

SDS-19

経済時系列の状態推定とマクロ指標

国友直人・佐藤整尚・櫻井智章

March 2021

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

経済時系列の状態推定とマクロ指標¹⁾

国友直人²⁾

佐藤整尚³⁾

櫻井智章⁴⁾

2021年3月10日

(要約)

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ているマクロ時系列をめぐる課題を解決する 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法を紹介する。主要なマクロ経済時系列では、公表系列が互いに矛盾することなくマクロ経済について適切な情報を提供することが望まれるが、SIML フィルタリング法は有用と思われる。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多次元時系列モデルを利用して全体として整合的に経済時系列データを扱う一般的方法を説明し、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報告する。

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が、特に主要なマクロ経済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、年次系列などが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では各系列は基礎とする統計調査にもとづき系列を作成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、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状況を理解するためには、早めに正確なデータを作成し、公表されることが望ましいが、国全体のデータを作成するには多くの情報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最速で四半期データ、しかもかなりの時間が経過した後で公表され、一次速報、二次速報、確報と改訂されている。早期にデータを収集するために消費系列や投資系列のように月次系列や四半期系列が公表された後により多くの情報を含み、より正確と思われるデータが利用可能となり、その結果として、しばらく後に四半期系列の公表数値が公表され、改訂される。こうした改訂で生じる誤差はデータ収集の過程で生じる観測誤差と考えられるが、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の情報と見かけ上で矛盾するように受けとめられ

1) 未定稿 2021-3-10. 本稿は学術研究であり、本稿の内容は著者が所属している職場の見解とは無関係である。

2) 明治大学政治経済学部

3) 東京大学経済学部

4) 総務省統計局

る事例なども散見される。さらに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るということも問題をより複雑にしていると考えられる。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、総務省統計局から家計調査などが従来からあるが、月次系列として統計局により数年前から消費動向指数 (CTI) が作成、公表されている⁵⁾。

本稿ではマクロ時系列を巡る課題に対して、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策を検討する。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説明する。

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事例についての問題を述べる。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説明し、さらに外れ値や変化点の処理などを含め、統計モデルに分析よるトレンド・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説明し、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。

2 研究の動機：マクロ消費の問題

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よる家計消費の動向と企業に対するサンプリング調査による商業動態統計などから得られる生産・販売の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のためもあり、標本数は 8,000~9,000 世帯から調査データが得られる。他方、企業側からの消費データとして商業動態統計は生産・販売の調査データであり、標本数が約 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ない数値なども算入されている。こうしたことからマクロ消費を計測する際には家計消費と企業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。ここで民間や政府のエコノミストが日本のマクロ消費動向として重視しているのは 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が、季節性の処理などにも注意する必要がある。⁶⁾さらに毎年作成されている GDP 推計の確報値の作成にあたって

⁵⁾ 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。

⁶⁾ 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 を参照されたい。

は、全体の金額は生産面についてより細か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マクロ消費動向を把握する上では重要な事実だろう。

さて GDP 消費の速報値の基礎となっている家計調査、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のも、近年における月次レベルでの消費動向を分析するには需要面と生産面におけるマクロ指標としてよく利用されている。(その動向については図 3, 図 5, 図 7 を参照されたい。)ところが、これらの月次データの変動が同一とは限らないことが少なくないことから、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は十分ではなく、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、生産・販売面における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、月次で得られる生産・販売面の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面などでの問題も無視はすべきでない。加えて、速報的に月次レベルで集計されている基本的な消費の動向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 消費の動向を統合的に理解し、消費動向を的確に把握する必要性があると云えよう。

ここで近年になり良く利用されるようになってきている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推定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) の問題として考える。GDP 最終消費は四半期ベースで様々なデータの集計値として収集するので、月次データとしては欠損値が多く存在する場合と考えるのである。消費については同時に月次データとして利用可能な情報があるので同時に考えることで多次元時系列データの一部分が欠損しているとみて、欠損値の状態を推定することが考えられる。

このとき重要なことは、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の理解が必要となることである。こうした現実のマクロ時系列を構成する幾つかの要因が重なる為、時系列分析では古くから統計的方法としては移動平均法が知られているが、それを単純に利用するだけでは、現実的に利用可能な分析を行うことが困難であることを指摘しておこう。また、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しておこう。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ことが重要である。さらに実務的には異なる系列に対して異なる Reg-ARIMA モデルを利用していることには重要な意味がある。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と考えられるが、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なのであるが、複数の時系列が統合的に処理されているとは限らないのである。ここでは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次に説明しよう。

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考察する。時刻 t の多次元経済時系列を $p \times 1$ 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であり、 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 $s = 4$) とすると、 i 年 j 期は時刻 $t = (i - 1)s + j$ ($i = 1, \dots, n; j = 1, \dots, s$)、初期値ベクトルは \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

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

を考察する⁷⁾。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 (四半期データは $s = 4$, 月次データは $s = 12$) とする。次にトレンド・循環成分 \mathbf{x}_t ($t = 1, \dots, T$) (確率変数列) は非定常 I(1) 過程として

