

研究ノート

季節調整法プログラム センサス局法 X-12-ARIMAのversion 0.3について

奥本佳伸

1. はじめに

本稿は2007年5月にアメリカの商務省センサス局から公表された、季節調整法のプログラムX-12-ARIMAの新しいversionであるversion 0.3について紹介するものである。X-12-ARIMAは、現在、我が国の官庁統計においても広く使われている。この新しいversionのX-12-ARIMAは、従来のversion 0.2を改良したものであり、いくつかの点で、ユーザーにとって、より使いやすく、高度な手法が使えるものとなっている。以下においては、まず季節調整とは何かということから説明を始め、季節調整法としてのセンサス局法の歴史を簡単に振り返った後、このX-12-ARIMAのversion 0.3の新しい機能について概観することにした。

2. 季節調整法プログラム センサス局法X-12-ARIMAについて

まず初めに、季節調整法プログラム センサス局法X-12-ARIMAについての説明から始めたい。

(1) 季節調整について

「季節調整」とは時系列の統計データから季節変動を除去することである。時系列の統計データに含まれる季節変動とは、1年を周期として、

比較的規則正しく繰り返す変動である。その要因としては、自然的要因（暑さ、寒さ、農作物の収穫期など）や社会的制度・慣習による要因（お中元、お歳暮、クリスマス、ゴールデンウィーク、夏休み、企業の決算期、年末決済など）などがある。統計データの変動を観察する場合（特に、前月比（月次データの場合）や前期比（四半期データの場合）で見える場合）には、こうした季節変動を除いて観察した方がよいと考えられている。その理由は、季節変動以外の変動（それは傾向変動、循環変動、不規則変動の3つが合わさった変動ということができる。）が、季節変動におおい隠されなくて観察できるためである。

統計データから季節変動を除去する方法は、1910年代からこれまでにいろいろと考案されてきたが、現在、我が国の官庁統計などで一般的に用いられている方法は、アメリカの商務省センサス局が開発した季節調整法のコンピューター・プログラムであるセンサス局法X-12-ARIMAを用いる方法である。

(2) センサス局法の歴史

センサス局法の歴史について簡単に見てみると、アメリカ商務省センサス局では1950年代前半にシスキン（Julius Shiskin）を中心に移動平均法を基礎とした季節調整法の研究が行われ、1954年に最初の季節調整法のコンピューター・プログラムが開発された。このプログラムはセンサス局法Ⅰと呼ばれた。1956年には、これをさらに改良したセンサス局法Ⅱと呼ばれるプログラムが開発された。この後、センサス局法は、文字Xと一連の番号が付けられて、順次改良版が開発されていった。このうち最初に公表されたものがX-3（1961年）で、その後、順次X-8（1960年）、X-9（1961年）、X-10（1961年）、X-11（1965年）が公表された。このうちX-11は、この時期までのセンサス局におけり季節調整法の開発過程において1つの頂点に達するものであった。

センサス局では、X-11の公表後もさらに季節調整法の研究・開発が進められ、X-11の改良版であるセンサス局法X-12-ARIMAが1996（平成8）年6月に公表された。この時、公表されたのは評価目的用のベータ・バージョン（Beta version）と呼ばれるものであった。番号で表すバージョンでは、version 0.1であった。その後、1998（平成10）年2月にはX-12-ARIMAのファイナル・バージョン（Final version）が公表された。これは、version 0.2であった。

さらに、2007年5月にはX-12-ARIMAのversion 0.3が公表された。これが、本稿で紹介しようとするX-12-ARIMAの新しいversionである。

(3) X-12-ARIMAプログラムの基本的な構成

季節調整法としてのX-12-ARIMAは、次の3つのプロセスから成っている。

まず第1には、季節調整を行う前にデータの事前調整を行う。この事前調整を行うために、REGARIMA（レグアリマ）と呼ばれる時系列モデルにより、データに含まれる異常値やレベルシフト、曜日変動等を回帰式を用いて推計し、これらをあらかじめ原系列から除去する。こうすることは、季節調整した後のデータ（季節調整系列）が不安定になることを防ぐ効果がある。REGARIMAとは、Regression（回帰）とARIMAモデルを組み合わせたモデルのことである。

第2には、REGARIMAを用いて原系列の予測値を推計した上で、この予測値と原系列をつないだ系列に対してX-11を用いて季節調整を行う。これにより、データの末端部分についても、片側移動平均ではなく、両側のデータを用いた移動平均をすることができる。その結果、末端部分での移動平均によるゆるがみが少なくなり、季節調整系列を安定化させる効果がある。

