of *filtering* in this paper.

There exists vast published research on the use of statistical time series analysis for macro-economic time series, which have the non-stationary trend, cycle, seasonal. and measurement errors. For statistical filtering and smoothing methods, Kitagawa (2010) discussed the standard statistical methods already known, including the Kalman-filtering and particle-filtering methods. Although many studies have examined statistical filtering theories, we must exercise caution in analyzing non-stationary multivariate time series. The existing methods often depend on the underlying distributions such as the Gaussian distributions for the Kalman-filtering, and the procedure essentially depends on the dimension of state variables, there may be some difficulty in extending the existing methods to high-dimension cases, even when the dimension is about 10. On the other hand, we expect that our method is simple and has some merits when handling small sample economic times series with non-stationarity and seasonality with many variables, because our method does not depend on the specific distributions as well as the dimension of the underlying random variables. See Kunitomo, Awaya and Kurisu (2017) for a comparison of small sample properties of the ML (maximum likelihood) and SIML estimation methods for the non-stationary errors-in-variables models, and Nishimura, Sato and Takahashi (2019) for an application of financial data smoothing. The most important feature of the present procedure is that it may be applicable to small sample time series data.

In Kunitomo and Sato (2019) there is an implicit assumption that we can handle the initial value problem in smoothing or filtering. In the non-stationary time series, however, the initial value of state estimate plays a crucial role in the resulting estimates of unobservable state vectors subsequently and then we need to investigate this problem in a systematic way. The initial motivation of the present paper to resolve this problem rather nicely. In fact we shall show that it is possible to develop backward smoothing and iterative smoothing procedures, and we have their convergence. Furthermore, some related issues arise and we shall develop the multi-step smoothing and band-smoothing, which seem to be new in this paper. As we will see, it seems that they are related to the general problem in the analysis of non-stationary time series data.

In Section 2, we explain the non-stationary errors-in-variables model and the SIML method. Then in section 3, we develop the SIML-filtering methods including the forward, backward, and multi-step smoothing procedures. We give a theoretical result of convergence of the smoothing or filtering method for the initial value problem and discuss the evaluation criteria. Then in section 4, we discuss generalizations of the non-stationary errors-in-variables model and a mathematical interpretation of our procedure. In Section 5, we give some numerical example and in section 6, some conclusions are given. Some details of mathematical derivations and figures are given in Appendix.

2. Non-stationary Errors-in-variables models

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

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

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

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

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

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

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

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

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

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

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

where

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

and

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

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

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

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

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

Given the initial condition, the log-likelihood function except some constants when the underlying distributions are Gaussian can be written as

$$L_n(\boldsymbol{\theta}) = \sum_{k=1}^{n} \log|a_{kn}^*\mathbf{\Sigma}_v + \mathbf{\Sigma}_x|^{-1/2} - \frac{1}{2}\sum_{k=1}^{n} \mathbf{z}_k^{'}[a_{kn}^*\mathbf{\Sigma}_v + \mathbf{\Sigma}_x]^{-1}\mathbf{z}_k \tag{2.10}$$

and

$$(-2)L_n(\boldsymbol{\theta}) = \sum_{k=1}^{n} \log|a_{kn}^*\mathbf{\Sigma}_v + \mathbf{\Sigma}_x| + \sum_{k=1}^{n} \mathbf{z}_k^{'}[a_{kn}^*\mathbf{\Sigma}_v + \mathbf{\Sigma}_x]^{-1}\mathbf{z}_k , \tag{2.11}$$

where $\boldsymbol{\theta}$ is a vector of parameters.

The model of (2.1) and (2.2) can be generalized to the cases when we have cycle and seasonal components, and when $\mathbf{v}_t$ and $\Delta \mathbf{x}_t$ are auto-correlated (see Kunitomo and Sato (2019)). In this paper, however, we first focus on the smoothing or filtering procedure of the simple non-stationary multiple time series. We shall discuss several extensions in Section 4 briefly.

3. SIML Smoothing and Backward Smoothing

3.1 Forward SIML Smoothing

Kunitomo and Sato (2019) investigated the general filtering procedure based on the $\mathbf{K}_n-$transformation. When we interpret that the elements of the resulting $n \times p$ random matrix $\mathbf{Z}_n$ by this transformation take real values in the frequency domain, it is easy to understand their roles. Since $\mathbf{P}_n$ is a kind of real-valued discrete Fourier transformation, vectors $\mathbf{z}_k$ $(k = 1, \cdots, n)$ in $\mathbf{Z}_n$ are asymptotically uncorrelated. We consider the partial inversion of the transformed orthogonal processes. Let an $n \times p$ matrix

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

and

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

where $\mathbf{X}_n^* = (\mathbf{x}_i^{*'})$ $\mathbf{V}_n = (\mathbf{v}_i^{'})$ are $n \times p$ matrices, $\mathbf{x}_i^* = \mathbf{x}_i - \mathbf{x}_0$ and 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 $\mathbf{Y}_n$ consist of non-stationary time series, we need a special form of transformation $\mathbf{K}_n$ in (3.13). We give explicit form for the trend smoothing (or filtering) procedure. Let an $m \times n$ choice matrix $\mathbf{J}_m = (\mathbf{I}_m, \mathbf{O})$, and let also $n \times p$ matrix

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

and an $n \times n$ matrix $\mathbf{Q}_n = \mathbf{J}_m^{'} \mathbf{J}_m$.

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 deleting the estimated noise parts (see Nishimura, Sato and Takahashi (2019)). We denote the hidden trend state as

(3.4) $$\mathbf{X}_n(m) = \mathbf{C}_n \mathbf{P}_n \mathbf{J}_m^{'} \mathbf{J}_m \mathbf{P}_n \mathbf{C}_n^{-1} \mathbf{X}_n \ .$$

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, which is different from the measurement noise component $\mathbf{v}_i$ $(i = 1, \cdots, n)$ of (3.1) and (3.3). We try to estimate the trend component of $\mathbf{x}_i$ by using (3.3) and recover the trend component of $\mathbf{X}_n$ near at the zero frequency because the effects of differenced measurement error noises $(\mathbf{v}_i - \mathbf{v}_{i-1})$ are negligible around at zero frequency. This method differs from some existing procedures that consider the decomposition of 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), and we will discuss this issue in Section 4 (see Section 5.2 of Kunitomo and Sato (2019)).