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

ラグ作用素 $\mathcal{L}\mathbf{x}_t = \mathbf{x}_{t-1}$, $\Delta = 1 - \mathcal{L}$ を利用して不規則成分 \mathbf{v}_t (確率変数列) およびトレンド・循環成分を生成する確率変数列 $\mathbf{v}_t^{(x)}$ は移動平均表現

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

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

を持つとする。ここで各成分を生成する (ベクトル) 確率変数列 $\mathbf{e}_t^{(v)}$, $\mathbf{e}_t^{(x)}$, 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 $\mathbf{E}(\mathbf{e}_t^{(v)}) = \mathbf{0}$, $\mathbf{E}(\mathbf{e}_t^{(x)}) = \mathbf{0}$, $\mathbf{E}(\mathbf{e}_t^{(v)} \mathbf{e}_t^{(v)'}) = \Sigma_e^{(v)}$ (正定符号), $\mathbf{E}(\mathbf{e}_t^{(x)} \mathbf{e}_t^{(x)'}) = \Sigma_e^{(x)}$ (非負定符号) とする。ここで移動平均表現における係数について条件 $\mathbf{C}_j^{(v)} = O(\rho^j)$, $\mathbf{C}_j^{(x)} = O(\rho^j)$ ($0 \leq \rho < 1$) を仮定しよう。このとき係数 and 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。(なおここで $p \times p$ 行列 $\mathbf{C} = (c_{ij})$ に対するノルムは $\|\mathbf{C}\| = \max_{i,j} |c_{ij}|$ とする。)

季節成分は正整数 s ($s > 1$), $n, T = sn$ として確率変数列を \mathbf{s}_t ($t = 1, \dots, T$) は定常過程

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

⁷⁾ 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型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。

という表現を持ち, $\mathbf{e}_t^{(s)}$ は i.i.d. 確率変数列, 期待値と共分散行列は $\mathbf{E}(\mathbf{e}_t^{(s)}) = \mathbf{0}$, $\mathbf{E}(\mathbf{e}_t^{(s)} \mathbf{e}_i^{(s)'}) = \Sigma_e^{(s)}$ 、行列係数について条件 $\mathbf{C}_{j_s}^{(s)} = O_p(\rho^j)$ ($0 \leq \rho < 1$) を満足すると仮定する。

ここで $\mathbf{f}_{\Delta x}(\mu)$, $\mathbf{f}_s(\mu)$, $\mathbf{f}_v(\mu)$ 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と, (3.3)-(3.5) のように移動平均表現を持つとの仮定の下でそれぞれ

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

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

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

で与えられる⁸⁾. ここでは基準化として $\mathbf{C}_0^{(x)} = \mathbf{C}_0^{(s)} = \mathbf{C}_0^{(v)} = \mathbf{I}_p$, $i^2 = -1$ とした. 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

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

で与えられる. ここでトレンド・循環成分, ノイズ成分の長期分散共分散行列をそれぞれ

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

$$(3.11) \quad \mathbf{\Omega}_v = f_v(0) = (\omega_{gh}^{(v)}).$$

で表そう. 特にトレンド・循環成分の長期分散共分散が重要であり, Kunitomo-Sato (2017) は (3.10) の SIML 推定量を提案している⁹⁾.

\mathbf{K}_n 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・循環要素を状態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) $\mathbf{y}_t = (y_{kt}) (k = 1, \dots, p; t = 1, \dots, T)$ より観測データ行列 ($T \times p$) $\mathbf{Y}_T = (\mathbf{y}_t')$ を構成する. ここで $T \times p$ 行列 $\mathbf{Y}_0 = \mathbf{1}_T \mathbf{y}_0'$ は固定された初期値ベクトル \mathbf{y}_0 , $\mathbf{1}_T' = (1, \dots, 1)$ より作る. このデータ行列に対して \mathbf{K}_n -変換 \mathbf{Y}_T より $\mathbf{Z}_T (= (\mathbf{z}_k'))$ は

$$(3.12) \quad \mathbf{Z}_T = \mathbf{K}_T (\mathbf{Y}_T - \bar{\mathbf{Y}}_0), \mathbf{K}_T = \mathbf{P}_T \mathbf{C}_T^{-1},$$

により定義する. ただし変換行列は

$$(3.13) \quad \mathbf{C}_T^{-1} = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & 0 & \cdots \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}_{n \times n},$$

⁸⁾ 例えば Anderson (1971) を参照のこと. なお本稿の記号とは若干異なっている.

⁹⁾ 数理的付論の (A.6) を参照.

$$(3.14) \quad \mathbf{P}_T = (p_{jk}^{(T)}), p_{jk}^{(T)} = \sqrt{\frac{2}{T + \frac{1}{2}}} \cos \left[\frac{2\pi}{2T + 1} \left(k - \frac{1}{2}\right) \left(j - \frac{1}{2}\right) \right].$$

この変換のスペクトル分解により $\mathbf{C}_T^{-1} \mathbf{C}_T'^{-1} = \mathbf{P}_T \mathbf{D}_T \mathbf{P}_T' \mathbf{D}_T$ 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 を用いたトレンド・循環要素の状態推定を考えよう。データ行列から関心のある周波数成分のみを取り出すために $0 < m < T$ 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

$$(3.15) \quad \hat{\mathbf{X}}_T(m) = \mathbf{C}_T \mathbf{P}_T \mathbf{Q}_T^{(1)} \mathbf{P}_T' \mathbf{C}_T^{-1} (\mathbf{Y}_T - \bar{\mathbf{Y}}_0)$$

