

## 経済時系列の状態推定とマクロ指標

国友 直人<sup>1)</sup>櫻井 智章<sup>2)</sup>佐藤 整尚<sup>3)</sup>

A Filtering Method of Economic Time Series and a Macro-Index

KUNITOMO Naoto

SAKURAI Tomoaki

SATO Seisho

月次時系列、四半期時系列、年次系列など異なる頻度で公表されるマクロ時系列の扱いをめぐる実際的課題を解決するため、SIML フィルタリングと呼ぶ多次元非定常時系列のスペクトルに基づく分析を提案する。公的統計での主要なマクロ経済時系列では、公表系列を統合的に処理し、マクロ経済の動向について適切な情報を提供することが望まれる。本稿で説明する SIML フィルタリング法はノイズを含む複数の非定常時系列から真の状態を推定する統計的方法である。応用例として日本の四半期および月次の消費系列のマクロ指標の作成という問題を取りあげる。ノイズを含む非定常時系列モデルを利用して全体として統合的に経済時系列データを扱う方法を説明し、複数のマクロ消費系列からマクロ指標を構成する方法と分析例を例示する。

**キーワード** マクロ経済時系列, ノイズを含む非定常多次元時系列, トレンド・循環と季節性, 変化点と外れ値, マクロ指標, SIML フィルタリング

We apply the SIML filtering method to handle macro time series with different frequencies occurring in official statistics. It is based on the spectral decomposition of noisy non-stationary multivariate time series in the frequency domain. It is important to construct the published official time series in a unified way and to estimate the true state variables from non-stationary economic time series. We illustrate the SIML filtering method by constructing a monthly macro-consumption index from several monthly and quarterly consumption data.

**Key Words** Macro-Economic Time Series, Noisy Non-stationary Multivariate Time Series, Trend-Cycle-Seasonals, Change Points and Outliers, Filtering, Macro-Index, SIML Filtering.

---

<sup>1)</sup> 統計数理研究所 E-mail: kunitomo@ism.ac.jp

<sup>2)</sup> 総務省統計局統計調査部消費統計課

<sup>3)</sup> 東京大学経済学部 E-mail: seishou@e.u-tokyo.ac.jp

## 1 はじめに

時間の経過とともに多数の経済時系列が観察されている。特に主要なマクロ経済公的データの場合にはデータの収集上・作成上の理由から日々の系列、月次系列、四半期系列、年次系列など様々な頻度や異なる時間的タイミングで計測され、政府の各部署から公表されている。消費・投資・政府支出・輸出入など主要なマクロ時系列の各系列はそれぞれ基礎となる統計調査にもとづき系列が作成されるなどの事情からそれぞれの特徴があり、必ずしも相互に調整されて公表されるとは限らない。他方、経済全体の動向を理解し、データ分析を行い、また政策判断を行うなどの目的からはこうしたデータの作成・公表を巡る事情は望ましいものとは必ずしもいえない。経済全体の動向を示すマクロ系列の場合には、直近のマクロ経済状況を理解するため、なるべく早めに正確なデータを作成し、公表することが望ましいが、国全体のデータを作成するには基礎となる多くの調査データを収集する必要がある。例えば日本の GDP やその主要な項目はその基礎データが時間差をともない次々に公表されることから、一次速報に続き、二次速報、確報と改訂される。例えば投資系列のデータは GDP 一次速報が公表された後に、より正確と思われる基礎データが利用可能となるので、その結果、しばらく後に四半期系列の一次公表値が改訂されることになる。こうしたデータの改訂で生じる誤差を本稿では真の状態変数の推定の過程で生じる観測誤差とするが、公表される四半期データが直近で得られる月次系列など基礎データの情報と見かけ上で矛盾する事例なども散見される。また日本の公的統計では各系列を調査する担当部署が分かれているので、公表する責任がそれぞれ異なっていることも問題をより複雑にしている。例えば本稿で事例としてとりあげるマクロ消費については、四半期で内閣府から公表される GDP 最終消費系列、月次に総務省統計局が公表している家計調査や経済産業省の商業動態統計などがあるが、さらに月次系列として数年前から消費動向指数 (CTI) が作成、公表されるようになって<sup>4)</sup>。

本稿ではマクロ時系列を巡る課題について、統計的時系列分析の立場より SIML フィルタリング・アプローチと呼ばれる新しい分析法に基づく一つの解決法を提案する。日本の主要なマクロ経済時系列の処理については、現状に比べてより矛盾が少なくなるような方策が望まれるが、ここで提案する方法、すなわち SIML フィルタリング法 (あるいはスムージング法) が有効で役立つと思われる。本稿では特に近年における日本の四半期および月次の消費系列を実例としてとりあげ、ノイズを含む非定常多次元時系列モデルを用いることにより、かなり容易に全体として整合的に経済時系列データを扱うことができることを例示する。

まず第 2 節ではこの研究の動機である日本のマクロ消費を巡る問題を述べる。第 3 節ではノイズを含む非定常多次元時系列モデルの概略を述べ、さらに外れ値や変化点の処理などを含め、非定常時系列分析よるトレンド・循環・季節性など多次元時系列の構成要素の状態推定法を説明する。次に 4 節では統計分析にもとづく月次の経済指標の作成問題をとりあげ、実際にマクロ消費系列の試算例を説明する。最後に第 5 節では本稿のまとめを述べる。幾つかの基礎となる関連する数理的補論および利用した R プログラム例を付論にまとめておく。

<sup>4)</sup> この系列の推定には時系列 `decomp` モデルを利用、かなり複雑な統計的操作により CTI 系列を作成している。詳しくは総務省統計局 (2021) を参照されたい。

## 2 研究の動機：マクロ消費動向の計測問題

日本における近年のマクロ消費動向を理解する上では、家計に対するサンプリング調査に基づく家計調査の消費動向と企業に対するサンプリング調査に基づく商業動態統計などから得られる販売における消費動向の乖離についての理解と解決が重要な課題である。一方で家計調査データは世帯をベースにした標本調査の集計値であることから、家計調査で得られるデータを解釈するときには近年の世帯数の変化など世帯の動向を考慮する必要がある。家計調査は詳細な項目を含む月次調査であり、標本サイズは8,000～9,000世帯、家計の詳細なデータが得られる。他方、企業側から得られる消費データとしての商業動態統計は生産・販売など供給面での調査データであり、標本サイズが約20,000のサンプリング調査、事業所ベースで調査されることから、企業消費・政府消費やインバウンド消費などGDPにおける家計消費概念とは必ずしも整合的ではない。こうした事情からマクロ消費を計測するには需要サイドと供給サイドの消費という特性を十分理解し、消費概念の整合性を考慮する必要がある。さて民間や政府のエコノミストが日本におけるマクロ消費動向としてもっとも重視しているのはGDP最終消費の数値であるが、四半期GDP速報の推計では最終消費は家計面と企業面における消費の情報を統合した数値を四半期ベースで作成している。ここで月次系列や四半期系列の場合には季節性の処理が行われているのでその処理にも注意する必要がある<sup>5)</sup>。また毎年作成されているGDP推計の確報値の作成では、供給面におけるより詳細な推計値を利用しているが、その際には年次推計値から四半期分割により四半期系列を構成するときのベンチマークなどにGDP速報の数値を利用していることも重要である。

ところでGDP消費の速報値の基礎となっている家計調査、また商業動態統計、第三次産業活動指数などは月次調査であり、これらのデータは直近の情報として利用できるため、近年における月次レベルでの消費動向を分析する際には需要面と供給面におけるマクロ指標としてよく利用されている。(各原系列の最近の動向については4節の図1(GDP最終消費)、図3(家計調査)、図5(商業動態)、図7(第三次産業)の実線を参照されたい。)これらの月次データの変動は同一とは限らないことからその乖離について注目されることがある。例えば家計の消費行動では従来の標本調査では十分に把握することが困難なネット販売など新しい形態の消費行動の補足が十分ではなく、消費全体に与える影響も無視できない水準になっている。他方、販売面における消費データではこうした新たな形態の消費はある程度までは補足可能であるが、月次で得られる販売面での消費データもサンプリングで得られていることもあり、データ調査上のサンプリング精度の問題などでの問題も無視すべきでない。すなわち速報的に月次レベルで集計されている月次データとより広範なデータを利用する四半期レベル、年次レベルにより推計されているGDP消費の動向を整合的に理解し、マクロ消費動向を的確に把握する必要性があると云える。

ここで近年になり利用可能となった統計的時系列分析を利用すると、月次レベル及び四半期レベルで得られる情報を矛盾なく活用できる可能性があることを指摘しておこう。本稿ではマクロ指数の構成問題を多次元時系列における一部のデータが不完全な場合の欠損値(missing observation)の問題ととらえる。つまりGDP最終消費は四半期ベースで様々なデータの集計値

<sup>5)</sup> 四半期GDPの詳細は内閣府・国民所得部(2010)、あるいは内閣府・経済社会総合研究所のHP情報を参照されたい。

として公表されるが、月次データとしては欠損値が多く存在する系列と考えるのである。このとき消費については月次データとして利用可能な情報があるので多次元時系列データの一部分が欠損しているとみると、欠損値の状態を推定することは可能と考えられる。