3.2 A Backward Smoothing

We reconsider the role of initial condition in the non-stationary process. We take $n \times p$ 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

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

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

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

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

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

and

$$\text{(3.8)} \qquad \mathbf{P}_n^* = (p_{jk}^{*(n)}) , \ p_{jk}^{*(n)} = \sqrt{\frac{2}{n + \frac{1}{2}}} \sin\left[\frac{2\pi}{2n+1}(k - \frac{1}{2})j\right] .$$

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

$$\text{(3.9)} \qquad a_{kn}^* \ (= d_k) \ = 4 \sin^2\left[\frac{\pi}{2}\left(\frac{2k-1}{2n+1}\right)\right] \ (k = 1, \cdots, n) .$$

(See Appendix for derivations.)

[1] Given the initial condition $\mathbf{y}_n$ we consider the joint distribution of $(\mathbf{y}_{n-1}^{'}, \cdots, \mathbf{y}_0^{'})^{'}$, where we take (2.1) for $i = 0, \cdots, n-1$.

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

$$\hat{\mathbf{X}}_n^*(\mathrm{Q}_n) = \mathbf{C}_n^{'}\mathbf{P}_n^{*'}\mathbf{Q}_n\mathbf{P}_n^*\mathbf{C}_n^{'-1}(\mathbf{Y}_n^* - \bar{\mathbf{Y}}_n^*) \tag{3.10}$$

and

$$\mathbf{Z}_n^* = \mathbf{P}_n^*\mathbf{C}_n^{'-1}(\mathbf{Y}_n^* - \bar{\mathbf{Y}}_n^*) \ , \mathbf{Y}_n^* = \bar{\mathbf{Y}}_n^* + \mathbf{X}_n^* + \mathbf{V}_n^* \ , \tag{3.11}$$

where $\mathbf{X}_n^* = (\mathbf{x}_{i-1}^{*'})$ and $\mathbf{V}_n^* = (\mathbf{v}_{i-1}^{*'})$ are $n \times p$ matrices.
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 $\mathbf{K}_n^{**}$. We give explicit form for the trend filtering procedure and let also $n \times p$ matrix

$$\hat{\mathbf{X}}_n^*(m) = \mathbf{C}_n^{'}\mathbf{P}_n^{*'}\mathbf{J}_m^{'}\mathbf{J}_m\mathbf{P}_n^*\mathbf{C}_n^{'-1}(\mathbf{Y}_n^* - \bar{\mathbf{Y}}_n^*) \tag{3.12}$$

and an $n \times n$ matrix $\mathbf{Q}_n = \mathbf{J}_m^{'}\mathbf{J}_m$.
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 deleting the estimated noise parts. We denote the hidden trend state as

$$\mathbf{X}_n^*(m) = \mathbf{C}_n^{'}\mathbf{P}_n^{*'}\mathbf{J}_m^{'}\mathbf{J}_m\mathbf{P}_n^*\mathbf{C}_n^{'-1}\mathbf{X}_n^* \ . \tag{3.13}$$

3.3 Initial Value Problem and Convergence

When we have non-stationary time series observations that follows a random walk as statistical model, the role of initial value is important because of non-stationarity. This aspect is different from stationary time series models, in which the effects of initial value are negligible when the sample size is large. Hence it is important to have smoothing or filtering procedure of non-stationary time series, that does not depend much on the initial value. As the initial value, there can be two possibilities as $\mathbf{y}_1$ and $\mathbf{y}_n$ when we have n observations $\mathbf{y}_i$ $(i = 1, \cdots, n)$. On this problem we have an interesting useful result.

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

$$T_{2k+1}^{(m,n)}(\mathbf{y}) = \mathbf{C}_n\mathbf{P}_n\mathbf{Q}_n^{(m)}\mathbf{P}_n\mathbf{C}^{-1}[\mathbf{y} - \mathbf{1}_n(\mathbf{e}_1^{'}T_{2k}^{(m,n)}(\mathbf{y}))^{'}] + \mathbf{1}_n(\mathbf{e}_1^{'}T_{2k}^{(m,n)}(\mathbf{y}))^{'}, \tag{3.14}$$

and

$$T_{2k}^{(m,n)}(\mathbf{y}) = \mathbf{C}_n^{'}\mathbf{P}_n^{*}{}^{'}\mathbf{Q}_n^{(m)}\mathbf{P}_n^*\mathbf{C}^{'-1}[\mathbf{y} - \mathbf{1}_n(\mathbf{e}_n^{'}T_{2k-1}^{(m,n)}(\mathbf{y}))^{'}] + \mathbf{1}_n(\mathbf{e}_n^{'}T_{2k-1}^{(m,n)}(\mathbf{y}))^{'}, \tag{3.15}$$

where $\mathbf{Q}_n^{(m)} = \mathbf{J}_m^{'}\mathbf{J}_m$, $\mathbf{e}_1^{'} = (1, 0, \cdots, 0)$ and $\mathbf{e}_n^{'} = (0, \cdots, 0, 1)$ are unit vectors.
The operator $T_{2k+1}^{(m,n)}$ $(k = 1, 2, \cdots)$ is the SIML filering with the initial value $\mathbf{y}_1$ and

$T_{2k}^{(m,m)}$ $(k = 1, 2, \cdots)$ is the revesed filtering. For non-stationary time series two operators have different meaning.
Then we have the next proposition on the convergence of smoothing rocedures. The proof is given in the Appendix.

Theorem 3.1 : As $k \to \infty$, there exists n_0 such that for $n_0 < n$ and $m < n$, we have

$$T_{2k+1}^{(m,n)} \to T_{1*}^{(m,n)} = \sum_{s=0}^{\infty} (\mathbf{A}_2^{(m,n)})^s \mathbf{A}_1^{(m,n)} , \tag{3.16}$$

and

$$T_{2k}^{(m,n)} \to T_{2*}^{(m,n)} = \sum_{s=0}^{\infty} (\mathbf{A}_{2*}^{(m,n)})^s \mathbf{A}_{1*}^{(m,n)} , \tag{3.17}$$

where

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

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

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

Since the initial value is the starting point of non-stationary time series, we need to develop some smoothing procedure, which does not depend on the initial value. For the practical purpose, often we do want to use the procedure, which does not depend on the latest observation $\mathbf{y}_n$. In this case, it may be reasonable to use the $T_2^{(n)}$.
The formulation of two filtering in this subsection is slightly different from Section 3.2 (we use $n + 1$ observations) because we have used n observations. It may be natural for repeating smoothing and this aspect of difference is negligible when n is large.

