

SSE-DP-2023-1

日本の公的統計と季節調整

- X-13ARIMA-SEATS と労働力調査を題材に -

国友直人(編集責任)

統計数理研究所

2023年3月

SSE-DP(ディスカッションペーパー・シリーズ)は以下のサイトから無料で入手可能です。

<https://stat-expert.ism.ac.jp/training/discussionpaper/>

このディスカッション・ペーパーは、関係者の討論に資するための未定稿の段階にある草稿である。著者の承諾なしに引用・複写することは差し控えられたい。

SSE-DP-2023-1

Official Statistics in Japan and Seasonal Adjustment

-X-13ARIMA-SEATS and Labor Force Statistics in Japan-

edited by

Naoto Kunitomo

The Institute of Statistical Mathematics

March 2023

(Summary)

In recent years, many official statistics agencies in governments have been using X-13ARIMA-SEATS program developed by U.S. Census Bureau while major official statistics agencies in Japan have been using X-12-ARIMA. In this report, we explain the main contents of X-11, X-12-ARIMA and X-13ARIMA-SEATS from the perspective of statistical time series analysis. We compare the Decomp program by Kitagawa (2020) and the X12SIML program by Sato (2023) with X-13. We also applied several seasonal adjustment programs to the real data Labor Force Survey (Rodo-ryoku-chosa) by Statistics Bureau of Japan. We discuss some historical thoughts and statistical problems in the analysis of macroeconomic data and statistical seasonal adjustment.

日本の公的統計と季節調整

X-13ARIMA-SEATS と労働力調査を題材に*

国友直人（編集責任）†

2023年3月

鍵言葉 (Key Words)：時系列の季節性, 季節調整法, ARIMA のスペクトル分解, X-11, X-12-ARIMA, X13ARIMA-SEATS, DECOMP, X12SIML 労働力調査, 産業別就業者数, 月次の季節調整系列

要約

統計数理研究所が推進している統計エキスパート養成事業の一環として2022年度のコンサルテーション演習「公的統計と季節調整」が実施された。この演習では総務省統計局で労働力調査を実際に扱っている担当者から季節調整の課題についての説明を受け、題材として「労働力調査・産業別就業者数」と米国センサス局が公開している「X-13ARIMA-SEATS」などの検討が行われた。

本報告ではまず X-11, X-12, X-13ARIMA-SEATS, DECOMP, X12SIML、など2023年2月の時点で利用可能な統計的な季節調整法の内容を説明すると共に、季節調整法の X-13ARIMA-SEATS の理解や運用に必要となるデータのスペクトル分解などの統計的時系列解析の基本についても若干ではあるが解説した。また実例として労働力調査における月次の産業別雇用者数についての分析結果も報告した。さらに、実務で利用されることがある前年同月比系列巡る間を解釈上の問題および月次の季節調整系列の利用、不規則変動部分を除いた指標の有用性などについても議論した。

*統計エキスパート養成事業におけるコンサルテーション演習「公的統計と季節調整」の報告書。演習の参加メンバーは国友直人, 湯浅良太, 趙宇, 総務省統計局の担当者（内田翔太, 谷道正太郎）の他、高岡慎（琉球大学）、佐藤整尚（東京大学）が協力者として参加した。原案に対する千野雅人教授（統計数理研究所）のコメントに感謝する。なおこの報告書の内容は総務省統計局の見解を反映するものではない。

†統計数理研究所

目次

はじめに

第I部: 季節調整法と公的統計 (国友直人)

第II部: X-13ARIMA-SEATS について

(i) Kolmogorov-Wiener フィルター (国友直人)

(ii) X-13ARIMA-SEATS (高岡慎)

第III部: 労働力調査・産業別系列の季節調整について (国友直人・
佐藤整尚)

おわりに

はじめに

日本をはじめ公的統計では多くのマクロ経済時系列の季節調整値を定期的に公表している。これは多くの経済データの原系列では季節性が顕著に観察されるために、経済の現況を正しく理解、政策立案を科学的に行う為には原系列が含む季節性の除去が必要不可欠なことが理由である。日本を含め主要国の公的統計における季節調整法としては米国センサス局が開発してきた X-11, X-12-ARIMA を利用されていたが、2022 年時点では X-12-ARIMA のメンテナンスは中止、センサス局では X-13ARIMA-SEATS の利用を推奨している。ところが X-12-ARIMA と比べてもさらに X-13ARIMA-SEATS には統計的時系列分析の方法が多用されていることもあり、その内容の検討は特に日本の公的統計ではあまり進んでいない。

こうした現状のなか、統計数理研究所が推進している統計エキスパート養成プロジェクトでは 2022 年度のコンサルテーション演習「公的統計と季節調整」が実施された。この演習では総務省統計局で労働力調査を実際に作成している担当者から季節調整の現状と課題についての説明を受けた。その後、コンサルテーション演習ではボランティアとして季節調整問題に詳しい佐藤整尚(東京大学)と高岡慎(琉球大学)のご参加もいただき活発に自由な議論を行った。なお国友、佐藤、高岡は統計エキスパート養成プロジェクトで 2022 年度に実施した講義「統計的時系列分析」の展開編で季節調整法についての特別講義を行ったが、この報告書ではその講義内容の一部を利用している。内容的には若干の重なりがあるが、講義内容については各担当者の責任であり、あえて相互に調整はしていないことをお断りしておく。

コンサルテーション演習での議論を受けて、この報告書ではまず X-11, X-12, X-13ARIMA-SEATS, DECOMP, X12SIML、など 2023 年 2 月の段階で利用可能な主な統計的季節調整法の内容を説明すると共に、季節調整法の X-13ARIMA-SEATS の理解や運用に必要なデータのスペクトル分解などの統計的時系列解析の方法についても若干の解説を述べることにした。これらの内容はおおよそ講義

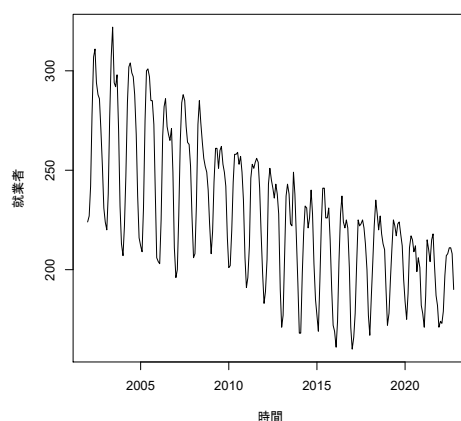
「統計的時系列分析」展開編の内容の一部をまとめたものであるが、現時点では季節調整をめぐる理論と実務に関する近年の展開について、統計学的な観点から詳しく議論している文献が見当たらないのでまとめて報告する意味があると判断したためである。この報告ではさらに実務で利用されることがある前年同月比系列巡る問題を解釈上の問題と月次の季節調整系列の利用についての意見も述べる。また、この報告書では、実例としてコンサルテーション演習において議論した労働力調査における月次の産業別雇用者数の分析結果も合わせて報告する。

なお、2023年初頭の時点においてはX-13ARIMA-SEATSを季節調整法として利用していると説明している日本の公的統計は皆無と思われる。また本報告書で実例として取り上げた産業別雇用者数の季節調整値も公表されていないと思われる。したがって今後の日本の公的統計における季節調整をめぐる諸問題を検討する際にこの報告書の内容が議論の参考になれば幸いである。

第I部：季節調整法と公的統計（国友）¹

1 経済時系列の季節性

時間的経過とともに観察される経済時系列の中で観測周期が1年以内の時系列、例えば四半期系列や月時系列などでは季節的な変動はしばしば観察される。身近な一例として図1に日本の総務省統計局が毎月公表している労働力調査における農林業部門の就業者数の時間的変化を示す時系列を挙げておく。この系列では長期的趨勢（トレンド）、景気変動、不規則変動などとともにほぼ1年間を周期とする季節的変動が無視できないほどに大きいことを示している。



< 図1：農業・林業部門の就業者(労働力調査)>

2002年1月～2022年10月, 労働力調査, 長期時系列表1, c-3 主な産業別就業者, 全国, 月別, 総務省統計局

農林業部門の就業者数に限らず、労働、消費、生産、売上高、投資といった重要な経済時系列の多くにおいて、顕著な季節的変動が観察

¹ 「季節調整法」(国友, 2012年, 朝倉書店, 「経済時系列分析ハンドブック」収録)の元原稿の改訂版。2022年5月～7月に統計エキスパート養成における講義「統計的時系列分析」展開編で行った授業の基礎資料を修正した原稿である。

され、経済時系列における重要な変動要因であることについては古くから注目されている。経済学の文献の中ではこの問題は少なくとも19世紀の英国の経済学者の議論までは容易に遡ることができる。例えばS. Jevons (スタンレー・ジェボンズ)をはじめ、当時の著名であった英国の経済学者は経済における景気変動を分析することを試みたが、経済変動の中で好況・不況の循環の波よりもかなり短い1年周期の変動が大きな役割を演じていることに気がつき、主として記述統計的な方法で分析したことが知られている(例えばNerlove et. al.(1979)を参照)。

近代的な数理統計学の成立とともに時系列データの分析方法が研究され、統計的時系列分析 (statistical time series analysis) と呼ばれる分野も発展した。経済時系列における変動の中では比較的周期が明確に確認できる、という意味では統計的分析の対象として扱いやすいと思われたようである。1950年代頃より移動平均法や多項式の当てはめなどの時系列分析における記述統計的手法、さらにはピリオドグラム (periodogram) 解析やスペクトル分析 (spectral analysis) と呼ばれている統計的時系列解析の方法が開発され、経済時系列の季節性分析に応用されはじめた。こうした初期の経済時系列分析における季節変動を巡る問題は単に経済現象を巡る学問的関心ということに留まらずに、重要な経済時系列については中央政府の統計部門における季節調整値の作成、公表という実務的な意味で重要な意味を持っていた。2023年の時点においてもこうした公的統計の実務的要請という事情は変わっていない。

ここでは議論をわかりやすくする為に、時刻 t に観察可能な1次元の経済時系列 Y_t をトレンド (趨勢項)・循環変動項 TC_t , 季節変動項 S_t , 曜日・休日効果 TD_t , 不規則変動項 I_t により分解する統計モデルを考えよう。原系列がこれらの要素の和

$$(1) \quad Y_t = TC_t + S_t + TD_t + I_t$$

とするのが時系列の加法的成分モデルである。近年での時系列計量経済分析では TC_t 項をさらにトレンド項 T_t と循環項 C_t に分解し、 $TC_t = T_t + C_t$ とすることがより一般的である。ここで公的統計に

における季節調整ではトレンド（趨勢）項として時間の多項式関数など大域的トレンドではなく、循環的変動を含む経済時系列をより柔軟に扱うことができる局所的トレンド (local trend) 項を多くの場合に採用していると解釈できる。なお、実際の季節調整ではより一般的に乗法的成分モデル $Y_t = TC_t \times S_t \times TD_t \times I_t$ を仮定して時系列を処理することが多いが、乗法的成分モデルは対数変換により加法的成分モデルの議論に帰着することができる。加法的成分モデルを用いて説明する方が特に移動平均にもとづく季節調整の議論を理解しやすい。統計的には、より一般にした加法的成分モデルと乗法的成分モデルを組み合わせる可能性などもありうるが、こうした混合型モデルはこれまであまり実用化はされていない。ただし、ここでは公的統計でよく利用されている X-11 法, X-12, X-13ARIMA-SEATS 法 (以下では X-13 とする) などでは、乗法的成分モデルの仮定のもとに移動平均法を適用するなど、経済時系列の統計理論と季節調整の実際とは整合的とは必ずしも限らないことにも注意しておく。

経済時系列における季節変動はエコノミストや政府統計の担当者にとっては経済の長期的趨勢や景気変動を理解する上では不必要な規則的変動と見なされることが多い。したがって、実際に観察される経済時系列 $\{Y_t\}$ (原系列と呼ぶ) に何らかの変換をほどこすことにより季節性を取り除き、別の時系列 $\{Y_t^*\}$ (季節調整済系列と呼ぶ) を作成する統計的方法である季節調整法が古くから研究されているのである。仮にここで議論している季節成分 S_t が任意の t について関数関係 $S_t = S_{t-s}$ (s は季節周期) で定義できたり、確定的な関係としてあらかじめ分かっている場合を大域的季節性 (global seasonality) と呼ぶと、その統計的処理は比較的単純である。例えば、今日でもエコノミストや経済学者の中には実証的研究で時系列を利用する方法として、季節ダミー変数 $D_{ti} = 1$ (t が第 i 季節), $D_{ti} = 0$ (それ以外) を定義し、観測系列をダミー変数に回帰することにより季節成分を推定し、取り除こうとしていることをよく見かける。数理的に

は類似の方法であるが、古典的な統計的時系列分析²では、有限個の三角関数から確定的な周期を持つ周期関数を構成して最小二乗法を利用して周期性を分析する方法などが利用されることもある。

ところで、統計的時系列解析と呼ばれている研究分野では、1960年頃から時系列データにフーリエ変換をほどこし周波数領域（あるいは周期成分）の分析を行うことが行われ始めた。数理的には定常確率過程にしたがう離散時間間隔の確率過程における自己共分散関数のフーリエ変換をスペクトル密度関数、離散時間の時系列データのフーリエ変換はピリオドグラム、とそれぞれ呼ばれているが、大域的季節性が適切な場合には季節周波数（および調和周波数）のみにおいて鋭い山（ピーク）を持つことになる。しかしながら、現実に観察されている経済時系列においては、近年の年始でのコンビニの営業や夏期休業のあり方の変化など、経済をとりまく環境が時間とともに変化するとともに季節性も時間的に変化している可能性、すなわち局所的季節性 (local seasonality) の扱いを考慮する必要がある。こうした季節性が時系列データに存在する場合には、スペクトル密度関数は季節周波数（および調和周波数）の周辺でよりなだらかな山（ピーク）を持つと考えられる（例えばハーベイ (1986) 3章の議論を参照）。また周波数領域の統計的分析は観測される時系列が定常過程の実現値と見なすことが妥当であれば正当化の数理的根拠ははっきりしているが、実際の時系列データが定常的でない場合には様々な統計的問題が生じる。

統計的時系列解析の理論はさておき、既に言及したようにエコノミストや公的統計の担当者は実際に観測される経済時系列（原系列）から季節要因に起因しない時系列の真の状態について理解する必要がある。このことを簡単に実行する方法としては原系列の前年同期比、前年同月比を求める方法が知られているが、この方法は今でも利用されることがある。前年との比率をとり季節要因を除去する方法の一つの正当化としては、乗法的成分モデル $Y_t = TC_t \times S_t \times TD_t \times I_t$ において季節成分が一定 $S_t = S_{t-s}$ (s は季節周期)、曜日効果が無視で

²例えば Anserson (1971) 第 2 章に詳しい説明がある。

きる場合には、 $Y_t/Y_{t-s} = [TC_t/TC_{t-s}] \times [I_t/I_{t-s}]$ となるので多くの場合には原系列の変動についての重要な情報が得られる、というものである。経済の景気分析でよく知られている問題点としては例えば TC 成分をトレンド成分 T_t と循環成分 C_t に乗法的に $TC_t = T_t \times C_t$ にさらに分割できると考えると、循環部分の山(ピーク)と谷(ボトム)の位置が時間差が比をとることによりズレてしまうこと、などである³。

さらに経済の季節性もある程度の時間が経てば変化する⁴ などの議論もあり、何らかの統計的方法により局所的にはほぼ一定の季節性を推定し、原系列より推定された季節性を除去して、季節調整系列を作成することが必要になると考えられてきた。こうした原系列に対してある種の操作を加えることは、周波数領域(すなわち波動成分)の時系列分析より、何らかの方法(しばしばフィルターと呼ばれる)で季節周波数に対応するピークをデータより除去する変換を施していると解釈できる。ここでは観測される時系列より季節成分のみを除去することを目的とした経済時系列データの変換方法を季節調整法(seasonal adjustment method)と呼ぶことにする。今日でも日本を含めた先進諸国の中央政府においてよく用いられている季節調整法としては、米国センサス局で開発された X-11 法, X-12-ARIMA 法, X-13ARIMA-SEATS 法などが知られている。他方、経済時系列分析に関心のある統計学の研究者の間では統計数理研究所が開発した DECOMP 法など幾つかの統計的方法が知られている。

2 季節調整法 X-11 法

2.1 歴史的経緯

米国センサス局で開発している一連の季節調整法はセンサス局法と呼ばれているが、1950年代に J. Shiskin(シスキン)を中心として当時ようやく利用可能となりつつあった(メインフレーム)電子計算機を利用してセンサス局 I 法を開発したことが現代的な季節調

³こうした古くから知られている問題については例えば溝口・刈屋(1983)第4章に説明がある。

⁴例えば日本の年末年始、夏休みのあり方などは時間とともに変化するもので、そうした変化に伴い生産や消費活動も変化する、という説明がある。

整法の始まりである。このセンサス局法は当時の記述的時系列分析でよく用いられていた移動平均法 (moving average) を基礎としつつも、実際の時系列データの処理に際して生じる問題を実務的に処理するために、途中の計算アルゴリズムを次々に複雑化して開発されたプログラムである。1950年代から1960年代にかけて国際的にも官庁統計家の間では季節調整法を開発しようとする機運が高まり、ヨーロッパ主要国の政府統計の担当者や日本の政府統計の担当者によっても幾つかの季節調整プログラムが開発されている。日本では当時の通商産業省がMITI法と呼ばれる季節調整法、当時の経済企画庁はEPA法⁵という季節調整法を開発し、それぞれ実際に運用していた。これら二つの季節調整プログラムもまたセンサス局I法と同様に、移動平均法を中心のアイデアとしつつも、独自に様々な工夫を施した方法であったが、EPA法は1978年に運用が中止、MITI法は1998年に運用が中止されている。ヨーロッパでもドイツの Bundes・バンク法や英国中央銀行の季節調整法などが開発されていたが、最近までヨーロッパ統計局を中心に研究・開発が行なわれていたが、ここではヨーロッパの政府統計における研究例として、スペイン中央銀行の A.Maravel(マラベル)により統計的信号抽出法 (signal extraction method) を理論的背景とした TRAM-SEATS と呼ばれる季節調整法が開発されていたことに言及しておく⁶。

再び話題を米国センサス局における1960年代の開発動向に戻すと、センサスI法を開発した後も改良を重ね、改訂版が開発されたが一連の方法は実験用 (Experimental methods) という意味でのセンサス局II法のXシリーズという名前になり、改良結果として1965年 (解説マニュアル Shiskin et. al. (1967) は内部資料として発行) にセンサス局X-11法と呼ばれる季節調整法が開発された。このX-11は当時、ようやく一般に利用可能となってきた (メイン・フレーム) 計算機の利用を想定したものでフォートラン (FORTRAN) 言語でプログラムが書かれたが、そのプログラムを関係者が配布されて利用する形式をとった。こうした開発の動きの中で、1970年頃からわ

⁵例えば経済企画庁 (1971) を参照。

⁶TRAMO-SEATS 法についてはのちの章で説明する予定である。

が国では日本銀行をはじめ季節性を持つ経済データを公表している経済関係の部局が、このセンサス局法 X-11 法を採用して季節調整系列を公表することが多くなった。なお、センサス局 X-11 法はその内部に(担当者が指定することができる)かなり複雑なオプションを含んでいたが、多くの日本の政府当局では季節調整法として X-11 法を採用したが、各責任者がどの様に X-11 法のオプションを利用していたかについては、一般には公開されてはいなかった⁷。

2.2 移動平均法と X-11 法

センサス局 X-11 法における統計的手続きはかなり複雑であるが、基本的には移動平均 (moving average) に基づき時系列データを平滑化 (smoothing) することにより季節性を除去することが意図されている。一般に原系列に移動平均をかけるとは、原時系列 $\{Y_t\}$ から調整系列 $\{Y_t^*\}$ を

$$(2) \quad Y_t^* = \sum_{i=-m}^n w(i)Y_{t+i}$$

により変換する方法である。ここでウェイト係数(あるいはフィルタ関数) $\{w(i), i = -m, \dots, n\}$ (m と n は正整数) は通常は観測される時刻 t には依存させないが、基準化規則 $\sum_{i=-m}^n w(i) = 1$ を満たすとしても、様々な移動平均フィルターが考えられる。

ここで月次データを取って説明しよう。仮に 2000 年 1 月より 2000 年 12 月のデータの 12ヶ月平均値 $(1/12) \sum_{i=1}^{12} Y_{t+i}$ ($t=1999$ 年 12 月にとる) は 12 が偶数なので、2000 年 6.5 月が対応する月と考えられる。また 2000 年 2 月より 2001 年 1 月のデータの 12ヶ月平均値 $(1/12) \sum_{i=1}^{12} Y_{t+1+i}$ は 2000 年 7.5 月に対応すると考えられよう。さらに二つの平均値を平均すると

$$(3) \quad \frac{1}{2} \left[\frac{1}{12} \sum_{i=1}^{12} Y_{t+i} + \frac{1}{12} \sum_{i=1}^{12} Y_{t+1+i} \right] = \frac{1}{24} Y_{(t+7)-6} + \frac{1}{12} \sum_{i=-5}^5 Y_{(t+7)+i} + \frac{1}{24} Y_{(t+7)+6}$$

⁷当時のオプションについては黒川 (1979), 溝口・刈屋 (1983) などを参照されたい

と表現できる。したがって、 $t = 1999$ 年12月とすると $t + 7 = 2000$ 年7月に対応する移動平均値と解釈できる。この13項移動平均であるフィルタ関数は二つの12ヶ月平均の平均であるから中心化12項 (centered 12 terms) 移動平均と呼ばれている。ここで、ラグ (遅れ) 作用素の記号 B を用いると、例えば時系列 Y_t に対して $BY_t = Y_{t-1}$, $B^2Y_t = BY_{t-1} = Y_{t-2}$, $B^{-1}Y_t = Y_{t+1}$ などとなる。ラグ作用素を用いると中心化12項移動平均は $M_{2,12}(B) = (1/2)[(1/12)\sum_{i=-6}^5 B^i + (1/12)\sum_{i=-5}^6 B^i]$ で与えられる。さらに、 3×3 移動平均と 3×5 移動平均はそれぞれ $M_{3 \times 3}(B) = (1/9)(B^{-12} + 1 + B^{12})(B^{-12} + 1 + B^{12})$, $M_{3 \times 5}(B) = (1/15)(B^{-12} + 1 + B^{12})(B^{-24} + B^{-12} + 1 + B^{12} + B^{24})$ により定める。

次に X-11 における移動平均操作による時系列データの処理の概略を要約しておこう⁸。ある時刻 t における原系列 Y_t を構成するトレンド・循環変動項 TC_t , 季節変動項 S_t , 曜日効果 TD_t , 不規則変動項 I_t の推定アルゴリズムではまず次の4ステップを考える。

- (1) 中心化12項移動平均でトレンド・循環成分の推定値 \hat{TC}_t を抽出する。
- (2) 季節・不規則成分 $S_t \times I_t$ の推定値を Y_t/\hat{TC}_t で求める。
- (3) 季節・不規則成分の推定値より 3×3 移動平均により季節成分の推定値 \hat{S}_t を求める。
- (4) 季節調整値を Y_t/\hat{S}_t により求める。

さらに、同様の4ステップを中心化12項移動平均を Henderson の13項移動平均⁹、 3×3 移動平均を 3×5 移動平均に変更して実行することで合計8ステップからなる基本処理操作が得られる。次に X-11 プログラムではこうした8ステップからなるアルゴリズムをはずれ値の補正と最終的季節調整値の作成のためにそれぞれ適用し、細部は多少の差異はあるものの、合計3回適用される。こうした複雑な移動平均と幾つかの段階でのはずれ値の処理が X-11 プログラムの