第3には、季節調整をした結果について、統計的な分析などにより、

適切に季節調整が行われたかどうかを診断する機能が付いていることである。

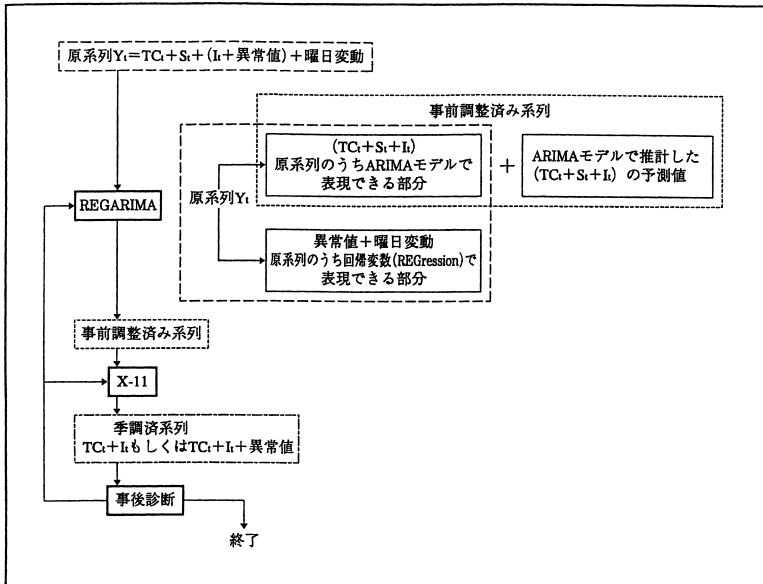
こうしたX-12-ARIMAの3段階の季節調整の手順は、

- ① REGARIMAによる原系列の事前調整パート
- ② 従来のX-11による季節調整パート
- ③ 事後診断パート

の3つのパートから成り立っている（図1参照）。なお、図1における TC_t は傾向・循環変動、 S_t は季節変動、 I_t は不規則変動を表す。

それぞれのパートについて、改めてもう少し詳しく見ると、まず第1のパートでは、REGARIMAを用いて、原系列をARIMAモデルで表現できる部分と、異常値、曜日変動等により説明される部分に分解する。

図1 X-12-ARIMAの推計フロー



(出所) 日本銀行調査統計局「季節調整法について」『日本銀行月報』1996年5月，p. 83，図表5。

その上で、ARIMAモデルを用いて推計した予測値（特に指定しなければ、先行き1年分）と、ARIMAモデルで表現された実績データとをつなぎ合わせた「事前調整済み系列」を作成する。

第2のパートでは、このようにして得られた「事前調整済み系列」に対して、従来のX-11による季節調整を行う。ここでは、データの末端でも先行きの予測値を用いた両側のデータによる移動平均が可能であり、また、異常値や曜日変動等による攪乱を受けることがないので、移動平均による調整で季節変動をより適切に除去できると考えられている。

第3のパートでは、季節変動が過不足なく除去されているかを統計的な手法によりチェックするとともに、季節調整系列の安定性についての診断を行う。季節調整系列の安定性とは、ある統計データに、時間の経過とともに新しいデータが付け加わったとき、そのデータを含めて季節調整を行った場合、これまでの最近時点での季節調整値がどの程度変化するかを見て、変化が小さければ安定性が高いと判断するものである。この場合、季節調整値の安定性が高いほど、望ましい季節調整の方法であると評価される。第3のパートの結果次第では、REGARIMAにおけるモデル化の方法や異常値、曜日変動等の調整の仕方を再検討して、変更することになる。

3. X-12-ARIMAのversion 0.3の特徴

X-12-ARIMAのversion 0.3については、センサス局のBrian C. Monsell氏が執筆した“Release Notes for Version 0.3 of X-12-ARIMA” (Monsell (2007)) という文書がセンサス局のWEBサイトで公表されている。この文書に従って、version 0.3の特徴を見た場合、version 0.3のそれ以前のversion 0.2からの変更点としては、主なものとして以下のような点がある。