We have the next result on the backward SIML smoothing. The proof is given in the Appendix.

Theorem 3.2 : Assume $m/n \to 0$ as $n \to \infty$. Then,

$$\|\mathbf{C}_n^{'} \mathbf{P}_n^{*'} \mathbf{Q}_n^{(m)} \mathbf{P}_n^{*} \mathbf{C}^{'-1} - \mathbf{P}_n \mathbf{Q}_n^{(m)} \mathbf{P}_n\| \to 0 \tag{3.18}$$

$n \to \infty$, where the norm of a matrix $\mathbf{A} = (a_{ij})$ $(n \times n)$ is defined by $\|\mathbf{A}\| = \max_{i,j=1,\cdots,n} |a_{ij}|$.

From this representation, the reversed SIML smoothing is essentially the same smoothing with a real (finite, and discrete) Fourier transformation if we take the time is reversed from n to 1 instead of 1 to n.

3.4 Band Smoothing

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

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

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, 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 quarterly data and $s = 12$ for monthly data.
As in the trend smoothing problem, the SIML-filtering value for

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

and it is an estimate of some frequency components of $\mathbf{x}_i$ $(i = 1, \cdots, n)$.
In this case we can define T_{2k-1} and T_{2k} for $k = 1, \cdots M$ as (3.14) and (3.15) by using $\mathbf{J}_{m_1,m_2}$ instead of $\mathbf{J}_m$. Then it is straightforward to to find the next proposition on the convergence of smoothing rocedures and the proof is in the Appendix.

Theorem 3.3 : As $k \to \infty$, there exists n_0 such that for $n_0 < n$ we have

$$T_{2k+1}^{(m_1,m_2,n)} \to T_{1*}^{(m_1,m_2,n)} = \sum_{s=0}^{\infty} (\mathbf{A}_2^{(m_1,m_2,n)})^s \mathbf{A}_1^{(m_1,m_2,n)} \ , \tag{3.21}$$

and

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

where

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

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

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

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

Theorem 3.1 is a special case of Theorem 3.3 when $m_1 = 0$ and $m_2 = m$.

3.5 Multi-Step Smoothing

In the forward and backward smoothing procedures, it is important to choose an appropriate m. In some applications, however, the problem becomes difficult when there are some seasonal components. However, it may be possible to run the forward and backward smoothing several times, which may be called *multi-stage smoothing*.

Let $T_{1*}^{(m_1,n)}$ be the first stage forward smoothing with a specific choice of m_1. Then we can define the double-stage forward smoothing by

$$
T_{1,1}^{(m_1,m_2,n)} = T_{1*}^{(m_1,n)}T_{1*}^{(m_2,n)} \ . \tag{3.23}
$$

Similarly, we can define the double-stage backward smoothing by

$$
T_{2,2}^{(m_1,m_2,n)} = T_{2*}^{(m_1,n)}T_{2*}^{(m_2,n)} \ . \tag{3.24}
$$

There can be more complicated smoothing procedures. Then we need some criterion to find an appropriate smoothing procedure for applications. It may be possible to deal with complicated seasonal patterns in the frequency domain because we first take a rather large m_1 and then we take a smaller m_2, for instance.

For real applications, it may not be clear to find an appropriate m or m_1 and m_2 at the beginning. One strategy in the trend estimation would be to choose a relatively large m_1, which should be less than the seasonality frequency, at the initial stage. Then, at the second stage, we choose m_2, which is smaller than m_1 and use the following evaluation criterion.

3.6 On 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. Since our procedure does not assume a particular distribution such as

Gaussianity and semi-parametric, it looks a challenging one. As we shall discuss in the next section (see Kunitomo and Sato (2019)), there is a natural way to handle the problem, which will be illustrated by the forward filtering case. (There is a similar argument for the backward smoothing case.)

Let $\mathbf{r}_j^{(n)} = \mathbf{y}_j^{(n)} - \mathbf{y}_{j-1}^{(n)}$ $(j = 1, \cdots, n)$ and we write

$$\hat{\mathbf{r}}_j^{(n)} = \sum_{k=1}^{n} p_{jk}\mathbf{z}_k \ , \tag{3.25}$$

where $\mathbf{z}_k^*$ is the orthogonal process at the frequency $\lambda_k^{(n)} = (k - 1/2)/(2n+1)$ $(k = 1, \cdots, n)$ (see Section 5 of Kunitomo and Sato (2019)).
Then for $h \geq 1$, it may be natural to use the predictor

$$\hat{\mathbf{r}}_{n+h}^{(n)}(m) = \sum_{k=1}^{m} p_{n+h,k}\mathbf{z}_k \ , \tag{3.26}$$

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

$$\hat{\mathbf{x}}_{n+h}^{(n)}(m) = \sum_{s=h+1}^{n+h} \hat{\mathbf{r}}_s^{(n)}(m) = \sum_{s=h+1}^{n+h} \sum_{k=1}^{m} p_{sk}\mathbf{z}_k \ . \tag{3.27}$$

By using (3.1) and (3.2), the prediction error can be written as

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

We use an elementary relation that

$$\sum_{s=h+1}^{n+h} p_{sk} = \frac{1}{\sqrt{2n+1}} \frac{\sin\frac{2\pi}{2n+1}(n+h)(k-\frac{1}{2}) - \sin\frac{2\pi}{2n+1}h(k-\frac{1}{2})}{\sin\frac{2\pi}{2n+1}\frac{1}{2}(k-\frac{1}{2})} \ .$$

Then, when $p = 1$ for instance, by using a_{kn}^* $(k = 1, \cdots, m)$ in (2.8), we can derive the prediction MSE as

$$\begin{aligned} MSE(m) \quad &= \quad \frac{4\sigma_v^2}{2n+1} \sum_{k=1}^{m} [\sin\frac{2\pi}{2n+1}(n+h)(k-\frac{1}{2}) - \sin\frac{2\pi}{2n+1}h(k-\frac{1}{2})]^2 \\ & \quad + \frac{\sigma_x^2}{2n+1} \sum_{k=m+1}^{n} [\frac{\sin\frac{2\pi}{2n+1}(n+h)(k-\frac{1}{2}) - \sin\frac{2\pi}{2n+1}h(k-\frac{1}{2})}{\sin\frac{2\pi}{2n+1}\frac{1}{2}(k-\frac{1}{2})}]^2 \ . \end{aligned} \tag{3.28}$$

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 notice that the first term is an increasing function of m while the second term

is a decreasing function of m. There can be a point of m^* such that $MSE(m)$ is minimized. There can be several criteria, which are based on the prediction MSE. Because the prediction error depends on the unknown parameters of $\mathbf{\Sigma}_x$ and $\mathbf{\Sigma}_v$, we need to replace them in a simple way. When $p=1$, we need the ratio of estimated variances, which can be constructed by the discussion of Section 3 of Kunitomo, Awaya and Kurisu (2019).