⁸ここで説明した X-11 の計算アルゴリズムの詳細については経済企画庁 (1971)、黒川 (1979) が説明している。

⁹より正確には計算される SI 比に応じて9項、13項、23項のいずれかの Henderson の移動平均が選択されるが、X-11 法の原プログラムとを X-12-ARIMA 法の改良 X-11 プログラムで多少の変更がある。(Findley et. al. (1998) を参照。)

基本的な構成となっている。したがって、X-11 アルゴリズムとは時系列データに移動平均操作を繰り返し施すことにより、原系列を平滑化している操作と解釈できる。

例えば計算アルゴリズムで登場している Henderson の $2m + 1$ 項移動平均とは

$$(4) \quad w(i) = c_m [(m+1)^2 - i^2][(m+2)^2 - i^2][(m+3)^2 - i^2] \\ \times [3(m+2)^2 - 16 - 11i^2]$$

で与えられる線形フィルタ関数である。ただし c_m は m に依存する定数であり制約条件 $\sum_{i=-m}^m w(i) = 1, \sum_{i=-m}^m i^2 w(i) = 0, w(i) = 0 (m+1 \leq |i| \leq m+3)$ のもとで3次階差 $\sum_{i=-m}^m (\Delta^3 w(i))^2$ を最小化する解として得られる。(ここで階差記号 Δ は $\Delta w(i) = w(i) - w(i-1)$ を意味し、3次階差は $\Delta^3 w(i) = \Delta^2(w(i) - w(i-1)) = w(i) - 3w(i-1) + 3w(i-2) - w(i-3)$ である。) このような操作により滑らかに変動するトレンド成分と循環成分を局所的に推定し、逐次的に妥当な季節調整値を構成していると解釈できるだろう。

ここで原系列 $\{Y_t\}$ に対して移動平均を用いる時には結果として得られる調整系列では末端の部分が欠けることになる。例えば、原系列 $\{Y_t\}$ が期間 $t = 1, \dots, T$ に得られるときに $2m + 1$ 項移動平均を用いると最初と最後の m 時点では移動平均値を計算することができない。すなわち、時刻 $t = T - k (k = 0, \dots, m - 1)$ における移動平均は $X_t = \sum_{i=-m}^k w(i) Y_{t+i}$ としてしか計算することができないのである。したがって、この場合には線形フィルタ関数を用いても対称性の条件 $w(i) = w(-i)$ を満たすことができないので非対称移動平均を用いることになる。センサス局 X-11 法では期初及び期末の原データに対しては、ここで仮に Musgrave 移動平均と呼ばれる非対称フィルターを用いているが、この線形フィルタ関数は次のようにして対称フィルター $\{w(i)\}$ から導かれると考えられる。原系列 $\{Y_t\}$ に対して時間の1次関数 $P_1(t)$ のトレンドと互いに独立で正規分布にしたがう期待値ゼロ、分散一定の誤差 $\{a_t\}$ から構成される加法

的モデル $Y_t = P_1(t) + a_t$ を仮定した上で損失関数

$$(5) \quad E\left[\sum_{i=-m}^m w(i)y_{t+i} - \sum_{i=-m}^{m-d} v_d(i)y_{t+i}\right]^2$$

を最小にするようにフィルタ関数 $\{v_d(i)\}$ を決めている。この Musgrave 移動平均と呼ばれるフィルターでは非対称フィルタ関数が対称移動平均に対するある種の最適近似となることで正当化されている（詳しくは Findley et. al. (1998) を参照）。ここではこの移動平均が必ずしも Henderson のフィルタ関数の議論と整合的か否かははっきりしないことに注意しておく。統計的時系列の理論を用いると、X-11 法で利用されている移動平均フィルターを用いると原系列における季節周波数近くでのスペクトル密度関数のピークが平滑化されることが分かるが、複雑な線形移動平均の非線形的な組み合わせであるので、その妥当性については必ずしも明らかではない。また、X-11 アルゴリズムでは曜日変動項 TD_t や不規則変動項 I_t については、季節成分を一度推定した上でさらに不規則成分に対して回帰分析を利用した幾つかの付加的処理を行うことが、オプションとして実行することができる。実際の季節調整では異常値や曜日効果の処理などの処理も重要となるが、実際の政府統計の作成過程における各統計担当者が実行していた処理の詳細に不透明な要素があったように判断できる。

ここで要約しておく、季節調整法 X-11 は移動平均など 1950 年頃に時系列の処理に関して知られていた古典的な統計的分析法を（統計学の言葉を使えば「ノンパラメトリック的」に）利用して開発された。X-11 が実行する季節調整の妥当性については様々な議論がありうるが、政府統計という実用性を重視した方法である、と解釈できるだろう。

3 X-12-ARIMA 法

3.1 X-12-ARIMA の開発

統計的時系列分析の分野では 1970 年代にはボックス・ジェンキンス (Box and Jenkins (1976)) が提唱した線形定常時系列モデルとしての ARMA (autoregressive-moving average, 自己回帰移動平均) モデル, その非定常版である ARIMA (autoregressive-integrated-moving average, 自己回帰和文移動平均) モデル、にもとづく予測の方法が実用化されたことが重要である。数理的にはフーリエ解析という一見すると高度に見えるが経済時系列への応用としては実用性が疑問視されていた周波数領域分析に対し、時間領域分析において扱いやすい線形時系列モデルが実用化された。こうした時系列分析を巡る新しい機運の中で、カナダ・センサス局のダグム (E.B. Dagum) を中心にセンサス局 X-11-ARIMA と呼ばれる季節調整法プログラムが 1975 年頃に開発された。この方法が X-11 法と異なる主要な機能は ARIMA モデルを用いた予測系列を利用する事で季節調整の際に生じる直近の末端処理問題を改善しようとしたことにあった。この X-11-ARIMA 法はカナダ統計局ではかなり実用的にも用いられていたが、米国や日本を始めとする他の国々の公的統計ではあまり用いられなかった。

1980 年代後半になると計算機を取り巻く環境が劇的に変化するとともに、米国センサス局においても再び季節調整法の検討が始まり、時系列研究グループにより X-12-ARIMA プログラムが開発された。このプログラムは X-11-ARIMA 法をさらに改良し、次に説明する RegARIMA モデルを原系列に適用し、時系列の将来の予測値を利用して平滑化により季節調整値を得ようとする方法である。すなわち、X-12-ARIMA 法では X-11-ARIMA 法に回帰分析を利用した統計的機能を付け加えることで、さらに利用者が様々なオプションを比較的簡単な操作で実行する事ができることが特長である。さらに、X-12-ARIMA は 1990 年代になって実現し始めたインターネットを利用した不特定多数のユーザーに対するプログラム配布という形で 1996 年に X-12-ARIMA(β -Version) がセンサス局 HP 上で公開

(<http://www.census.go.jp>)、配布され始めた。このとき配布されたものは β -Version という名前が付いているように実験用プログラムであり、その後しばらく頻繁に改良が重ねられ、2000年頃から X-12-ARIMA という名前になった。

なお、X-12-ARIMA において利用する可能な様々なオプションを応用するには様々な統計的問題が付随することを付け加えておく（国友(2001)や国友編(2004, 2006)を参照）。日本の公的統計における X-12-ARIMA の導入を巡っては日本銀行の木村(1997)、統数研の川崎・佐藤(1997)などにより季節調整の妥当性の規準をめぐる論争が起こったが、実務的には1996年頃の総務庁・統計審議会により設置された季節調整検討小委員会での検討を経て、「季節調整法の適用について(指針、平成9年6月20日)」が出され、各統計作成の担当者が X-11, X-12-ARIMA, DECOMP などを注意深く利用し、さらに適用するオプションを開示することが義務づけられた。この仕組みにより公表系列の季節調整値を巡る不透明さを解消する、という意味では一歩前進と言えよう。

3.2 RegARIMA モデルの利用を巡る問題

X-12-ARIMA プログラムによる季節調整値の作成法はおおよそ次のような3ステップに要約することができる。

(1) 観察される時系列の原データ $\{Y_t, 1 \leq t \leq T\}$ からまず RegARIMA モデルと呼ばれている統計的モデリングを用いて将来の予測値 (forecasts), 過去の逆予測値 (backcasts) を作り出す。同じモデルの中の回帰 (regression) モデルを用いて季節調整に先だって様々な事前調整を行い、事前調整系列 $\{Y_t^*, -H + 1 \leq t \leq T + H\}$ を作り出す。ここで $H (\geq 1)$ は事前に設定する予測期間である。この事前調整としては異常値・変化点の検出, 閏年調整, 曜日効果調整, 休日調整などの項目を挙げることができる。RegARIMA モデルを用いる統計的モデリングではモデル診断と呼ばれている一連の操作により様々なモデルの選択を行う。

(2) 次に予測値と逆予測値を含む事前調整された系列 $\{Y_t^*\}$ に対し

改良された X-11 プログラムにより季節調整を行い、季節調整系列 $\{X_t, 1 \leq t \leq T\}$ が作り出される。この部分は従来から利用可能な X-11 法プログラムを手直した改良 X-11 プログラムにより実行されるが、基本的には X-11 法の計算アルゴリズムと同一の手続きで行われる。

(3) 最後に診断と呼ばれる部分により季節調整を行った結果を評価する。この為に季節調整系列の推定から求められた不規則変動の推定値としての残差系列から、スペクトル密度関数や調整系列の安定性に関する新しい指標などいくつかの統計量が計算される。

以上の説明から明らかなように RegARIMA モデルと呼ばれる統計モデルの利用が X-12-ARIMA の特長と云える。RegARIMA モデルでは 1 次元の時系列データを生成する確率変数の組 $\{Y_t\}$ とすると、 $\{Y_t\}$ が r 個の説明変数 $\{X_{it}\}$ を伴い線形時系列表現

$$\phi_p(\mathcal{B})\Phi_P(\mathcal{B}^s)(1-\mathcal{B})^d(1-\mathcal{B}^s)^D(Y_t - \sum_{i=1}^r \beta_i X_{it}) = \theta_q(\mathcal{B})\Theta_Q(\mathcal{B}^s)a_t$$

を持つと仮定される。ここで \mathcal{B} は時系列 Y_t に対してラグ記号、季節周期を表す $s = (4 \text{ あるいは } 12)$ 、係数の次数 p, d, q, P, D, Q はあらかじめ決められた非負整数である。また $\phi_p(z) = 1 - \phi_1 z \cdots - \phi_p z^p$, $\Phi_P(z) = 1 - \Phi_1 z \cdots - \Phi_P z^P$, $\theta_q(z) = 1 - \theta_1 z \cdots - \theta_q z^q$, $\Theta_Q(z) = 1 - \Theta_1 z \cdots - \Theta_Q z^Q$ は z についての多項式を表し、 $\Phi(\mathcal{B}^s), \Theta(\mathcal{B}^s)$ 例えば \mathcal{B}^s ($\mathcal{B}^s y_t = y_{t-s}$) の多項式となる。係数 β_i ($i = 1, \dots, r$), ϕ_i ($i = 1, \dots, p$), Φ_i ($i = 1, \dots, P$), θ_i ($i = 1, \dots, q$), Θ_i ($i = 1, \dots, Q$) は利用可能な時系列データから統計的に推定される未知母数である。誤差項 $\{a_t\}$ は期待値ゼロ、分散 σ^2 (σ は未知母数とする) であり、各 t について互いに独立な確率変数列と仮定する。

この RegARIMA モデルは統計モデルとしては線形回帰 (linear regression) モデルと季節 ARIMA(時系列) モデルの混合型統計モデルの一つとして理解できる。ここで ARIMA とは ARIMA (autoregressive integrated moving average 自己回帰和分移動平均) モデルの略であるが、同時に季節 (seasonal) ARIMA モデルをも含んでいると見なすことができる。例えば $D = 0, \Phi_P(z) = \Theta_Q(z) = 1$ とおけば ARIMA

モデルとなるので、季節 ARIMA モデルは季節性を含む ARIMA モデルの特殊な場合であり、より少ない数の母数で季節性を含む経済時系列の変動を表現する意味での節約型の時系列モデルと解釈できる。この季節 ARIMA モデルはしばしば $(p, d, q) \times (P, D, Q)_s$ と表される。季節 ARIMA モデルでは AR 部分と MA 部分の固有方程式をそれぞれ $|\lambda^p - \sum_{i=1}^p \lambda^{p-i} \phi_i| = 0$, $|\lambda^q - \sum_{i=1}^q \lambda^{q-i} \theta_i| = 0$, および $|\lambda^P - \sum_{i=1}^P \lambda^{P-i} \Phi_i| = 0$, $|\lambda^Q - \sum_{i=1}^Q \lambda^{Q-i} \Theta_i| = 0$ とすると、固有方程式を満足する固有値を λ_i ($i = 1, \dots, \max\{p, q, P, Q\}$) について定常条件と反転条件 $|\lambda_i| < 1$, 及び「固有方程式に共通根は無い」ことを仮定する必要がある。

ここで注意すべき点としては、標準的な統計的時系列分析では回帰モデルには定常時系列構造を導入する¹⁰ことがより一般的な方法であるが、RegARIMA モデルでは非定常時系列モデルをそのまま導入していることである。このことから統計学上では非正則問題が発生する。この問題については例えば国友・高岡 (2005) が議論している¹¹。この RegARIMA モデルは階差系列に対して最尤推定法 (maximum likelihood method) により推定されるが、推定された母数を使えば統計的時系列モデルにもとづく予測を容易に行うことができる。したがって予測値を利用することで直近のデータに対しても対称移動平均を直接に適用することができることになる。ただし時系列データへのフィルターは推定された季節 ARIMA モデルに依存するので、結局は原系列に対して非対称移動平均を行っているが、「推定された季節 ARIMA モデルの予測値が適切であれば」Musgrave 移動平均よりもよい結果が得られる可能性がある。

さらに、X-12-ARIMA プログラムの特長としては、RegARIMA モデルの回帰部分を利用して様々な処理が容易に実行できることに言及する必要がある。X-12-ARIMA 法では、異常値 (outlier) や変化点の処理、曜日効果 (trading days effects) や休日数 (holidays effects)・

¹⁰統計的時系列解析の標準的教科書としては、例えば Anderson (1971), Fuller (1996) を挙げておく。

¹¹通常の統計的時系列解析では多くの場合には回帰モデルに定常確率過程が想定されている。

閏年効果の処理を予測値を作成する段階で RegARIMA モデルを利用して行うことが特徴である。むろん、X-11 パートでも類似の処理が可能であるが、RegARIMA における回帰変数の定義など詳しいオプションは国友編 (2004) に説明がある。

4 X-12-ARIMA・DECOMP・最適な季節調整？

季節調整法としてはセンサス局法以外にもこれまでに様々な統計的方法が開発されている。ここでは特に X-12-ARIMA でのアルゴリズムを理論的に解釈する上で有用であるので DECOMP 法について言及しておこう。この時系列解析プログラム開発の経緯としては、統計数理研究所の赤池弘次・石黒真木夫により 1980 年代に季節調整法 BAYSEA が開発されたことが挙げられる。BAYSEA は伝統的な季節調整法の基礎をなす移動平均法とは異なり、時系列において季節性に関する滑らかさの事前情報¹² を直接的に活用して時系列成分分解を行うことにより季節性を推定する方法である。季節調整プログラム Baysea はその後、北川源四郎（北川 (2020)）により改良され、状態空間モデルに基づく季節調整プログラム DECOMP が開発された。2023 年の時点ではこのプログラムは R ベースに翻訳され、RS-DECOMP (<https://jasp.ism.ac.jp/RS-Decomp/>) がネット上から利用可能になっている¹³。

ここでは季節調整問題のために DECOMP において循環成分と曜日効果をゼロとして¹⁴、加法的な時系列成分の分解モデル $Y_t = TC_t + S_t + I_t$ を考察しよう。このとき、トレンド成分の差分 $\Delta TC_t = TC_t - TC_{t-1}$ の変動が小さく、 d 階差分 $\Delta^d TC_t = \Delta^{d-1} [TC_t - TC_{t-1}] = u_{1t}$ が $N(0, \tau_1^2)$ の互いに独立な確率変数、季節成分は制約 $\sum_{j=0}^{s-1} S_{t-j} = u_{2t}$ (季節ラグ s は 4 か 12) が $N(0, \tau_2^2)$ の互いに独立な確率変数と見なすことになる。したがって、トレンド成分の階差成分が徐々に変動

¹²時系列解析プログラム BAYSEA と DECOMP については、開発者による統計学的アイデアの説明としては、赤池 (1989) や北川 (2020) があるが、本稿での説明は少し異なることに注意しておく。

¹³R 上ではパッケージ TSSS 上でも動かすことができる。TSSS について詳しくは北川 (2020) を参照されたい。

¹⁴四半期 GDP の場合には曜日効果はあまり検出されないが、循環成分は重要である。他方、X-11、X-12-ARIMA ではトレンド・循環成分は区別されない。

し $\Delta^d TC_t \cong 0$ 、同時に季節成分には $\sum_{j=0}^{s-1} S_{t-j} \cong 0$ となる制約条件を課すことを意味している。次に不規則変動成分を u_{4t} を $N(0, \tau_4^2)$ の互いに独立な確率変数と見なせば、観測時系列 Y_t は観測方程式 (measurement equation) は

$$(6) \quad Y_t = \mathbf{H}\mathbf{x}_t + u_{4t},$$

($1 \times (d+s-1)$) 状態ベクトルは $\mathbf{x}'_t = (TC_t, \dots, TC_{t-(d-1)}, S_t, \dots, S_{t-(s-2)})$, $\mathbf{H} = (1, 0, \dots, 0, 1, 0, \dots, 0)$ と表現される。ここで、状態ベクトルについては適当に誤差ベクトル $\mathbf{v}'_t = (u_{1t}, u_{2t})$, (要素にゼロ成分を含む) 既知行列 \mathbf{F} と \mathbf{G} を適切に選べば

$$(7) \quad \mathbf{x}_t = \mathbf{F}\mathbf{x}_{t-1} + \mathbf{G}\mathbf{v}_t$$

という状態方程式 (state equation) を得ることができる。例えばここで2次のトレンド階差を四半期データに適用すれば、 $d = 2, s = 4$ 及び $\mathbf{H} = (1, 0, 1, 0, 0)$, $\mathbf{x}'_t = (TC_t, TC_{t-1}, S_t, S_{t-1}, S_{t-2})$ を用いて

$$\begin{pmatrix} TC_t \\ TC_{t-1} \\ S_t \\ S_{t-1} \\ S_{t-2} \end{pmatrix} = \begin{pmatrix} 2 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & -1 & -1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} TC_{t-1} \\ TC_{t-2} \\ S_{t-1} \\ S_{t-2} \\ S_{t-3} \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix}$$

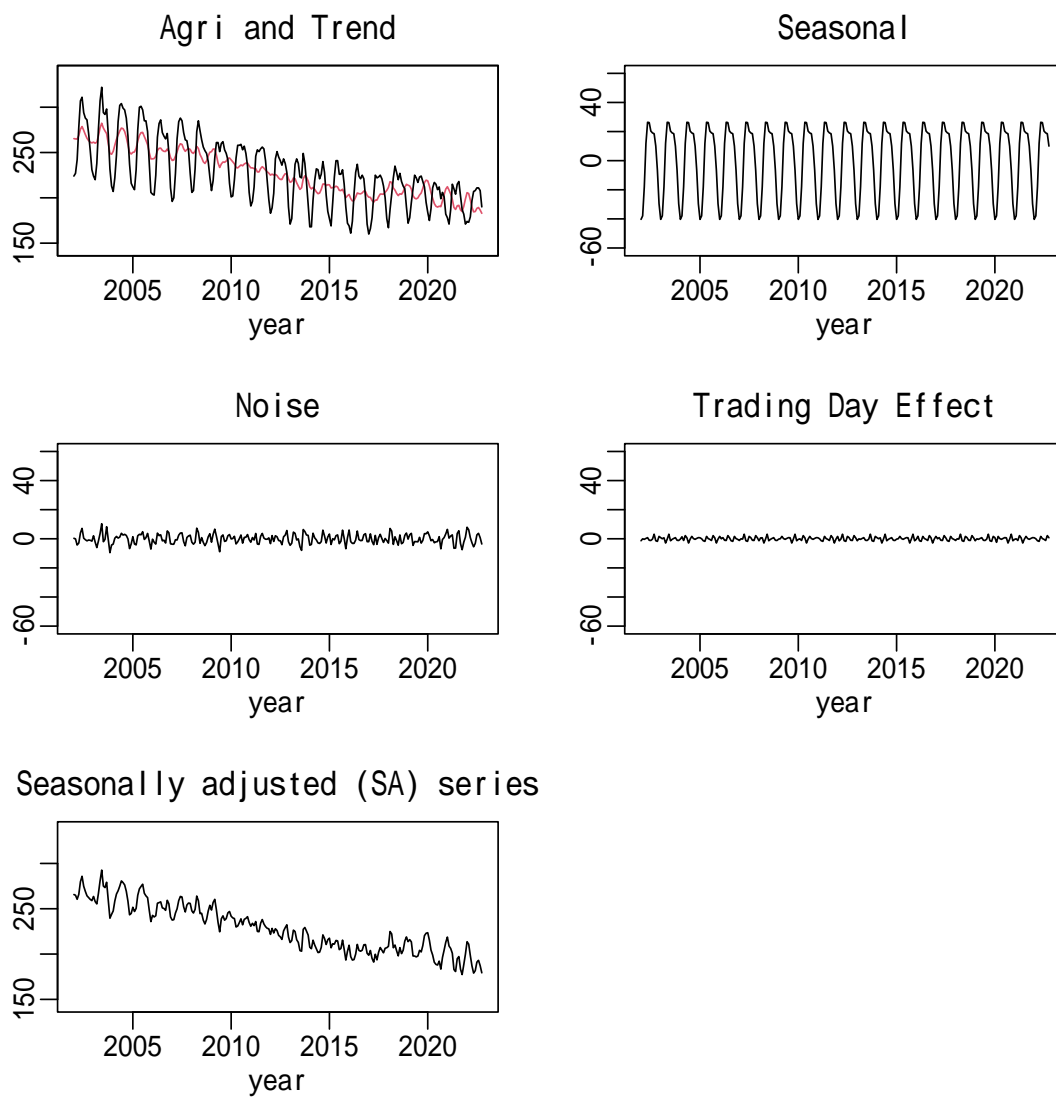
という状態空間表現が得られる。

さらに、不規則変動 (u_{4t}) の分散及び2つの誤差分散比を母数 (ハイパー・パラメーターと呼ばれる) としてガウス過程を仮定した上で、状態方程式で表現された制約条件付き最尤推定 (maximum likelihood) 法を実行して推定を行うのが DECOMP の基本的計算アルゴリズムである。季節成分はフィルタリングによる状態変数ベクトルの最適な推定値として得られるので、観測値を季節成分の推定値で割れば、季節調整系列を求めることができる。ここで二つの線形方程式からなる状態空間 (state space) モデルでは、状態方程式と観測方程式が線形であるので、北川 (2020) が説明しているようにカルマン・フィルターを利用することが可能である。

DECOMP モデルでは時系列に循環成分が存在すると考えられる場合には状態変数をさらに $TC_t = T_t + C_t$ と分解して、 T_t にはトレンドモデル、循環成分 C_t には自己回帰モデル AR(p) を導入（循環成分のノイズを u_{3t} とする）、また曜日効果としては月内での曜日日数を全体の係数和をゼロとする制約条件付きの回帰モデルを (6)-(7) の状態空間表現に組み込むことにより実装している¹⁵。