- (1) ARIMAモデルの自動選択機能が組み込まれたこと

- (2) 季節調整値の年合計値を元の原数値の系列の年合計値に一致させることができる新しい機能が加わったこと
 - (3) 合成されたデータを季節調整する場合に適用される手続きの改訂
 - (4) 事後診断において、季節調整の診断結果のデータファイルとモデルの診断結果のデータファイルが分かれていたのを統合したこと
- このうち、筆者が重要な変更点と思うものは(1)と(2)である。

以下、次の4節で、このうち(1)の点について、さらに詳しく見ていくことにする。(2)の点については、5節で見ることとする。

4. ARIMAモデルの自動選択機能

(1) 自動選択機能の概要

X-12-ARIMAのversion 0.3で新たにARIMAモデルの自動選択機能が組み込まれたのであるが、以前のversionでは、この点はどのようになっていたかについて述べる。

Version 0.2までのX-12-ARIMAのプログラムでは、季節調整に用いるARIMAモデルを選ぶのに、2つの方法があった。その1つの方法は、季節調整をしようとするデータの自己相関と偏自己相関を計算し、そのデータをプロットした形状を観察して、ARIMAモデルを決める方法である。この方法は、Box=Jenkins (1970) が提唱したARIMAモデルの同定の方法である。例えば、日本語で書かれた時系列分析についての教科書的な本である山本 (1988) でも、ARIMAモデルの同定方法として、このような方法が述べられている。しかし、この方法はデータの自己相関と偏自己相関がプロットされたものの形状から判断するのがなかなか難しく、特に時系列分析の初心者にとっては難しく感じられる点である。候補となるARIMAモデルがいくつか考えられる場合には、それらについてAIC (赤池情報量基準) を計算し、その値の最も小さいものを選ぶという方法がよく用いられてきた。

もう1つのARIMAモデルを選ぶ方法として用いられてきたのは、version 0.2までのX-12-ARIMAに組み込まれていたコマンド（英文マニュアルの用語ではspec）であるAUTOMDLを用いる方法である。このAUTOMDLを用いる方法は、あらかじめ用意された候補となるARIMAモデルの中から、次の統計的基準に合うものを選ぶという方法である。（X-12-ARIMA Reference Manual, Version 0.2.10, p. 51）

- ① データの最近3年間以内の期間についての、モデルによる外挿予測値の絶対平均誤差率（absolute average percent error）が15%未満であること
- ② モデルの推計値の残差に自己相関がないこと（Ljung=Box検定におけるQ統計量のp値が5%よりも大きいこと）
- ③ 過剰階差になっていないこと（季節変動以外の部分を表現する（non-seasonalの部分の）ARIMAモデルのMAパートの係数の和が0.9を超えないこと）

このAUTOMDLを用いる方法を使うためには、あらかじめ候補となるARIMAモデルを用意しておかなければならない。

X-12-ARIMAのversion 0.2では、ユーザーが特に候補となるARIMAモデルを指定していない場合に適用されるモデル（defaultのモデル）として、次の5つのモデルが用意されている。

(0 1 1)(0 1 1), (0 1 2)(0 1 1), (2 1 0)(0 1 1),
(0 2 2)(0 1 1), (2 1 2)(0 1 1)

なお、ここで乗法型季節ARIMAモデルを $(p\ d\ q)(P\ D\ Q)$ と表しているが、各パラメーターの意味は次のとおりである。

p = ARモデルの階数

d = 原データを定常化するための差分の階数

q = MAモデルの階数

P = 季節変動ARモデルの階数

D = 季節階差の数

Q = 季節変動MAモデルの階数

この場合は、AUTOMDLを実行すると、上記の5つの候補となるモデルから、3つの基準のいずれをも満たすモデルが選択されることになる。また、5つの候補となるモデルのいずれもが、その時系列データについては3つの統計的基準を満たさなかった場合には、最初のモデルである $(0 \ 1 \ 1)(0 \ 1 \ 1)$ が選ばれることになっている。

筆者の場合は、ARIMAモデルのdとDをどちらも1として、p, q, P, Qがそれぞれ0, 1, 2という値をとる場合の組み合わせを考えて、 $(0 \ 1 \ 0)(0 \ 1 \ 0)$ から $(2 \ 1 \ 2)(2 \ 1 \ 2)$ までの81とおりのARIMAモデルを用意し、AUTOMDLを適用して、上の3つの基準のあてはまるモデルのうち、サンプル期間内で予測誤差が最も小さいモデルを選ぶという方法を用いていた。この方法は、X-12-ARIMAのスペック・ファイルの書き方では、

```
automdl { method=best }
```