4. Discussions

4.1 An Extended Errors-in-Variables Model

There are possible generalizations of the basic model in Section 3. We consider the additive decomposition model

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

where we take positive integers s ($s>1$), N, and $n=sN$ for the resulting simplicity and arguments, and $\mathbf{s}_i$ ($i=1,\cdots,n$) is a sequence of non-stationary process, which satisfies

$$\Delta \mathbf{s}_i = (1-\mathcal{L})\mathbf{s}_i = \mathbf{v}_i^{(s)} \ , \tag{4.2}$$

where with the lag-operator $\mathcal{L}^s \mathbf{s}_i = \mathbf{s}_{i-s}$, $\Delta_s = 1-\mathcal{L}^s$,

$$\mathbf{v}_i^{(s)} = \sum_{j=0}^{\infty} \mathbf{C}_{sj}^{(s)} \mathbf{e}_{i-sj}^{(s)} \ , \tag{4.3}$$

and $\mathbf{e}_i^{(s)}$ is a a sequence of i.i.d. random vectors with $\mathbf{E}(\mathbf{e}_i^{(s)}) = \mathbf{0}$ and $\mathbf{E}(\mathbf{e}_i^{(s)}\mathbf{e}_i^{(s)\prime}) = \mathbf{\Sigma}_e^{(s)}$ (a non-negative definite matrix). The $p\times p$ coefficient matrices $\mathbf{C}_j^{(s)}$ are absolutely summable ($\sum_{j=0}^{\infty} \|\mathbf{C}_{sj}^{(v)}\| < \infty$, $\|\mathbf{C}_{sj}^{(s)}\| = \max_{k,l=1,\cdots,p} |c_{sk,sl}^{(s)}(j)|$ with $\mathbf{C}_{sj}^{(s)} = (c_{sk,sl}^{(s)}(j))$).
Let $\mathbf{f}_{\Delta x}(\lambda)$, $\mathbf{f}_{\Delta s}(\lambda)$, and $\mathbf{f}_v(\lambda)$ be the spectral density ($p\times p$) matrices of $\Delta \mathbf{x}_i$, $\Delta \mathbf{s}_i$ and $\mathbf{v}_i$ ($i=1,\cdots,n$) and

$$\mathbf{f}_{\Delta s}(\lambda) = (\sum_{j=0}^{\infty} \mathbf{C}_{sj}^{(s)} e^{2\pi i\lambda sj}) \mathbf{\Sigma}_e^{(s)} (\sum_{j=0}^{\infty} \mathbf{C}_{sj}^{(s)\prime} e^{-2\pi i\lambda sj}) \quad (-\frac{1}{2} \le \lambda \le \frac{1}{2}) \ , \tag{4.4}$$

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

$$\mathbf{f}_{\Delta y}(\lambda) = \mathbf{f}_{\Delta x}(\lambda) + \mathbf{f}_{\Delta s}(\lambda) + (1-e^{2\pi i\lambda}) f_v(\lambda)(1-e^{-2\pi i\lambda}) \ . \tag{4.5}$$

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

$$(4.6) \qquad \mathbf{\Sigma}_x = \mathbf{f}_{\Delta x}(0)\ (= (\sigma_{gh}^{(x)}))\ ,\ \mathbf{\Sigma}_s = \mathbf{f}_{\Delta s}(\frac{1}{s})\ (= (\sigma_{gh}^{(s)}))\ ,$$

and

$$(4.7) \qquad \mathbf{\Sigma}_v = f_v(0)\ = (\sigma_{gh}^{(v)})\ .$$

Let $f_v^{(SR)}(\lambda_k)$, $f_{\Delta s}^{(SR)}(\lambda_k)$ 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, $f_v^{(SR)}(\lambda_k) = (1/2)[f_v^{(SR)}(\lambda_k) + \bar{f}_v^{(SR)}(\lambda_k)]$, $f_{\Delta s}^{(SR)}(\lambda_k) = (1/2)[f_{\Delta s}^{(SR)}(\lambda_k) + \bar{f}_{\Delta s}^{(SR)}(\lambda_k)]$ and $f_{\Delta x}^{(SR)}(\lambda_k) = (1/2)[f_{\Delta x}^{(SR)}(\lambda_k) + \bar{f}_{\Delta x}^{(SR)}(\lambda_k)]$.
Theorem 5.1 of Kunitomo and Sato (2019) gives the condition that the orthogonal processes are approximately distributed as the Gaussian distribution. Then, (-2) times the log-likelihood function in the general model can be approximated as

$$(4.8) \qquad (-2)L_n(\boldsymbol{\theta}) = \sum_{k=1}^{n} \log |a_{kn}^* f_v^{(SR)}(\lambda_k) + f_{\Delta s}^{(SR)}(\lambda_k)) + f_{\Delta x}^{(SR)}(\lambda_k))| + \sum_{k=1}^{n} \mathbf{z}_k^{'}[a_{kn}^* f_v^{(SR)}(\lambda_k) + f_{\Delta s}^{(SR)}(\lambda_k)) + f_{\Delta x}^{(SR)}(\lambda_k)]^{-1}\mathbf{z}_k\ .$$