として、初期状態ベクトル $\mathbf{x}_0 = \mathbf{y}_0$, $\mathbf{Y}_0 = \mathbf{X}$ としたときの状態変化 $\mathbf{x}_t^* = \mathbf{x}_t - \mathbf{x}_{t-1}$ ($t = 1, \dots, T$) を推定することを考えよう。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、 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) 法を発展させ、Nishimura, Sato and Takahashi (2019), Kunitomo-Sato (2021), Sato and Kunitomo (2020a, b) などが導入、展開したものである。ここで、変換 \mathbf{C}_T 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で定義される。 \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直交増分過程に対応している。以上の説明を利用してここでの議論を直観的に要約すると次のようになる：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種の実フーリエ変換を行い、直交分解を利用してトレンド・循環成分のみを抽出、さらに元に戻して状態推定を行うという解釈になる。

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 $\mathbf{Q}_T^{(1)}$ を置き換えることで (\mathbf{Q}_T とする) より一般化することできる。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

$$(3.16) \quad \hat{\mathbf{X}}_n(m_1, m_2) = \mathbf{C}_T \mathbf{P}_T \mathbf{J}_{m_1, m_2}' \mathbf{J}_{m_1, m_2} \mathbf{P}_T' \mathbf{C}_T^{-1} (\mathbf{Y}_n - \bar{\mathbf{Y}}_0)$$

と置けば $m_1 + 1 \sim m_1 + m_2$ に対応する周波数で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

$$(3.17) \quad \mathbf{Q}_T = \mathbf{J}_{m_1, m_2}' \mathbf{J}_{m_1, m_2}.$$

で与えられる。

データ周波数の解釈

観測データから推定した直交過程を $\mathbf{z}_{Tk} = \mathbf{z}_n(\lambda_k^{(T)})$ と表現しよう。このとき変換されたベクトルの各成分は周波数 $\lambda_k^{(T)} = (k - \frac{1}{2}) / (2T + 1)$ ($k = 1, \dots, T$) に対応している。例えばトレンド

ド・循環要素にのみ関心があれば、 $\lambda_k^{(T)} = (k - \frac{1}{2}) / (2T + 1)$ ($k = 1, \dots, m$) ととれば直交過程の最大周波数は大体

$$(3.18) \quad \lambda_{max}^{(T)} = \frac{m}{2T} .$$

となる。例えば 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。また 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。

なお季節周期 $s (> 1)$ に対して $m_1 = [2n/s] - [m/2]$, $m_2 = m$ とするとこの周波数成分周辺の成分のみを推定することも可能である。ただし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 $2/12, 3/12, 4/12, 5/12, 6/12$ などにもその影響が表れることに注意する必要がある。月次データから計算される直交過程からはでは 1 年間のサイクルと半年のサイクルを区別できない。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¹⁰⁾。

3.2 外れ値・変化点・モデル選択を巡る問題

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見極めるのはかなり困難である。最近の日本のマクロ消費に関する実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、などが実例として考えられよう。こうした経済の変動、循環を巡る問題とはとらえにくい時系列における一時的な変化をどうとらえたらよいであろうか。また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。こうした効果を総称して曜日効果と呼ぶと、変化点の扱いと同様の問題と理解できる。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 $\mathbf{c}_t(\mathbf{w})$ $t = 1, \dots, T$ を加えて加法分解モデル

$$(3.19) \quad \mathbf{y}_t = \mathbf{x}_t + \mathbf{s}_t + \mathbf{c}_t(\mathbf{w}) + \mathbf{v}_t \quad (t = 1, \dots, T),$$

を考察しよう。

ここでは操作変数の効果を $\mathbf{c}_t(\mathbf{w})$ とおいたが、例えば外れ値モデルによる季節調整とは元のノイズ \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 を推定して除去する問題と理解できる。また、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できる¹¹⁾。

¹⁰⁾ 例えば四半期データで 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。

¹¹⁾ 通常の季節調整では季節調整値系列は $\mathbf{y}_t^* = \mathbf{y} - \hat{\mathbf{s}}_t$ ($t = 1, \dots, T$) 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。

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¹²⁾。

ダミー回帰と SIML 平滑化

Sato-Kunitomo (2020b) はここでの問題に対して周波数領域における Dummy-Regression-SIML 法を用いて一般的に対処する方法を提案しているので、その概略を説明しておこう。

まず $\mathbf{W}_T = (\mathbf{w}'_t)$ を説明変数のベクトル \mathbf{w}_t ($q \times 1$ からなる観測行列 ($n \times q$) とする。予め基準化して

$$(3.20) \quad \mathbf{W}_T^* = [\mathbf{I}_T - \mathbf{1}(\mathbf{1}'\mathbf{1})^{-1}\mathbf{1}']\mathbf{W}_T$$

として、この基準化されたダミー変数行列から操作変数行列

$$(3.21) \quad \hat{\mathbf{W}}_T = \mathbf{J}_W \mathbf{P}_T \mathbf{C}_n^{-1} \mathbf{W}_T^*,$$