このとき重要なこととしては、実際に観察されるマクロ経済時系列ではエコノミストや政策担当者が主に関心のあるトレンド成分や循環成分ばかりではなく、季節成分や不規則成分、さらには外れ値や変化点などがあり得る、との理解が必要である。すなわち現実のマクロ時系列は系列を構成する幾つかの要因が重なって観測される為、時系列分析において古くから知られている移動平均法などを単純に利用するだけでは、現実的に利用可能なデータの適切な分析を行うことは困難なのである。さらに、日本を含む多くの国の公的統計では移動平均に基づく季節調整法 X-12-ARIMA が利用されていることにも注意する必要がある。X-12-ARIMA では Reg-ARIMA モデルという一変量時系列回帰モデルを内部の計算では利用しているが、実務的には公表されている各系列に対して担当者が別個に Reg-ARIMA モデルを適用、運用している。したがって、マクロ時系列の場合には時系列間には様々な関係があるが、異なる構成要素が混在する為に時系列の成分分解が必要であるにもかかわらず、複数の時系列が整合的に季節調整が行われているとは限らない。本稿ではここで述べた諸問題を解決する方法として、複数の非定常時系列の構成成分における関係を分析する為に開発された SIML(分離情報最尤) 法の考え方を利用し、実務にも耐えうる頑健な状態推定の結果が得られる実用的な統計的分析法を展開する。

### 3 非定常時系列の状態推定

#### 3.1 ノイズを含む非定常多次元時系列

##### ノイズを含む非定常多次元時系列モデル

不規則変動(ノイズ)を含む非定常多次元時系列モデルを導入しよう。時刻  $t$  の多次元時系列を  $p \times 1$  ベクトル  $\mathbf{y}_t = (y_{kt})$ 、各観測系列データは  $p$  次元ベクトルの要素、 $y_{kt}$  は時刻  $t$  における  $k$  番目の変数とする。ここで季節周期  $s$ (月次なら  $s = 12$ 、四半期なら  $s = 4$ ) のとき  $l$  年  $j$  期は時刻  $t = (l - 1)s + j$  ( $l = 1, \dots, n; j = 1, \dots, s$ )、初期値ベクトル  $\mathbf{y}_0$  は固定する。次に時刻  $t$  における時系列  $\mathbf{y}_t$  に対する加法的分解モデル

$$\mathbf{y}_t = \mathbf{x}_t + \mathbf{s}_t + \mathbf{v}_t \quad (t = 1, \dots, T), \quad (3.1)$$

を考察しよう<sup>6)</sup>。ここでは議論の単純化のためにデータ数は  $T = sn$ 、 $s \geq 2$  の正整数(四半期データでは  $s = 4$ 、月次データは  $s = 12$ ) とする。トレンド・循環成分  $\mathbf{x}_t$  ( $t = 1, \dots, T$ ) は非定常 I(1) 過程とする。すなわち、ラグ作用素  $\mathcal{L}\mathbf{x}_t = \mathbf{x}_{t-1}$  ( $\Delta = 1 - \mathcal{L}$ ) を利用して、

$$\Delta \mathbf{x}_t = (1 - \mathcal{L})\mathbf{x}_t = \mathbf{v}_t^{(x)} \quad (3.2)$$

とする。さらに不規則成分ベクトル  $\mathbf{v}_t$  および  $\mathbf{v}_t$  とは独立なトレンド・循環成分を生成する確率変数ベクトル  $\mathbf{v}_t^{(x)}$  は定常的で移動平均表現

$$\mathbf{v}_t = \sum_{j=0}^{\infty} \mathbf{C}_j^{(v)} \mathbf{e}_{t-j}^{(v)}, \quad (3.3)$$

<sup>6)</sup> 加法的分解モデルは扱いが容易である上に非線形分解モデルを加法的に帰着できることも少なくない。例えば X-12-ARIMA は乗法モデルが基本形であるが、対数変換により加法モデルに変換が可能である。

$$\mathbf{v}_t^{(x)} = \sum_{j=0}^{\infty} \mathbf{C}_j^{(x)} \mathbf{e}_{t-j}^{(x)}, \quad (3.4)$$

を持つとしよう。各成分を生成する  $\mathbf{e}_t^{(v)}$ ,  $\mathbf{e}_t^{(x)}$  は独立・同一分布 (i.i.d.) の確率変数ベクトル, 期待値と共分散行列は  $\mathbf{E}(\mathbf{e}_t^{(v)}) = \mathbf{0}$ ,  $\mathbf{E}(\mathbf{e}_t^{(x)}) = \mathbf{0}$ ,  $\mathbf{E}(\mathbf{e}_t^{(v)} \mathbf{e}_t^{(v)'}) = \boldsymbol{\Sigma}_e^{(v)}$  (正定符号),  $\mathbf{E}(\mathbf{e}_t^{(x)} \mathbf{e}_t^{(x)'}) = \boldsymbol{\Sigma}_e^{(x)}$  (非負定符号) である。移動平均表現における係数について、ある正値  $\rho$  ( $0 \leq \rho < 1$ ) が存在して  $\mathbf{C}_j^{(v)} = O(\rho^j)$ ,  $\mathbf{C}_j^{(x)} = O(\rho^j)$  を仮定すると、係数和の絶対値収束性  $\sum_{j=0}^{\infty} \|\mathbf{C}_j^{(v)}\| < +\infty$ ,  $\sum_{j=0}^{\infty} \|\mathbf{C}_j^{(x)}\| < +\infty$  が成り立つ。(ここで  $p \times p$  行列  $\mathbf{C} = (c_{ij})$  に対するノルムは  $\|\mathbf{C}\| = \max_{i,j} |c_{ij}|$  とする。)

季節成分としては様々な定式化が可能であるが、本稿では正整数  $s$  ( $s > 1$ ),  $n$ ,  $T = sn$  として  $\mathbf{s}_t$  ( $t = 1, \dots, T$ ) は定常過程で

$$\mathbf{s}_t = \sum_{j=0}^{\infty} \mathbf{C}_j^{(s)} \mathbf{e}_{t-sj}^{(s)}, \quad (3.5)$$

という表現を持つとする。ここで  $\mathbf{e}_t^{(s)}$  は i.i.d. 確率変数ベクトル, 期待値と共分散行列は  $\mathbf{E}(\mathbf{e}_t^{(s)}) = \mathbf{0}$ ,  $\mathbf{E}(\mathbf{e}_t^{(s)} \mathbf{e}_t^{(s)'}) = \boldsymbol{\Sigma}_e^{(s)}$ , 行列係数について条件  $\mathbf{C}_j^{(s)} = O_p(\rho^j)$  ( $0 \leq \rho < 1$ ) を満足することを仮定する。

ここで  $p \times p$  行列  $\mathbf{f}_{\Delta x}(\mu)$ ,  $\mathbf{f}_s(\mu)$ ,  $\mathbf{f}_v(\mu)$  をそれぞれ  $\Delta \mathbf{x}_t, \mathbf{s}_t, \mathbf{v}_t$  ( $t = 1, \dots, T$ ) の ( $p \times p$ ) スペクトル密度行列とする。移動平均表現 (3.3)-(3.5) からそれぞれ

$$\mathbf{f}_{\Delta x}(\mu) = \left( \sum_{j=0}^{\infty} \mathbf{C}_j^{(x)} e^{2\pi i \mu j} \right) \boldsymbol{\Sigma}_e^{(x)} \left( \sum_{j=0}^{\infty} \mathbf{C}_j^{(x)' } e^{-2\pi i \mu j} \right) \quad \left( -\frac{1}{2} \leq \mu \leq \frac{1}{2} \right), \quad (3.6)$$

$$\mathbf{f}_s(\mu) = \left( \sum_{j=0}^{\infty} \mathbf{C}_{js}^{(v)} e^{2\pi i \mu js} \right) \boldsymbol{\Sigma}_e^{(s)} \left( \sum_{j=0}^{\infty} \mathbf{C}_{js}^{(v)' } e^{-2\pi i \mu js} \right) \quad \left( -\frac{1}{2} \leq \mu \leq \frac{1}{2} \right), \quad (3.7)$$

$$\mathbf{f}_v(\mu) = \left( \sum_{j=0}^{\infty} \mathbf{C}_j^{(v)} e^{2\pi i \mu j} \right) \boldsymbol{\Sigma}_e^{(v)} \left( \sum_{j=0}^{\infty} \mathbf{C}_j^{(v)' } e^{-2\pi i \mu j} \right) \quad \left( -\frac{1}{2} \leq \mu \leq \frac{1}{2} \right) \quad (3.8)$$

で与えられる<sup>7)</sup>。ここで (3.6)-(3.8) では基準化として  $\mathbf{C}_0^{(x)} = \mathbf{C}_0^{(s)} = \mathbf{C}_0^{(v)} = \mathbf{I}_p$ , 虚数  $i$  は  $i^2 = -1$  を満たすとする。このとき原系列から階差変換を施した系列  $\Delta \mathbf{y}_t (= \mathbf{y}_t - \mathbf{y}_{t-1})$  の  $p \times p$  スペクトル密度行列は

$$\mathbf{f}_{\Delta y}(\mu) = \mathbf{f}_{\Delta x}(\mu) + (1 - e^{2\pi i \mu}) [f_s(\mu) + f_v(\mu)] (1 - e^{-2\pi i \mu}) \quad (3.9)$$

で与えられる。さらにトレンド・循環成分の長期分散共分散行列は

$$\boldsymbol{\Omega}_x = \mathbf{f}_{\Delta x}(0) (= (\omega_{gh}^{(x)})), \quad (3.10)$$

で表される。このトレンド・循環成分の長期分散共分散の推定法について Kunitomo and Sato (2017) は SIML 推定量を提案している<sup>8)</sup>。