4.2 On a Frequency Interpretation

At the first glance, the SIML smoothing method might be regarded as an *ad-hoc* statistical procedure without any mathematical foundation. However, on the contrary, there is a rather solid statistical foundation. Section 5 of Kunitomo and Sato (2019) has discussed a justification of the SIML forward-smoothing and it is different from the standard explanation of time series analysis in the frequency domain (Doob (1953), and Brockwell and Davis (1990), and some extensions to non-stationaru process (see Brillinger and Hatanaka (1969), Brillinger (1980)). We can proceed a similar argument of Kunitomo and Sato (2019) on the backward smoothing.
For $\lambda_k^{(n)} = (k - 1/2)/(2n+1)$ $(k = 1, \cdots, n)$, we write

$$(4.9) \qquad \mathbf{z}_n^*(\lambda_k^{(n)}) = \sum_{j=1}^{n} \mathbf{r}_j^{(n)*}[\frac{2}{\sqrt{2n+1}} \sin[2\pi\lambda_k^{(n)} j]\ \ (k = 1, \cdots, n),$$

where $\mathbf{r}_j^{(n)} = \mathbf{y}_{j-1}^{(n)} - \mathbf{r}_j^{(n)}$ $(j = 0, \cdots, n-1)$.
Then, by using the inversion transformation with $\mathbf{P}_n^*$, we can confirm that

$$(4.10) \qquad \mathbf{r}_s^{(n)} = \sum_{k=1}^{n} p_{sk}^* \mathbf{z}_n^*(\lambda_k^{(n)})\ \ (s = 1, \cdots, n).$$

It is another representation of $\mathbf{R}_n^* = (\mathbf{r}_{i-1}^{*(n)'}) = \mathbf{C}_n^{'-1}\hat{\mathbf{X}}_n^*(\mathrm{Q}^*)$ when $\mathbf{Q}_n^* = \mathbf{I}_n$. For any s $(s = 1, \cdots, n)$, $\mathbf{r}_s^{(n)}$ can be recovered as the weighted sum of othogonal processes

$\mathbf{z}^{*(n)}(\lambda_k^{(n)})$ at frequency $\lambda_k^{(n)}$ $(k = 1, \cdots, n)$. We then, by using $\mathbf{Y}_n^* = \mathbf{C}_n'\mathbf{R}_n'$, recover the non-stationary process $\mathbf{y}_t^{(n)}$ $(t = 0, \cdots, n-1)$ given the initial condition $\mathbf{y}_0$ as

$$\mathbf{y}_t^{(n)} = \mathbf{y}_n + \sum_{s=1}^{t} \mathbf{r}_{n-s}^{*(n)} \ . \tag{4.11}$$

Let

$$\alpha_n(\lambda_m^{(n)}, j) = \frac{1}{n}\sum_{k=1}^{m}[2\sin 2\pi\lambda_k^{(n)} j] \tag{4.12}$$

Then, when $\lambda_m^{(n)} \to \lambda$ as $n \to \infty$ $(0 < \lambda < \frac{1}{2})$, we find

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

If we set the uncorrelated stochastic process of uncorrelated increments with continuous parameter λ $(0 \leq \lambda \leq \frac{1}{2})$ as $B_n(\lambda) = \sum_{j=1}^{n} \beta(\lambda, j)\mathbf{r}_j^{*(n)}$, then we find

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

This corresponds to the continuous representation of a discrete (real-valued) stationary time series in the frequency domain (see Chapter 7.4 of Anderson (1971)). If we write the limit of $\mathbf{B} = \lim_{n\to\infty} \mathbf{B}_n(\lambda)$ (assuming it exists), the (real-valued) spectral distribution matrix F_{RS} for any $0 \leq \lambda_1 < \lambda_2 \leq 1/2$ can be defined as

$$F_{RS}(\lambda_2 - \lambda_1) = \mathbf{E}[(\mathbf{B}(\lambda_2 - \lambda_1)\mathbf{B}(\lambda_2 - \lambda_1)'] = \int_{\lambda_1}^{\lambda_2} f_{RS}(\lambda)d\lambda \tag{4.14}$$

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.
If we set $\hat{\mathbf{R}}_n^*(m) = (\hat{\mathbf{r}}_i^{*(m,n)'}) = \mathbf{C}_n'^{-1}\hat{\mathbf{X}}_n^*(m)$, where $\mathbf{r}_i^{*(m,n)}$ are $p \times 1$ vectors for $i = 1, \cdots, n$. If we write

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

it is the trend SIML-smoothing value for $\mathbf{r}_s^{*(m,n)}$. It is $\hat{\mathbf{X}}_n^*(m)$ $(= \mathbf{C}_n'\mathbf{R}_n^*(m))$. and

$$\mathbf{r}_s^{*(m,n)} = \sum_{k=1}^{m} p_{sk}^* \mathbf{z}_n^{**}(\lambda_k^{(n)}) \ (s = 1, \cdots, m; 0 < m < n), \tag{4.16}$$

where $\mathbf{z}^{**(n)}(\lambda_k^{(n)})$ are constructed from the $n \times p$ hidden states matrix $\mathbf{X}_n^*$ instead of the observed $n \times p$ matrix data $\mathbf{Y}_n^*$. Hence it is the same as the element of $\mathbf{C}_n'^{-1}\hat{\mathbf{X}}_n^*(m)$, and for $\lambda_m^{(n)} = m/n$ in the frequency domain it is a discrete version of

$$\hat{\mathbf{r}}_s^{*(n)}(\lambda_m^{(n)}) = \int_0^{\lambda_m^{(n)}} \sin[2\pi\lambda\, s]dB_n(\lambda) \ . \tag{4.17}$$

Then, it is the same as the element of $\mathbf{C}_n^{\prime -1}\hat{\mathbf{X}}_n^*(m)$, and its has the corresponding (continuous) version in the frequency domain.
Similarly, $\hat{\mathbf{r}}_s^{*m_1,m_2,n)} = \sum_{k=m_1+1}^{m_1+m_2} p_{sk}\Delta_\lambda \mathbf{z}^{*(n)}(\lambda_k^{(n)}) = \hat{\mathbf{r}}_s^{*(m_2,n)} - \hat{\mathbf{r}}_s^{*(m_1,n)}$ $(s = 1, \cdots, m; 0 < m_1 < m_2 < n)$ can be regarded as a discrete version of

$$\hat{\mathbf{r}}_s^{*(n)}(\lambda_{m_1}^{(n)}, \lambda_{m_2}^{(n)}) = \int_{\lambda_{m_1}^{(n)}}^{\lambda_{m_2}^{(n)}} \sin[2\pi\lambda\ s] dB_n(\lambda)\ . \tag{4.18}$$

5. A Numerical Example

We illustrate the use of the SIML-forward smoothing and SIML-backward smoothing for real data. We have used the monthly US Manufacturers' New Orders Data during 2010-2020 because it has been known that this time series data has trend, wild seasonal fluctuation and noise components.