を作る (ここで \mathbf{J}_W は $m_q \times T$ 選択行列, $m_q < T$ で m_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 q の $m_q \times m_q$ 行列) を

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

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

$$(3.23) \quad \hat{\mathbf{R}}_T = \mathbf{C}_T \mathbf{P}_T \mathbf{Q}_W \mathbf{P}_T \mathbf{C}_T^{-1} (\mathbf{Y}_T - \bar{\mathbf{Y}}_0)$$

を考えることができる。ここで二つのスムージング法を考えよう。ただし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る、次にタイプ-Iとタイプ-IIと呼ぶ方法を考察しよう。予め二つのタイプを考える理由に言及しておく、タイプ-II 平滑化は一次的変動ノイズの一部とも考えられる外れ値処理に相当であり、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相当と考えられる。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、適切に利用することが重要と判断している。

(i) **タイプ-I 平滑化** : (3.21) における観測行列 \mathbf{Y}_T の回帰部分を構成するために $\mathbf{J}_W = \mathbf{I}_m$ ($m_q = m$) ($\hat{\mathbf{W}}_T$ は $m \times q$ 行列, $\mathbf{J}'_m = (\mathbf{I}_m, \mathbf{O})'$ は $T \times m$ 行列), さらに $T \times T$ 行列

$$(3.24) \quad \mathbf{Q}_T^{(2)} = \mathbf{J}'_m \hat{\mathbf{W}}_T (\hat{\mathbf{W}}_T' \hat{\mathbf{W}}_T)^{-1} \hat{\mathbf{W}}_T' \mathbf{J}_m$$

とする。

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

$$(3.25) \quad \mathbf{Q}_T^{(3)} = \mathbf{J}'_m \mathbf{J}_m - \mathbf{Q}_T^{(2)} = \mathbf{J}'_m [\mathbf{I}_m - \hat{\mathbf{W}}_T (\hat{\mathbf{W}}_T' \hat{\mathbf{W}}_T)^{-1} \hat{\mathbf{W}}_T'] \mathbf{J}_m$$

¹²⁾ 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推定している。

と置いた。このとき時系列のトレンド・循環成分はさらに分解されて

$$(3.26) \quad \begin{aligned} \hat{\mathbf{R}}_T &= \mathbf{C}_T \mathbf{P}_T \mathbf{J}'_m \mathbf{J}_m \mathbf{P}_T \mathbf{C}_T^{-1} (\mathbf{Y}_T - \bar{\mathbf{Y}}_0) \\ &= \mathbf{C}_T \mathbf{P}_T \mathbf{J}'_m [\mathbf{Q}_T^{(2)} + \mathbf{Q}_T^{(3)}] \mathbf{J}_m \mathbf{P}_T \mathbf{C}_T^{-1} (\mathbf{Y}_T - \bar{\mathbf{Y}}_0) \end{aligned}$$

と表現できる。ここでベキ等行列の性質 $\mathbf{Q}_T^2 = \mathbf{Q}_T = \mathbf{Q}_T^{(2)} + \mathbf{Q}_T^{(3)} = \mathbf{J}'_m \mathbf{J}_m$ 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、この平滑化は簡単に解釈できるよう。この分解では m 個の低周波数部分の回帰によりダミー回帰部分を推定したり、取り除くことが可能となる。ダミー変数をトレンド・循環に加える場合には $\hat{\mathbf{R}}_T$ を利用し、取り除く場合は次の操作を行えばよい。まず $\mathbf{J}_W = \mathbf{I}_T$, $m = T$ として \mathbf{Y}_T から回帰部分を取り除き、

$$(3.27) \quad \mathbf{X}_T^{(1)} = \mathbf{C}_T \mathbf{P}_T [\mathbf{I}_T - \mathbf{Q}_T^{(2)}] \mathbf{P}_T \mathbf{C}_T^{-1} (\mathbf{Y}_T - \bar{\mathbf{Y}}_0)$$

とする。さらに二回目の平滑化を $\mathbf{X}_T^{(1)}$ から

$$(3.28) \quad \mathbf{X}_T^{(2)} = \mathbf{C}_T \mathbf{P}_T \mathbf{J}'_m \mathbf{J}_m \mathbf{P}_T \mathbf{C}_T^{-1} \mathbf{X}_T^{(1)}$$

とすることが考えられる（ただし $m < T$ とする）。

なおこの種の繰り返し操作は Sato-Kunitomo (2020a) が導入した繰り返し平滑化と見ることも可能であるが、この種の繰り返し計算は早い収束が保証されるので何度か繰り返して実行することも可能である。

(ii) **Type-II スムージング** : タイプ II 平滑化は観測行列 \mathbf{Y}_T に対して変換 \mathbf{J}_m ($\hat{\mathbf{W}}_T$ は $T \times m_q$ 行列, $\mathbf{J}'_m = (\mathbf{I}_m, \mathbf{O})'$ は $T \times m$ 行列), $T \times T$ 行列

$$(3.29) \quad \mathbf{Q}_T^{(4)} = (\mathbf{J}'_m, \hat{\mathbf{W}}_T) [(\mathbf{J}'_m, \hat{\mathbf{W}}_T)]^{-1} (\mathbf{J}'_m, \hat{\mathbf{W}}_T)'$$