と指定するものである。ここで、method=bestと記述しないと、defaultはmethod=firstと指定したと見なされることになっており、この場合は、候補となるモデルのうち3つの基準に合った最初のモデルが選ばれて、そこでモデル選択のプロセスは終わってしまう。

次に、X-12-ARIMAのversion 0.3において組み込まれたARIMAモデルの自動選択機能について述べる。以下の説明は、Monsell (2007) に基づくものである。

このARIMAモデルの自動選択機能は、スペイン銀行のVictor Gómez と Agustín Maravall (Gómez and Maravall (1997)) が開発したTRAMO-SEATSプログラムのうちのTRAMOという時系列モデリングプログラムの手続きに基づいたものである。X-12-ARIMAのversion 0.3に組み

込まれたARIMAモデルの自動選択機能は、TRAMOの手続きと非常によく似ているが、X-12-ARIMAのモデル推計の手続きやREGARIMAのモデル作成のオプションなどが利用できるように若干の修正が加えられている。その結果、X-12-ARIMAのversion 0.3が選んだモデルがTRAMOの選んだモデルと異なるということもありうる。

TRAMO-SEATSについて簡単に説明すると、TRAMOはTime series Regression with ARIMA noise, Missing observations, and Outliersの略であり、SEATSはSignal Extraction in ARIMA Time Seriesの略である。TRAMOはARIMAモデルの自動選択を行うプログラムであり、SEATSはこのモデルを用いて、シグナル抽出法により時系列データをトレンドや季節変動成分などその変動成分に分解するプログラムである。シグナル抽出法とは、ノイズが混入した観測値から未観測値であるシグナルを推定する手法である。TRAMO-SEATSのこの分解構造については、日本語文献では東(2003)において、詳しく説明されている。

TRAMOにおけるARIMAモデルの自動選択機能の手順については、Monsell(2007)で次の5つのステップがあることが述べられている。

① default (標準設定) のモデルの推計

defaultのモデルが推定され、最初の異常値の検出、異常値や曜日変動などの説明変数のテストが行われ、残差についての診断が行われる。defaultのモデルとしては、airline modelと呼ばれる $(0 \ 1 \ 1)(0 \ 1 \ 1)$ のモデルが用いられる。

② 階差の次数の決定

モデルの階差の次数を決めるために、単位根検定が行われる。

③ ARMAモデルの次数の決定

Schwarz(1978)が考案したベイズ統計によるモデル比較基準であるBICを用いて、ARMAモデルの次数を決めるために繰り返し計算が行われる。

- ④ 同定されたモデルとdefaultのモデルとを比較する。

同定されたモデルとdefaultのモデルとを一定の統計的基準に基づいて比較し、より望ましい特性を持つモデルを選択する。

- ⑤ 最終的なモデルのチェック

④で選ばれたモデルについて、適切かどうかについてのさらに一連の統計的なテストを行い、最終的に選ぶモデルを決める。

以上の手続きのうち、②のステップはオプションになっている。ユーザーは②のステップを省略して、モデルの季節変動部分とそれ以外の(通常の)変動部分の次数を指定することができる。

以上の①から⑤までのステップについての詳しい説明は、X-12-ARIMA Reference Manual, Version 0.3のpp. 69-72に述べられている。

X-12-ARIMA version 0.3では、このARIMAモデル自動選択がAUTOMDLというコマンドで利用できるようになっている。他方で、version 0.2でAUTOMDLというコマンドで利用することができたARIMAモデル自動選択機能は、version 0.3ではPICKMDLと名称を変えて組み込まれており、従来version 0.2で用いていたAUTOMDLの機能も、version 0.3においても用いることができるようになっている。

(2) ARIMAモデルの自動選択機能についてversion 0.2と0.3との比較

Dent et al. (2007) では、ARIMAモデルの自動選択機能についてversion 0.2と0.3との比較を行っている。ここでは、その結果の主な点を紹介する。

この論文では、アメリカの小売り、輸入、輸出から季節変動のある307の時系列データを選び、それらにX-12-ARIMAのversion 0.2と0.3を適用し、どのようなARIMAモデルが選ばれるかを調べている。その結果は、表の1～3に示されている。

まず、表1では、307の系列のうち、version 0.2と0.3で同じモデル

表1 同じモデルが選ばれた頻度

時系列の種類	頻 度
小 売 り	17/43 (39.5%)
輸 入	74/122 (60.6%)
輸 出	79/142 (55.6%)
合 計	170/307 (55.4%)