Red Curve in Figure 1 shows the forward smoothing given the first observation as the initial condition with $m = 5$. Green Curve in Figure 1 shows T_1^* as the limit of the forward-backward iterations. Violet Curve in Figure 1 shows the two-step forward filtering with $m_1 = 15$(the first smoothing) and $m = 5$(the second smoothing). Blue Curve in Figure 2 shows the backward smoothing given the last observation as the initial condition with $m = 5$. Skyblue Curve in Figure 2 show T_2^* as the limit of the backward-backward iterations. Violet Curve in Figure 2 shows the two-step backward filtering with $m_1 = 15$(the first smoothing) and $m = 5$(the second smoothing).

As we have expected, the initial value of both forward and backward smoothers have significant effects around the initial values at which we start smoothing. The effects of choosing the initial value become negligible either repeating smoothing (or filtering) and multi-step smoothing. The resulting differences in two procedures after a few steps are small for practical purposes.

6. Conclusions

When the observed non-stationary multivariate time series contain noises, it may be difficult to disentangle the effects of trends and noises. This study is a subsequent one to Kunitomo and Sato (2019), which investigated a new statistical procedure to decompose time series into non-stationary trend component, seasonal component and stationary noise (or measurement errors) component. The resulting smoothing or filtering method for non-stationary multivariate series is simple and free from the underlying distributions of noise and state vector and therefore it is robust against possible misspecification in the non-stationary multivariate time series.

There are several interesting problems developed by our approach in this study. Our method framework gives some earlier studies on the filtering methods in the time domain and frequency domain. Although our method is a non-parametric

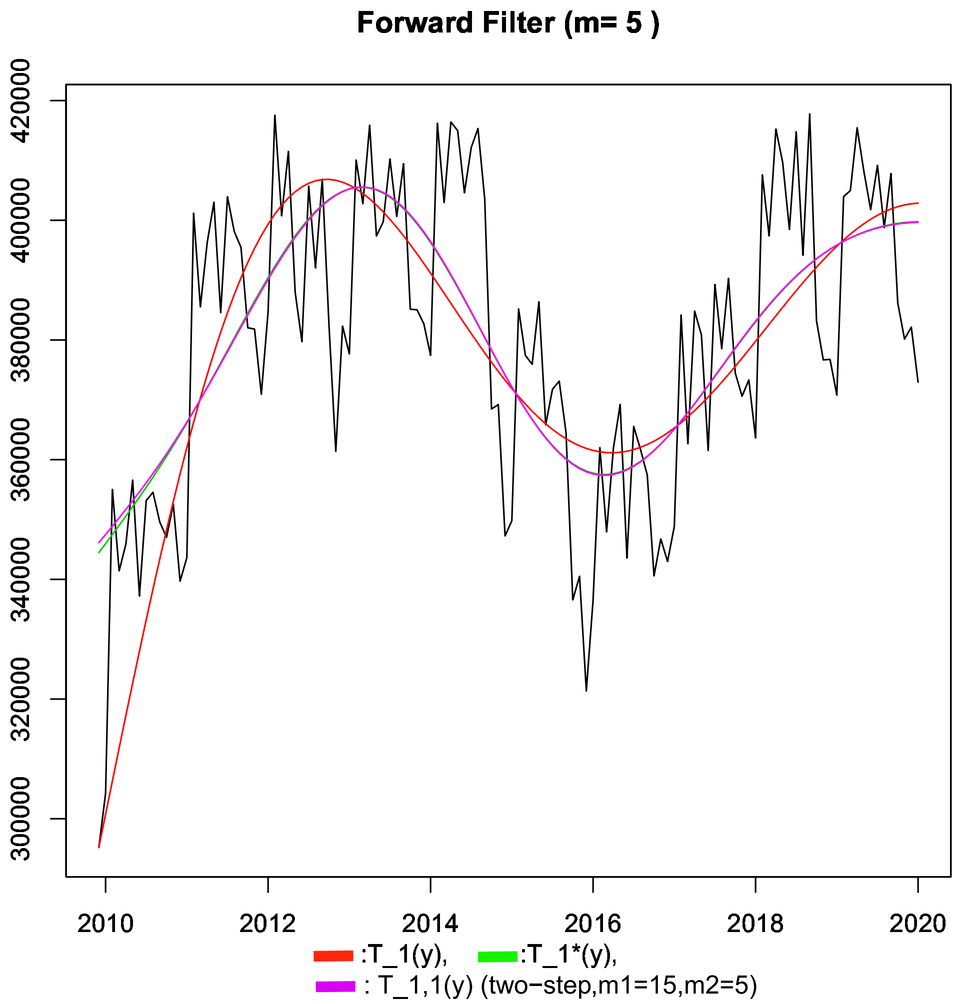

Figure 1: The forward filtering results for US monthly US Manufacturers' New Orders from 2010 to 2020. (https://www.census.gov/manufacturing/m3/index.html)