により定める ($0 < m + m_q < T$ とする)。ここで逆行列のよく知られた表現からこのベキ等行列は二つの和に分解でき

$$(3.30) \quad \begin{aligned} \mathbf{Q}_T^{(4)} &= \mathbf{J}'_m \mathbf{J}_m + \hat{\mathbf{W}}_T^* (\hat{\mathbf{W}}_T^{*'} \hat{\mathbf{W}}_T^*)^{-1} \hat{\mathbf{W}}_T^{*'} \\ &= \hat{\mathbf{W}}_T (\hat{\mathbf{W}}_T' \hat{\mathbf{W}}_T)^{-1} \hat{\mathbf{W}}_T' + \mathbf{J}'_m (\mathbf{J}_m^* \mathbf{J}_m^*)^{-1} \mathbf{J}_m^* , \end{aligned}$$

となる。ただし

$$\hat{\mathbf{W}}_T^* = [\mathbf{I}_T - \mathbf{J}'_m \mathbf{J}_m] \hat{\mathbf{W}}_T .$$

and

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

である。この分解は一見すると平滑化操作は二つの表現を持つが同一であることを意味している。第一の表現では最初に m 個の低周波数部分と操作変数に回帰してトレンドを推定する。第二の表現では最初に m 個の低周波数の成分に回帰、次に残りの $T - m$ 個の周波数に回帰することを意味している。この二つの操作は数理的には同一となる。

特に $\mathbf{J}_W = (\mathbf{O}, \mathbf{I}_{T-m})$ とすると, $\mathbf{J}_m \mathbf{J}'_W = \mathbf{O}_m$ となり直交性が得られ, また $\mathbf{J}'_{T-m} = (\mathbf{O}, \mathbf{I}_{T-m})'$ ($T \times (T-m)$ 行列) であるから

$$(3.31) \quad \mathbf{Q}_T = \mathbf{J}'_m \mathbf{J}_m + \mathbf{J}'_{T-m} \mathbf{Q}_W \mathbf{J}_{T-m}$$

と分解される. この表現から分解は m 個の低周波数によりトレンド要素を推定, 残りの $T-m$ の周波数によりダミー変数を用いた非トレンド要素を推定する操作に対応している.

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¹³⁾. ダミー変数は w_s ($s = 1, \dots, T$) とする.

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 $w_t = 1$ ($s \geq t, s = 1, \dots, T$) により定められる.

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) : ダブル・ランプ変数は $w_s = 1$ ($s < t_0$), $w_s = 1 - (t - t_0)/(t_1 - t_0)$ ($t_0 \leq t \leq t_1$), $w_s = (t - t_1)/(t_2 - t_1)$ ($t_1 \leq t \leq t_2$), $w_t = c$ ($s \geq t_2$) により定める.

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が, 新たなダミー変数を含めて佐藤 (2020) は X12SIML プログラムにより実現している. 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よう.

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データと見なし, マクロ月次指標の状態を推定する方法を考察する. 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る. 四半期データの最終消費系列は概念的に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

¹³⁾ 佐藤 (2020) は様々なダミー変数を含む R プログラムを開発している.

データに対応する隠れた月次系列が存在する、と仮定するのが妥当だろう。

ここでは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を考察する。マクロ指標として作成されている四半期データを y_{1t} ($t = 4(i-1) + 3j; i = 1, \dots, n; j = 1, 2, 3, 4; T = 12n$) 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} ($t = 12(i-1) + j, i = 1, \dots, n; j = 1, \dots, 12; T = 12n$) 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} ($t = 12(i-1) + j, i = 1, \dots, n; j = 1, \dots, 12$) を推定することとなる。

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推定法について既存の計算アルゴリズムは知られていないと思われる。ここでは比較的容易に実現する一般的な方法を提案する。既に前節でトレンド・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おいてトレンド要因を状態推定する比較的簡単でロバストな方法を説明した。ここでは不完全観測である目標変数と呼ぶべき時系列から状態変数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を説明しよう。全体の変数の個数 $p = (p_1 + p_2)$ 、目標変数の個数 $p_1 = 1$ 、説明変数の個数 p_2 として一般化することも可能だが、特に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) 次に四半期データ \mathbf{y}_{1t} ($t = 12(i-1) + 3j, i = 1, \dots, n; j = 1, 2, 3, 4$) より状態変数 \mathbf{x}_{1t} ($t = 12(i-1) + 3j, i = 1, \dots, n; j = 1, 2, 3, 4$) を推定できるのでターゲット変数を \mathbf{x}_{1t} ($t = 12(i-1) + 3j, i = 1, \dots, n; j = 1, 2, 3, 4$) 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) に対しては月次データ $t = 12(i-1) + j, i = 1, \dots, n; j = 1, \dots, 12$ より状態変数 x_{2t}, x_{3t}, x_{4t} ($t = 12(i-1) + j, i = 1, \dots, n; j = 1, \dots, 12$) 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、このとき \mathbf{x}_t が非定常変数であることに注意する必要がある変数 x_{1t} ($t = 12(i-1) + 3j, i = 1, \dots, n; j = 1, 2, 3, 4$) の状態推定値を 3 次元ベクトル (一般には $p-1$ 次元) 説明変数ベクトル $\mathbf{x}_t^{(d)} = (x_{2t}, x_{3t}, x_{4t})'$ ($t = 12(i-1) + 3j, i = 1, \dots, n; j = 1, 2, 3, 4$) 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ベクトル $\mathbf{b} = (b_1, b_2, b_3)'$ により推定できる。ここで真の係数は $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3)' = \boldsymbol{\beta}'_x$ であり、