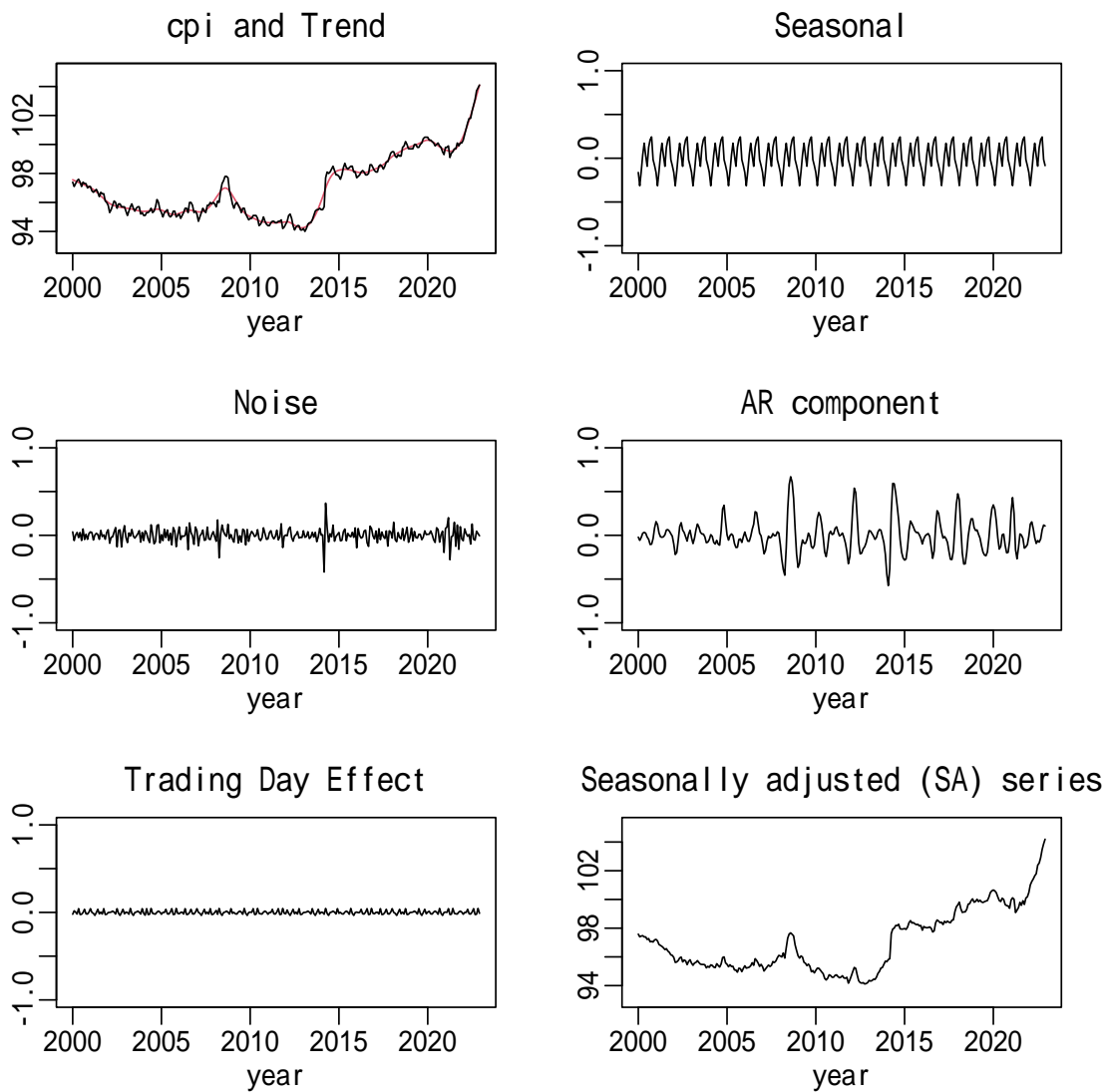
¹⁵こうした状態空間表現を利用した統計的モデル分析については北川 (2020) が説明している。状態空間表現はこうした追加の情報を比較的自然的に組み込むことが可能となるという長所がある。

< 図2：就業者数 (農林業, 月次, 労働力調査) の分解, DECOMP >
2002年1月～2022年10月, 原系列 (Agri) とトレンド (trend), 季節成分 (Seasonal), ノイズ成分 (Noise), 曜日成分 (Trading Day Effect), 季節調整系列 (SA).



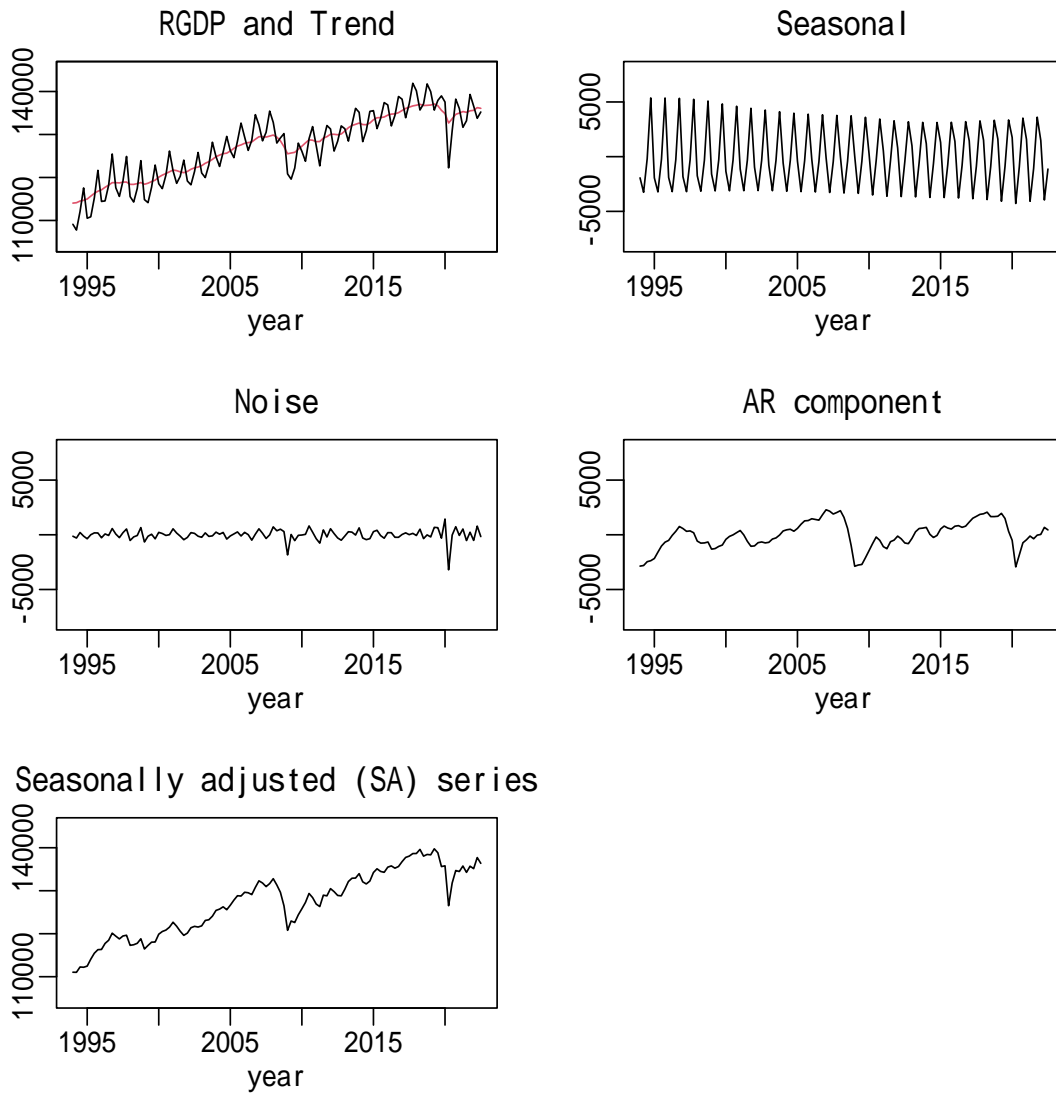
< 図3 : CPI, 全国, 総合, DECOMP >

2000年1月~2022年12月, 統計局原系列 (Aggri) とトレンド (trend), 季節成分 (Seasonal), ノイズ成分 (Noise), 曜日成分 (Trading Day Effect), 季節調整系列 (SA).



< 図4：RGDP データ (四半期) の分解, DECOMP >

1994Q1～2021年Q4, 内閣府・経済社会総合研究所, 原系列 (Aggri) とトレンド (trend), 季節成分 (Seasonal), ノイズ成分 (Noise), 循環成分 (AR), 曜日成分 (Trading Day Effect), 季節調整系列 (SA).



ここで図1で例として用いた就業者数データに対して DECOMP を利用して季節成分を推定し、得られた季節調整値を図2に示しておく。ここでのデータ分析では原系列を対数変換し、トレンド次数 $d=2$ 、ar 次数はゼロ、曜日効果あり¹⁶ と設定して計算した結果を示しておく。図3には最近注目されている日本の CPI, 図4には比較のため四半期データである実質 GDP に対する AR(2) を組み込んだ DECOMP の計算結果も示しておく¹⁷。こうして DECOMP を用いて求めた季節調整値と X-12, X-13 法に基づいて計算される数値を比べることは容易である。経済分析では水準そのものよりも増加率の数値に関心が集中するので、公表季節調整値による前期比と DDECOMP 季節調整値による前期比を比べてみると、両方の数値はそれほど異ならない。ただし、高岡・国友(2010)が報告している貿易統計の結果と比較すると、季節調整系列における不規則変動が比較的大きい。一つの解釈は、貿易統計データは業務における集計データであり、労働力調査はサンプリングに基づく調査による誤差の可能性があると思われる。かなり規則的な季節成分が検出されているが、その場合には結果として計算される季節調整済み系列は短期的な変動を含んで推定される。こうした場合には不規則変動を取り除いたトレンド・循環成分を利用する方が妥当ではないかと推察されるが、あるいは本報告書の最後の章で示すように季節成分の変動をより大きく認めれば、影響は小さくなる

さらに図3には四半期データとして実質 GDP 系列に DECOMP を利用して季節成分を推定し、得られた季節調整値を示しておく。ここでのデータ分析では原系列を対数変換し、トレンド次数 $d=2$ 、AR 次数は 2、曜日効果なしを用いた。この例では 2008 年ごろ 2019 年ごろの変化をトレンド成分、AR 成分、Noise 成分に分散して検出していることが分かる。

¹⁶ここでは統計数理研究所 Web から利用可能な RS-DECOMP (<https://jasp.ism.ac.jp/RS-Decomp/>) を用いた。なお、DECOMP では AR モデルによる循環変動を含めることも実用化されている。なお(故)赤池弘次氏などにより開発された時系列分析プログラム”TIMSAC” (<https://jasp.ism.ac.jp/ism/timsac/>) は R 上で利用可能な TIMSAC for R package となっているが RS-Decomp も含まれている。X-11 と DECOMP の比較については例えば「統計数理(1997) 季節調整特集などを参照されたい。

¹⁷むろん図2-図4は単なる例示である。産業別就業者数の季節調整値は公表されていない。また GDP 公表値を巡る最近の動向については高岡(2022)、山岸・高井・清水(2022)を参照されたい。

国友(2001)が指摘しているように、ここで利用した DECOMP モデルはラグ作用素により

$$(8) \quad Y_t = \frac{1}{(1 - \mathcal{B})^2} u_{1t} + \frac{1 - \mathcal{B}}{1 - \mathcal{B}^s} u_{2t} + u_{4t}$$

で与えられる。したがって、確率過程 Y_t の表現

$$(1 - \mathcal{B})(1 - \mathcal{B}^s)Y_t = \left(\sum_{i=0}^{s-1} \mathcal{B}^i \right) u_{1t} + (1 - \mathcal{B})^2 u_{2t} + (1 - \mathcal{B})(1 - \mathcal{B}^s) u_{4t}$$

が得られる。ここで右辺 $v_t = (\sum_{i=0}^{s-1} \mathcal{B}^i) u_{1t} + (1 - \mathcal{B})^2 u_{2t} + (1 - \mathcal{B})(1 - \mathcal{B}^s) u_{4t}$ は MA(s+1) 過程となるので、確率過程 Y_t は $ARIMA(0, 1, s+1) \times (0, 1, 0)_s$ 表現を持つと解釈できる。すなわち、DECOMP モデルは観測時系列が季節的線形構造を持つ d 次和分過程 (integrated process) $I(d)$ の実現値であることを仮定した上で、状態推定に最適フィルタリングを適用したと解釈できるのである。

ここで季節性を巡る議論における評価が困難な問題に言及しておこう。そもそも固定的な季節性から変化する季節性が現実的であると、季節性の定義、したがって季節調整値の妥当性が必ずしも明確でないことが指摘できる。直観的には時系列にトレンド成分が存在しない場合にはスペクトル密度関数が季節周期の周辺で山が存在し、その山を取り去る手続きが季節調整法という解釈がありうる。実際、ノーベル賞経済学者となった Granger(1978) は季節調整法の妥当性は季節調整系列の季節周期に山や谷が存在しないことと主張したが、この論点は経済学者 Sims(1978) により理論的には妥当でないことが指摘されている。例えば回帰モデルにおいて標準的な仮定の下で回帰成分を最適に最小二乗法により推定すると、データから推定されたノイズ成分には必ず自己相関が生じることが知られているが、類似の議論が周波数領域というより複雑な統計量について生じ得ることが知られているのである¹⁸。例えば Decomp プログラム

¹⁸この問題は日本銀行の木村(1997)が実務的観点より日本の公的統計における X-12-ARIMA による季節調整を巡る議論として再び提起した。統計学的な観点から川崎・佐藤(1987)はこの問題の解釈について詳しく議論している。

を用いるとしばしば季節調整系列のスペクトルには溝 (dip) が生じ、X-12-ARIMA プログラムを用いると季節調整系列のスペクトルに溝が生じにくいこと、などが報告されることがある。例えこうした指摘が事実であったとしても、X-12-ARIMA プログラムの適切性を示すことにはならないことには注意する必要があるだろう。

5 X-13ARIMA-SEATS

日本を含め多くの国の公的統計における季節調整法の利用については様々な経緯がある。公的統計における公表値の作成に関わることもあり、学術的な議論ではない各国の統計担当者間による対応の温度差などもある。特に公的統計が分散型である日本では担当部局間や政策当局での経済情勢の判断基準を巡って学術的でない議論や混乱などもあり、結果としては日本独自に開発されていた MITI 法や EPA 法は実務から姿を消したという経緯もある。また例えば米国や日本などでは 2010 年代末頃までは米国センサス局が開発してきた X-11, X-12-ARIMA が主に利用されているが、EU 諸国ではかなりスペイン中央銀行が開発した TRAMO-SEATS などが採用されていた経緯があるようである。

こうした公的統計における季節調整の潮流の中、2020 年代になり X-12-ARIMA のサポート・メンテナンスは既に中止、米国センサス局でも X-13ARIMA-SEATS の利用を推奨するようになっている。X-11, X-12-ARIMA は Fortran ソースプログラムの配布という形式をとっていたので、既存の季節調整の実務に大きな支障が生じるということはないという意味では大きな問題が生じたわけではない。

さて X-13ARIMA-SEATS の特長は X-12-ARIMA に TRAMO-SEATS が組み込まれたことである。TRAMO-SEATS はスペイン中央銀行の Gometz, Maraval により開発されたが、回帰モデルに基づく TRAMO と Kolmogorov-Wiener フィルターにより季節 ARIMA モデルのスペクトル分解を利用する SEATS という統計モデルに依存する Model-Based な季節調整法である。この TRAMO-SEATS、および X-13ARIMA-SEATS の詳細については、前者については Gomez and Marvall (2001),

Findley, Lytras and Maravall (2014), 高岡慎 (2015)、後者については X-13 マニュアル (Census Bureau (2020)) などにそれぞれ説明がある。さらに X-13ARIMA-SEATS の概略および評価については本報告書の他の解説を参照されたい¹⁹。

なお X-13ARIMA-SEATS は X-12-ARIMA と同様に Fortran で書かれているが、スペックファイルの作成方法などは同一であるから、SEATS 部分を除き X-11, X-12-ARIMA のコマンドは実行可能である。さらに Sax and Eddelbuettel (2018) が解説しているように、X-13ARIMA-SEATS は R 上のライブラリー”’seasonal’ (ウェブサイトは <http://seasonal.website/>) として利用可能となっているので、R から Fortran のオリジナルソースの大部分は実行可能になっている。ただし、X-13ARIMA-SEATS プログラムは基本的には X-11, X-12-ARIMA, TRAMO-SEATS などの FORTRAN プログラムを呼び出す形式になっているが、X12-ARIMA と同様に RegARIMA による前処理など、計算プログラム上で細かな修正が施され、定期的に Version-Up が行われていることには実務的運用の際には十分に注意する必要がある。

6 季節性と季節調整を巡る話題と X12SIML

日本においては 2008 年～2010 年頃に発生したマクロ経済の変動を巡る議論の中で、突然に発生する大きな経済変動と季節調整を巡る問題が時系列の季節調整という実務界においてもかなり大きな問題を投げかけた。さらに 2019 年から始まったコロナ問題や 2022 年からのウクライナ問題など世界経済、マクロ経済への影響が大きいこともあり、日本の GDP 公表値を巡っても問題が生じ得る可能性がある。その詳細についてはなお様々な議論があるが、2023 年初頭の段階で見るとリーマンショックの扱いを巡る教訓が生かされたようで、これまでに大きな問題は生じてはいないようである。

ここで以前に議論された論点、特に定期的に政府が公表している

¹⁹例えば X-13ARIMA-SEATS におけるスペクトル分析は赤池・石黒の BAYSEA のソースコードが利用されているようである。

マクロ経済指標の解釈やその評価をめぐり、政府当局者やエコノミストの間において季節調整値を巡る意見の相違は見かけないないようである²⁰。

ここでは特に2010年頃の大きな変動が観察される場合を貴重な事例として見てみよう。この時期の趨勢的変動、急激な変動、季節変動、景気循環、あるいは不規則なノイズ、などの構成要素を観測データから分解しようとする統計的問題となるので、注意深く統計的分析を行う必要がある²¹。また2009年頃と2019年頃の経験を踏まえて、幾つかの論点を述べておこう。季節調整の実務的方法として広く利用されているX-12-ARIMAでは統計学的には非標準的な非定常RegARIMAモデル利用しているが、デフォルトで用意されている変化点分析の方法は統計学的には妥当とは見なせないことを注意すべきである。例えば異常値処理コマンドoutlierなどを利用してダミー変数(LS, TSなど)による異常値の検出を行うことなどが可能であるが、その妥当性については今日的な統計学からはかなりの議論が必要である。また関連した問題としては、情報量規準に基づく統計モデルの選択を基礎づける議論は一時的な異常値とは見なせない大きな変化点が存在する場合などについては様々な統計学的な議論もある。観察される時系列に変化点があると、変化点の位置、変化点の数、などを考慮すると標準的な統計理論が成り立たない可能性が高い。統計学ではこうした変化点問題(change point analysis)については様々な研究が行われてきているが、例えばここで取りあげた季節性を含む非定常時系列についての実用的な結果は得られていない。

そこで教科書にない試行錯誤的な検討がいずれにしても必要となる。そこで日本のマクロ時系列の実例として貿易統計の分析例を挙げておく。高岡・国友(2010)は期間1999年7月から2009年9月のデータをかなり詳細に分析して「構造変化と季節調整の実務的問題」を検討している。彼らが事例研究として用いた貿易統計の時系列で

²⁰日本のGDP統計における季節調整問題の現状については高岡(2022)、山岸・高井・清水(2022)が議論している。

²¹こうしたマクロ経済統計を巡る幾つかの問題については国友・佐藤(2010)を参照されたい。

は2008年後半に米国・ヨーロッパの金融市場に端を発したマクロ経済ショックがその後世界的に大きな影響を及ぼし、日本の輸出入データは前年同月比で見ると一時期には30パーセント前後の落ち込みを記録した。2010年後半の時点では原系列は2009年より反転の兆しもあるが今後の変動を大きく見誤る可能性も少なくない。そこでRamp変数

$$(9) \quad Ramp[t_0, t_1]_t = \begin{cases} -1 & t \leq t_0 \\ \frac{t-t_0}{t_1-t_0} - 1 & t_0 < t < t_1 \\ 0 & t \geq t_1 \end{cases}$$

により定義、利用することなどが考えられた。この場合にはDECOMPによる分析ではその柔軟性のためと思われるが、不必要な季節性の変化は生じなかったことから高岡・国友(2010)はDECOMPによるX-12-ARIMAの数値の妥当性があることを主張している。いずれにしても実例として分析した2008年から2010年にかけて観察された主要な日本のマクロ経済変数の変動を単なる異常値と理解するには変動幅が大きい上に、広範囲にその影響が及んでいる。例えばこの変動を構造変化としてとらえ構造変化関数

$$(10) \quad f_t(t_0, t_1, t_2) = c_0 + c_1(-Ramp[t_0, t_1]_t) + c_2(Ramp[t_1, t_2]_t + 1)$$

などにより処理することなどが考えられた。(なおここで c_0, c_1, c_2 は定数、 t_0, t_1, t_2 は変化点である。)

こうした複数の変化点を含むような時系列分析の理論は未発達であり、理論的正当化の議論はなお今後の検討が期待される。X-12-ARIMAを利用する範囲内では変化点分析(change point analysis)を行う必要があり、さらに変化の始期や終期を確定する(あるいは推定する)必要が生じる。一つの暫定的試みとして高岡・国友(2010)は変化点を赤池情報量規準(AIC)最小化により選択することを提唱している。ここで(i)情報量規準の選択の議論の中では $AIC=(最大尤度)+2(母数の数)$ では第二項は小さすぎるのでペナルティ項2をより大きくすべきという議論や(ii)通常の変化点の選択では変化点を多く取りやすい、という論点を無視できないことにも注意してお

く。いずれにしても、大きな変化点があるにもかかわらずその存在を無視すると、季節調整値、季節調整値から計算される前期比などの公表値に不規則な影響を及ぼす可能性が大きい、と考えられる。

季節性の分析と季節調整の方法についての最近の研究動向の一つとして、X12SIML法にも言及しておく。日本における幾つかのマクロ経済時系列の原系列、季節調整系列を分析した経験、国友・櫻井・佐藤(2022)、Kunitomo and Sato (2021, 2023), Sato and Kunitomo (2021) などによる非定常時系列の周波数分析(スペクトル分解)の考察から佐藤(2023)はR上で動く季節調整プログラムX12-SIMLを開発している。基本的には季節性とは時系列における季節周期という特定の周期成分を意味すると考えられるが、厳密な三角関数で表現されるような決定論的に同一の季節性が繰り返されているわけではない。観測系列にトレンドなどの非定常性もあると考えられる場合に、離散時間の観測時系列のフーリエ型変換系列から特定の周期の周りの影響を取り除くことが考えられる。また時系列の離散的な観測値からフーリエ変換によりスペクトルを求めるときには、例えば月次データの場合、季節性はデータ上では6, 5, 4, 3, 2ヶ月周期に影響が出ることが知られている²²。X12SIML法はこれまで説明してきたX-11系統やDECOMPなどの季節調整法とは視点が異なっている。X-11, X12-ARIMA, X-13ARIMA-SEASTについての基本的疑問は季節性とは何か実ははっきり定義しないままに季節調整値を求めている、という基本的問題があり、X12SIML法はそうした疑問への解答を与える可能性がある。また既に議論したように、適切に季節性が除去されると結果としては季節調整系列には季節周期のあたりにdipが生じることが知られている。X12SIML法は周波数バンドにもとづくフィルタリング、という新たなアイデアのもとに実用化されつつある分析法であるが、R上で開発されたプログラムによりPC上で簡単に利用できる。マクロ経済データの季節性を分析する際に今後、応用可能性がかなりあると考えられるので、X12-SIML

²²例えば月次系列から12ヶ月周期の波動と6ヶ月周期の波動を区別することはできない。同様に四半期系列から4期の波動と2期の波動は区別できない。

法の一つの応用例を本稿の最後の章で述べておく。

最後に蛇足であるが一言。残念ながら学術的な意味での統計的時系列解析 (statistical time series analysis) の中で季節調整の問題に関心を持つ研究者はそれほど多くない。学術的な議論では理論的根拠の明晰性や独自性が評価されるが、実務的観点からは目前にある時系列データから信頼される公表値の作成には猶予は許されないことも一因であろう。実務家からの観点での季節調整法の説明も文献としては散見されるが、既存の季節調整法の計算プログラムの説明や利用法の説明が中心なことが多く、実は学術的検証に耐えうる内容はそう多くないとは理論的根拠を求める統計家の偏見かもしれない²³。2023年初頭においても学術的、実務的の両面において検討課題は少なくない。実務界と学術界の共同作業が必要であり、実務的に評価され得る学術的研究は重要な課題であり続けている。

参考文献

赤池弘次(1989),「事前分布の選択とその応用」,鈴木・国友編「ベイズ統計学とその応用」東京大学出版会.(収録)

<http://www.kunitomo-lab.sakura.ne.jp/Bayes1987.pdf>

経済企画庁経済研究所(1971),「季節変動調整法」研究シリーズ22.