<sup>7)</sup> 定常過程のスペクトル分析については、例えば Anderson (1971)7章を参照のこと。なお本稿の記号とは若干異なっている。

<sup>8)</sup> 数理的付論の (A.6) を参照。

### $\mathbf{K}_T$ 変換と SIML 状態推定

ここで  $p$  次元非定常過程  $\mathbf{y}_t$  ( $t = 1, \dots, T$ ) の観測データよりトレンド・循環要素の状態を推定する方法を説明しよう。時刻  $t$  の多次元データ ( $p \times 1$  ベクトル)  $\mathbf{y}_t = (y_{kt})(k = 1, \dots, p; t = 1, \dots, T)$  より  $(T \times p)$  観測データ行列を  $\mathbf{Y}_T = (\mathbf{y}'_t)$ 、 $T \times p$  初期行列  $\mathbf{Y}_0 = \mathbf{1}_T \mathbf{y}'_0$ 、 $\mathbf{y}_0$  は固定された初期値ベクトル、横ベクトル  $\mathbf{1}'_T = (1, \dots, 1)$  とする。観測データ行列  $\mathbf{Y}_T$  に対する  $\mathbf{K}_T$  変換により変換データ行列  $\mathbf{Z}_T (= (\mathbf{z}'_k))$  を

$$\mathbf{Z}_T = \mathbf{K}_T (\mathbf{Y}_T - \bar{\mathbf{Y}}_0), \quad \mathbf{K}_T = \mathbf{P}_T \mathbf{C}_T^{-1}, \quad (3.11)$$

により定義する。ただし変換行列は二つの  $T \times T$  行列

$$\mathbf{C}_T^{-1} = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & 0 & \cdots \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}_{T \times T}, \quad (3.12)$$

$$\mathbf{P}_T = (p_{jk}^{(T)}), \quad \left( (j, k) \text{ 要素は } p_{jk}^{(T)} = \sqrt{\frac{2}{T + \frac{1}{2}}} \cos \left[ \frac{2\pi}{2T + 1} \left( k - \frac{1}{2} \right) \left( j - \frac{1}{2} \right) \right] \right). \quad (3.13)$$

この変換のスペクトル分解は  $\mathbf{C}_T^{-1} \mathbf{C}'_T^{-1} = \mathbf{P}_T \mathbf{D}_T \mathbf{P}'_T$ 、ここで  $\mathbf{D}_T$  は対角行列であり、第  $(k, k)$  要素は  $d_k = 2[1 - \cos(\pi(\frac{2k-1}{2T+1}))] = 4 \sin^2(\frac{\pi}{2}(\frac{2k-1}{2T+1}))$  ( $k = 1, \dots, T$ ) となる。

次に  $\mathbf{K}_T$  変換された系列  $\mathbf{Z}_T$  から逆変換によるトレンド・循環要素の状態推定を考える。データ行列から関心のある周波数成分のデータを取り出すために  $0 < m < T$  となる  $m$  を選び、 $m \times T$  選択行列  $\mathbf{J}_m = (\mathbf{I}_m, \mathbf{O})$ 、 $\mathbf{Q}_T^{(1)} = \mathbf{J}'_m \mathbf{J}_m$  に対して  $T \times p$  行列

$$\hat{\mathbf{X}}_T(m) = \mathbf{C}_T \mathbf{P}_T \mathbf{Q}_T^{(1)} \mathbf{P}'_T \mathbf{C}_T^{-1} (\mathbf{Y}_T - \bar{\mathbf{Y}}_0) \quad (3.14)$$

とする。このとき初期状態ベクトル  $\mathbf{x}_0 = \mathbf{y}_0$ 、とすると、状態変化  $\mathbf{x}_t^* = \mathbf{x}_t - \mathbf{x}_0$  ( $t = 1, \dots, T$ ) を推定していることになる。この操作は状態推定の議論では平滑化 (smoothing) とも呼ぶことができるが、ここで  $p$  次元のデータ直交過程は  $\mathbf{Z}_T = \mathbf{P}_T \mathbf{C}_T^{-1} (\mathbf{Y}_T - \bar{\mathbf{Y}}_0)$  で与えられる。

この非定常多次元時系列におけるトレンド・循環要素の状態推定法は Kunitomo, Sato and Kurisu (2018) が開発した SIML(separating information maximum likelihood) 法を利用、Nishimura, Sato and Takahashi (2019), Kunitomo, Awaya and Kurisu (2020), Kunitomo and Sato (2021), Sato and Kunitomo (2020, 2021) などが導入、展開した方法である。ここで、変換  $\mathbf{C}_T^{-1}$  は階差変換、変換  $\mathbf{P}_T$  は一種の実フーリエ変換であり、変換  $\mathbf{K}_T$  はこれら二つの変換の積 ((3.11)) で定義した。 $\mathbf{P}_T$  は実フーリエ変換であるから変換された変数  $\mathbf{Z}_T = (\mathbf{z}'_t)$  は時系列のスペクトル分解と呼ばれる周波数領域におけるデータ直交増分過程に対応している。以上の説明を利用すると、直観的には次のように要約できる。非定常時系列として得られる多次元時系列を階差をとり、ある実フーリエ変換を行いデータを直交分解、トレンド・循環成分の推定部分のみを抽出、これを反転させて状態推定を行うことによりトレンド・循環成分を推定するのである。

ここで導入した状態推定の方法はもっとも単純な場合であり、ベキ等行列  $\mathbf{Q}_T^{(1)}$  を他の変換に置き換えることで (これを  $\mathbf{Q}_T$  で表す) より一般化できる。例えば正整数  $0 < m_1 < m_1 + m_2 < T$

に対して  $m_2 \times [m_1 + m_2 + (T - m_1 - m_2)]$  選択行列  $\mathbf{J}_{m_1, m_2} = (\mathbf{O}, \mathbf{I}_{m_2}, \mathbf{O})$  として  $T \times p$  行列

$$\hat{\mathbf{X}}_n(m_1, m_2) = \mathbf{C}_T \mathbf{P}_T \mathbf{J}'_{m_1, m_2} \mathbf{J}_{m_1, m_2} \mathbf{P}_T \mathbf{C}_T^{-1} (\mathbf{Y}_n - \bar{\mathbf{Y}}_0) \quad (3.15)$$

と置けば変換された  $m_1 + 1$  番目～ $m_1 + m_2$  番目に対応する周波数の状態成分のフィルタリングが可能となる。ここで  $\mathbf{Q}_T$  は  $T \times T$  行列

$$\mathbf{Q}_T = \mathbf{J}'_{m_1, m_2} \mathbf{J}_{m_1, m_2} \cdot \quad (3.16)$$

で与えられる。

### データ周波数の解釈

観測データを変換したデータ直交過程を  $\mathbf{z}_k^{(T)} = \mathbf{z}_k(\lambda_k^{(T)})$  と表現しよう。この変換された  $(p \times 1)$  ベクトルの各成分はデータ周波数  $\lambda_k^{(T)} = (k - \frac{1}{2}) / (2T + 1)$  ( $k = 1, \dots, T$ ) に対応する確率変数である。例えばトレンド・循環要素にのみ関心があれば、 $\lambda_k^{(T)} = (k - \frac{1}{2}) / (2T + 1)$  ( $k = 1, \dots, m$ ) からデータ直交過程の最大周波数を

$$\lambda_{max}^{(T)} = \frac{m}{2T} \quad (3.17)$$

とすることが考えられる。20年間の月次データがあれば  $T = 240$ ,  $s = 12$  であり、1.5年以上の周期に関心があれば  $\lambda_{max}^{(T)} = 1.5/24$  とすると  $m^* = 480 \times (1.5/24) = 30$  となる。30年の四半期データであれば  $T = 120$ ,  $s = 4$  よりデータ最大周波数は  $\lambda_{max}^{(T)} = 1.5/8$ ,  $m^* = 240 \times (1.5/8) = 45$  である。

また (3.15) を利用して季節周期  $s (> 1)$  に対して  $m_1 = [2n/s] - h$ ,  $m_2 = 2h + 1$  ( $h$  は正整数) とするとこの周波数  $\lambda_s^* = 1/s$  の周辺成分のみを推定することが可能である。例えば月次データの場合には  $1/s = 1/12$  に対応する基本周期とともに季節周期は  $2/12, 3/12, 4/12, 5/12, 6/12$  などにデータ上では影響が表れることに注意する必要がある。なお、時間軸上の月次データの時系列プロットからは例えば1年間のサイクルと半年のサイクルを区別しにくいことに注意しておく。また、季節調整とはこうした季節周波数成分とその波及成分をデータから推定、その成分を取り除いた部分を季節調整系列と呼ぶ平滑化法と解釈できる<sup>9)</sup>。

## 3.2 外れ値・変化点を巡る問題

### 外れ値・変化点の処理問題

マクロ時系列では時々かなり大きな変動が観察される。しかし、特に直近に大きな変化が観察される場合には単なる一時的な変動(外れ値と呼ばれることがある)、あるいは大きな景気循環の一部と判断すべきなのかすぐに見極めるのはかなり困難となる。最近の日本のマクロ消費に関する事例としてはリーマンショック、東日本大震災、消費税、(そして多分)コロナショック、などが事例として考えられよう。こうした経済の変動、循環においてはとらえにくい時系列における一時的な変化をどう分析したらよいであろうか? さらに月次時系列では構成する曜日による変動、閏年効果、営業日効果(trading-day effects)などの影響も無視できないことが少なくない。

<sup>9)</sup> また四半期データで1年周期の季節性と半年周期の季節性は時間領域では区別できない。季節調整プログラム X-12-ARIMA は移動平均変換を基本としているので周波数領域での変換として解釈することは困難であり、季節調整値の妥当性は主に時間領域における季節調整系列の妥当性の判断に依存している。

こうした効果を総称して曜日効果と呼ぶと、本稿では変化点の扱いと同様の問題と理解して操作変数法により処理する。ここで基本モデル (3.1) に対しさらに説明変数 (操作変数と呼ぶ)  $\mathbf{w}_t$  の効果を時系列の構成要素  $\mathbf{c}_t(\mathbf{w})$  ( $t = 1, \dots, T$ ) としてさらに加えて加法分解モデル

$$\mathbf{y}_t = \mathbf{x}_t + \mathbf{s}_t + \mathbf{c}_t(\mathbf{w}) + \mathbf{v}_t \quad (t = 1, \dots, T) \quad (3.18)$$