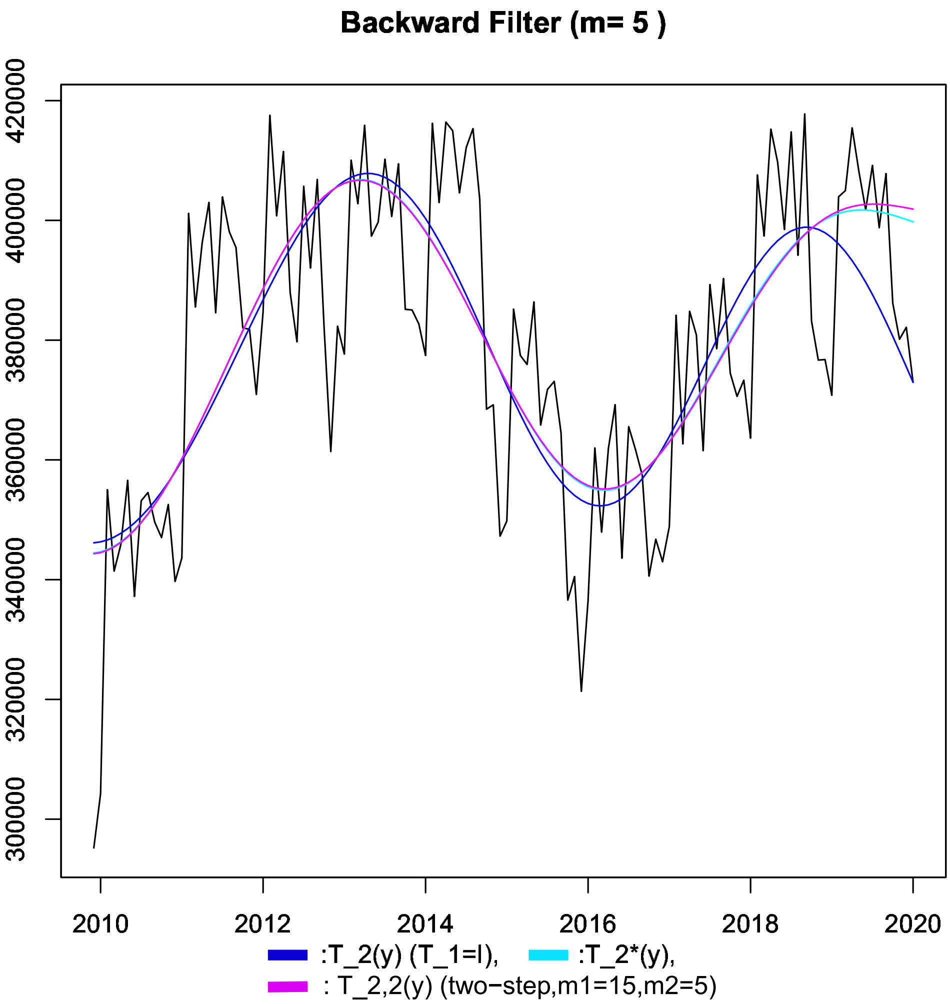

Figure 2: The backward filtering results for US monthly US Manufacturers' New Orders from 2010 to 2020. (https://www.census.gov/manufacturing/m3/index.html)

smoothing method, there is a close relationship with the existing smoothing and filtering methods such as Decomp by Kitagawa (2010), which was a subsequent method of Akaike (1980). (See some discussions in Kunitomo and Sato (2019).) There would be many empirical applications, which will be reported in another occasion.

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," J. Japan Statistical Society, 38-1, 41-49.

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

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

[10] Kunitomo, N. and S. Sato (2019), "A Robust-filtering Method for Noisy Non-Stationary Multivariate Time Series," SDS-9, MIMS, Meiji University, http://www.mims.meiji.ac.jp/publications/2017-ds, revised.

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

[12] Nishimura, G. K., Seisho Sato, Akihiko Takahashi (2019), "Term Structure Models During the Global Financial Crisis: A Parsimonious Text Mining Approach," *Asic-Pacific Financial Markets.*

APPENDIX A : Mathematical Derivations

We now present some details of derivations that we have omitted in the previous sections. Most of our derivations is to apply trigonometric relations, which are elementary and straightforward in a sense. We show only the essential parts of derivations.

(i) **Characteristic Roots and Vectors** :

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

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

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

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

(ii) We have the spectral decomposition

$$\mathbf{C}_n^{'-1}\mathbf{C}_n^{-1} = \mathbf{P}_n^*\mathbf{D}_n\mathbf{P}_n^{*'} = 2\mathbf{I}_n - 2\mathbf{A}_n^* , \tag{A.3}$$

where $\mathbf{D}_n$ is a diagonal matrix with the k-th element

$$d_k = 2\left[1 - \cos(\pi(\frac{2k-1}{2n+1}))\right] \ \ (k = 1, \cdots, n) , \tag{A.4}$$

$$\mathbf{C}_n^{'-1} = \begin{pmatrix} 1 & -1 & \cdots & 0 & 0 \\ 0 & 1 & -1 & \cdots & 0 \\ 0 & 0 & 1 & -1 & \cdots \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{A.5}$$

and

$$\mathbf{P}_n^* = (p_{jk}^*) \ , \ p_{jk}^* = \sqrt{\frac{2}{n+\frac{1}{2}}}\sin\left[\frac{2\pi}{2n+1}(k-\frac{1}{2})\ j\right] . \tag{A.6}$$

Proof of Lemma A.1 : (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

$$\text{(A.7)} \qquad \frac{x_2}{2} = \lambda x_1 \ ,$$

$$\text{(A.8)} \qquad \frac{x_{t-1}+x_{t+1}}{2} = \lambda x_t \ (t = 2,\cdots,n-1) \ ,$$

$$\text{(A.9)} \qquad \frac{1}{2}x_{n-1} + x_n = \lambda x_n \ .$$

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 $\xi_1^{2n+1} = -1$. Therefore,

$$\text{(A.10)} \qquad \lambda_k = \cos[\pi\frac{2k-1}{2n+1}] \ (k = 1,\cdots,n) \ .$$

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

$$\text{(A.11)} \qquad x_t = \frac{1}{2i}\left[\xi_1^t - \xi_1^{-t}\right] = \sin\left[\pi\frac{2k-1}{2n+1}t\right] \ .$$

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

Proof of Theorem 3.1 :

(i) We consider the case of T_{2k+1} $(k \geq 1)$. By using the recursive relations, for $k \geq 1$ we can represent

$$\text{(A.12)} \qquad T_{2k+1} = A_1^{(m,n)} + A_2^{(n)}T_{2(k-1)+1} \ ,$$

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

$$\text{(A.13)} \ A_1^{(m,n)} = (\mathbf{I}_n - \mathbf{C}_n\mathbf{P}_n\mathbf{J}_m^{'}\mathbf{J}_m\mathbf{P}_n\mathbf{C}^{-1})\mathbf{1}_n\mathbf{e}_1^{'}(\mathbf{I}_n - \mathbf{C}_n^{'}\mathbf{P}_n^{*'}\mathbf{J}_m^{'}\mathbf{J}_m\mathbf{P}_n^*\mathbf{C}^{'-1})\mathbf{1}_n\mathbf{e}_n^{'} \ .$$