$$(4.1) \quad x_{1t} = \beta_1 x_{2t} + \beta_2 x_{3t} + \beta_3 x_{4t} + u_t$$

が対応する。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はない回帰、係数ベクトルは共和分ベクトル、と解釈でき、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 Δx_{1t} ($t = 12(i-1) + 3j, i = 1, \dots, n; j = 1, 2, 3, 4$) 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 $\Delta \mathbf{x}_t^{(d)} = (\Delta x_{2t}, \Delta x_{3t}, \Delta x_{4t})'$ ($t = 12(i-1) + 3j, i = 1, \dots, n; j = 1, 2, 3, 4$) に回帰、回帰係数ベクトルを $(b_1^*, b_2^*, b_3^*)'$ とすることも考えられる。しかしながら我々の実例では幾つかの問題によりあまりうまく行かなかったので詳細は省略する。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よう。(iii) さらに本稿で議論した外れ値と変化点処理を行えば、ターゲット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

$$(4.2) \quad \hat{x}_{1t} = \hat{\beta}_1 x_{2t} + \hat{\beta}_2 x_{3t} + \hat{\beta}_3 x_{4t} \quad (t = 12(i-1) + j; i = 1, \dots, n; j = 1, \dots, 12)$$

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ので、TC 成分の推定を利用すること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よう。すなわち回帰係数はかなり安定してはいるものの時間の推移と共に一定とは見なしにくい。また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。

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、妥当性を利用することは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¹⁴⁾。

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赤の曲線で示しておく。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から、

¹⁴⁾ 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和分ベクトルの 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 $\mathbf{w}_t (= (x_{2t}, x_{3t}, x_{4t})')$ だけではなく $\mathbf{w}_{t-1}, \mathbf{w}_{t+1}$ などに回帰、バイアスを補正するというものである。本稿の例では結果はあまり変わらなかったため省略する。

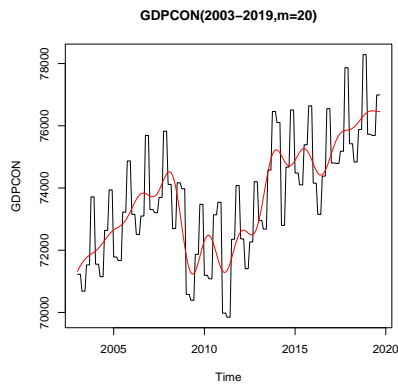


図 1 GDP-最終消費支出 (GDP 実質最終支出の四半期データから月次系列を仮置きで補完。m=20としてトレンド・循環部分を状態推定した系列が赤の曲線。)

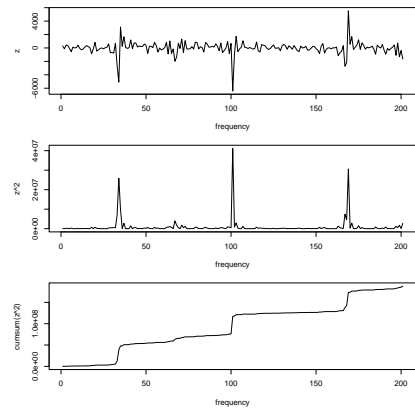


図 2 GDP-最終消費支出のスペクトル情報 (GDP 実質最終支出の四半期データから構成した月次系列のスペクトル情報。上から $\{z_k\}$ 系列, $\{z_k^2\}$ 系列, $\{z_k^2\}$ の累積系列を示している。)

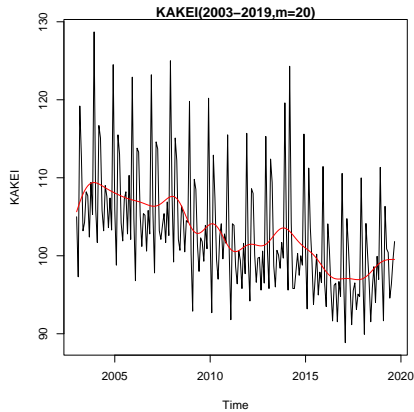


図 3 家計支出系列 (家計調査から計算された家計支出の月次データ。m=20としてトレンド・循環部分を状態推定した系列が赤の曲線。)