黒川恒雄(1979),「経済時系列の分析とその季節変動の調整」(1)~(12)、統計,日本統計協会.

ハーベイ(1986),「時系列モデル入門」(国友直人・山本拓訳)東京大学出版会.

<http://www.kunitomo-lab.sakura.ne.jp/Harvey1985.PDF>

川崎能典・佐藤整尚(1997)“季節調整の最適性について,”季節調整法特集,245-263,統計数理研究所.

²³そうした中では例えば、赤池弘次、石黒真木夫、北川源四郎など統計数理研究所の研究者たちの研究は今なお実務に有用な研究として例外的に寄与してきているというのが妥当な評価と思われる。

- 木村武 (1997),”季節調整に関する実務的諸問題,”季節調整法特集, 181-216, 統計数理研究所.
- 国友直人 (2001),”季節調整法 X-12-ARIMA(2000) の利用 : 法人企業統計の事例,” 経済学論集 Vo.67-3.
- 国友直人編 (2004),”解説 X-12-ARIMA(2002),” 研究報告 CIRJE-R-1, 東京大学経済学研究科・日本経済国際共同研究センター (CIRJE).
<http://www.cirje.e.u-tokyo.ac.jp/research/dp/2001/2001cj47.pdf>
- 国友直人・高岡慎 (2005),”経済季節性と季節転換時系列モデル,” 日本統計学会誌, Vol. 35-1, 1-26.
- 国友直人編 (2006),”季節調整法 X-12-ARIMA と日本の官庁統計,” 研究報告 CIRJE-R-5, 東京大学経済学研究科・日本経済国際共同研究センター (CIRJE).
- 国友直人・佐藤整尚 (2010),”日本のマクロ経済統計の課題 : 季節性・構造変化・平滑化問題,” Discussion Paper CIRJE-J-230, 日本経済国際共同研究センター (CIRJE),
<http://www.cirje.e.u-tokyo.ac.jp/research/dp/2010/2010cj230.pdf>.
- 国友直人・櫻井智章・佐藤整尚 (2022),”経済時系列の状態推定とマクロ指標,” 統計研究彙報, 79, 1-20, 統計局.
<https://www.stat.go.jp/training/2kenkyu/ihou/79/pdf/2-2-791.pdf>
- 北川源四郎 (2020),”R による時系列モデリング入門,” 岩波書店.
- 溝口敏行・刈屋武昭 (1983), 「経済時系列入門」日本経済新聞社.
- 高岡慎 (2002),”曜日効果の識別と確率モデル,” 日本統計学会誌 Vo.32-2, 327-343.
- 高岡慎 (2015), 「経済時系列と季節調整法」, 朝倉書店.
- 高岡慎 (2022),”GDP 関連系列の季節調整における外れ値処理の妥当性について,” 内閣府社会経済研究所, 国民経済計算関連論文, (記

- 号/sna_に注意)
https://www.esri.cao.go.jp/jp/sna/data/sna_ronbun/menu.html
- 高岡慎・国友直人 (2010),”最近のマクロ経済変動と季節調整：貿易統計を題材に,” 経済学論集 (東京大学経済学部) Vo.76-1, 56-74.
- 佐藤整尚 (1997)”Web-Decomp の紹介,” 季節調整法特集, 233-243, 統計数理研究所.
- 佐藤整尚 (2023) 「X12SIML を使った季節調整の実際」.
<http://www.kunitomo-lab.sakura.ne.jp/x12simldoc92.pdf>
- 山岸圭輔, 高井健作, 清水美絵 (2022), ”四半期別 GDP 速報における季節調整方法について—大きな経済変動に対する異常値処理,” 内閣府社会経済研究所, 国民経済計算関連論文, (記号/sna_に注意)
https://www.esri.cao.go.jp/jp/sna/data/sna_ronbun/menu.html
- Anderson, T.W. (1971), *The Statistical Analysis of Time Series*, John-Wiley.
- Census Bureau (2020), ”X-13ARIMA-SEATS Reference Manual”(Accessible HTML Outout Version), U.S. Census Bureau.
- Findley, D., Monsell, B., Bell, W., Otto, M. and Chen, B. (1998),”New Capabilities and Methods of the X-12-ARIMA Seasonal Adjustment Program,” *Journal of Business and Economic Statistics*, 16, 127-177.
- Findley, D., D.P. Lytras and A. Maravall (2014),”Illuminating Model-Based Seasonal Adjustment with the First Order Seasonal Autoregressive Model, Bureau of Census.
- Fuller, W. (1996) *Introduction to Statistical Time Series*, 2nd Edition, John-Wiley.
- Gomez, V. and Maravall, A. (2010), ”Seasonal Adjustment and Signal Extraction in Economic Time Series,” Chapter 8 of *A Course on Time Series Analysis*, by Pena, Tiao and Tsay ed., Wiley.

- Granger, C. (1978), "Seasonality : Causation, Interpretation and Implications," In *Seasonal Analysis of Economic Time Time Series*, Proceedings of on the Coference of Seasonal Analysis of Economic Time Time Series, Census Bureau, Department of Commerce, U.S. Goverment.
- Kunitomo, N. and S. Sato (2021), "A robust-filtering method for noisy non-stationary multivariate time series with econometric applications," *Japanese Journal of Statistics and Data Science*, 4, 373-410.
<https://link.springer.com/article/10.1007/s42081-020-00102-y>
- Kunitomo, N. and S. Sato (2023), "Frequency Regression and Smoothing for Noisy Non-stationary Time Series," Discussion Paper, Expert Project (ISM, forthcoming).
<https://stat-expert.ism.ac.jp/training/discussionpaper/>
- Nerlove, M, D. Grether, and Carvalho, J. (1979), *Analysis of Economic Time Series*, Academic Press.
- Sato, S. and N, Kunitomo (2021), "On Backward Smoothing for Noisy Non-stationary Time Series","
<http://www.mims.meiji.ac.jp/publications/2021-ds/SDS-15.pdf>
- Shiskin, J., A. H. Young, and J.C. Musgrave (1967), "The X-12 Variant of the Census Method II Seasonal Adjustment Program," Technical Report No. 15, U.S. Census Bureau, Department of Commerce.
- Sax, C. and Eddelbuettel, D. (2018), "Seasonal Adjustment by X-13ARIMA-SEATS in R," *Journal of Statistical Software*, 87-11, 1-17.
- Sims, C. (1978), "Comments on Granger's Seasonality : Causation, Interpretation and Implications," In *Seasonal Analysis of Eco-*

conomic Time Time Series, Proceedings of on the Coference of Seasonal Analysis of Economic Time Time Series, Census Bureau, Department of Commerce, U.S. Goverment.

第II部: X-13ARIMA-SEATS

(i) Kolmogorov-Wiener フィルター (国友)²⁴

1. 簡単な例示

X-13ARIMA-SEATS 季節調整法では TRAMO-SEATS 法と呼ばれていた季節調整プログラムが利用可能である。X-13-ARIMA 法の詳細については次章の説明に譲るが、TRAMO-SEATS 法は第I部で議論した X-11, X-12 の基礎となっていた移動平均平均法とはかなり異なる確率論・統計的時系列分析の原理から開発されていたと云う経緯がある。

そこで本稿では TRAMO-SEATS 法の基礎となっている Kolmogorov-Wiener フィルターについて公的統計・経済統計における時系列の応用に関心のある応用家を対象に直観的な説明を試みる。なお Kolmogorov、Wiener はともに数学系の研究者であり、一般的状況では数理的に様々な問題と関連している。本稿では季節調整への応用という観点より定常的な 1 次元離散時系列の線形予測理論という観点からできるだけ分かりやすい内容に限定して説明を試みる。多次元確率過程、連続時間の確率過程など一般の場合の説明については例えば Hannan (1971) が知られている²⁵。

まず簡単な例により離散時間の時系列における問題と数理的理論の必要性を説明しよう。観測可能な時系列を分解して (時系列) = (シグナル) + (ノイズ)、と見なし、シグナルが定常 AR(1) $X_t = aX_{t-1} + U_t$, ノイズ V_t が i.i.d. 系列, 観測系列 $Y_t = X_t + V_t$ としよう。ここで $U_t \sim N(0, \sigma_u^2)$, $V_t \sim N(0, \sigma_v^2)$ の互いに独立な確率変数列とする。 Y_t

²⁴2022 年 6 月 6 日に統計エキスパート養成講座「時系列と季節調整」の為に準備した内容をまとめた原稿である。

²⁵日本の大学の理系学部のカリキュラム数学・物理学で学ぶ Fourier 解析の基礎は連続時間の関数を連続観測できるという設定での議論が大部分と思われる。統計的時系列分析の標準的内容は離散時間にデータ観察されることを想定したスペクトル理論であり、データ分析における連続時間・離散時間の違いを認識することが時系列のデータ分析の出発点と云える。統計的時系列の古典的教科書は Anderson (1971), Brockwell and Davis (1990) などである。Hilbert 空間上の (物理学を想定していると思われる連続パラメータの) スペクトル分解については藤田・黒田・伊藤 (1991) が詳しいが、離散時間パラメータの定常過程のスペクトル分解については Brockwell and Davis (1990) chapter 4 が比較的詳しく議論している。マクロ経済時系列はトレンドがあり非定常性が顕著であるとともに循環成分・季節成分・ノイズ成分なども混在し、観測数がそれほど多くないがデータであり、その統計的分析には実はなお学問的課題も少なくない。

の自己共分散を求めると

$$\begin{aligned} \text{Var}(Y_t) &= \text{Var}(X_t) + \sigma_v^2, \text{Cov}(Y_t, Y_{t-1}) = \text{Cov}(X_t + V_t, X_{t-1} + V_{t-1}) = \\ &= \text{Cov}(aX_{t-1} + U_t + V_t, X_{t-1} + V_{t-1}) = a\text{Var}(X_{t-1}) = a[\text{Var}(Y_t) - \sigma_v^2], \\ \text{Cov}(Y_t, Y_{t-s}) &= a^s \text{Var}(X_{t-s}) = a^s [\text{Var}(Y_{t-s} - \sigma_v^2)] = a\text{Cov}(Y_t, Y_{t-s+1}) \\ &(s \geq 2) \end{aligned}$$

となる。このことから、さらに若干の計算（例えば自己共分散関数は $C(s) = aC(s-1)$ ($s \geq 2$), $C(1) = a\sigma_u^2/(1-a^2)$, $C(0) = \sigma_u^2/(1-a^2) + \sigma_v^2$ となる）より定常 ARMA(1,1) モデルの自己共分散の構造と一致することが確認できる²⁶。すなわち、ある i.i.d. 確率変数列 w_t が存在して母数 θ と σ_w^2 を適切に選ぶと、 $Y_t = aY_{t-1} + w_t - \theta w_{t-1}$ ($|a| < 1$, $|\theta| < 1$) という反転可能な (invertible) 定常解の表現を持つことが分かる。こうした単純な例でも計算はかなり複雑化するので、ここで扱っている問題は自明でないことが確認されよう。

ここでの問題は与えられた観測系列 Y_t がしたがう確率過程が分かっているときにノイズ部分を取り除き、シグナル部分を抽出することである。この問題は自明ではないが、工学系の信号処理の問題としてはかなり古くから研究されている。

2, Hilbert で行こう！

実際の時系列が離散的に観測されるとして、離散時間にデータが得られる離散時間の 1 次元確率変数列 $X_t, t = \dots - 2, -1, 0, 1, 2, \dots$ を考える。また数理的に扱うには都合の良い条件、分散が存在する弱定常過程の実現値がデータとして得られる状況を想定しよう。

こうした設定の下である時刻 t において過去の情報 X_{t-1}, X_{t-2}, \dots が与えられた時に X_t の予測問題を考えよう。ここではまず次のように問題を設定する。

(i) 不確実な予測の良さを測る評価関数として予測の平均二乗誤差 MSE を採用する。

(ii) 確率変数は分散・自己共分散が存在する二乗可積分全体、すな

²⁶例えば山本 (1986) には ARMA(1,1) の自己共分散関数が与えられているが、 $Y_t - aY_{t-1} = w_t - \theta w_{t-1}$ (w_t は i.i.d. 系列で分散 σ_w^2) の自己共分散関数は $C(s) = aC(s-1)$ ($s \geq 2$), $C(1) = \sigma_w^2(1-a\theta)(a-\theta)/(1-a^2)$, $C(0) = \sigma_w^2(1-2a\theta+\theta^2)/(1-a^2)$, とする。 $C(0), C(1)$ の条件から $\sigma_w^2(1-2a\theta+\theta^2) = \sigma_u^2 + (1-a^2)\sigma_v^2$, $\sigma_w^2(1-a\theta)(a-\theta) = a\sigma_u^2$ を解くと $\theta^2 - [(1+\lambda+a^2\lambda)/(a\lambda)]\theta + 1 = 0$ が得られるので $|\theta| < 1$ となる反転可能な (invertible) 実根を得ることができる。

わち L^2 空間 (Hilbert 空間) の元として、(原理的には可算) 無限個の確率変数が存在することを想定する²⁷。

(iii) このとき最適な予測量は条件付期待値

$\mathbf{E}[X_t | X_{t-1}, X_{t-2}, \dots]$ で与えられる。ただし一般には条件付期待値は線形予測量とは限らないので問題はかなり複雑化する。

(iv) そこで比較的簡単に構成することが可能な予測量として、線形予測量を検討する。ここで線形予測量とは

$$(11) \quad L(t, m, n) = X_{t+m|t} = \sum_{k=1}^n a_k X_{t-k}$$

と表現できることを意味する (ただし係数は $\sum_{k=1}^{\infty} a_k^2 < \infty$ とする)。

(iv) この問題は無限次元 L^2 空間における最小化問題となり、最小二乗法により解が求まる。この最小二乗解は Komogorov-Wiener 解と呼ばれている。ここで注意すべきことは弱定常確率過程という限定的な範囲において内積に基づく解が得られることである。

ヒルベルト空間における最小二乗法 : 本稿で利用する数理的道具立ての設定に馴染みのない諸君も少なくないと思われるの直観的に Hilbert 空間を説明しておこう。直観的にはヒルベルト空間は比較的なじみ深いユークリッド空間 R^n ($n \geq 2$) に似た性質を備えている応用上で有用な空間であるので、まずはユークリッド空間を考えよう。実数上の有限次元 R^n とは n 個の実数を並べたベクトル全体からなる空間であり、 $n = 3$ の場合は高校の幾何で既に扱われている。3次元空間ではベクトルの角度、ベクトルと平面までの距離などがイメージできるだろう。そこで m 個の n 次元縦ベクトルで張られる n 次元ユークリッド空間における二つの元 $X_j = (X_{1j}, \dots, X_{nj})'$, $Y_k = (Y_{1k}, \dots, Y_{nk})'$ ($j, k = 1, \dots, m$) に対して内積を $\langle X_j, Y_k \rangle = \sum_{i=1}^n X_{ij} Y_{ik}$ により定義しよう。

ベクトル Y を予測する線形予測量 $L(X, k) = \sum_{j=1}^k a_j X_j$ とすると予測誤差は $U = Y - \sum_{j=1}^k a_j X_j$ となる。予測されるベクトル Y と予

²⁷ 本稿では時間パラメーターが離散なので可算個の確率変数の議論で十分である。時間パラメーターが連続なら非可算個の確率変数を考察する必要が生じる。

測量との距離の二乗

$$(12) \quad \langle U, U \rangle = \|Y - \sum_{j=1}^k a_j X_j\|^2$$

を係数 a_j について最小にする解は正規方程式

$$(13) \quad \langle Y - L(X, k), X_j \rangle = 0 \quad (j = 1, \dots, k)$$

の解として与えられるが、この回帰方程式は $n \times k$ 行列 $X = (X_{ij})$, k 次元ベクトル $a = (a_j)$ を用いて $X' X a = X' Y$ と表現できることはよく知られ、統計学では正規方程式 (normal equation) と呼ばれている。

最小二乗解は n 次元空間においてベクトルから超平面までの距離を最小にするという幾何学的解釈が可能である。この場合をしばしば図1のように表現される。すなわち有限次元のベクトル Y を k 個の説明変数ベクトル x_j が張る超平面 X に対して射影されたベクトルが最小二乗法により得られる解と一致するのである。ここでベクトルの次元は n としたがここでの幾何学的なイメージはベクトル間の距離が適切に定義されれば数理的に拡張できることが想像される。

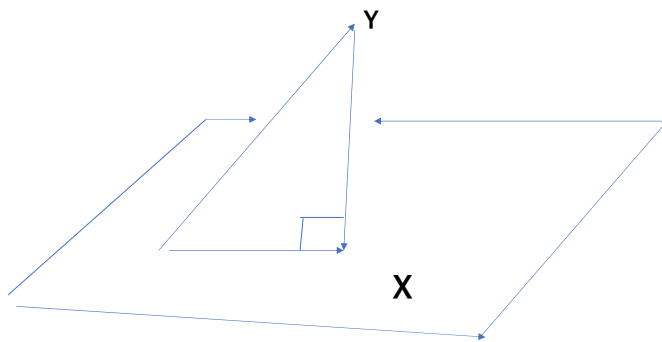
すなわち、ここで述べたよく知られた最小二乗法の議論は、 n 次元ユークリッドの幾何学であるが、ほぼそのまま次元数が可算無限個のヒルベルト空間に拡張することができるのである。

ここで二乗可積分な確率変数からなる空間

$$H = \{X_t, \mathbf{E}[X_t] = 0, \mathbf{E}[X_t^2] < \infty, t = \dots, -1, 0, +1, \dots\}$$

を考察する。数理的議論の詳細は教科書を参照してもらうこととして²⁸、ここでは完備距離空間 (complete metric space) である標準的場合を扱い、内積は $\langle X_s, X_t \rangle = \mathbf{E}[X_s X_t]$ により定義する。ここで時刻 s (および時刻 t) に等しいかそれより前の確率変数全体を H_s (および H_t) とそれぞれ置こう。これらの部分空間 $H_s \subset H_t$ ($s < t$) として全体の空間を $H = H_\infty = \bigcup_t H_t$ としよう。

図1:最小二乗法の幾何学



²⁸例えば藤田・黒田・伊藤 (1991) は基本的・標準的内容が説明されている。ここではあくまで応用上で必要な最小限で直観的な説明にとどめる。

3. 定常過程の最小二乗法

弱定常確率過程については古典的な時系列解析の基礎としてかなりの研究の蓄積がある。それらについては例えば Anderson (1971), Brockwell and Davis (1990) などの教科書で詳しく説明されている。ここでは様々な議論の出発点として Anderson (1971) Section 7.6.3 で詳しく説明されているウォルド (Wold) 分解に言及しておく。直観的にはある時刻 t における確率変数 X_t を過去の部分空間 H_{t-1}, H_{t-2}, \dots で予測した時の予測誤差を表す確率変数列により X_t の Causal (因果的) な表現が得られる、ということの意味する。

Wold の分解定理 : 弱定常 1 次元確率過程 X_t は期待値ゼロ, 分散が存在するとき表現

$$(14) \quad X_t = \sum_{s=0}^{+\infty} b_s Z_{t-s} + V_t, \quad \sum_{s=0}^{+\infty} b_s^2 < +\infty$$

を持つ。ただし $\mathbf{E}[Z_t] = \mathbf{E}[V_t] = \mathbf{E}[Z_s Z_t] = \mathbf{E}[Z_t V_s] = 0$ ($s \neq t$), $\mathbf{E}[Z_t^2] = \sigma^2$, $Z_t \in H_t$, $V_t \in \cap_{s=0}^{+\infty} H_{t-s}$ とする。ここで $V_t = 0$ となる場合を正則過程 (regular) と呼ぶ。

(無限の過去に決まっている項 V_t を含まない普通の) 正則過程の例としては自己回帰 (autoregressive, AR) モデル, 移動平均 (moving average, MA) モデル, 自己回帰移動平均 (ARMA) モデルなど時系列解析ではよく知られた時系列モデルがある。通常の時系列解析の応用では正則過程を扱うので以下では正則性を仮定する。

次に期待値ゼロの定常確率過程において基本的で重要な概念を定義しておく。自己共分散関数は

$$(15) \quad C(k) = \mathbf{E}[X_t X_{t-k}], \quad k = 1, 2, \dots$$

とするが、この量は $\int_{-1/2}^{1/2} \exp(2\pi i \lambda k) f(\lambda) d\lambda$ に一致する。ただし $i^2 =$

-1, $f(\lambda)$ はスペクトラム

$$(16) \quad f(\lambda) = \sum_{k=-\infty}^{+\infty} C(k) \exp(-2\pi i \lambda k) \\ = \sum_{k=-\infty}^{+\infty} C(k) \cos(2\pi \lambda k) \quad , -1/2 \leq \lambda \leq 1/2$$

で与えられる。(スペクトル密度関数の存在を仮定する。) 同様に相互共分散関数は

$$(17) \quad C_{xy}(k) = \mathbf{E}[X_t Y_{t-k}] \quad , k = 1, 2, \dots$$

および相互スペクトラム $f_{xy}(\lambda)$ は

$$(18) \quad f(\lambda) = \sum_{k=-\infty}^{+\infty} C_{xy}(k) \exp(-2\pi i \lambda k)$$

で与えられる。ここで z 変換は

$$(19) \quad g(z) = \sum_{k=-\infty}^{+\infty} C(k) z^k$$

とするが、 z は複素数である。このとき

$$f(\lambda) = g(e^{-2\pi i \lambda})$$

と表現できる。

予測問題の解 : ここで導入した概念を利用して弱定常過程の予測問題の最小二乗解を求めよう。簡単化の為に観測誤差は存在せずに確率過程がそのまま観測されると仮定する。Wold 表現の 1 次元定常過程 $X_t = \sum_{s=0}^{+\infty} b_s Z_{t-s}$, $\sum_{s=0}^{+\infty} b_s^2 < +\infty$ における時刻 $t+m$ の X_t (m 期先の予測) に対する予測量を $L(t, m) = \sum_{k=1}^{\infty} a_k X_{t-k}$ とする。この予測量が $L(t, m) = \sum_{k=1}^{\infty} c_k Z_{t-k}$ と表現できるとする。ここで後退作用素 (バックワードオペレーター) B を用いて $\sum_{k=1}^{\infty} a_k B^k X_t$, $\sum_{k=1}^{\infty} c_k B^k Z_t$ より母関数 $a(z) = \sum_{k=1}^{\infty} a_k z^k$, $c(z) = \sum_{k=1}^{\infty} c_k z^k$ と表すと、形式的に

$$(20) \quad c(z) = a(z)b(z)$$

となる。このことから最小二乗解は

$$(21) \quad a(z) = \frac{[b(z)z^{-m}]_+}{b(z)}$$

で与えられる。ただし $[\]_+$ は z の正ベキ項のみをとる操作とした。

導出の概略 1: 予測の平均二乗誤差を評価すると

$$(22) \quad \begin{aligned} \mathbf{E}[X_{t+m} - L(X, t, m)]^2 &= \mathbf{E}[b(B)Z_{t+m} - c(B)Z_t]^2 \\ &= \mathbf{E}[b(B)Z_{t+m} - c(B)B^m Z_{t+m}]^2 \\ &= \mathbf{E}[p(B)Z_{t+m}] \end{aligned}$$