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

$$\begin{aligned}\text{(A.14)} \ a_{2n} &= \mathbf{e}_n^{'}(\mathbf{I}_n - \mathbf{C}_n\mathbf{P}_n\mathbf{J}_m^{'}\mathbf{J}_m\mathbf{P}_n\mathbf{C}^{-1})\mathbf{1}_n\mathbf{e}_1^{'}(\mathbf{I}_n - \mathbf{C}_n^{'}\mathbf{P}_n^{*'}\mathbf{J}_m^{'}\mathbf{J}_m\mathbf{P}_n^*\mathbf{C}^{'-1})\mathbf{1}_n \\ &= [1 - \mathbf{1}_n\mathbf{C}_n\mathbf{P}_n\mathbf{J}_m^{'}\mathbf{J}_m\mathbf{P}_n\mathbf{C}^{-1}\mathbf{e}_n][1 - \mathbf{1}_n\mathbf{C}_n^{'}\mathbf{P}_n^*\mathbf{J}_m^{'}\mathbf{J}_m\mathbf{P}_n^*\mathbf{C}^{'-1}\mathbf{e}_1] \ .\end{aligned}$$

By using the relation

$$1 - \mathbf{1}_n^{'}\mathbf{C}_n\mathbf{P}_n\mathbf{J}_m^{'}\mathbf{J}_m\mathbf{P}_n\mathbf{C}^{-1}\mathbf{e}_1 = \mathbf{1}_n^{'}\mathbf{C}_n\mathbf{P}_n\mathbf{J}_{n-m}^{'}\mathbf{J}_{n-m}\mathbf{P}_n\mathbf{C}^{-1}\mathbf{e}_1 \ ,$$

where $\mathbf{J}_{n-m} = (\mathbf{O}, \mathbf{I}_{n-m})$ is a $(n-m) \times [m + (n-m)]$ matrix. Because $\mathbf{C}_n^{-1}\mathbf{e}_1 = \mathbf{e}_1$ and the last term 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. It is because by using Lemma 5.1 of Kunitomo et al. (2018), the above term is

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

Because

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

it 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 fact that $\sin\pi(k-\frac{1}{2})$ takes $+1$ and -1 alternatively, we evaluate the difference of

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

We can take $n > n_0$ such that $\sin\frac{\pi}{2n+1}$ and $1-\cos\frac{\pi}{2n+1}$ being sufficient small. Then each term becomes small and finally $\mathbf{1}_n^{'}\mathbf{C}_n\mathbf{P}_n\mathbf{J}_{n-m}^{'}\mathbf{J}_{n-m}\mathbf{P}_n\mathbf{C}^{-1}\mathbf{e}_1$ is less than one. Similarly,

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

Because $\mathbf{1}_n^{'}\mathbf{C}_n^{'} = \mathbf{1}_n^{'}$ and $\mathbf{C}_n^{'-1}\mathbf{e}_n = \mathbf{e}_n$, the last term 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 have utilized the relation that

$$\text{(A.15)} \qquad \sum_{k=1}^{n} \sin\frac{2\pi}{2n+1}(k-\frac{1}{2})j = \frac{1}{2i}\frac{e^{i\frac{2\pi}{2n+1}jn} + e^{-i\frac{2\pi}{2n+1}jn} - 2}{e^{i\frac{2\pi}{2n+1}j\frac{1}{2}} - e^{-i\frac{2\pi}{2n+1}j\frac{1}{2}}} = \frac{1}{2}\frac{1-\cos\frac{2\pi}{2n+1}jn}{\sin\frac{2\pi}{2n+1}j\frac{1}{2}} \ .$$

By using the elementary relation on trigonometric functions that for $2(n-1/2) = (2n+1)-2$

$$\sin\frac{2\pi}{2n+1}(n-\frac{1}{2})j = \sin\frac{2\pi}{2n+1}j = 2\sin\frac{\pi}{2n+1}j\cos\frac{\pi}{2n+1}j \ ,$$

we find that $|a_{2n}| < 1$ and we have convergence of T_{2k+1} as $k \to \infty$.
(ii) We can apply the similar arguments to T_{2k} $(k \geq 1)$. By using the recursive relations, for $k \geq 1$ we can represent

$$T_{2k} = A_{1*}^{(m,n)} + A_{2*}^{(n)} T_{2(k-1)+1} \ , \tag{A.16}$$

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

Proof of Theorem 3.2 :

We need to evaluate

$$\mathbf{F}_n = (f_{ab}) = \mathbf{C}_n^{'}\mathbf{P}_n^{*'}\mathbf{J}_m^{'}\mathbf{J}_m\mathbf{P}_n^{*}\mathbf{C}^{'-1} \ . \tag{A.17}$$

We write

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

We use the relations

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

and

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

Then $2n+1)/4 \times f_{ab}$ becomes

$$\begin{aligned}
&\sum_{j=1}^{m}[\cos\frac{2\pi}{2n+1}(a-1)j - \cos\frac{2\pi}{2n+1}nj]\cos\frac{2\pi}{2n+1}j(b-1) \\
= &\sum_{j=1}^{m}[\cos\frac{2\pi}{2n+1}(a-1)j\cos\frac{2\pi}{2n+1}j(b-1) - \cos\frac{2\pi}{2n+1}nj\cos\frac{2\pi}{2n+1}j(b-1)] \ .
\end{aligned}$$

By utilizing the relation

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

and elementary calculations, we find that

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

When $m/n \to 0$ as $n \to \infty$ for $a, b = 1, \cdots, m$, this quantity converges to 0. Hence as $n \to \infty$ and $m/n \to 0$, f_{ab} $(a, b = 1, \cdots, m)$ is asymptotically equivalent to

$$h_{ab} = \frac{4}{2n+1}\sum_{j=1}^{m}[\cos\frac{2\pi}{2n+1}j(a-\frac{1}{2})\cos\frac{2\pi}{2n+1}j(b-\frac{1}{2})] \ , \tag{A.18}$$

which is the $(a,b)-$th element of $\mathbf{H}_n = \mathbf{P}_n\mathbf{J}_m^{'}\mathbf{J}_m\mathbf{P}_n$. Hence 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 $\mathbf{Q}_n^{(m_1,m_2)} = \mathbf{J}_{m_1,m_2}^{'}\mathbf{J}_{m_1,m_2}$. Then by using the similar arguments as the proof of Theorem 3.1, we find that the absolute values of the eigenvalues of $\mathbf{A}_2^{(m_1,m_2,n}$ and $\mathbf{A}_{2*}^{(m_1,m_2,n}$ are less than one and we have the convergence of the repeated smoothing procedures.
(**Q.E.D.**)