を考察しよう。

ここではさらに操作変数の効果を  $\mathbf{c}_t(\mathbf{w}) = \mathbf{B}' \mathbf{w}_t$  ( $\mathbf{B}$  は  $r \times p$  係数行列) とする。本節では不規則変動 (ノイズ) 成分に対する外れ値モデルによる季節調整は元のノイズ  $\mathbf{v}_t$  から新ノイズ項  $\mathbf{v}_t^* = \mathbf{c}_t(\mathbf{w}) + \mathbf{v}_t$  として処理、季節性  $\mathbf{s}_t$  を推定してその影響を除去する問題と理解する。さらにトレンド・循環関数に変化点がある場合の季節調整とは、元のトレンド関数  $\mathbf{x}_t$  を新トレンド関数  $\mathbf{x}_t^* = \mathbf{x}_t + \mathbf{c}_t(\mathbf{w})$  としてから季節性  $\mathbf{s}_t$  を推定して除去する問題と理解する<sup>10)</sup>。なお観察される原系列  $\mathbf{y}_t$  は非定常であるから、非定常性を考慮して説明変数の効果を推定する必要があることに注意しておこう<sup>11)</sup>。

### 周波数回帰と SIML 平滑化

Sato and Kunitomo (2021) はここで説明した表現を利用して周波数領域における回帰 (Regression)SIML 法を提案している。ここでは応用上で必要となる事項のみを説明しておく。

まず  $T \times r$  行列  $\mathbf{W}_T = (\mathbf{w}_t')$  を説明変数のベクトル  $\mathbf{w}_t$  ( $r \times 1$ ) からなる観測行列とする。(なお予め  $\mathbf{W}_T^* = [\mathbf{I}_T - \mathbf{1}(\mathbf{1}'\mathbf{1})^{-1}\mathbf{1}']\mathbf{W}_T$  と基準化することも考えられる。) この操作変数行列から

$$\hat{\mathbf{W}}_T = \mathbf{J}_W \mathbf{P}_T \mathbf{C}_n^{-1} \mathbf{W}_T, \quad (3.19)$$

を構成する (ここで  $\mathbf{J}_W$  は  $q \times T$  選択行列,  $0 < q < T$  として  $q$  を選択する)。最小二乗法を適用して回帰部分として求められるベキ等行列 (ランク  $r$  の  $q \times q$  行列) を

$$\mathbf{Q}_W = \hat{\mathbf{W}}_T (\hat{\mathbf{W}}_T' \hat{\mathbf{W}}_T)^{-1} \hat{\mathbf{W}}_T' \quad (3.20)$$

とする。このとき平滑化 (smoothing) 行列  $\mathbf{Q}_W$  を利用すると変換

$$\hat{\mathbf{R}}_T = \mathbf{C}_T \mathbf{P}_T \mathbf{Q}_W \mathbf{P}_T \mathbf{C}_T^{-1} (\mathbf{Y}_T - \bar{\mathbf{Y}}_0) \quad (3.21)$$

を考えることができるが、以下では二つのタイプのスムージング法を考える。ここで平滑化行列はベキ等行列にとり  $\mathbf{Q}_W$  は  $\mathbf{Q}_W^2 = \mathbf{Q}_W$  を満足する必要があり、タイプ-I とタイプ-II と呼ぶ二つの方法を考察する。予め二つのタイプを考える理由に言及しておく、タイプ-I 平滑化は基本変換の応用であり、トレンド・循環要素の変化点処理などに適切、タイプ-II 平滑化は一次の変動ノイズの一部とも考えられる外れ値処理に適切と考えられる。周波数領域から見ると、この両方のタイプの意味はかなり異なるので、応用上では適切に利用することが重要となる。

(i) **タイプ-I 平滑化** : (3.21) を用いて観測行列  $\mathbf{Y}_T$  の回帰部分を構成するために  $\mathbf{J}_W = \mathbf{J}_m$  ( $\hat{\mathbf{W}}_T$  は  $T \times r$  行列,  $\mathbf{J}_m' = (\mathbf{I}_m, \mathbf{O})'$  は  $T \times m$  行列),  $T \times T$  行列

$$\mathbf{Q}_T^{(1)} = \mathbf{J}_m' \hat{\mathbf{W}}_T (\hat{\mathbf{W}}_T' \hat{\mathbf{W}}_T)^{-1} \hat{\mathbf{W}}_T' \mathbf{J}_m \quad (3.22)$$

<sup>10)</sup> 通常の季節調整では季節調整値系列は  $\mathbf{y}_t^* = \mathbf{y}_t - \hat{\mathbf{s}}_t$  ( $t = 1, \dots, T$ ) として構成されている。不規則変動を除くことも考えられるが、その場合には外れ値処理の効果を考慮する必要がある。

<sup>11)</sup> 例えば公的統計でよく利用されている季節調整 X-12-ARIMA の Reg-ARIMA モデルでは階差操作を行ってから回帰部分を最尤推定している。



とする。ここで操作変数による回帰の影響を除くため  $\mathbf{J}_W = \mathbf{J}_m$ ,  $\mathbf{J}_m = (\mathbf{I}_m, \mathbf{O})$  ( $m \times T$  選択行列,  $m \leq T$ ) および

$$\mathbf{Q}_T^{(2)} = \mathbf{J}_m' \mathbf{J}_m - \mathbf{Q}_T^{(1)} = \mathbf{J}_m' [\mathbf{I}_m - \hat{\mathbf{W}}_T' (\hat{\mathbf{W}}_T' \hat{\mathbf{W}}_T)^{-1} \hat{\mathbf{W}}_T'] \mathbf{J}_m \quad (3.23)$$

と置いた。このとき時系列のトレンド・循環成分をさらに分解して

$$\begin{aligned} \hat{\mathbf{R}}_T &= \mathbf{C}_T \mathbf{P}_T \mathbf{J}_m' \mathbf{J}_m \mathbf{P}_T \mathbf{C}_T^{-1} (\mathbf{Y}_T - \bar{\mathbf{Y}}_0) \\ &= \mathbf{C}_T \mathbf{P}_T \mathbf{J}_m' [\mathbf{Q}_T^{(1)} + \mathbf{Q}_T^{(2)}] \mathbf{J}_m \mathbf{P}_T \mathbf{C}_T^{-1} (\mathbf{Y}_T - \bar{\mathbf{I}}_0) \end{aligned} \quad (3.24)$$

と表現できる。ここでベキ等行列の性質  $\mathbf{Q}_T^{(2)} = \mathbf{Q}_T = \mathbf{Q}_T^{(1)} + \mathbf{Q}_T^{(2)} = \mathbf{J}_m' \mathbf{J}_m$  を満たし、トレンド・循環部分と回帰部分に分割されているが、この平滑化も対応する周波数により解釈できる。この分解により  $m$  個の低周波数成分による回帰からダミー回帰部分を推定したり、高周波成分を取り除くことが可能となるのである。

(ii) **Type-II スムージング** : 季節性の状態推定を組み込むときにはより複雑な変換が必要となる。  $q \times T$  行列  $\mathbf{F}_q$  を後述のように構成して変換した観測行列を

$$\mathbf{Y}_m^* = \mathbf{F}_q \mathbf{P}_T \mathbf{C}_T^{-1} (\mathbf{Y}_T - \bar{\mathbf{Y}}_0) \quad (3.25)$$

として、変換した操作変数

$$\mathbf{W}_m^* = \mathbf{F}_q \mathbf{P}_T \mathbf{C}_T^{-1} (\mathbf{W}_T - \bar{\mathbf{W}}_0), \quad (3.26)$$

に回帰する。タイプ II 平滑化は  $\mathbf{F}_q$  (以下で定める  $\mathbf{F}_q^Q$  あるいは  $\mathbf{F}_q^M$  とおく)

$$\mathbf{Q}_T^{(3)} = \mathbf{W}_m^* (\mathbf{W}_m^{*'} \mathbf{W}_m^*)^{-1} \mathbf{W}_m^{*'} \quad (3.27)$$

および

$$\mathbf{Q}_T^{(4)} = \mathbf{F}_q' \mathbf{F}_q - \mathbf{Q}_T^{(3)} \quad (3.28)$$

とすると、成分分解

$$\begin{aligned} \hat{\mathbf{X}}_T^* &= \mathbf{C}_T \mathbf{P}_T \mathbf{F}_q' \mathbf{F}_q \mathbf{P}_T \mathbf{C}_T^{-1} (\mathbf{Y}_T - \bar{\mathbf{Y}}_0) \\ &= \mathbf{C}_T \mathbf{P}_T [\mathbf{Q}_T^{(3)} + \mathbf{Q}_T^{(4)}] \mathbf{P}_T \mathbf{C}_T^{-1} (\mathbf{Y}_T - \bar{\mathbf{Y}}_0) \end{aligned} \quad (3.29)$$

が得られる。この場合  $\mathbf{Q}_T^{(3)} + \mathbf{Q}_T^{(4)} = \mathbf{F}_q' \mathbf{F}_q$  よりトレンド・循環成分と回帰成分が分解される。

特に四半期データの場合には  $s = 4$ 、周波数  $\lambda_s = 1/4, 1/2$  ( $1/2$  は 2 四半期周期,  $1/4$  は 4 四半期周期) 対応するが、四半期データから 12 ヶ月周期と 6 か月周期の波動を区別するのは困難である。そこで  $m_1 = [2n/s]$  として  $(T - 2h - 1) \times T$ ,  $(T - 3h - 2) \times (T - 2h - 1)$  選択行列

$$\mathbf{J}_1^Q = \begin{bmatrix} \mathbf{I}_{m_1 - (h+1)} & \mathbf{O} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & \mathbf{I}_{T - m_1 - h} \end{bmatrix}, \quad \mathbf{J}_2^Q = [\mathbf{I}_{T - 3h - 2}, \mathbf{O}], \quad (3.30)$$