と表現できる。ここで $p(B) = b(z) - c(B)z^m$ である。 $p(B)$ として特に $p_j = b_j$ ($j = 0, \dots, m$), $p_{m+j} = b_{m+j} - c_j$ ($j \geq 1$) とおくと

$$c(z) = \left[\frac{b(z)}{z^m}\right]_+ = a(z)b(z)$$

となるので解 $a(z)$ が得られる。

導出の概略 2: 最小二乗解の正規方程式より求める方法も（よく知られている有限次元の最小二乗法から類推できるので統計的には）興味深いので述べておく。正規方程式は

$$(23) \quad \mathbf{E}[X_{t+m} - L(X, t, m), X_{t-k}] = 0 \quad (k = 1, 2, \dots)$$

で与えられる。（ $L(X, t, m) = a(B)X_t$ である。）上の式は

$$\mathbf{E}[L(X, t, m)X_{t-k}] = \mathbf{E}[X_{t+m}X_{t-k}] \quad , (k = 1, 2, \dots)$$

であるから

$$(24) \quad \sum_{l=1}^{\infty} a_l C_{xx}(k-l) = C_{xx}(k+m) \quad (k = 1, 2, \dots)$$

と表現できる。ここで両辺に z^{k+l} を乗じて

$$z^m \sum_{l=1}^{\infty} a_l z^l z^{k-l} C_{xx}(k-l) = z^{k+m} C_{xx}(k+m) \quad (k = 1, 2, \dots)$$

となることに注意して l, k について無限和をとると右辺は $\sum_{p=1}^{\infty} z^p C_{xx}(p) - \sum_{p=1}^m z^p C_{xx}(p)$ と分解できるので

$$z^m a(z) g_{xx}(z) = g_{xx}(z) + h(z),$$

ただし $h(z)$ は m ないし m 以下の多項式となる。ここで $g_{xx}(z) = \sigma^2 b(z) b(z^{-1})$ となることを利用すると

$$\sigma^2 a(z) b(z) = \frac{g_{xx}(z)}{z^m b(z^{-1})} + \frac{h(z)}{z^m b(z^{-1})}$$

となる。左辺は負のべきを含まず、右辺第1項は非負のべき、第2項は負のべきであるから右辺は左辺の第1項と一致するので、結局 $a(z) = [1/b(z)][b(z)/z^m]_+$ が得られる。

例1 : シグナルが AR(1) のとき $X_t = aX_{t-1} + Z_t$ とすると $(1 - aB)X_t = Z_t$ であるから $b(z) = (1 - az)^{-1} (|a| < 1)$ より

$$(25) \quad a(z) = (1 - az) \left[\frac{1}{z^m (1 - az)} \right]_+$$

となるので、予測量は

$$(26) \quad L(X, t, m) = a^{m+1} X_{t+1}$$

となる。この場合には予測の平均二乗誤差は $\sigma^2 \sum_{j=0}^m a^{2j}$ で与えられる。(ただし $\mathbf{E}[Z_t] = 0, \mathbf{E}[Z_t^2] = \sigma^2$ である。)

例2 : シグナルが MA(1) のとき $X_t = Z_t - bZ_{t-1}$ とすると $b(z) = (1 - bz) (|b| < 1)$ となる。 $m > 0$ なら

$$a(z) \left[\frac{1}{1 - bz} \right] [1 - bz]_+$$

となるので予測量は

$$(27) \quad L(C, t, m) = - \sum_{j=0}^{\infty} b^{j+1} X_{t-(j+1)}$$

で与えられる。

例3 : シグナルが ARMA(1,1) の時には $X_t = aX_{t-1} + Z_t - bZ_{t-1}$

とすると $(1 - aB)X_t = (1 - bB)Z_t$ より同様に解を求めることができるが、若干複雑になる。

観測誤差とフィルタリング : 観測するデータに観測に伴うノイズが混入している状況では観測データから真の状態を推定することが重要な問題であり、一種の予測問題と見なすことができるが、状態推定、フィルタリング問題とも呼ばれている。単純な設定の観測誤差モデルでは観測値は状態変数 X_t , 観測ノイズ V_t の互いに独立な確率変数として、 $Y_t = X_t + V_t$ ($t = 1, 2, \dots, n$) と表現することが考えられる。このとき観測系列 Y_{t-1}, Y_{t-2}, \dots から状態変数 X_{t+m} ($m > 0, m = 0, m < 0$) の推定問題を考察しよう。1次元弱定常過程 $X_t = \sum_{s=0}^{+\infty} b_s Z_{t-s}$, $\sum_{s=0}^{+\infty} b_s^2 < +\infty$, ノイズ系列 V_t として時刻 $t; m$ の X_t に対する推定量を $L(t, m) = \sum_{k=1}^{\infty} a_k Y_{t-k}$ とすると、最小二乗解は

$$(28) \quad a(z) = \frac{1}{\sigma^2 b(z)} \left[\frac{g_{xx}(z)}{b(z^{-1})} z^m \right]_+$$

で与えられる。

導出の概略 : (予測の場合の導出2を少し一般化して Y_{t-k} を X_{t-k} の代わりにして定式化すると) 正規方程式は

$$(29) \quad \mathbf{E}[X_{t+m} - L(Y, t, m), Y_{t-k}] = 0, (k = 1, 2, \dots)$$

で与えられる。($L(Y, t, m) = a(B)Y_t$ である。) 上の式は

$$\mathbf{E}[L(Y, t, m), Y_{t-k}] = \mathbf{E}[X_{t+m}, Y_{t-k}] = 0, (k = 1, 2, \dots)$$

であるから

$$(30) \quad \sum_{l=1}^{\infty} a_l C_{yy}(k-l) = C_{xy}(k+m), k = 1, 2, \dots$$

となるが、両辺に $z^{k+m-l+l}$ を乗じて導出2と同様に無限和をとると

$$(31) \quad z^m a(z) g_{yy}(z) = g_{xy} + h(z)$$

となる。だし $h(z)$ は m ないし m 以下の多項式となる。ここで $g_{yy}(z) = \sigma^2 b(z)b(z^{-1})$ となることを利用すると

$$(32) \quad a(z) = \frac{1}{\sigma^2 b(z)} \left[\frac{g_{xy}(z)}{z^m b(z^{-1})} \right]_+$$

となる。ここで X_t と V_t は仮定から独立なので、 $g_{yy}(z) = g_{xx}(z) + g_{vv}(z)$ となる。

例4 : 真の状態変数 $X \sim AR(1)$, 状態とは独立な項 V は i.i.d. 系列とすると $Y \sim ARMA(1, 1)$ となる。この場合には

$$g_{xx}(z) = \frac{\sigma^2}{(1-az)(1-az^{-1})}, \quad g_{vv}(z) = \lambda\sigma^2,$$

$$g_{yy}(z) = \omega^2 \frac{(1-\eta z)(1-\eta z^{-1})}{(1-az)(1-az^{-1})}$$

より $m > 0$ のとき

$$a(z) = \frac{a\sigma^2}{\omega^2} \frac{a^m}{1-a\eta} \frac{z}{1-\eta z}$$

となる。したがって

$$(33) \quad a_j = ca^m \eta^{j-1} \quad (j = 1, 2, \dots), \quad c = \frac{\eta}{\lambda(1-a\eta)}$$

となる。(c は定数である。)

観測誤差と平滑化 : 最後に状態変数 X_t , 観測誤差 V_t とする観測誤差モデル $Y_t = X_t + V_t$ ($t = 1, 2, \dots, n$) の下で観測値

$Y_1, \dots, Y_{t-1}, X_t, Y_{t+1}, \dots, Y_n$ から状態変数 X_t を推定する問題を考える。この場合には与えられた情報は前後に利用可能な観測データであり、平滑化 (smoothing) 問題と呼ばれている。1次元弱定常過程 $X_t = \sum_{s=0}^{+\infty} b_s Z_{t-s}$, $\sum_{s=0}^{+\infty} b_s^2 < +\infty$, ノイズ系列 V_t として時刻 t の X_t に対する平滑化推定量を $L(t) = \sum_{j=-\infty}^{\infty} a_j Y_{t-j}$ とする。最小二

乗解は

$$\begin{aligned}
 (34) \quad 0 &= \mathbf{E}[(X_t - L(t))Y_{t-k}] \\
 &= \text{Cov}(X_t, Y_{t-k}) - \text{Cov}\left(\sum_{-\infty}^{+\infty} a_j Y_{t-j}, Y_{t-k}\right) \\
 &= \int_{-1/2}^{1/2} e^{2\pi i k \lambda} [f_{xy}(\lambda) - \sum_{-\infty}^{+\infty} a_j e^{-2\pi i j \lambda} f_{yy}(\lambda)] d\lambda
 \end{aligned}$$

で与えられる。ここで

$$f_{xy}(\lambda) - \sum_{k=-\infty}^{\infty} a_k e^{-2\pi i k \lambda} f_{yy}(\lambda) = 0 \quad a.e.$$

とすると最小二乗解は

$$(35) \quad a(z) = \frac{g_{xy}(\lambda)}{g_{yy}(\lambda)} = \frac{g_{xy}(\lambda)}{g_{xx}(\lambda) + g_{vv}(\lambda)}$$

で与えられることが分かる。

例5 : 真の状態変数 X_t が $AR(1)$ モデル (1次自己回帰モデル) にしたがって、 V_t は i.i.d. 確率変数系列で互いに独立とする。この場合には

$$\begin{aligned}
 g_{xx}(z) &= \frac{\sigma^2}{(1-az)(1-az^{-1})}, \quad g_{vv}(z) = \lambda\sigma^2, \\
 g_{yy}(z) &= \omega^2 \frac{1}{1-az} \frac{1}{1-az^{-1}}
 \end{aligned}$$

より平滑化の係数は

$$(36) \quad a_j = c\eta^{|j|+1} \quad \left(c = \frac{\eta}{\lambda a(1-a^2)}\right)$$

で与えられることが導ける。

以上の議論では簡単な時系列モデルを用いてその導出を例示した。こうした導出は定常 ARMA モデルなどを含む確率過程へ一般化が可能であり、次章で説明されるように季節調整法 Tramo-Seats では実際に ARMA モデルのスペクトル分解が利用されている。

4. 季節性と非定常性を巡って

本稿では季節調整法の理解に必要な伝統的な離散時間の統計的時系列解析の基礎と Kolmogorov-Wiener フィルターの基礎的議論を例を用いて簡潔に説明した。公的統計の季節統計に関係する論点について、詳しい議論は別の機会とするが、ここでは気が付いたコメントを幾つか述べておくことにとどめる。

(1) 古典的 Kolmogorov-Wiener 理論は数理的には厳密に展開されているが、季節調整法である Tramo-Seats 法の基礎となっているのは離散定常時系列の予測理論であるから Kolmogorov フィルターと呼ぶべきだろう。

(2) 時系列の予測理論は定常確率過程にもとづく数理的理論であるが、これを直接に応用しようとした当時の数学系の統計家は現実の経済データが有限、非定常であることに十分に注意していたとは言いがたい。数理的には定常確率過程の下では最適予測の解が求まるが、非定常過程におけるスペクトルについては2023年時点においても数理的に明快な解が知られているとは言えないと考えられる。すなわち数理的には未解決な課題がある。非定常性が顕著である現実の時系列に予測理論を適用しようとするには、季節調整の問題についての理解が十分とはいえず、現時点でも統計理論的にはなお大きなギャップがあると云える²⁹。関連する重要な統計的問題としては、定常時系列では無限個の確率変数が想定できるが、現実の観測データは有限(例えばGDPは100程度が限度)であるので理論と実際のギャップは小さくない。

(3) 季節調整法を開発した当時、1950年代には Kolmogorov-Wiener の Signal Extraction 法は(通信工学などの信号抽出問題)最新の理論であったこともあり、スペイン中央銀行やEU諸国の公的統計では季節調整法としてかなり利用されたと思われる。ヨーロッパではかなり影響のあったこの方向は米国センサス局で開発された移動平均に基づく X-11 とは根本的に異なる考え方であると考えられる。

²⁹例えば Bell (1984) は初期値問題について検討している。非定常時系列の場合には定常過程とは異なり初期値問題の処理は今でも重要な課題である。TRAMO-SEATS では季節 ARIMA モデルという範囲内で解を求めている。その解が実務にも有用ということはある種の可能性を示唆していると言えるかもしれない。

(4) 工学分野においてはその後、1960年代になり制御工学を中心としてKalmanフィルターの議論を発端に非定常でもfiltering法が可能なが理解され、実装されることとなった。この研究動向は統計科学にも大きな影響があり、1970年代末ごろから有限個のデータを扱う統計的時系列解析においてKalmanフィルターによる分析方法が取り入れられた。例えば1990年代までにはBrockwell-Davis (1990)などの大学院・教科書などにも取り入れられるようになっていた。学問的な研究動向としては、非ガウスFilterなどで知られていたG. Kitagawa(北川源四郎)はMonte Carlo Filterを提唱(J. Comp. & Graph. Stat. (1996))、国際的に研究者間で普及、今では一般に粒子フィルター(Particle Filter)と呼ばれるようになっていた³⁰。

(5) 現時点(2023年1月)で評価すると、定常・非定常、Gaussian・non-Gaussian、線形・非線形に関わらず状態空間表現によるデータ分析が季節調整モデルを含め統計的時系列分析では一般的かつ広汎に応用可能となっている。しかしながら、これまでに開発された時間領域(time domain)における統計的モデリングと周波数領域(frequency domain)における統計的モデリングの関係は十分に解明されているとは言えない。例えば経済マクロ時系列の場合には季節周波数での影響が顕著にみられる場合が多いが、同時にトレンド・循環成分も存在すると考えられるので、数理的には非定常性と定常性が混在していると見ることができる。また自然科学や工学データと異なり利用可能なデータ数はそれほど多くないことが一般的である³¹。非定常時系列の周波数領域分析の試みとしては例えば第一章で言及したKunitomo-Sato(2021), Sato-Kunitomo(2021a,b)などがある。時系列のデータ数がそれほど大きくない場合にも適用可能なX12SIML(佐藤 2022)は周波数分析に基づく季節調整プログラムとなっている。

(6) 1990年代に開発されたセンサス局法X-12-ARIMAは以前からヨーロッパで開発されているTRAMO-SEATSを取り入れて2020年頃からX-13として公表、メンテナンスされるようになっていた。X-

³⁰フィルタリング・アプローチについては例えば北川(2020)の説明を参照されたい。

³¹データが仮に長期に得られる場合でもその間、経済構造が非常に長期にわたって一定な変動を繰り返す、と見なすことは無理がある場合が多い。さらにデータの定義も定期的に変更される場合が多い。

12-ARIMA のメンテナンスは既に終了しているので、これまでの慣行のもとで季節調整を行い続けることは困難であり、今後、日本の公的統計において季節調整を扱うに当たっては、X-13ARIMA-SEATS の内容を正しく理解し、プログラム上で選択する時系列モデリング上で必要となるオプションの意味と結果について理解する必要があるだろう。

したがって、離散時間の時系列データに対する時間領域と周波数領域の統計的時系列分析の基礎、すなわち本稿の内容の周辺を理解することも季節調整の実務への応用などにおいても意味があると思われる。

参考文献

Anderson, T.W. (1971), *The Statistical Analysis of Time Series*, John-Wiley.

Bell, W. (1984), "Signal Extraction for Non-stationary Time Series," *The Annals of Statistics*, 12-2,.

Brockwell, P. and Davis, R. (1990), "Time Series : Theory and Methods," Springer.

Hannan, E. J. (1971) "Multiple Time Series" Wiley.

藤田宏・黒田成俊・伊藤清三(1991)「関数解析」岩波書店.

北川源四郎(2020)「Rによる時系列解析入門」, 岩波書店.

高岡慎(2015)「経済時系列と季節調整法」朝倉書店.

付論

この付論では定常性とスペクトル密度関数の定義を述べておく。(なお記号は Kitagawa (2020) と整合的であるが、Anderson (1971), Brockwell-Davis(1990) などの定義とほんの少し異なる。)

時間間隔 $t_i - t_{i-1} = 1$ ($i = 1, 2, \dots$) として確率過程 X_t ($t = 0, 1, \dots$) を考察する。定常性が満たされていれば X_t ($t = 0, -1, \dots$) として現時点 $t = 0$ より過去に遡る確率変数列を考えることができる。離散時間の確率変数列に対して期待値と分散が存在する場合には $\mathbf{E}[X_t] = \mu_t$ および $\mathbf{Cov}[X_t, X_s] = \mathbf{E}[(X_t - \mu_t)(X_s - \mu_s)]$ とすると、定常性を次のように定義される。

定義 1: (i) 確率過程 X_t が定常過程 (stationary process, あるいは強定常過程) であるとは、任意の (t_1, \dots, t_k) に対して $(X_{t_1}, \dots, X_{t_k})$ の確率分布と $(X_{t_1+h}, \dots, X_{t_k+h})$ の確率分布が等しくなる。

(ii) 確率過程 X_t が (弱) 定常過程 (weakly stationary, あるいは 2 次定常過程) であるとは、任意の t, s に対して $\mathbf{E}[X_t] = \mathbf{E}[X_s] = \mu$, 分散は有限で共分散関数 (autocovariance) は $\mathbf{Cov}[X_t, X_s] = C(|t - s|)$ で与えられる。

例 1: 実数列 λ_j ($j = 1, \dots, m$) に対して

$X_t = \sum_{j=1}^m [A_j \cos 2\pi \lambda_j t + B_j \sin 2\pi \lambda_j t]$ を考える。このとき共分散関数は σ_j^2 を適当に選んで $\mathbf{E}[X_t X_s] = \sum_{j=1}^m \sigma_j^2 \cos 2\pi \lambda_j (t - s)$ と表現される。

例 2: (AR(p) モデル, MA(q) モデル, ARMA(p,q) モデル) 確率過程 X_t が確率的定差方程式 $X_t = \sum_{j=1}^p \beta_j X_{t-j} + v_t - \sum_{j=1}^q \theta_j v_{t-j}$ を満たすとき ARMA(p,q) モデルと呼ぶ。ここで $\{v_t\}$ は期待値ゼロ, 分散一定の i.i.d. 系列とする。(Causal モデルに限定すれば) 定常性を満足する一つの十分条件は特性方程式 $\lambda^p - \sum_{j=1}^p \beta_j \lambda^{p-j} = 0$ の根の絶対値が 1 より小さいことである。

自己共分散関数 $\sigma(t)$ が与えられたときにそのフーリエ変換はスペクトル密度関数と呼ばれている。スペクトル密度関数は常に存在する

とは限らないが

$$f(\lambda) = C(0) + 2 \sum_{h=1}^{\infty} C(h) \cos 2\pi\lambda h = \sum_{h=-\infty}^{\infty} C(h) e^{-i2\pi\lambda h}$$

により定める³²。このとき

$$(37) \quad \sigma(h) = \int_{-1/2}^{1/2} \cos 2\pi\lambda h f(\lambda) d\lambda = \int_{-1/2}^{1/2} e^{i2\pi\lambda h} f(\lambda) d\lambda$$

となる。

特に ARMA(p,q) モデルのスペクトル密度関数は

$$(38) \quad f(\lambda) = \sigma^2 \frac{\|1 - \sum_{j=1}^q \theta_j e^{i2\pi j\lambda}\|^2}{\|1 - \sum_{j=1}^p \beta_j e^{i2\pi j\lambda}\|^2}$$

で与えられる。(σ²は v_t の分散とする。)

³²ここで $i^2 = -1, C(-h) = C(h)$ より $e^{i\lambda h} = \cos \lambda h + i \sin \lambda h$ の $\sin \lambda h$ 項はなくても良いが、複素数の表現により計算が簡便化されることが多い。

(ii) X-13ARIMA-SEATS (高岡慎, 琉球大学国際地域創造学部)

1. はじめに

本稿は、米国商務省センサス局によって開発された季節調整プログラムである X-13ARIMA-SEATS の概要をまとめたものであり、統計数理研究所における統計エキスパート養成プログラムの一貫として 2022 年 5 月 25 日・6 月 1 日・6 月 22 日に実施された季節調整に関する講義の内容に基づいている。

現在日本国内の主要な公的統計の季節調整は、ほぼ X-12-ARIMA によって実施されているが、X-13ARIMA-SEATS は X-12-ARIMA の後継ソフトウェアの最新版である。センサス局は 2015 年にはウェブサイトでの X-12-ARIMA の配布を停止し、サポートを打ち切っている。このため、今後官公庁などで新たに公的統計の季節調整を算出する必要がある場合は、基本的に X-12-ARIMA ではなく X-13ARIMA-SEATS を使用する必要がある。本稿は、上記のような事情を踏まえ、主に官公庁で統計関連業務に従事する担当者の実務に資することを期して書かれている。

X-13ARIMA-SEATS は X-12-ARIMA の後継と位置づけることができるが、X-12-ARIMA の処理を発展ないし改善したというよりは、スペイン銀行によって開発された季節調整プログラム TRAMO-SEATS と統合させ、機能を拡張したものとなっている。このように、独立に開発されたプログラムをやや強引に統合しているため、X-13ARIMA-SEATS では両プログラムの関係性や機能が一目分かりにくくなっている。

本稿では、両プログラムの基本的な機能について数理的な側面にやや力点をおいて順に概観し、関係性について解説する。

2. X-12-ARIMA の概要

X-12-ARIMA はセンサス局によって開発された季節調整プログラムであり、1998 年に正式版が公開されている (Findley, et al.(1998))。

X-12-ARIMA は、X-11 プログラムの機能を大幅に拡張し、内部の

時系列モデルによる予測、事後診断プロシージャなどを実装したソフトウェアで、日本では現在に至るまで主要な公的統計の季節調整で標準的な方法として利用されている。

本節では、X-12-ARIMA の内部のフィルタである X-11 フィルタや、RegARIMA モデルの機能を含めた全体的なプログラムの構成について概観する。

2.1 X-11 フィルタによる季節調整

X-11 フィルタは、複数の移動平均フィルタの連続的な適用によって、時系列をトレンド、季節性、不規則変動といった複数の成分に分解するための一連の手続きである。

使用される移動平均フィルタは、トレンド抽出ためのフィルタと季節変動を平滑化するためのフィルタに分類される。月次系列に対応するフィルタの例をいくつか挙げておく。ここで B はバックシフトオペレータを表す。

トレンド抽出ためのフィルタ

- 2×12 移動平均フィルタ

$$\begin{aligned}u(B) &= \frac{1}{24} (B^{-6} + B^{-5}) (1 + B + B^2 + \dots + B^{11}) \\ &= \frac{1}{24} (B^{-6} + 2B^{-5} + 2B^{-4} + \dots + 2 + \dots + 2B^5 + B^6)\end{aligned}$$