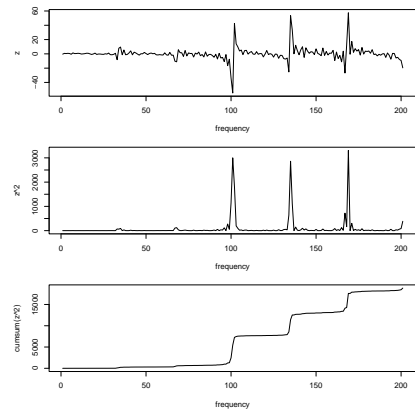


図 4 家計支出系列のスペクトル情報 (家計支出系列の月次系列のスペクトル情報。上から $\{z_k\}$ 系列, $\{z_k^2\}$ 系列, $\{z_k^2\}$ の累積系列を示している。)

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の推定されることが興味深い。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), 消費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・循環成分の状態推定値を赤の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), 消費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いた。

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) 家計調査の水準はややマイナスのトレンド成分と循環成分, 消費動向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

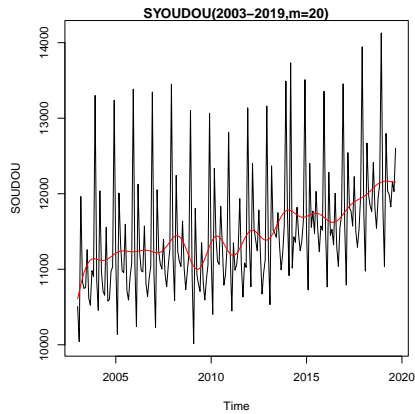


図5 消費動向系列（消費動向調査から得られた消費動向系列の月次データ。m=20としてトレンド・循環部分を状態推定した系列が赤の曲線。）

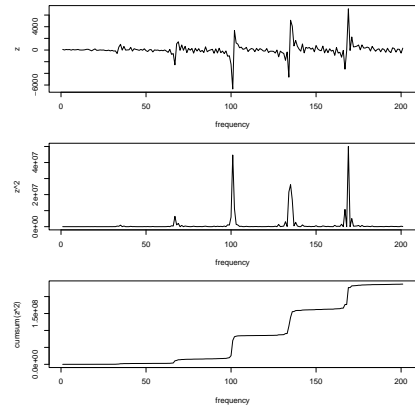


図6 消費動向系列のスペクトル情報（消費動向系列の月次系列のスペクトル情報。上から $\{z_k\}$ 系列、 $\{z_k^2\}$ 系列、 $\{z_k^2\}$ の累積系列を示している。）

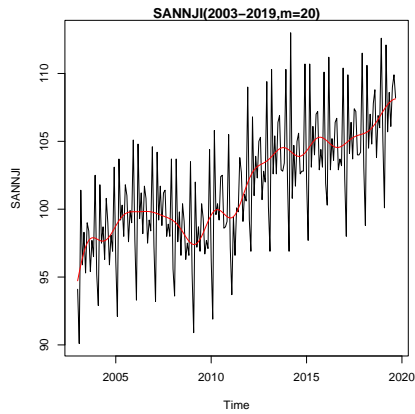


図7 第三次産業指数系列（商業動態統計から得られた第三次産業指数系列の月次データ。m=20としてトレンド・循環部分を状態推定した系列が赤の曲線。）

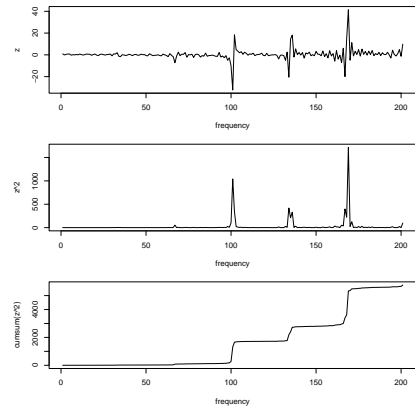


図8 第三次産業指数系列のスペクトル情報（第三次産業指数系列の月次系列のスペクトル情報。上から $\{z_k\}$ 系列、 $\{z_k^2\}$ 系列、 $\{z_k^2\}$ の累積系列を示している。）

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、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には季節成分や不規則成分の影響はあまり見られない。

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。

ここで月次データ（2002年1月～2019年12月の原系列，2015年基準）を説明変数、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にまとめておく。通常の回帰モデルとの整合性を持たせる意味から階差系列による結果であるが、回帰の適合度 ($R^2 = 0.858$) であり良好と云えるだろう。(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

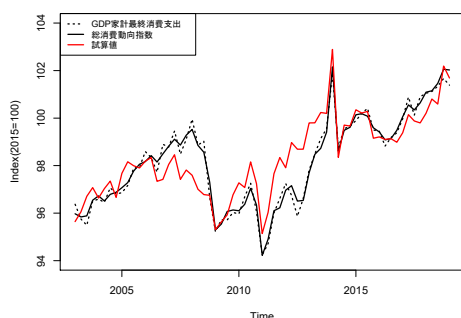


図9 消費指数(試算系列, 四半期)

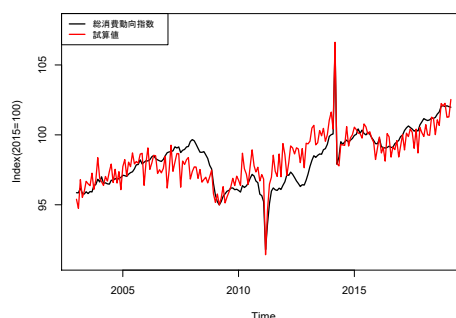


図10 消費指数(試算系列, 月次)