より  $q \times T$  行列

$$\mathbf{F}_q^Q = \mathbf{J}_2^Q \mathbf{J}_1^Q \quad (3.31)$$

を構成する。 ( $h > 0$  は小さな正整数、 $q = T - 3h - 2$  とする。) 月次データ ( $s = 12$ ) では 12 ヶ月周期の波動と 6, 4, 3, 2.4, 2 月周期の波動を区別することは困難である。そこで  $m_l = l[2n/s]$

として  $(T - l(2h + 1)) \times (T - (l - 1)(2h + 1))$  行列 ( $l = 1, \dots, 5$ ),  $(T - 5(2h + 1) - (h + 1)) \times (T - 5(2h + 1))$  行列をそれぞれ

$$\mathbf{J}_l^M = \begin{bmatrix} \mathbf{I}_{m_l - (l-1)(2h+1) - (h+1)} & \mathbf{O} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & \mathbf{I}_{T - m_l - h} \end{bmatrix}, \mathbf{J}_6^M = [\mathbf{I}_{T - 11h - 6}, \mathbf{O}], \quad (3.32)$$

より構成して

$$\mathbf{F}_q^M = \prod_{j=1}^6 \mathbf{J}_{7-j}^M \quad (3.33)$$

とすればよい。( $q = T - 11h - 6$  となる。)

### ダミー変数の例：

異常値処理や変化点処理のために様々なダミー変数を考えることができる。なお非定常時系列を扱う場合には何らかの基準化を行うことが重要と考えられることに注意しておく。様々なダミー変数を利用することが可能であるが、ここでは季節調整などに役立つ典型的な例を幾つか述べておくことにとどめる<sup>12)</sup>。ダミー変数ベクトルは  $\mathbf{w}_s$  ( $s = 1, \dots, T$ ) で表す。

**例 1**(AO 変数, Type-II)：外れ値変数は  $w_s = 1$  ( $s = t$ ) および  $w_t = 0$  ( $s \neq t, s = 1, \dots, T$ ) で定める。

**例 2**(LS 変数, Type-I)：レベル・シフト変数は  $w_s = 0$  ( $s < t$ ) および  $w_t = 1$  ( $s \geq t, s = 1, \dots, T$ ) によって定める。

**例 3**(Ramp 変数, Type-I)：ランプ変数は  $w_s = 1$  ( $s < t_0$ ),  $w_s = 1 - (t - t_0)/(t_1 - t_0)$  ( $t_0 \leq t \leq t_1$ ),  $w_t = 0$  ( $s \geq t_1$ ) により定める。

**例 4**(Double-Ramp 変数, Type-I)：二重(ダブル)ランプ変数は  $w_s = 1$  ( $s < t_0$ ),  $w_s = 1 - (t - t_0)/(t_1 - t_0)$  ( $t_0 \leq t \leq t_1$ ),  $w_s = (t - t_1)/(t_2 - t_1)$  ( $t_1 \leq t \leq t_2$ ),  $w_t = c$  ( $s \geq t_2$ ) によって定める。

実際の非定常経済時系列の分析では様々なダミー変数を考えることができる。以下の分析では幾つかのダミー変数を利用している。また、経済時系列データが与えられたときのダミー変数の選択は係数の有意性をはじめ、ガウス分布の仮定に基づく AIC(赤池情報量基準) を利用すること、などが考えられる (Sato and Kunitomo (2021)6 節を参照)。

## 4 応用例

### 4.1 マクロ指標の構成問題

欠損値のある不完全時系列を観測可能な非定常多次元時系列の実現値と見なし、マクロ月次指標の状態を推定する方法の応用例を考察する。例えば消費系列については需要側の家計調査による時系列データ、供給側の生産動態統計による時系列データはそれぞれ月次データとして利用可能である。これに対して目的変数である GDP 最終消費系列は四半期データのみが利用可能であ

<sup>12)</sup> 様々なダミー変数を含む R プログラムを開発中である。

る。現行の四半期データの最終消費系列は需要側と供給側の月時系列から推計されているので、四半期データに対応する隠れた真の月次系列が存在する、と仮定するのが妥当だろう。

そこで非定常多次元時系列表現を応用してマクロ消費指標の状態推定問題として考察しよう。マクロ指標として作成されている観測原四半期データを  $y_{1t}$  ( $t = 12(l-1) + 3j; l = 1, \dots, n; j = 1, 2, 3, 4; T = 12n$ ) とする。これに対して四半期データより高頻度な月次データとして利用可能なデータを  $p-1$  次元ベクトル  $\mathbf{y}_{2t}$  ( $t = 12(l-1) + j, l = 1, \dots, n; j = 1, \dots, 12; T = 12n$ ) とすると、ここでの統計的問題は利用可能な月次データ情報の下で観測不能な真のトレンド・循環成分の状態  $x_{1t}$  ( $t = 12(l-1) + j, l = 1, \dots, n; j = 1, \dots, 12$ ) を推定することである。

### 実用的な月次指標の状態推定

ノイズを含む非定常多次元時系列において一部の変数の観測が不完全な場合の最適な状態推定の計算アルゴリズムはよく知られていない。そこで比較的容易に実現できる方法を提案するが、前節のトレンド・循環要因、不規則変動要因、季節変動要因などを含む時系列に対してトレンド・循環成分を状態推定するロバストな方法である。不完全観測である目的系列の状態を他の完全観測な時系列によるトレンド・循環成分の回帰による推定法は、全体の変数の個数  $p = (p_1 + p_2)$ 、目標変数の個数  $p_1$ 、説明変数の個数  $p_2$  として一般化することも可能であるが、ここでは実際に応用した  $p = 4, p_1 = 1, p_2 = 3$  の場合を例示として説明する。

まず利用可能な四半期データよりトレンド・循環変数間の関係を推定するには、原データより季節変動成分・ノイズ成分を取り除く必要がある。

(i) TC 成分として1年以上の周期成分を SIML フィルタリングにより抽出する。具体的には  $m$  を季節周期  $2T/12$  よりも十分に小さく設定する必要がある。この操作により季節変動と不規則成分を取り除くことが可能となる。この際、四半期データのみが利用可能な GDP 消費は月次情報がないが月次系列として対応する四半期系列を仮置きすることにより月次系列から季節成分と不規則成分を取り除くことが可能である。

(ii) 次に観測される四半期データ  $\mathbf{y}_{1t}$  ( $t = 12(l-1) + 3j, l = 1, \dots, n; j = 1, 2, 3, 4$ ) より状態変数  $\mathbf{x}_{1t}$  ( $t = 12(l-1) + 3j, l = 1, \dots, n; j = 1, 2, 3, 4$ ) を推定する。(目的状態変数を  $\mathbf{x}_{1t}$  ( $t = 12(l-1) + 3j, l = 1, \dots, n; j = 1, 2, 3, 4$ ) とする。) 月次データが利用可能な系列  $y_{2,t}, y_{3,t}, y_{4,t}$  (一般には  $p_2$  ベクトル) から状態変数  $x_{2t}, x_{3t}, x_{4t}$  ( $t = 12(l-1) + j, l = 1, \dots, n; j = 1, \dots, 12$ ) を推定する。

ここで月次系列の TC 成分間の関係を推定する必要があるが、特に  $\mathbf{x}_t$  が非定常変数であることに注意しよう。変数  $x_{1t}$  ( $t = 12(l-1) + 3j, l = 1, \dots, n; j = 1, 2, 3, 4$ ) の状態推定値を 3次元 (一般には  $p-1$  次元) 説明変数ベクトル  $\mathbf{x}_t^{(d)} = (x_{2t}, x_{3t}, x_{4t})'$  ( $t = 12(l-1) + 3j, l = 1, \dots, n; j = 1, 2, 3, 4$ ) に回帰、回帰係数ベクトルを  $(b_2, b_3, b_4)'$  とする。このとき状態変数ベクトル間の関係は最小二乗法を利用して得られる係数ベクトル  $\mathbf{b} = (b_2, b_3, b_4)'$  により推定できる。ここで真の係数は  $\boldsymbol{\beta} = (\beta_2, \beta_3, \beta_4)' = \boldsymbol{\beta}'_x$  であり、

$$x_{1t} = \beta_2 x_{2t} + \beta_3 x_{3t} + \beta_4 x_{4t} + u_t \quad (4.1)$$

と表現する。ここで目的変数が月次指標なので、説明変数を基準化して定数項が無い回帰モデルであり、係数ベクトルは共和分ベクトル、との解釈が可能となり、誤差項  $u_t$  は定常系列として処理できる。なお、もう一つの定式化としては初期値をゼロとして、階差変数  $\Delta x_{1t}$  ( $t = 12(l-1) + 3j, l = 1, \dots, n; j = 1, 2, 3, 4$ ) の推定値を 3次元ベクトル (一般には  $p-1$  次元) 説明

変数ベクトル  $\Delta \mathbf{x}_t^{(d)} = (\Delta x_{2t}, \Delta x_{3t}, \Delta x_{4t})'$  ( $t = 12(l-1) + 3j, l = 1, \dots, n; j = 1, 2, 3, 4$ ) に回帰、回帰係数ベクトルを  $(b_2^*, b_3^*, b_4^*)'$  とすることも考えられる。しかしながら、我々の事例では幾つかの問題から実用性に乏しかったのでその詳細の説明は省略する。またトレンドを水準・階差で説明する統計モデルも考えられる。

(iii) さらに外れ値と変化点の処理を行い、推定した係数を用いて目的の状態変数の推定が可能となる。誤差の評価は新たなデータが得られるたびに状態推定

$$\hat{x}_{1t} = b_2 x_{2t} + b_3 x_{3t} + b_4 x_{4t} \quad (t = 12(l-1) + j; l = 1, \dots, n; j = 1, \dots, 12) \quad (4.2)$$