2×12 移動平均フィルタは、月次の場合 13ヶ月に渡る中心化移動平均を表し、季節性を均す作用を持つ。

- Henderson 移動平均フィルタ

Henderson 移動平均フィルタはユーザーが自由に項数を設定できるフィルタで、トレンド抽出に利用される。フィルタの構成はやや詳しく後述する。

季節変動を平滑化するためのフィルタ

- 3×3 移動平均フィルタ

$$a_1(B) = \frac{1}{9} (B^{-12} + 1 + B^{12}) (B^{-12} + 1 + B^{12})$$

これらのフィルタは、複数年にわたって同月の値の加重平均を求めるフィルタであり、季節変動のばらつきを均す作用を持つ。

- 3×5 移動平均フィルタ

$$a_2(B) = \frac{1}{15} (B^{-12} + 1 + B^{12}) (B^{-24} + B^{-12} + 1 + B^{12} + B^{24})$$

Henderson 移動平均フィルタの概要

Henderson 移動平均フィルタは単純な中心化移動平均とは異なり、ある種の条件を最も良く満たす満たすウエイトとして、最適化計算から導出される。

まず、観測系列 $\{Y_t\}$ が

$$\begin{cases} Y_t = X_t + \epsilon_t, & \epsilon_t \sim N(0, \sigma^2) \\ X_t = \alpha_0 + \alpha_1 t + \alpha_2 t^2 \end{cases}$$

という構造を持つと仮定し、 Y_t から X_t を推定するフィルタ

$$\hat{X}_t = \sum_{j=-n}^n W_j Y_{t+j}$$

を構成することを考える。Henderson 移動平均フィルタは、ウエイトを

$$\min_{w_{-n}, \dots, w_n} E \left[\{(1-B)^3 \hat{X}_t\}^2 \right] \quad s.t. \quad \sum_{k=-n}^n W_k X_{t+k} = X_t$$

という最適化問題から導出する。ここで制約条件 $\sum_{j=-n}^n W_j X_{t+j} = X_t$ は、このフィルタが X_t に対しては影響を与えず、ノイズ ϵ_t に対してのみ作用すること表している。

目的関数を変形すると

$$\begin{aligned}
E \left[\{(1-B)^3 \hat{X}_t\}^2 \right] &= E \left[\left\{ (1-B)^3 \sum_{k=-n}^n W_k (X_{t+k} + \epsilon_{t+k}) \right\}^2 \right] \\
&= E \left[\left\{ (1-B)^3 \left(X_t + \sum_{k=-n}^n W_k \epsilon_{t+k} \right) \right\}^2 \right] \\
&= E \left[\left\{ (1-B)^3 \sum_{k=-n}^n W_k \epsilon_{t+k} \right\}^2 \right] \\
&= E \left[\left\{ \sum_{k=-n}^{n+3} (W_k - 3W_{k-1} + 3W_{k-2} - W_{k-3}) \epsilon_{t+k} \right\}^2 \right] \\
&= \sigma^2 \sum_{k=-n}^{n+3} (W_k - 3W_{k-1} + 3W_{k-2} - W_{k-3})^2
\end{aligned}$$

X_t が2次曲線なので $(1-B)^3 X_t = 0$ となることに注意しておく。

次に制約条件を変形すると

$$\begin{aligned}
\sum_{k=-n}^n W_k X_{t+k} &= \sum_{k=-n}^n W_k \{ \alpha_0 + \alpha_1(t+k) + \alpha_2(t+k)^2 \} \\
&= \alpha_0 \sum_{k=-n}^n W_k \\
&\quad + \alpha_1 t \left(\sum_{k=-n}^n W_k + \sum_{k=-n}^n kW_k \right) \\
&\quad + \alpha_2 \left(t^2 \sum_{k=-n}^n W_k + 2t \sum_{k=-n}^n kW_k + \sum_{k=-n}^n k^2 W_k \right) \\
&= \alpha_0 + \alpha_1 t + \alpha_2 t^2
\end{aligned}$$

となるが、この条件が特定の $\alpha_0, \alpha_1, \alpha_2$ によらず常に成り立つためには、辺々の比較により

$$\sum_{k=-n}^n W_k = 1, \quad \sum_{k=-n}^n kW_k = 0, \quad \sum_{k=-n}^n k^2 W_k = 0$$

であることが必要となる。以上を踏まえて

$$D = \begin{pmatrix} 1 & & & & & & & & & \\ -3 & 1 & & & & & & & & \\ & 3 & -3 & 1 & & & & & & \\ -1 & & 3 & -3 & 1 & & & & & \\ & & -1 & & 3 & -3 & \cdots & & & \\ & & & -1 & & 3 & \cdots & 1 & & \\ & & & & -1 & \cdots & -3 & & & \\ & & & & & \cdots & & 3 & & \\ & & & & & & & & -1 & \end{pmatrix}, \quad \mathbf{w} = \begin{pmatrix} W_{-n} \\ W_{-(n-1)} \\ \vdots \\ W_0 \\ \vdots \\ W_{n-1} \\ W_n \end{pmatrix}$$

$$M = \begin{pmatrix} 1 & 1 & \cdots & 1 & 1 \\ -n & -(n-1) & \cdots & n-1 & n \\ n^2 & (n-1)^2 & \cdots & (n-1)^2 & n^2 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

のように行列を定義すると、最初の条件付き最適化問題は

$$\min_{\mathbf{w}} \sigma^2 \mathbf{w}' D' D \mathbf{w} \quad s.t. \quad M \mathbf{w} = \mathbf{b}$$

というベクトル \mathbf{w} を求める条件付き最適化問題に書き直すことができる。この問題をラグランジュ未定乗数法を用いて解くと、

$$\mathbf{w} = (D'D)^{-1} M' \{M(D'D)^{-1} M'\}^{-1} \mathbf{b}$$

が得られる。この結果を用いて、いくつかの n に対するウェイトを具体的に求めると表 1 のようになる。

n についてはデフォルトの値は設定されているが、ユーザーが自由に設定することもできる。

2.2 X-11 による時系列の分解

X-11 はここまでに説明した各種のフィルタを用いて次のような手順で時系列の分解を行う。

- (1) $\hat{N}_t^{(1)} = (1 - u(B))Y_t$ によってトレンドを除いた系列 $\hat{N}_t^{(1)}$ を得る。

表 1: Henderson 移動平均のウエイト

	$n = 2$	$n = 3$	$n = 4$	$n = 6$	$n = 7$	$n = 11$
-11						-0.0043
-10						-0.0109
-9						-0.0157
-8						-0.0145
-7					-0.0137	-0.0049
-6				-0.0193	-0.0245	0.0134
-5				-0.0279	-0.0141	0.0389
-4			-0.0407	0.0000	0.0240	0.0683
-3		-0.0587	-0.0099	0.0655	0.0829	0.0974
-2	-0.0734	0.0587	0.1185	0.1474	0.1459	0.1219
-1	0.2937	0.2937	0.2666	0.2143	0.1937	0.1383
0	0.5594	0.4126	0.3311	0.2401	0.2115	0.1441
1	0.2937	0.2937	0.2666	0.2143	0.1937	0.1383
2	-0.0734	0.0587	0.1185	0.1474	0.1459	0.1219
3		-0.0587	-0.0099	0.0655	0.0829	0.0974
4			-0.0407	0.0000	0.0240	0.0683
5				-0.0279	-0.0141	0.0389
6				-0.0193	-0.0245	0.0134
7					-0.0137	-0.0049
8						-0.0145
9						-0.0157
10						-0.0109
11						-0.0043

- (2) $\widehat{S}_t^{(1)} = a_1(B)\widehat{N}_t^{(1)}$ により暫定的な季節成分 $\widehat{S}_t^{(1)}$ を得る.
- (3) $\widehat{A}_t^{(1)} = Y_t - \widehat{S}_t^{(1)}$ によって暫定的な季節調整済系列 $\widehat{A}_t^{(1)}$ を得る.
- (4) $\widehat{N}_t^{(2)} = Y_t - H(B)\widehat{A}_t^{(1)}$ によって2段階目のトレンド除去系列 $\widehat{N}_t^{(2)}$ を得る.
- (5) $\widehat{S}_t^{(2)} = a_2(B)\widehat{N}_t^{(2)}$ によって2段階目の季節成分 $\widehat{S}_t^{(2)}$ を得る.
- (6) $\widehat{A}_t^{(2)} = Y_t - \widehat{S}_t^{(2)}$ によって2段階目の季節調整済系列 $\widehat{A}_t^{(2)}$ を得る.
- (7) 最終的な季節調整済系列 \widehat{A}_t 、トレンド \widehat{T}_t 、季節成分 \widehat{S}_t 、不規則成分 \widehat{I}_t を、

$$\begin{aligned}\widehat{A}_t &= \widehat{A}_t^{(2)} \\ \widehat{T}_t &= H(B)\widehat{A}_t^{(2)} \\ \widehat{S}_t &= \widehat{S}_t^{(2)} \\ \widehat{I}_t &= \widehat{A}_t^{(2)} - \widehat{T}_t\end{aligned}$$

とする。

以上より季節性、トレンドを抽出するフィルタをそれぞれ $W_S(B)$ 、 $W_T(B)$ とすると、

$$\begin{aligned}W_S(B) &= a_2(B) \left[1 - H(B) \left\{ 1 - a_1(B)(1 - u(B)) \right\} \right] \\ W_T(B) &= H(B)(1 - W_S(B))\end{aligned}$$

と書くことができる。また季節調整フィルタ $W_A(B)$ は

$$W_A(B) = 1 - W_S(B)$$

となる。

フィルタの端点の問題

ここまでに現れたフィルタはいずれも左右対称であるが、実際のデータに適用する場合、端点付近ではそのままの形では適用ができないという問題がある。これに対して、X-11では**Musgrave**移動平均と呼ばれる非対称フィルタを用いる仕組みが用意されている。

$\{W_k\}$ を対称なウエイト、 $\{V_k\}$ を非対称なウエイトとし、原系列 $\{Y_t\}$ が

$$Y_t = \alpha + \beta t + \epsilon_t, \quad \epsilon_t \sim NID(0, \sigma^2)$$

であると仮定した場合に、

$$E \left[\left(\sum_{j=-m}^m W_j Y_{t+j} - \sum_{j=-m}^{m-d} V_{d,j} Y_{t+j} \right)^2 \right]$$

を最小化する $\{V_{d,j}\}$ が Musgrave 移動平均のウエイトとなる。

これは、対称移動平均を非対称な移動平均で近似するアイデアの一つで、Henderson 移動平均と同様に、最適化問題の解として導出される。

2.3 RegARIMA モデル

X-12-ARIMA では、調整の安定性を向上させるために、RegARIMA モデルと呼ばれる時系列モデルを利用し、様々な事前調整を実施した上で X-11 フィルタを適用している。

RegARIMA モデルは季節 ARIMA モデルを回帰変数を含む形に拡張したモデルで、一般に

$$\phi_p(B)\Phi_P(B^s)(1-B)^d(1-B^s)^D \left(y_t - \sum_{j=1}^r \beta_j z_{jt} \right) = \theta_q(B)\Theta_Q(B^s)u_t$$

のように表される。回帰変数には外れ値、レベルシフト、曜日効果、休日効果などに対応する変数が用いられる。推定された季節 ARIMA

部分は予測値の算出に利用され、予測値を外挿することによって延長された原系列がX-11フィルタで処理される。

RegARIMA モデルの推定

RegARIMA モデルの推定は次のような手順によって行われる。

1. 階差次数と ARMA 次数の設定
2. 回帰係数の推定
 - ARMA 部分のパラメータには適当な初期値を設定
⇒2回目以降の反復では前回の推定結果を利用
 - 一般化最小二乗法による回帰係数の推定
 - 誤差項が自己相関を持つ場合の回帰モデル
3. ARMA 部分の推定
 - 回帰係数を1で得られた値に固定
 - 回帰モデルで表される変動を差し引いた残差系列に対して ARMA モデルを適用
 - 最尤法による ARMA 部分のパラメータの推定
4. 2と3のプロセスを収束するか繰り返し回数の上限に達するまで反復

2.4 X-12-ARIMA の構成

X-12-ARIMA は、以上の処理を図1のような流れにしたがって実行する。

3. TRAMO-SEATS の概要

3.1 TRAMO-SEATS(スペイン銀行) の構成

TRAMO-SEATS はスペイン銀行(Gómez and Maravall (1997))により開発された季節調整プログラムである。TRAMO-SEATS の季節調整はX-12-ARIMA と同様に事前調整パートと季節調整パートに分かれており、事前調整パートを担うプロシージャがTRAMO("Time

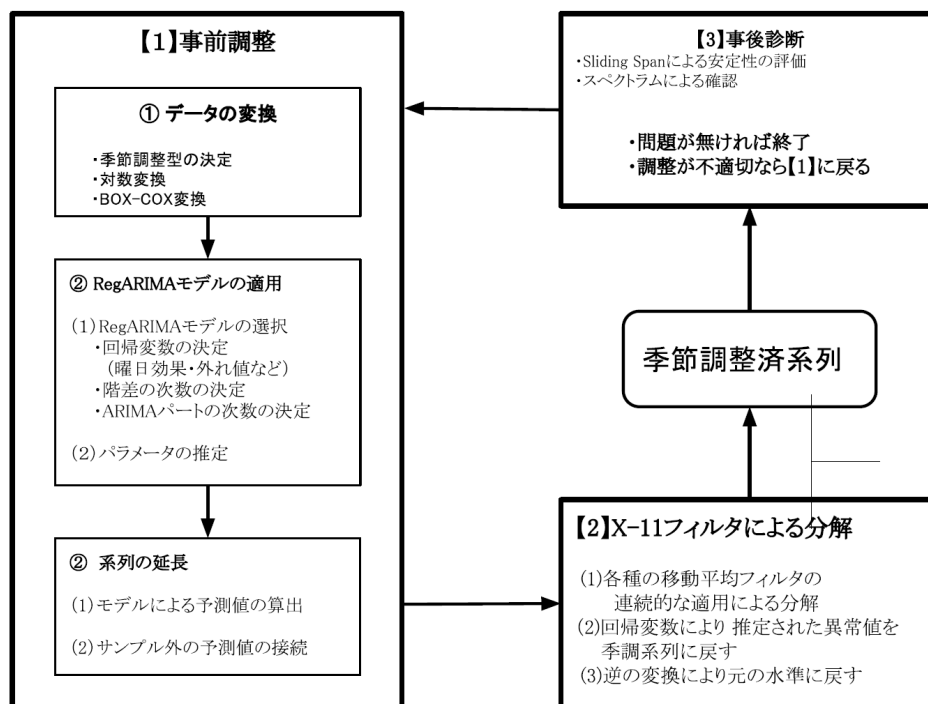


図 1: X-12-ARIMA の処理

series Regression with ARIMA noise, Missing Observation and Outliers”)、季節調整パートのプロシージャがSEATS(”Signal Extraction in ARIMA Time Series”) と名付けられている。

TRAMO-SEATS では、まず TRAMO によってデータに最も適合する RegARIMA モデルが選択される。このプロセスは、外れ値、欠損値の処理や前方予測および後方予測の追加による時系列の延長などの処理を含んでいる。

次に SEATS 部分の処理を通して、選択された RegARIMA モデルと統合的なフィルタが構成される。ここでは AMB 分解 (ARIMA-model Based Decomposition) による時系列の分解と、それに基づく WK フィルタ (Wiener-Kolmogorov Filter) によって季節調整が実行される。

このように、2つのパートによる事前調整と季節調整という処理の流れは X-12-ARIMA に類似しているものの、数学的背景にはかなりの相違が存在する。

3.2 WK フィルタ

WK フィルタは、観測系列 Y_t が

$$Y_t = S_t(\text{シグナル}) + N_t(\text{ノイズ})$$

のように分解できると仮定した場合に、 Y_t から S_t を推定する方法の一つである。WK フィルタは、 Y_t に移動平均を適用して S_t を推定する場合に、平均2乗誤差を最小化するウェイトとして導出される。このようなフィルタは、 Y_t 、 S_t 、 N_t の自己共分散関数が分かっているならば、一定の条件のもとで導出することができる。ここでは行列表現による定式化を示す。

n 期間分の観測データが利用できるとき

$$\mathbf{Y} = (Y_1, \dots, Y_n)', \quad \mathbf{S} = (S_1, \dots, S_n)', \quad \mathbf{N} = (N_1, \dots, N_n)'$$

とすると、それぞれのベクトルの共分散行列は

$$\Sigma_{yy} = E[\mathbf{Y}\mathbf{Y}'], \quad \Sigma_{ss} = E[\mathbf{S}\mathbf{S}'], \quad \Sigma_{nn} = E[\mathbf{N}\mathbf{N}']$$

と表すことができる。 S_t と N_t が独立ならば

$$\Sigma_{yy} = \Sigma_{ss} + \Sigma_{nn}$$

が成り立つ。 $(n \times n)$ 行列 β を用いて \mathbf{S} の推定量を $\hat{\mathbf{S}} = \beta\mathbf{Y}$ と表現すると、平均2乗誤差

$$E[(\mathbf{S} - \hat{\mathbf{S}})'(\mathbf{S} - \hat{\mathbf{S}})] = E[(\mathbf{S} - \beta\mathbf{Y})'(\mathbf{S} - \beta\mathbf{Y})]$$

を最小にする β は、回帰分析と同様の計算により

$$\beta = \Sigma_{sy}\Sigma_{yy}^{-1} = \Sigma_{ss}\Sigma_{yy}^{-1}$$

となる ($\Sigma_{sy} = E[\mathbf{S}\mathbf{Y}'] = E[\mathbf{S}(\mathbf{S} + \mathbf{N})'] = \Sigma_{ss}$ を利用)。この β の各行ベクトルが移動平均のウェイトになっている。

3.3 AMB 分解 (ARIMA Model Based Decomposition)

定常過程の和として表される系列のモデルは、和の定理 (付録参照) に基づいて導出することができるが、ここでの問題は RegARIMA

モデルによって表される時系列がどのような系列の和になっているかを考える逆の問題になっている。このような問題は一般には一意に解くことはできないため、SEATSでは適当な追加的条件を課した上で、RegARIMAモデルを複数の成分に分解している。このような処理はAMB分解と呼ばれている。

SEATSのAMB分解では

- 各成分は全て互いに独立
- パラメータ数はできるだけ抑える（ケチの原理）
- トレンドと季節性は滑らかに変動する

といった制約条件が適用される。こうした条件にはやや恣意的な側面があり、必ずしも説得的ではないが、SEATSでは実用性の観点からこうした条件を用いた処理が実装されている。一連の分解の過程はかなり煩雑なので、簡単なケースについての例を以下に示しておく。

AMB分解の例

観測系列 Y_t が

$$Y_t = T_t + S_t + I_t$$

であり、観測データから Y_t のモデルが

$$(1 - B^2)Y_t = \epsilon_t, \quad \epsilon_t \sim N(0, \sigma^2)$$

であると特定されたとする。このとき $(1 - B^2)$ は $(1 - B)(1 + B)$ と因数分解できるが、 $(1 - B)$ と $(1 + B)$ はそれぞれトレンドと季節変動に関連するオペレーターと解釈することができる。

まず、成分 T_t が

$$(1 - B)T_t = T_t - T_{t-1} = u_t$$

のように表されると仮定すると

$$T_t = T_{t-1} + u_t$$

となるので、 $(1 - B)$ によって特徴づけられる成分 T_t は、1期前の状態に強く影響される変動となりトレンド的変動とみなすことができる。

他方、成分 S_t が

$$(1 + B)S_t = S_t + S_{t-1} = v_t$$

のように表されると仮定すると、 S_t は連続する2期間を合計すると平均的にはゼロになるので、 $(1 + B)$ によって特徴づけられる成分 S_t は、周期を2とする季節成分であると解釈できる。

以上を踏まえると、分解モデルの候補としては、まず

$$\begin{aligned} Y_t &= T_t + S_t + I_t \\ &= \frac{u_t}{1 - B} + \frac{v_t}{1 + B} + I_t \end{aligned}$$

が考えられる。 u_t 、 v_t がそれぞれ $MA(1)$ に従うと仮定すると、さらに

$$Y_t = \frac{1 - \alpha B}{1 - B} \epsilon_t^T + \frac{1 - \beta B}{1 + B} \epsilon_t^S + I_t$$

となる。ここで

$$\epsilon_t^T \sim WN(0, \sigma_T^2), \quad \epsilon_t^S \sim WN(0, \sigma_S^2), \quad I_t \sim WN(0, \sigma_I^2),$$

とし、これらは互いに独立と仮定しておく。

Y_t の (擬似) 自己共分散母関数 $G_{yy}(z)$ は、右辺の各成分が互いに独立であることから

$$\begin{aligned} G_{yy}(z) &= \frac{\sigma^2}{(1 - z^2)(1 - z^{-2})} \\ &= \frac{(1 - \alpha z)(1 - \alpha z^{-1})}{(1 - z)(1 - z^{-1})} \sigma_T^2 + \frac{(1 - \beta z)(1 - \beta z^{-1})}{(1 + z)(1 + z^{-1})} \sigma_S^2 + \sigma_I^2 \end{aligned}$$

となる。自己共分散母関数については付録を参照されたい。これより

$$\begin{aligned} \sigma^2 &= (1 - \alpha z)(1 - \alpha z^{-1})(1 + z)(1 + z^{-1}) \sigma_T^2 \\ &\quad + (1 - \beta z)(1 - \beta z^{-1})(1 - z)(1 - z^{-1}) \sigma_S^2 \\ &\quad + (1 - z^2)(1 - z^{-2}) \sigma_I^2 \end{aligned}$$

が得られる。

z^j ごとに係数をまとめ、等式が成立するための条件を求めると、

$$\begin{aligned}\sigma^2 &= 2\sigma_I^2 + 2(\alpha^2 - \alpha + 1)\sigma_T^2 + 2(\beta^2 - \beta + 1)\sigma_S^2 \\ (1 - \alpha)^2\sigma_T^2 &= (1 + \beta)^2\sigma_S^2 \\ \beta\sigma_S^2 &= \sigma_I^2 + \alpha\sigma_T^2\end{aligned}$$

が得られる。これらを σ_S^2 、 σ_T^2 、 σ_I^2 について解くと

$$\begin{aligned}\sigma_S^2 &= \frac{\sigma^2}{4(1 + \beta)^2} \\ \sigma_T^2 &= \frac{\sigma^2}{4(1 - \alpha)^2} \\ \sigma_I^2 &= \left(\frac{\beta}{4(1 + \beta)^2} - \frac{\alpha}{4(1 - \alpha)^2} \right) \sigma^2\end{aligned}$$

という関係式が得られる。これらの式は未知数が5つで3本の式からなる連立方程式なので、適当な制約条件がなければ解くことができないため、

- σ_S^2 、 σ_T^2 、 σ_I^2 がいずれも正
- MA 部分に関する条件 $|\alpha| \leq 1$ 、 $|\beta| \leq 1$ の下で σ_I^2 が最大

という制約を課すと、 $(\alpha, \beta) = (-1, 1)$ という解が得られる。このとき、 Y_t の自己共分散母関数は

$$G_{yy}(z) = \frac{(1+z)(1+z^{-1})}{(1-z)(1-z^{-1})}\sigma_T^2 + \frac{(1-z)(1-z^{-1})}{(1+z)(1+z^{-1})}\sigma_S^2 + \sigma_I^2$$

となり、 Y_t は

$$\begin{aligned}Y_t &= T_t + S_t + I_t \\ (1 - B)T_t &= (1 + B)\epsilon_t^T \\ (1 + B)S_t &= (1 - B)\epsilon_t^S \\ I_t &= \epsilon_t^I\end{aligned}$$