ある情報と理解すべきだろう。

表1：回帰の推定結果
($m = 20$)

	(家計変数)	(消動変数)	(第三次産業変数)
(係数)	125.474	3.063	40.462
(誤差)	24.585	0.1922	6.529
(t-値)	5.105	15.934	6.197

次に2002年～2019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、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。直近の推定値は公表値にかなり近くなり、ここで提案している方法は簡単ではあるが、結果は公表値系列の作成方法を再生していることが分かる。ノイズ系列を含む月次系列の場合にはノイズを反映した月次変動が見られる。また過去にさかのぼって指数を求めると、公表指数とある程度の乖離が見られることがある。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は回帰係数の定期的な更新が必要と考えられる。

なおダミー変数を利用しない、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差が見られた。公表系列では幾つかの理由からこのような二つのイベントをトレンド・循環成分とともに指標の中に組み入れている、と考えられる。こうした一時的効果、あるいは統計的には大きな変化点、などの効果を本稿で開発している指数に組み入れることは必要がある。既に説

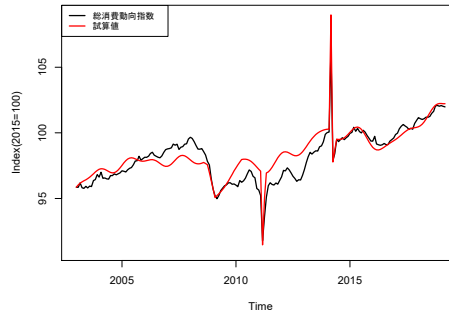
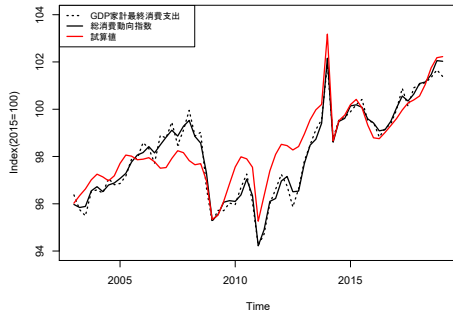


図 11 CTI との比較 (試算系列, 四半期) 図 12 CTI との比較 (試算系列, 月次)

明した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ることで回帰を行い、月次指数を推定することが可能であることが分かった。

最後に、説明したダミー変数を利用して TC 系列から状態を推定した消費系列と既存の CTI 系列を図 11 と図 12 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、消費税導入をダミー変数により処理、構成した消費系列である。四半期系列の比較からここで求めた四半期消費系列のトレンド・循環部分・不規則変動の妥当性を示していると評価できよう。このことから既存の消費系列は各月の消費における不規則変動を除いたトレンド・循環成分に加えて一時的変化点を含む系列である、と解釈できる。

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、改善可能性があるのではないかと考えられる。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強調しておこう。

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とが分かった。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、何よりも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が¹⁵⁾、例えば GDP 推計など利用されている。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整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・季節成分を整合的に行うこと、直近の年次データが得られた後に推計する、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-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いためだろうか、Hayashi (2000) に代表される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くなったようである。そうした中で近年でも Müller-Watson (2018) などによりトレンド・循環要素の分析が行われていて、今後、さらなる展開が期待できる。実は Müller-Watson (2018) の方法は Kunitomo-Sato (2017) の特殊な場合と解釈が可能である¹⁶⁾。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、本稿での議論が応用上の問題を解決するきっかけとなれば幸いである。

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 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), 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 (2019), "Some Properties of Estimation Methods for Structural Relationships in Non-stationary Errors-in-Variables Models," *Japanese*

¹⁵⁾ 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。

¹⁶⁾ 詳しくは Kunitomo-Sato (2021) を参照。

Journal of Statistics and Data Science, 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. (2020a), "On Backward Smoothing for Noisy Non-stationary Time Series," SDS-15, MIMS, Meiji University,
- [14] Sato, S. and Kunitomo, N. (2020b), "Frequency Regression and Smoothing for Noisy Non-stationary Time Series," Unpublished Manuscript.
- [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] 佐藤整尚 (2020) 「X12SIML を使った季節調整の実際」, 未定稿.
- [18] 総務省統計局 (2021) 「消費動向指数 (CTI)」, <https://www.stat.go.jp/data/cti/index.html>.

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なのか、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。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。この方法の中で, 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 $\theta_{jk} = \frac{2\pi}{2n+1}(j - \frac{1}{2})(k - \frac{1}{2})$ ($j, k = 1, \dots, T$) に対して

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

とする。このとき初期条件 \mathbf{y}_0 を所与として

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

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。

定理 A-1(離散時間) 確率過程 \mathbf{r}_t ($j = 1, \dots, T$) 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

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

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

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

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

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

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-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

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

により与えられる ($0 < m < T$)。の性質を考察している. Kunitomo-Sato (2017, 2021) はより一般的な場合にも分散共分散行列の推定量が一致性および漸近正規性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佐藤 (2020) で準備中である. 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無償でダウンロードできる.

```
# <2020-2-4 (Originally developed by Sato and modified by Kunitomo)>
# <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")

```