を繰り返し行えば目的変数の状態推定値が得られる。

こうして SIML フィルターで推定された状態変数を利用すれば、原理的にはトレンド・循環成分、季節成分、原系列-季節成分、などの状態推定が可能となる。しかし月次の季節成分の状態推定はかなり難しい。月次のノイズ成分を除いた TC 成分の推定を利用することの方が、よりロバストな指標の作成に向いているとも考えられる。すなわち回帰係数はかなり安定しているものの、時間の推移と共に一定とは見なしにくい。そこで利用可能な四半期データが更新されるたびに統計的フィルタリングとスムージングを繰り返し行えば、全体として整合的なマクロ指標を作成することができることを指摘しておこう。

なお注意事項としてはここで扱う非定常多次元時系列間の関係性を示す統計モデルは標準的な線形回帰モデルではなく、データが非定常かつ観測誤差が存在することである。したがって、通常の回帰モデルの推定法についての正当化は適用できないことに注意しておく。また、非定常系列の要因分解であるから階差系列を利用する場合にも誤差項に MA 単位根が混入してくる可能性などに注意が必要である。よく知られているように非定常系列を他の非定常系列に回帰すると見かけ上はある程度のフィットは得られるが推定結果が信頼できるとは限らないことにも注意する必要があるのである<sup>13)</sup>。

## 4.2 消費指数の推定例

これまでに説明した方法をマクロ消費指標の TC 月次系列の状態推定に応用した例を示しておく。図 1 には GDP 実質最終支出の四半期データから仮置きで月次系列を構成し、 $m=20$  としてトレンド・循環部分を状態推定した系列を滑らかな曲線で示しておく。実際には月次データが得られないのであるから、この曲線は月次消費の情報が何もないという条件下での月次のトレンド・循環部分の初期推定値と見ることができる。素朴に四半期データから月次データを仮置きで作成すると、3 節の方法により月次のトレンド・循環成分が滑らかに変動する状態が推定されていることが興味深い。この月次系列からのスペクトル情報を図 2 に示しておく。

次に補助系列として利用した月次消費系列として家計支出の月次系列 (図 3)、商業動態統計の月次系列 (図 5)、第三次指数の月次系列 (図 7) をそれぞれ示しておくが、各図ではトレンド・循

<sup>13)</sup> ここでの議論については例えば Hayashi (2000)10 章の議論が参考となる。ここで説明している最小二乗法は共積分ベクトルの SOLS と呼ばれる方法であり、より整合的かつ一般的には DOLS(Dynamic OLS) 法が開発されている。この方法では  $x_{1t}$  を  $\mathbf{x}_t^{(d)} (= (x_{2t}, x_{3t}, x_{4t})')$  だけではなく  $\mathbf{x}_{t-1}^{(d)}, \mathbf{x}_{t+1}^{(d)}$  などに回帰、バイアスを補正するというものである。本稿の例では推定結果はあまり変わらなかったため説明を省略する。

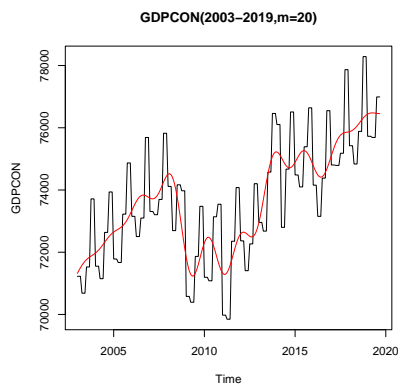


図1 GDP-最終消費支出 ( GDP 実質最終支出の四半期データから月次系列を仮置きで補完。m=20としてトレンド・循環部分を状態推定した系列が滑らかな曲線。)

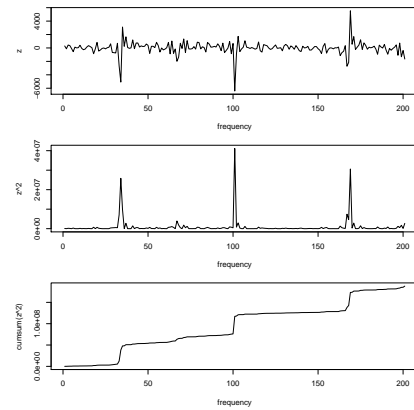


図2 GDP-最終消費支出のスペクトル情報 ( GDP 実質最終支出の四半期データから構成した月次系列のスペクトル情報。上から  $\{z_k\}$  系列,  $\{z_k^2\}$  系列,  $\{z_k^2\}$  の累積系列を示している。)

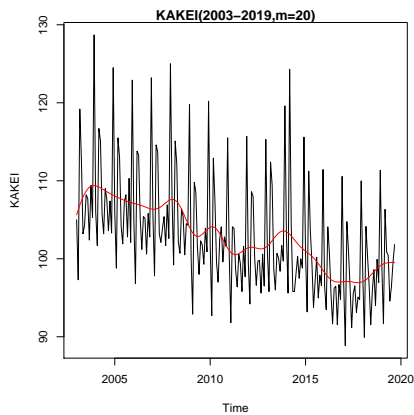


図3 家計支出系列 (家計調査から計算された家計支出の月次データ。m=20としてトレンド・循環部分を状態推定した系列が滑らかな曲線。)

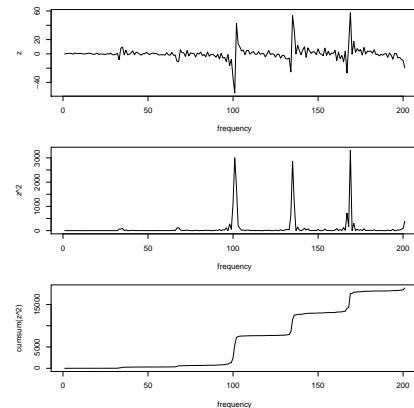


図4 家計支出系列のスペクトル情報 (家計支出系列の月次系列のスペクトル情報。上から  $\{z_k\}$  系列,  $\{z_k^2\}$  系列,  $\{z_k^2\}$  の累積系列を示している。)

環成分の状態推定値が滑らかな曲線で示されている。また SIML 変換を利用して作成した家計支出の月次系列からのスペクトル情報 (図 4)、商業動態統計の月次系列からのスペクトル情報 (図 6)、第三次指数の月次系列からのスペクトル情報 (図 8) をそれぞれ示しておく。

ここでこれらの系列についての観察事実を幾つか述べておこう。

- (i) 家計調査から計算された家計消費の水準はやや減少傾向のトレンド成分と循環成分、商業動態系列の水準はややプラスのトレンド成分と循環成分、第三次産業系列の水準はかなりのプラスのトレンド成分と循環成分の状態が推定されている。
- (ii) 月次 3 系列の循環成分は関連があるものの、各系列のデータ作成の事情 (例えばサンプリングに伴う季節変動成分や不規則変動成分はそれぞれのデータに固有な成分がある) を反映してか、互いに異なる側面を無視することはできない。四半期 GDP 最終消費系列はおおむね月次 3 系列の循環部分の合成として理解できる。
- (iii) トレンド・循環成分の状態推定では  $m=20$  としたためであるが、各系列共にトレンド・循環成分の状態推定値には季節成分や不規則成分の影響はあまり見られない。

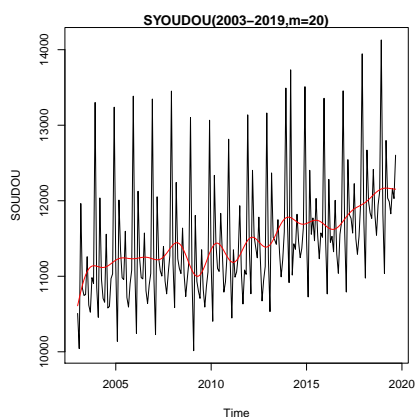


図5 商業動態系列（商業動態調査から得られた消費動向系列の月次データ。m=20としてトレンド・循環部分を状態推定した系列が滑らかな曲線。）

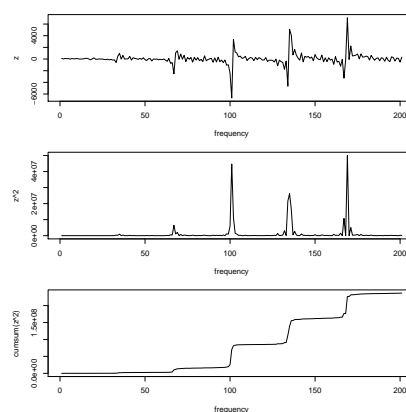


図6 商業動態系列のスペクトル情報（商業動態系列の月次系列のスペクトル情報。上から  $\{z_k\}$  系列、 $\{z_k^2\}$  系列、 $\{z_k^2\}$  の累積系列を示している。）

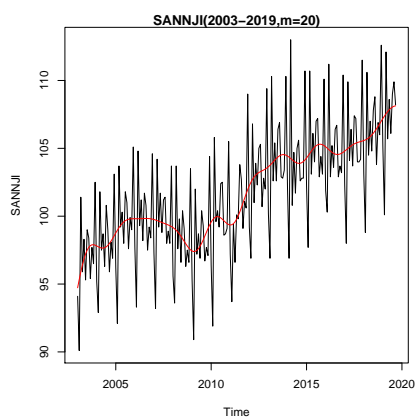


図7 第三次産業指数系列（商業動態統計から得られた第三次産業指数系列の月次データ。m=20としてトレンド・循環部分を状態推定した系列が滑らかな曲線。）

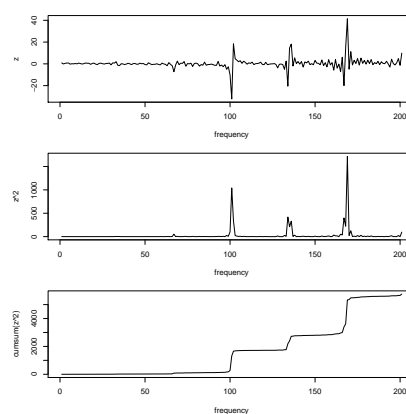


図8 第三次産業指数系列のスペクトル情報（第三次産業指数系列の月次系列のスペクトル情報。上から  $\{z_k\}$  系列、 $\{z_k^2\}$  系列、 $\{z_k^2\}$  の累積系列を示している。）