と分解される。この結果、トレンド成分と季節成分のモデルが特定され、WKフィルタを適用することができる。

AMB 分解の処理の流れ

例に示したプロセスをまとめると、

1. 観測系列 Y_t の AR 部分の B の多項式を、トレンドに関する部分と季節変動に関連する部分に分離し、それを基に T_t 、 S_t のモデルを設定
2. Y_t の AGF と $T_t + S_t + I_t$ の AGF の比較から、モデルに含まれるパラメータが満たすべき条件を導出
3. 必要に応じてモデルのパラメータを削減するなどの処理を行い、連立方程式を解く
4. 不規則成分の分散 σ_I^2 が最大になるように調整

となる。不規則成分の分散 σ_I^2 が最大になるようにするという調整により、トレンドおよび季節性が最も滑らかになる。

4. X-13ARIMA-SEATS の構成

X-13ARIMA-SEATS では、これまでに見た各処理を図 2 のように統合している。

モデルの自動選択については PICKMDL と AUTOMDL という 2 種類のスペックが用意されている。PICKMDL は X-11-ARIMA で導入されたモデル選択プロシージャであり、モデルの予測誤差に関する規準などの独自の複数のを用いて自動的にモデル選択が実行される。AUTOMDL は SEATS で実装されているモデル選択法を X-13 のスペックの一つとして統合したものである。AUTOMDL では、単位根検定や AIC などの情報量規準を段階的に用いてモデルの自動選択が実行される。

実際の季節調整に関しては、X11 と SEATS という 2 種類のスペックが用意されている。X11 スペックは中心化移動平均や Henderson 移動平均などを用いた X-11 法の成分分解を行う場合に使用する。X11 スペックを用いる場合には、RegARIMA モデルは回帰部分の分離と前方および後方の予測にのみ使用され、移動平均のウエイトの設定とは関係しない。他方、SEATS スペックを利用した場合は、選択

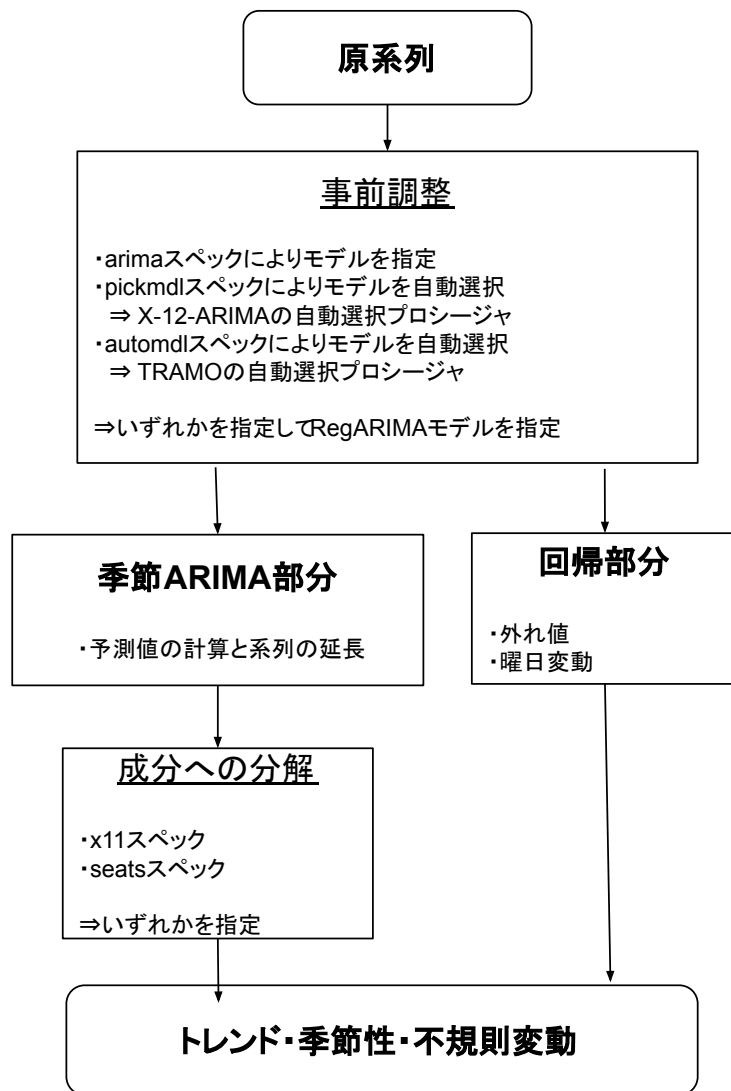


図 2: X-13ARIMA-SEATS の処理

された RegARIMA モデルに基づく AMB 分解と WK フィルタによる分解が適用され、RegARIMA モデルの選択がフィルタの構成にも影響する。

X-13ARIMA-SEATS では、季節調整パートで X11 スペックを選択すれば、従来の X-12-ARIMA と同一の処理を実行することができるため、X-12-ARIMA は X-13ARIMA-SEATS のサブセットとして包含されている。

5 まとめ

本稿では、X-13ARIMA-SEATS の構造と細かいスペックの機能の差異について、主に数理的側面に着目して概観した。以下では実務的な問題について若干のコメントを付してまとめとする。

2023 年時点における日本国内の官公庁の季節調整では、データに合わせた適当なスペックファイルを設定し、X-12-ARIMA によって季節調整を行っているケースが大半であると思われるが、X-12-ARIMA は X-13ARIMA-SEATS に包含されるプログラムであるため、X-12-ARIMA で使用可能なスペックファイルは、X-13ARIMA-SEATS でも基本的にはそのまま実行可能となっている³³。公的統計の処理のプロセスを性急に変更することは必ずしも好ましいとは限らないが、X-12-ARIMA のサポートが打ち切られていることや、X-13ARIMA-SEATS の開発・修正が続けられていることを考慮すると、徐々に X-13ARIMA-SEATS への切り替えを進めることが望ましいと思われる。

付論 1：自己共分散関数

定常過程 X_t の自己共分散関数 $\gamma_{xx}(\tau)$ に対して、自己共分散母関数は

$$G_{xx}(z) = \sum_{j=-\infty}^{\infty} \gamma_{xx}(j) z^j$$

と定義される。 z は複素数。

³³筆者が確認した範囲では、X-12-ARIMA 用のスペックが X-13ARIMA-SEATS でエラーになるケースが一部見られたが、マイナーバージョンアップによって解消されるなど、地道な修正が継続されていると思われる。また、問題なく計算が実行できるケースでは、同一のスペックファイルによる X-12-ARIMA と X-13ARIMA-SEATS の計算結果に差異は見られなかった。

時系列 X_t がラグ作用素 B の多項式 $\theta(B)$ によって

$$X_t = \theta(B)\epsilon_t, \quad \epsilon_t \sim WN(0, \sigma^2)$$

のように表現される場合（ウォルド表現）、 X_t の自己共分散母関数は

$$G_{xx}(z) = \sigma^2 \theta(z^{-1})\theta(z)$$

と書けることが分かっている。

付論 2 : 和の定理

独立な定常過程を同時点で合計した場合、各系列の自己共分散関数または自己共分散母関数を利用して合計系列が従うモデルを導くことができ、「和の定理」として知られている。簡単な例を以下に示しておく。

和の定理の例（MA(m) と MA(n) の和）

$$X_t = \theta_x(B)u_t \quad u_t \sim WN(0, \sigma_u^2)$$

$$Y_t = \theta_y(B)v_t \quad v_t \sim WN(0, \sigma_v^2)$$

とする。ただし $\theta_x(B), \theta_y(B)$ はそれぞれ B に関する m 次および n 次の多項式で、 u_t と v_t は独立とする。このとき、

$$W_t = X_t + Y_t$$

となる W_t のモデルを考える。

X_t と Y_t が独立なので、自己共分散関数は

$$\gamma_{ww}(\tau) = \gamma_{xx}(\tau) + \gamma_{yy}(\tau)$$

であり、従って自己共分散母関数も

$$G_{ww}(z) = G_{xx}(z) + G_{yy}(z)$$

となる。

MA 過程の自己共分散母関数は、何らかの偶関数 ψ により

$$G_{xx}(z) = \psi_{xx}(m)z^{-m} + \cdots + \psi_{xx}(0) + \cdots + \psi_{xx}(m)z^m$$

$$G_{yy}(z) = \psi_{yy}(n)z^{-n} + \cdots + \psi_{yy}(0) + \cdots + \psi_{yy}(n)z^n$$

と書けるので、 $G_{ww}(z)$ は $q = \max\{m, n\}$ として

$$G_{xx}(z) = \psi_{ww}(q)z^{-q} + \cdots + \psi_{ww}(0) + \cdots + \psi_{ww}(q)z^q$$

となる。よって

$$MA(m) + MA(n) \sim MA(\max(m, n))$$

が得られる。

参考文献

Akaike, H. (1973), “Information Theory and an Extension of the Likelihood Principle,” in the *Second International Symposium on Information Theory*, eds. B.N. Petrov and F. Czaki, Budapest: Akademia Kiado, 267-287.

Findley, D.F., B.C. Monsell, W.R. Bell, M.C. Otto, B.C. Chen (1998), “New Capabilities and Methods of the X-12-ARIMA Seasonal Adjustment Program,” *Journal of Business and Economic Statistics*, 16, 127-176 (with Discussion).

Gómez, Victor and Agustin Maravall (1997), “PROGRAMS TRAMO AND SEATS, INSTRUCTIONS FOR THE USER,” Banco de España.

U.S. Census Bureau (2022), “X-13ARIMA-SEATS Reference Manual, version 1.1”.

高岡慎 (2015), 「経済時系列と季節調整法」, 朝倉書店.

第 III 部：労働力調査産業別系列の季節調整について（国友・佐藤）

1. 問題の所在

労働力調査は日本の公的統計において労働力を調査している基幹統計の一つである。マクロ経済全体にとり雇用、労働力、失業といった数値は重要な要素であることからその重要性については説明するまでもないだろう。統計局の労働力調査は毎月の雇用状況を把握する統計であり、全国的なサンプリングを実施して調査されている重要な公的統計の一つである。より正確には、層化二段抽出法により全国から約4万世帯をサンプリングして得られるデータであるが、労働力調査のサンプリングの実施方法や精度設計などについて詳しくは総務省統計局(2019)を参照されたい³⁴。

労働力調査では全国的に集計した上で日本全体の労働力および失業率を毎月公表しているが、日本のマクロ経済における労働市場における重要な指標となっている。全国的な労働力や失業率についてはX-12-ARIMAを利用して季節調整値を公表、日本経済の状況についての重要な指標となっている。他方、産業別就業者については様々な経緯があったようで、現在は季節調整値の公表は行っていない。一部のエコノミストからは産業別に雇用情勢を知りたいとの根強い要望がかねてからあることから、今回、特に産業別の就業者の動向の季節性と季節調整の必要性について検討することとした。本稿はその分析結果の中間報告である³⁵。

2. 産業別就業者数の分析

既に農林業の就業者数の月次系列については本報告書の他のところで簡単な結果を報告している。本報告では非農林業部門、製造業部門などを例に季節性と季節調整の有用性について考察する。(就業者数)=(農林業就業者)+(非農林業就業者)、であるから、ここでの

³⁴労働力統計の全国の主要な系列についてはX-12-ARIMAプログラム(U.S.Census Bureau)を利用して季節調整値を公表している。詳しくは総務省統計局のホームページを参照されたい。
<https://www.stat.go.jp/data/roudou/index.html>

³⁵ここでの検討したデータは2002年1月～2022年10月のデータである。労働力調査は1946年から実施されているが、調査法は変遷している。(詳しくは統計局のWeb情報を参照されたい。)今回のデータ分析では現行の調査方法に基づくと考えられる期間のみの時系列データを用いている。

主な問題は季節調整値をどこまで細分化したらよいか?というより一般的な季節調整の実務を巡る問題に関係している。

前章の図1として既に掲載したが、DECOMP(trend=2,ar=2)により推定された農林業における就業者数の季節性の大きさは原水準と比較して無視できない大きさである。これは就業者数の定義、データの取り方に大きく依存すると考えられるが、日本全体で見ると、寒冷地を中心に冬に就業できない地域がかなりあり、労働力調査における就業者の定義から、非就業者となるからである。こうした事情を無視して労働市場を分析することに大きな問題があることは明らかであろう。就業者数という労働市場の性格から当然のことではあるがいずれの場合も曜日効果や祝祭日の効果はかなり小さく推定されている。

なお、DECOMPで得られた季節調整の結果と比較するためにX-13ARIMA-SEATSの中のSEATS(auto)およびX-11(RegARIMAモデルを含むので正確にはX-12-ARIMAと呼ぶべきだろう)による原系列および季節調整値を図4と図5に示しておく³⁶。いずれの場合も数値は省略するがうるう年効果と曜日効果が有意に検出されたことは興味深い。日本の農林業の就業者数における曜日効果を合理的に説明することは困難であり、例えば複雑なサンプリング方法による生じる計測誤差が理由ではないかとは想像される。X-13のRegARIMAモデルやモデルベースSEATSの場合、季節成分を抽出するためにノイズ成分の扱いにも注意が必要である。こうした論点についてはデータの作成者と議論することが良いのではと考えられる。いずれにしても、二つの図から事後的にSEATSとX-11から計算された季節調整値はDECOMPによる数値からの乖離は大きくないことは確認できる。

この間、日本では経済に占める農林業の比重は低下しているが、農

³⁶二つの図はRのlibrary("seasonal")よりコマンド *seas*(*.*)を用いて計算した。近年ではRS-DECOMPやRライブラリー *seasonal* により季節調整の複雑な数値計算を容易にネット上や個人PC上で実現できる。(例えば時系列 *ts* 形式の *data* 名を *agri* とすると *s=seas(agri);sagri=final(s);summary(s);ts.plot(agri,sagri)* とすればよい。ただしデフォルトはSEATSプログラムであり、X11の *auto* モデルは例えば *sears(agri,x11="* ")として実行する必要がある。)しかし適切な統計モデルを選択したり、得られた数値を解釈して実際に利用するには統計的時系列解析と季節調整法についての十分な理解と経験が必要であることは変わりはない。詳しくは *seasonal* のHP, Sax (2022)を参照されたい。

林業と対照的と思われる経済全体ではかなりな比重を占めている産業の典型例として製造業をとりあげてみよう。DECOMP(trend=2,ar=2)で推定した季節性は非常に小さいレベルであり、季節性をデータから取り除く季節調整はあまり意味がないことになろう。統計的には季節性がないという帰無仮説を棄却することはできない。農林業と比べると就業者数の水準はかなり大きく、製造業の就業者のみに関心がある場合には特に季節調整値を利用する必要性は乏しいと云えよう。

次に製造業は非農林業の一部でもあるので、さらに同じDECOMPモデルにより非農林業の季節性の推定値を見てみると、季節性の水準はそれほど大きくないが無視できないレベルであることがわかる。また(就業者数)=(農林業就業者)+(非農林業就業者)であるから全体の就業者に置ける季節性は農林業の就業者の季節性が最大要素、非農林業の就業者の季節性がそれほど大きくはないが寄与していることがDECOMP(trend=2, ar=2)による時系列分解により確認された。この利用を探るには製造業以外の非農林水産業の季節性を調べる必要があるだろう。

推定された全体の就業者数について推定された季節性を見ると、図からの観察では若干の季節性の大きさが時間的に変化していることを伺わせるが、農林業の就業者数がこの間のトレンドとしては減少傾向にある。農林業の全体の就業者数への影響は低下しているが、かなり明瞭な季節性が存在しているので、就業者数全体の季節性を無視することが適切でないことになっている、と解釈可能だろう。このように各産業の就業者数の集計値で観察される季節性の大部分が農林業というそれほど大きな比重を占めてはいない産業に起因する変動と見ることができる。ここで多くの季節調整法では特別な処理を施さなければ、個別系列の季節調整値の和は原系列の和の季節調整値に一致するとは限らないことにも注意する必要がある。

なお農林業の就業者数の季節性も近年には変化して若干、弱まりつつあるような印象である。しかしこうした季節性の変化についてはDECOM(trend=2, ar=2)による季節性の推定結果には十分に反

映されていないようであるので、他の統計的モデリングの検討が必要のように思われる。なお関連して詳しい議論は省略するが、X-12-ARIMA, X-13ARIMA-SEATS についても季節性の変化を検出することはかなり困難である。

また曜日効果ではないノイズが無視できないレベルであり、推定された季節調整値の月次レベルでのばらつきは無視できない水準となることは興味深い。実務的には厄介な問題があると考えられるが、要因として例えば時系列データのサンプリングによる影響かも知れない。こうした不規則変動の扱いについてより踏み込んだ分析が必要であろう。

最後に本稿の第一章の最後に言及した Kunitomo and Sato (2021, 2023), 国友・櫻井・佐藤 (2022), 佐藤 (2023) による X12SIML による分析結果を図6に挙げておく³⁷。農林業の就業者数の季節性は全体の就業者数に比較するとそれほど大きくないが、負のトレンド・循環成分が存在する中、季節性は変化、季節性の変動幅は近年になり若干小さくなっていると思われる。これは季節性が比較的明確に観察されている農林業部門の全体がトレンド的に低下していることを考慮すると自然に解釈できそうである。他方、SEATS(auto), X-11(auto), DECOMP などの季節調整により推定された季節成分は前期にわたって安定して推定されて見えるが、実は直近の季調値は相対的には乱高下しているが、毎月のノイズの影響とみることは困難なようである。因みにSEATSやX-11では就業者系列に曜日効果が有意となっている。それに対してX12-SIMLによる季節調整系列はまずまずの安定性を示している。これは季節周波数成分のみを抽出した為であるが、こうした現象が単なる一例に過ぎないのか、興味のあるところである。

3. 幾つかの暫定的結論

³⁷X12SIMLの説明は佐藤(2023)に譲るが、trendはデータ数の10%程度、sorderは季節性のバンド幅、lyはうるう年回帰を意味している。R上で
z=x12siml(agri,trend=25,sorder=4,frequency=12,pa=4,start=c(2002,1),ly=c(2004,2))というモデルによる結果を示しておく。例えばzはデータの差分を直交変換した周波数領域におけるデータの解析結果を表している。Rプログラムの一部は国友・櫻井・佐藤(2022)の付論に掲載されている。

労働量調査は全国的な規模での労働市場に関するデータとしてマクロ経済の動向を知る上で書くことのできない資料である。すべての産業の集計値に関する原データには様々な産業が直面する問題に対応して就業者数が時系列的に変化している。その中でトレンド・景気循環・季節性・不規則変動に分解することは重要であるが、伝統的な季節調整値の公表の形態は（季節調整値）＝（原系列）－（季節性の推定値）により行われている。

製造業の就業者のように季節性が統計的にはほとんど見受けられない業種について季節調整値を公表する意義は乏しいが、逆に農林業のように季節性がかなり大きく観察される場合には季節調整系列を公表する意義は小さくないと考えられる。全体の就業者数の推定された季節性の中身を見ると、農林業が大きな比重を占めているが、非農林業の比重はかなり小さいが全体としては無視できない水準であることが分かった。こうしたことは調べてみれば当然の帰結、担当者にとっては自明と云えるかもしれないが、季節調整値の作成や公表の際の説明には役立つのではないかと考えられよう。

最後になるが季節調整の公表値に関する根本的な問題を提起したい。月次データの場合には観察される時系列をトレンド成分、循環成分、季節性成分、不規則成分、などに分解すると不規則成分の影響が小さくない場合がある。これについては曜日効果・祝祭日効果などが考えられるが、DECOMPの曜日効果モデルによればそれほど大きな成分としては検出することができない。労働量調査の場合には月次調査であり、サンプリングの影響が無視できないのではないかと想像される。近年では公的統計においてはサンプリングの困難性、回収率の低下などが関係者の間では議論されている。こうした公的統計の質の問題が顕在化する中で月次の季節調整値の意味を再度、検討する必要がある。特に月次の季節調整値を求める意味として、季節性を除いて前月からの伸び率を推定するという目的が考えられるが、ノイズレベルが小さくない場合には伸び率の推定値の変動が無視できない場合があるように考えられる。

他方、第1章で議論したように時系列にトレンドや景気変動が存

在して、無視できない影響がある場合には前年同期比や前年同月比の数值は（例えば加法モデルを仮定すると、景気の上昇局面や下降局面などのタイミングがズれる）経済の実態を知る上では問題を生じることがあることが知られている³⁸。例えば最近（2023年1月）の新聞やTVニュースなどマスメディアによるCPI（消費者物価）に関する報道では前年同月比の数值に基づく議論がしばしば行われている。一つの解釈は近年の日本のようにトレンド要因がそれほど大きくないときにはトレンド・循環に関する問題が顕在化しにくい、のかもしれない。しかしながら、前月比の系列はかなり変動する可能性がある。例えば第一章の図3より明らかなように2021年～2022年における日本の月次CPI(消費者物価指数)の上昇は月次変化と前年同月比の変化は時間的ズレが生じている可能性が大きい。

こうした変動を考慮しつつマクロ経済動向の現状を正確に理解する為の一つの方法は季節調整値から不規則変動の推定値を除いて公表系列にすることなどが考えられるだろう。例えば（新季節調整値）＝（旧季節調整値）－（不規則変動推定値）、すなわち（新季節調整値）＝（トレンド成分＋循環成分の推定値）とすることが考えられよう。むろん推定されるノイズのレベルが全体に比べて小さく、比率の公表値を求めると影響が少ない場合も少なくないと思われる。前年同月比を利用する一つの理由が年率、という比較しやすい数值にしたいのであれば、この推定値を年率に換算し、年率換算推定値として利用することなどが考えられるかもしれない³⁹。なおここでの提案が従来の季節調整値と紛らわしく適切ではない場合には新季節調整値ではなく就業者数動向(参考値, 推定値)などとして徐々に変更していくことなども考える余地があるだろう。

参考文献

総務省統計局 (2019) 「労働力調査の解説 (第5版)」, 統計局労働力調査 HP. <https://www.stat.go.jp/data/roudou/index.html>

³⁸例えば溝口・刈屋 (1983) 第4章の説明を参照されたい。

³⁹例えば現行の四半期GDP統計では新聞やTVニュースなどのメディアでは年率換算値が最も注目されがちであるが、季節調整系列を4倍する現行の方式だと不規則変動項の影響も4倍されることになる。

国友直人・櫻井智章・佐藤整尚 (2022), ”経済時系列の状態推定とマ
クロ指標,” 統計研究彙報, 79, 1-20.

<https://www.stat.go.jp/training/2kenkyu/ihou/79/pdf/2-2-791.pdf>

佐藤整尚 (2023) 「X12SIML を使った季節調整の実際」.

<http://www.kunitomo-lab.sakura.ne.jp/x12simldoc92.pdf>

Sax, C. (2022), ”seasonal: R Interface to X-13ARIMA-SEATS,”

<http://www.seasonal.website/>

溝口敏行・刈屋武昭 (1983), 「経済時系列入門」日本経済新聞社.

Kunitomo, N. and S. Sato (2021), ”A robust-filtering method for
noisy non-stationary multivariate time series with econometric
applications,” *Japanese Journal of Statistics and Data Science*,
4, 373-410.