以上のような考察から、月次3系列を利用することで四半期データでは観察できないトレンド・循環変動に関する情報を得ることができると考えられる。

以下ではスペース上の制約もあり、2003年1月～2018年12月のデータを用いてTCI公表系列とかなり整合的となる  $m = 36$  のときの分析結果について報告する。（むろん観測期間やパラメーター  $m$  の選択などにより結果についてはかなりの差が生じ得ることにあらかじめ注意しておこう。）まず四半期GDPデータより求めた月次の系列を被説明変数とした回帰分析の推定結果を表1にまとめておいた。通常回帰モデルとの整合性を持たせる意味から水準系列を回帰分析した結果であるが、非定常性のためもあり回帰の適合度は良好である。（なお前節で述べたように状態推定した結果として求めた系列を利用しているので  $t$  値は  $t$ -分布に従うわけではないが、有意性を示す一つの意味がある情報と理解すべきである。）

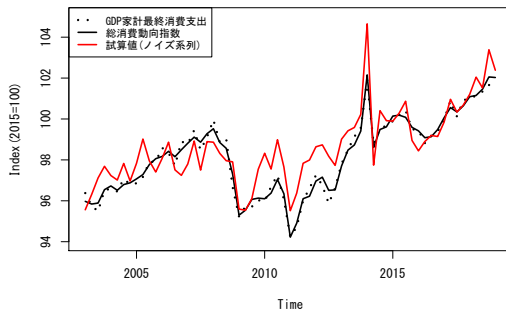


図9 消費指数(試算系列, 四半期, TC+ ノイズ)

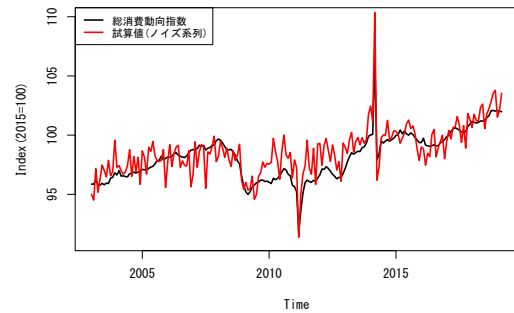


図10 消費指数(試算系列, 月次, TC+ ノイズ)

表1 : 回帰の推定結果

( $m = 36$ )

	(家計調査)	(商業動態)	(第三次産業)
(係数)	138.601	4.562	53.562
(誤差)	7.379	0.384	40.343
(t-値)	18.783	11.887	1.322

次に2002年～2018年のデータのトレンド・循環成分を状態推定、回帰分析で推定された係数を所与として2019年までの指数を求め、消費動向系列の公表指数の値と比較した。なおここでの回帰はレベルによる回帰であるが、階差系列による回帰から原系列を構成するには初期値などを処理した。ここでは(TC)+(ノイズ系列)に基づく2017年1月を基準点とする一つの例を検討したので図9, 図10として示しておく。四半期の推定値は公表値にかなり近くなるが、ノイズを含む月次系列は構成要素の異なる季節性によると思われるがかなりの不規則変動が見られる。いずれにせよ、ここで提案している方法により得られた結果の推定されたトレンド・循環成分は公表値系列に近いことが分かる。また異常値や変化点を無視した場合には各系列の不規則な変動を反映した月次変動が見られるが、そうした系列を過去に遡って系列を求めてみると、公表指数とはかなり乖離が見られる。そうした現象の原因は回帰係数の安定性及び異常値や変化点の扱いなどであると見られるので、実用的な一つの解決策として変化点や回帰係数の定期的な更新が必要と解釈できよう。

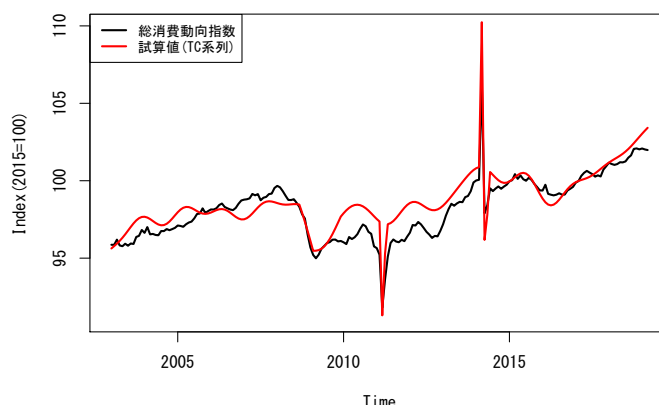


図 11 CTI との比較 (試算系列, 月次)

ここではデータ分析の詳細は省略するが、ダミー変数を利用しないで推定される消費系列と公表されている消費指数と比較すると、2011年3月に発生した東日本大震災、および2014年4月の消費税引き上げ、などの影響が顕著にあらわれる。公表系列ではこのような大きな変動をトレンド・循環成分とともに指標の中に組み入れていると解釈できよう。むしろ、こうした一時的効果や大きな変化点などの効果は本稿3節で説明した方法を利用することで解決できる。適切なダミー変数を利用して、外れ値と変化点の要素を組み入れて周波数回帰を行い、月次指数を推定すればよいのである。ただし今回の分析では単純なAOタイプの消費税ダミーを利用したが、多少、効果が大きく出すぎているようにも考えられる。こうした状況ではRamp変数などの利用をさらに検討する必要があるかもしれない。

最後に、図9と図10の系列からダミー変数を利用して月次TC系列の状態を推定した月次の消費系列と既存のCTI系列を図11に示しておく。この消費系列は $m=36$ (月次)として回帰によりTC系列に基づく系列を作成、さらにリーマンショック、東日本大震災、消費税導入を高岡ramp<sup>14)</sup>、AOなどのダミー変数により処理、作成した消費系列である。ここでは曜日効果を含め、様々な可能性を検討する必要がある。また、モデル選択やフィルター選択なども検討する必要がある、その意味ではなお多くの実務的課題があるが、その検討については別の機会としたい。今回の限られた分析を全体として評価すると、本稿で提案する方法で構成した四半期系列・月次系列と公表系列との比較からここで求めた消費系列におけるトレンド・循環部分・不規則変動の妥当性を示すことが可能となると評価できよう。

なお現行のCTI公式指数の計算はマニュアルによればかなり複雑であり、季節調整の扱い方を含め、改善可能性があるとも解釈できるだろう。ここでは本稿で提案しているアプローチはかなり単純な上に結果的には類似のマクロ時系列を容易に推定できることを最後に強調しておく。

<sup>14)</sup> ダミー変数 tkramp を利用したが詳しくは高岡 (2013) を参照。



## 5 まとめと展望

本稿ではノイズを含む非定常時系列が年次・四半期・月次など周期が異なる頻度で観測される時、互いに矛盾しないような状態変数の推定問題・平滑化問題を考察し、SIML フィルタリングが有効であることを確認した。また具体的な例として目的時系列が不完全観測の時に利用可能な月次時系列データを利用して月次のマクロ消費指標の作成を実例として示した。ここで提案している方法はかなり簡単な方法であるが、結果として月次消費の公表系列にかなり類似した系列を推定できる。ここで示した方法によれば多くのマクロ時系列で観察されているトレンド・循環要素、季節変動要因、不規則変動要因、外れ値や変化点要因など非定常時系列データに関わる様々な要因を統一的に扱うことができるので、季節調整系列などの意味もより明確になると考えられる。

なお本稿で考察した問題と関連する政府統計において開発されている方法としてベンチマーク法を挙げることができる<sup>15)</sup>。既存のベンチマーク法では年次推計と四半期推計を他の四半期推計の時系列パターンと統合的に推計することが主な内容となるが、トレンド・循環成分・季節成分を統合的に行うこと、直近の年次データが得られた後に推計すること、など検討すべき課題がある。こうした問題は、本稿で説明したようにノイズを含む多次元時系列のスペクトル分解を利用することで、かなり解決できると思われる。

さて本稿で利用した周波数領域 (frequency domain) の統計分析を経済時系列に適用する試みは古典的な Granger and Hatanaka (1964) まで遡ることができる。しかしながら、経済時系列が古典的な意味で定常過程とは見なしにくいためだろうか、Hayashi (2000) に代表される最近のマクロ計量経済分析ではスペクトル分析はほとんど議論されないようである。そうした中で例えば Müller and Watson (2018) などによりトレンド・循環要素の分析が提唱されている。実は Müller and Watson (2018) の方法は Kunitomo and Sato (2017) の特殊な場合と解釈が可能である<sup>16)</sup>。非定常多次元経済時系列を扱うことは現在においてもそれほど簡単な統計的方法は知られていないが、本稿での議論が実務的な応用上の諸問題を解決するきっかけとなれば幸いである。

## 参考文献

- [1] Anderson, T.W. (1971), *The Statistical Analysis of Time Series*, John-Wiley.
- [3] 内閣府・国民経済計算部 (2010) 「四半期別 GDP 速報の推計方法」, (平成 18 年 7 月改定).
- [5] Granger, C. and M. Hatanaka (1964), *Analysis of Economic Time Series*, Pinceton University Press.
- [6] Hayashi, H. (2000), *Econometrics*, Princeton University Press.
- [7] Kitagawa, G. (2010), *Introduction to Time Series Modeling*, CRC Press.
- [8] Kunitomo, N. and S. Sato (2017), "Trend, Seasonality and Economic Time Series : the

<sup>15)</sup> ベンチマーク法については例えば国友・川崎 (2011) を挙げておく。

<sup>16)</sup> 詳しくは Kunitomo and Sato (2021) を参照。

- Non-stationary Errors-in-variables Models,” SDS-4, MIMS, Meiji University, <http://www.mims.meiji.ac.jp/publications/2017-ds>.
- [9] Kunitomo, N. and S. Sato (2021), ”A Robust-filtering Method for Noisy Non-Stationary Time Series,” *Japanese Journal of Statistics and Data Science (JJSD, Springer)*, 4, 373-410, Open Access.
- [10] Kunitomo, N. , Sato and D. Kurisu (2018), *Separating Information Maximum Likelihood Estimation for High Frequency Financial Data*, Springer.
- [11] Kunitomo, N., N. Awaya and D. Kurisu (2020), ”Some Properties of Estimation Methods for Structural Relationships in Non-stationary Errors-in-Variables Models,” *Japanese Journal of Statistics and Data Science (JJSD)*,3, 73-101, Springer.
- [12] Nishimura, G. K., Seisho Sato, Akihiko Takahashi (2019), ”Term Structure Models During the Global Financial Crisis: A Parsimonious Text Mining Approach,” *Asia-Pacific Financial Markets*.
- [13] Sato, S. and Kunitomo, N. (2020), ”On Backward Smoothing for Noisy Non-stationary Time Series,” SDS-15, MIMS, Meiji University.
- [14] Sato, S. and Kunitomo, N. (2021), ”Frequency Regression and Smoothing for Noisy Non-stationary Time Series,” SDS-20, MIMS, Meiji University.
- [15] Müller, U. and M. Watson (2018), ”Long-run Covariability,” *Econometrica*, 86-3, 775-804.
- [16] 国友直人・川崎能典 (2011) 「景気判断と平滑化問題」, 経済学論集, 77-1, 2-19. (東京大学経済学部).
- [17] 総務省統計局 (2021) 「消費動向指数 (CTI)」, <https://www.stat.go.jp/data/cti/index.html>.
- [18] 高岡慎 (2013) 「X-12-ARIMA におけるモデル選択の安定性の改善について—法人企業統計の事例—」, 琉球大学経済研究 = Ryukyu University Economic Review(86): 53-77.

### 補論 A：数理的補論

この補論では 3 節で説明した変換によりノイズを含む非定常多次元データの分析が有効となる理由を追加的に説明する。ここで説明している内容の数学的証明は省略するが Kunitomo-Sato (2021) の Proposition 1, Proposition 2 の証明として与えられている。また統計的時系列解析の標準的文献として Anderson (1971), Kitagawa (2010) を挙げておく。本稿では非定常時系列として得られる多次元時系列を周波数領域に変換, 周波数領域における直交分解を利用, トレンド・循環成分を抽出して, 元に戻して状態推定・平滑化を行っている。この方法の鍵となるのが変換された離散確率過程  $\mathbf{Z}_T = (\mathbf{z}_t')$  の周波数領域における直交性である。

ここで  $\theta_{tk} = \frac{2\pi}{2T+1}(t - \frac{1}{2})(k - \frac{1}{2})$  ( $t, k = 1, \dots, T$ ) に対して

$$p_{tk}^{(T)} = \frac{1}{\sqrt{2T+1}}(e^{i\theta_{tk}} + e^{-i\theta_{tk}}) \quad (\text{A.1})$$

とする。(虚数  $i$  は  $i^2 = -1$  とする。) このとき初期条件  $\mathbf{y}_0$  を所与として

$$\mathbf{z}^{(T)}(\lambda_k^{(T)}) = \sum_{t=1}^T p_{tk}^{(T)} \mathbf{r}_t^{(T)}, \quad \mathbf{r}_t^{(T)} = \mathbf{y}_t - \mathbf{y}_{t-1}, \quad (\text{A.2})$$

と表現すると, 変換系列  $\mathbf{Z}_T$  はデータのある種の実フーリエ変換  $\mathbf{z}^{(T)}(\lambda_k^{(T)})$  ( $k = 1, \dots, T$ ) から成り, データ周波数  $\lambda_k^{(T)} (= (k - 1/2)/(2T + 1))$ , における実フーリエ変換による実データ直交増分過程の分解となる. このとき次の結果が成立する (Kunitomo and Sato (2021)).

**定理 A-1**(離散時間) 確率過程  $\mathbf{r}_t$  ( $t = 1, \dots, T$ ) は (エルゴード的) 定常確率過程, 期待値と共分散関数は  $\mathbf{E}[\mathbf{r}_t] = \mathbf{0}$ ,  $\mathbf{\Gamma}(h) = \mathbf{E}(\mathbf{r}_t \mathbf{r}_{t-h}')$  は

$$\sum_{h=0}^{\infty} \|\mathbf{\Gamma}(h)\| < \infty. \quad (\text{A.3})$$

を満たし, さらに4次モーメントの有界性を仮定する.

(i)  $\mathbf{z}^{(T)}(\lambda_k^{(T)}) = \sum_{t=1}^T p_{tk}^{(T)} \mathbf{r}_t^{(T)}$ ,  $\mathbf{r}_t^{(T)}$  が (エルゴード的) 定常確率過程で  $\mathbf{E}[\mathbf{r}_t] = \mathbf{0}$ , 対称化実スペクトル密度行列

$$\mathbf{f}_{SR}(\lambda) = \mathbf{\Gamma}(0) + \sum_{h=1}^{\infty} \cos(2\pi h \lambda) [\mathbf{\Gamma}(h) + \mathbf{\Gamma}(-h)], \quad (\text{A.4})$$

が正定符号, 有界性となることを仮定する.  $\lambda_k^{(n)} \rightarrow a$ ,  $\lambda_{k'}^{(n)} \rightarrow b$  ( $0 < a < b < \frac{1}{2}$ ).  $T \rightarrow \infty$  ( $k \rightarrow \infty$ ) のとき

$$\begin{bmatrix} \mathbf{z}^{(T)}(\lambda_k^{(T)}) \\ \mathbf{z}^{(T)}(\lambda_{k'}^{(T)}) \end{bmatrix} \xrightarrow{w} N_{2p} \left[ \mathbf{0}, \begin{bmatrix} \mathbf{f}_{SR}(a) & \mathbf{0} \\ \mathbf{0} & \mathbf{f}_{SR}(b) \end{bmatrix} \right]. \quad (\text{A.5})$$

となり, 漸近正規性が成立する。

なお第3節で用いた非定常多次元時系列モデル (3.1)-(3.5) では各ノイズ項  $\mathbf{e}_t^{(x)}$ ,  $\mathbf{e}_t^{(v)}$ ,  $\mathbf{e}_t^{(s)}$  の4次モーメントが有界なら上の条件を満足する。

トレンド・循環成分の長期的分散共分散行列は周波数0における状態変数  $\Delta \mathbf{x}_t$  のスペクトルであるのでその推定問題は Kunitomo and Sato (2017) が検討している。ノイズ  $\mathbf{v}_t$  と  $\mathbf{v}_t^{(x)}$  が互いに独立な確率変数列の場合, SIML 推定量は

$$\mathbf{G}_m = \frac{1}{m} \sum_{k=1}^m \mathbf{z}_k^{(T)}(\lambda_k^{(T)}) \mathbf{z}_k^{(T)}(\lambda_k^{(T)})' \quad (\text{A.6})$$

により与えられる ( $0 < m < T$ )。の性質を考察している. Kunitomo and Sato (2017, 2021), Sato and Kunitomo (2021) はより一般的な場合にも分散共分散行列の推定量が良い漸近的性質を持つことを示している。

## 補論 B：関連する R プログラム

ここで本稿で利用した状態推定の R プログラムの基本例を示しておく。より一般のプログラムは本論文の筆者の一人により準備中である。なお時系列データ（月次データを想定している）の読み込み、回帰分析などの R プログラムは R の入門的解説書にあるので省略する。R プログラムはフリーソフトであり R-Web サイトから無償でダウンロードできる。

```
# <2020-2-4 version >
# <Decomposition of Noisy Non-stationary Multivariate Time Series>
# <y is the data vector (monthly time series from 2002.1)>
n= length(y)
m= as.integer(n/10)+3
Ci= diag(rep(1,n))
diag(Ci[-1,-n]) = -1
C= solve(Ci)
h= 0
P= sqrt(2/(n+h+0.5))*cos((2*pi/(2*(n+h)+1))* outer(seq(n)-0.5,seq(n)-0.5) )
z= P %*% Ci %*% (y-y[1])
y2= ts(y,start=c(2002,1),frequency=12)
op= par(mfrow=c(1,1))
plot.ts(y2)
# <time domain decomposition>
# <trend and cycles>
z1= C%*% P[,1:10]%*% z[1:10]+y2[1]
lines(ts(z1,start=c(2002,1),frequency=12),col=2)
# <Cycles, Seasonals and Noise>
# <in the decomposition>
z2= C%*% P[,85:95]%*% z[85:95]+mean(y2)
z3= C%*% P[,96:180]%*% z[90:180]+mean(y2)
lines(ts(z2,start=c(2002,1),frequency=12),col=3)
lines(ts(z3,start=c(2002,1),frequency=12),col=4)
op= par(mfrow=c(3,1),mar=c(4.5,4,1,1))
ts.plot(y2,z1)
abline(v=c(2014,4))
ts.plot(ts(z2-mean(y2),start=c(2002,1),frequency=12))
ts.plot(ts(z3-mean(y2),start=c(2002,1),frequency=12))
# <frequency decomposition>
op= par(mfrow=c(3,1),mar=c(4.5,4,1,1))
ts.plot(z, xlab="frequency")
ts.plot(z^2, xlab="frequency")
ts.plot(cumsum(z^2), xlab="frequency")
```