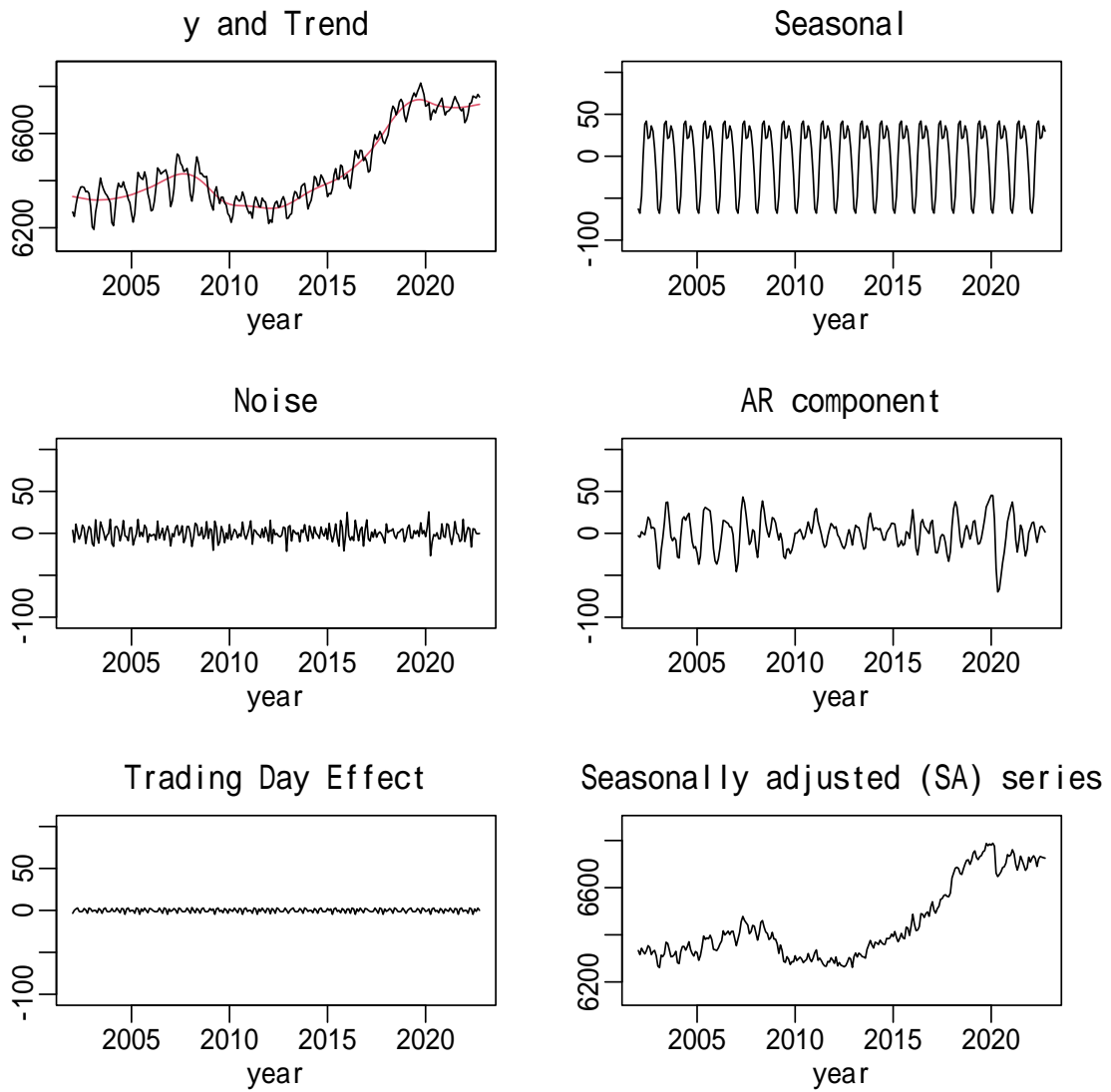
<https://link.springer.com/article/10.1007/s42081-020-00102-y>

Kunitomo, N. and S. Sato (2023), ”Frequency Regression and
Smoothing for Noisy Non-stationary Time Series,” Discussion Pa-
per, Expert Project (ISM, forthcoming).

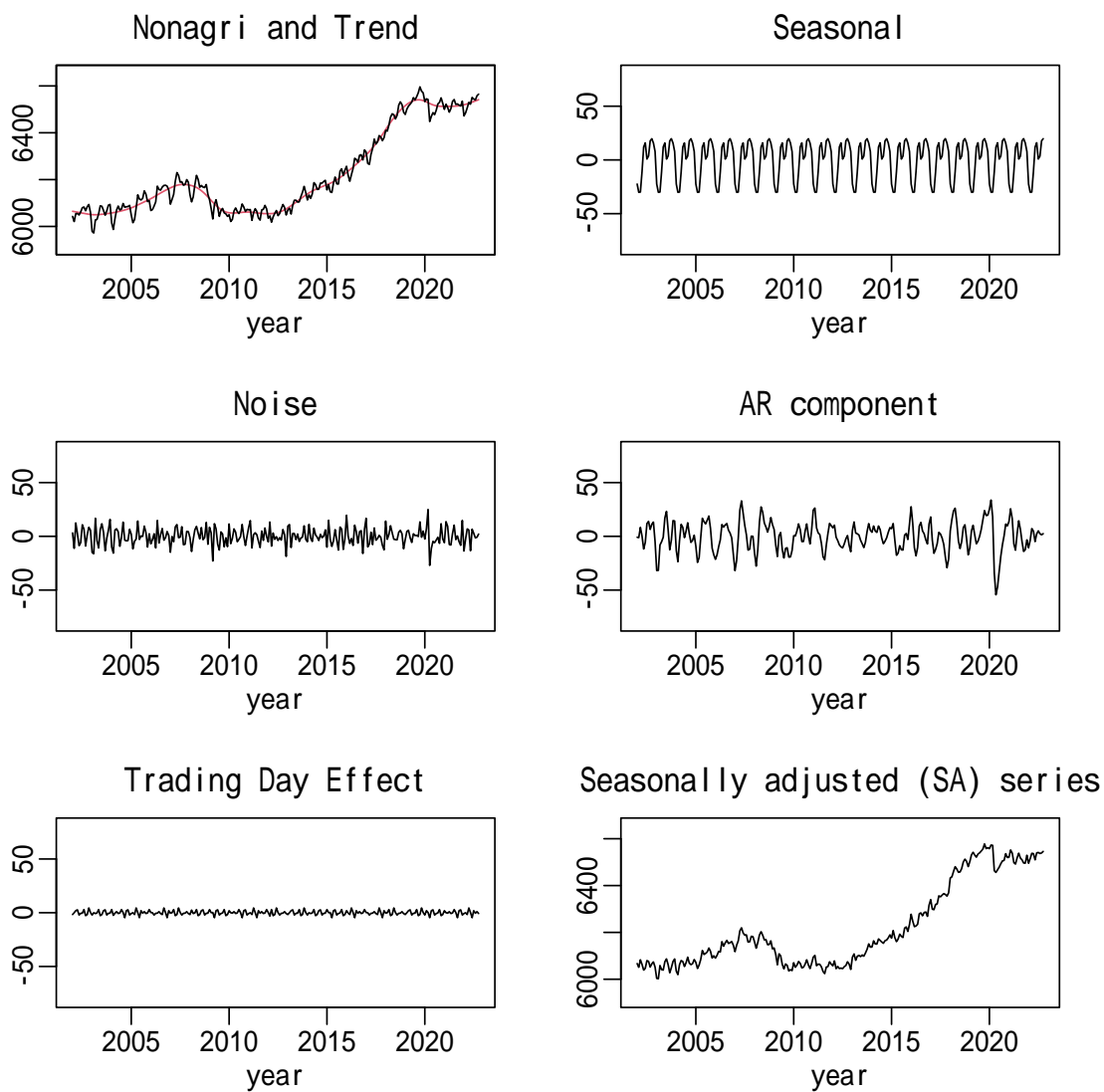
<https://stat-expert.ism.ac.jp/training/discussionpaper/>

< 図1：就業者数データ(全体, 月次)の分解, DECOMP >

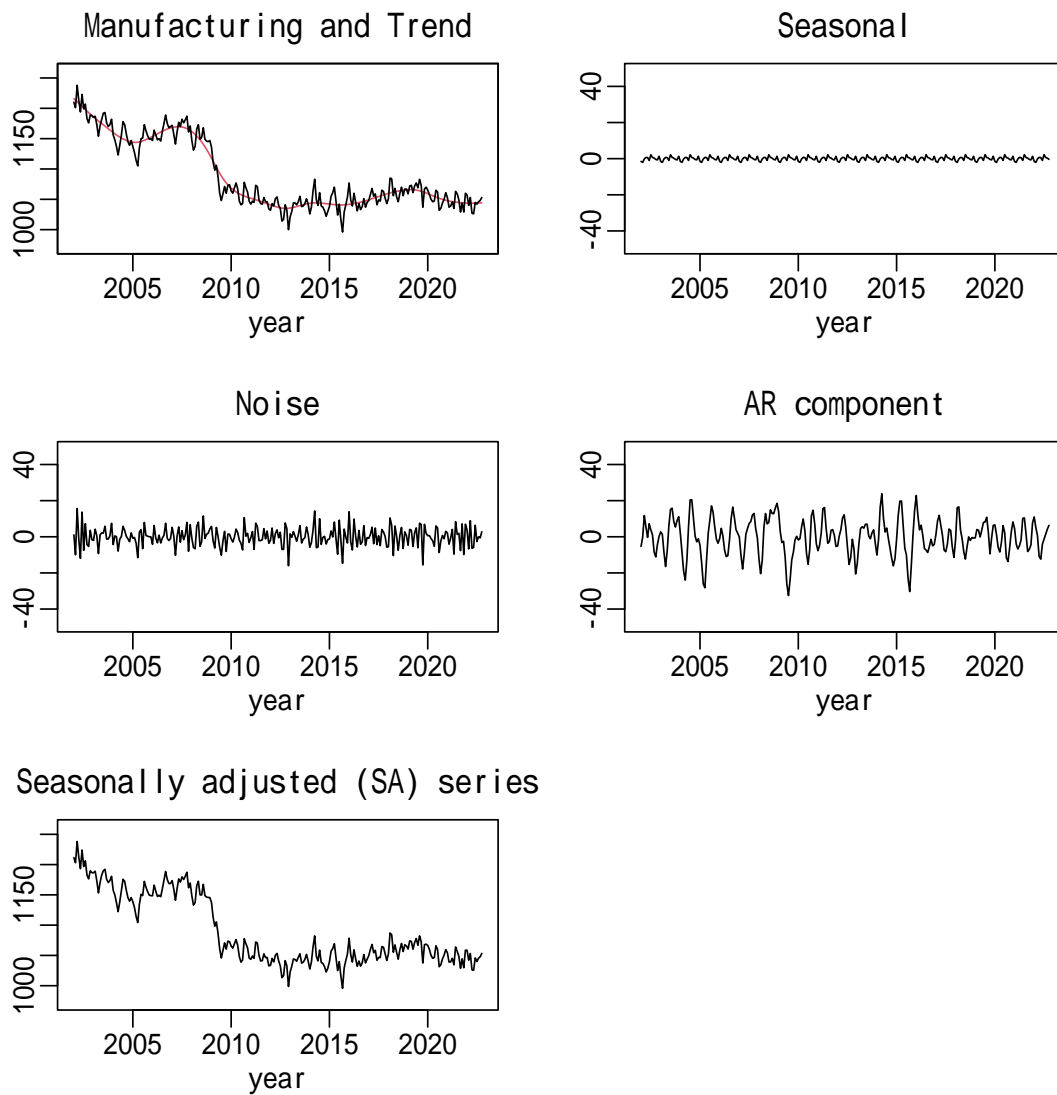
2002年1月～2022年10月, 原系列 (Aggri) とトレンド (trend), 季節成分 (Seasonal), ノイズ成分 (Noise), 曜日成分 (Trading Day Effect), 季節調整系列 (SA).



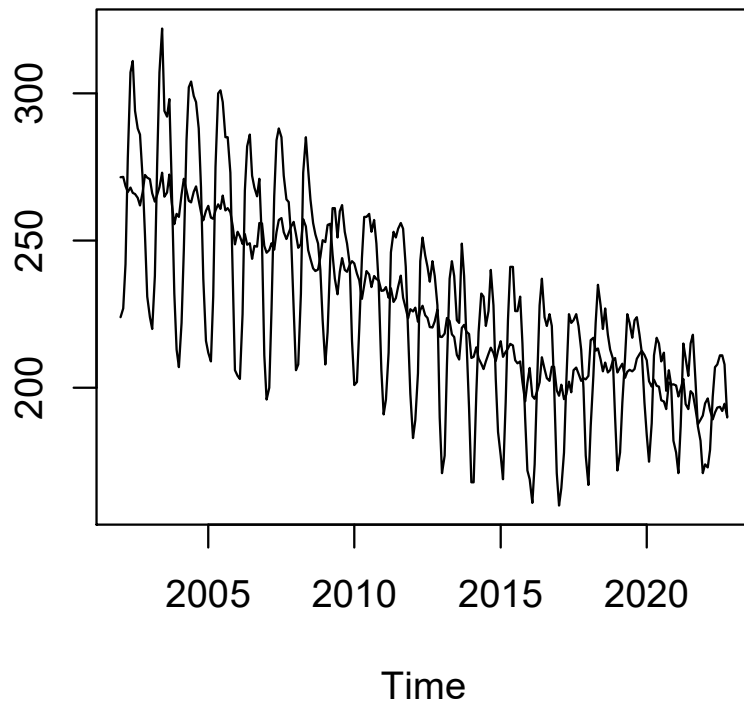
< 図2：就業者数データ (Non-Agri, 月次) の分解, DECOMP >
2002年1月～2022年10月, 原系列 (Agri) とトレンド (trend), 季節成分 (Seasonal), ノイズ成分 (Noise), 曜日成分 (Trading Day Effect), 季節調整系列 (SA).



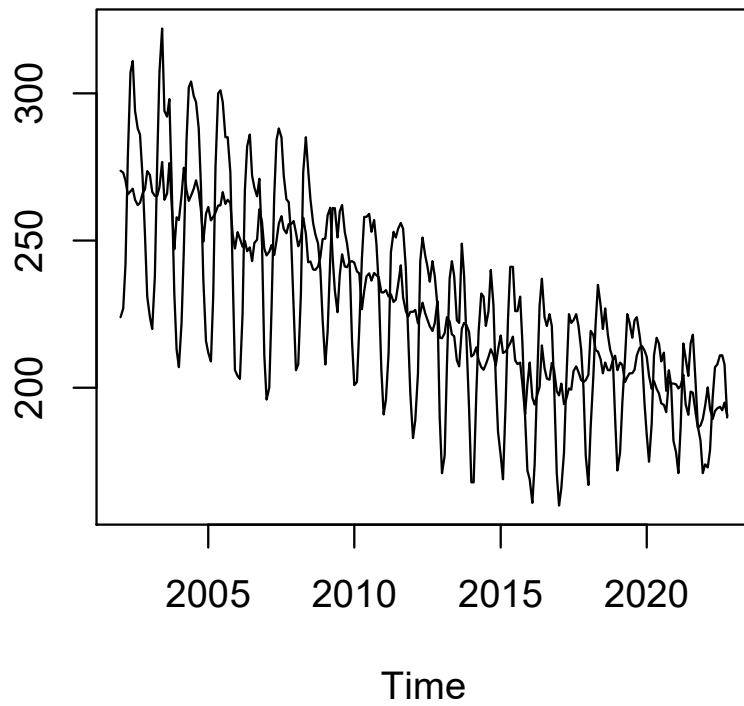
< 図3：就業者数データ (製造業, 月次) の分解, DECOMP >
2002年1月～2022年10月, 原系列 (Aggri) とトレンド (trend), 季節成分 (Seasonal), ノイズ成分 (Noise), 曜日成分 (Trading Day Effect), 季節調整系列 (SA).



< 図4 : 就業者数データ (農林業, Seats(auto モデル)) >
2002年1月~2022年10月, 原系列と季節調整系列, SEATS デフォルト (auto モデル) で作成.

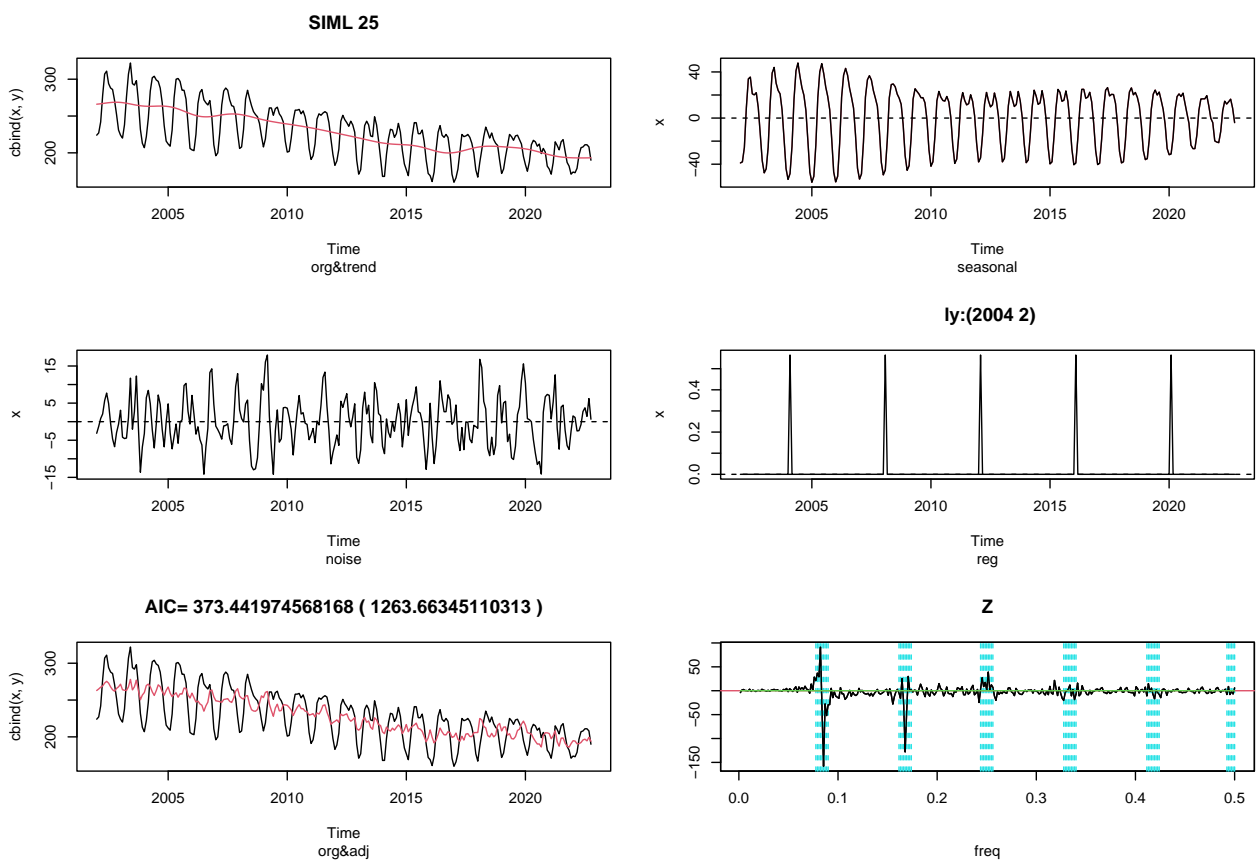


<図5：就業者数データ(農林業, X-11(autoモデル))>
2002年1月～2022年10月, 原系列と季節調整系列, X-13ARIMA-SEATS, X-11デフォルト
(autoモデル)で作成.



< 図6：就業者数データ (農林業, X12-SIML)>

2002年1月～2022年10月, 原系列 (org) とトレンド (trend), 季節成分 (seasonal), ノイズ成分 (noise), うるう年効果 (1 y), z 変換値 (freq).



おわりに

統計エキスパート養成プロジェクトの一環として企画された 2022 年度のコンサルテーション演習「公的統計と季節調整」では総務省統計局で労働力調査を実際に扱っている担当者からの説明を受け、経済時系列と実際の季節調整をめぐる具体的な課題について考察する機会を得た。この報告ではまず X-11, X-12, X-13ARIMA-SEATS, DECOMP, X12SIML などの統計的に季節性の分析に基づく季節調整法の内容を説明すると共に、センサス局が最新の季節調整法としてプログラムを公開している X-13ARIMA-SEATS の理解に必要な統計的時系列解析について（特に実務家はあまり耳にしないだろう）離散時間観測の時系列に対するスペクトルの予測問題への応用についても若干の解説を述べた。むろん時系列データのスペクトル分析になじみのない実務家も少なくないと思われるが、「季節性とは何か？」という実務的にも重要な問いについてのヒントにつながれば幸いである。また実例として労働力調査における産業別雇用者数の時系列についての若干の分析結果も報告した。異なる産業間で異なる季節性がある場合、実務的にどのように処理するかという問題は重要ではあるが今のところ明快な答えは存在しないので、産業における地道な季節性の分析が重要となる。産業の集計値の季節性は実にごく一部の特殊な産業の季節性による影響が大きい実例であろう。現時点では産業別雇用者数の季節調整値は公表されていないとのことであるから、今後の実務においてこの報告書が役立つことを期待したい。

X-13ARIMA-SEATS プログラムは X-11, X-12-ARIMA と TRAMO-SEATS という異なる考え方による処理方法が混在している総合的なプログラムであり、X-13 を利用することによりこれまでと同様に X-11, X-12-ARIMA のプログラムを実行することが可能である。この意味では X-13 自体には新鮮味が乏しいとも云えるが、新たに Wiener-Kolmogorov フィルターによる離散時間観測の時系列データから ARIMA モデルを利用するスペクトル分解に基づいた季節調整

法である SEATS コマンドが実行できるので、これまで TRAMO-SEATS を利用していたヨーロッパを中心とする公的統計においては有用だろう。他方、これまで TRAMO-SEATS を利用したことがない日本を含めた公的統計では今のところ特に SEATS を利用する必然性は特に見当たらないと考えられる。しかし国際的な協力の中で我が国の公的統計における季節調整を扱うに当たっては、今後は X-13ARIMA-SEATS の内容、時系列データの周波数分析の基本に基づく X-13 における様々なオプションの意味と結果について理解することは不可欠だろう。

しばらく前に日本の公的統計における新しい X-12-ARIMA が導入されようとした時点において本編集者は「X-12-ARIMA の利用する場合には DECOMP など信頼できる他の結果などと比較することなどを推奨する」と云う意見を表明した。本稿での考察を要約すると、「X-12-ARIMA」を「X-13ARIMA-SEATS」に、「DECOMP」を「DECOMP や X12SIML」などにそれぞれ置き換える必要がある、との結論を述べておく。

さらに本報告書では季節調整値から不規則変動の推定値を除くことの検討を提案した。こうした方法を日本で実現しているよく知られた公的統計はまだないようであるが、例えば（新季節調整値）＝（旧季節調整値）－（不規則変動推定値）、すなわち（新季節調整値）＝（時系列データのトレンド成分＋循環成分の推定値）とすることが考えられる。あるいはトレンド・循環成分の一定期間前（例えば3か月、6か月）からの伸び率を利用して経済の瞬間的な伸び率を推定し、年率推定値として毎月公表する、などの可能性もある。

実務的には最近（2023年1月20日統計局が公表）でも CPI（消費者物価指数）の変動を巡り、マスメディアなどで前年同月比系列の数値が大きく報道され、議論されているが、この前年同月比による経済の現状把握についてはかなり以前より望ましくない性質があることが議論されている。すなわちここで言及した方法により前年同月比に比べてよりよい年率換算値が得られる可能性があることをここ

で強調したい。なお従来の季節調整値と紛らわしく適切ではない場合には新季節調整値ではなく就業者数動向(参考値, 推定値)などとして徐々に変更することなども考える余地があるだろう。この問題はなぜ原系列から月次の季節調整系列を作成して公表値とするのか、と云う公的統計における根本的な問題と関連するので、あえてここで言及した。

最後に本報告書は比較的自由的な立場から、季節調整法についての開発の経緯や内容を説明し、月次系列の季節調整についての基本問題にまで言及した。本稿の議論が今後の日本の公的統計における公表系列の改善に役立つことを期待したい。