**表2 versin 0.2が選んだモデル
(パラメーターはdefaultを使用)**

モ デ ル	頻 度
(0 1 1)(0 1 1)	270 (87.9%)
(0 1 2)(0 1 1)	24 (7.8%)
(2 1 0)(0 1 1)	9 (2.9%)
(0 2 2)(0 1 1)	0 (0.0%)
(2 1 2)(0 1 1)	4 (1.3%)

表3 versin 0.3が選んだ上位5つのモデル

モ デ ル	頻 度
(0 1 1)(0 1 1)	185 (60.3%)
(0 1 0)(0 1 1)	20 (6.5%)
(0 1 2)(0 1 1)	10 (3.3%)
(1 1 0)(0 1 1)	9 (2.9%)
(2 1 1)(0 1 1)	9 (2.9%)

(注) 下の2つのモデル (1 1 0)(0 1 1), (2 1 1)(0 1 1) は, version 0.2では選べないモデルである。

を選んだのはどれくらいの数であったかを見ている。この表によると、小売りデータでは43系列中17系列で、割合では39.5%、輸入データでは122系列中74系列で60.6%、輸出データでは142系列中79系列で55.6%であった。合計では、307系列中170系列で、割合では55.4%であった。また、この2つのversionが同じモデルを選んだケースでは、1つの系列を除き、選ばれたモデルは $(0 \ 1 \ 1)(0 \ 1 \ 1)$ であった。この結果を見ると、半分以上の系列でversion 0.2と0.3は同じARIMAモデルを選んでおり、モデル選択の結果としては、一致する場合がかなりあることがわかる。

次に表2は、version 0.2が選んだモデルは、5つの候補となるモデルのうちどのモデルであったかを示している。これを見ると、最初のモデルである $(0 \ 1 \ 1)(0 \ 1 \ 1)$ が約88%とほぼ9割を占めている。このモデルはairline modelと呼ばれるモデルであり、多くの時系列データに当てはまりがよいモデルであることが知られている。これに次ぐのは $(0 \ 1 \ 2)(0 \ 1 \ 1)$ のモデルで、7.8%となっている。それ以外のモデルが選ばれる割合は非常に小さく、 $(0 \ 2 \ 2)(0 \ 1 \ 1)$ のモデルが選ばれる割合は0となっている。

表3は、version 0.3が選んだモデルを示したものである。ここでも最も多く選ばれたのは $(0 \ 1 \ 1)(0 \ 1 \ 1)$ であった。このモデルが選ばれた割合は60.3%であり、version 0.2の場合の87.9%よりは割合が低い、それでも6割というのは高い割合であると言えよう。この表に見られるように、 $(0 \ 1 \ 1)(0 \ 1 \ 1)$ 以外のモデルが選ばれる割合は、いずれも10%未満であった。

次に、この論文では、調査対象とした時系列データの各々について異常値 (outliers) を検出する数については、2つのversionの間で大きな違いはないという結果を得ている。ただし、次のような点にも注目している。それは、2つのversionの間である程度の差が見られた系列につ

いて見ると、version 0.2は相対的に多くの異常値を検出して、モデルはより単純な (simpler) モデルを選ぶ傾向がある。他方、version 0.3については、相対的に少ない数の異常値を検出し、mixed modelと呼ばれるARパートとMAパートの次数がどちらも0でないcomponentを含むモデルを選ぶことがしばしばあった、ということである。例えば、ある小売りのデータについては、version 0.2は10の異常値を検出して、 $(0 \ 1 \ 2)(0 \ 1 \ 1)$ のモデルを選んだのに対し、同じデータについてversion 0.3は3つの異常値を検出して、 $(2 \ 1 \ 1)(0 \ 1 \ 1)$ というモデルを選んだとのことである。

(3) 筆者によるテスト利用の結果

筆者は、日本の大口電力使用量のデータ（データ期間は、1990年1月～2007年12月）にX-12-ARIMAのversion 0.3のPICKMDLとAUTOMDLを適用して、そのモデル選択の結果を見てみた。ここでversion 0.3のPICKMDLというのは、version 0.2のAUTOMDLと全く同じモデル自動選択機能を持つコマンドである。

ここでは、最初に $(0 \ 1 \ 1)(0 \ 1 \ 1)$ のモデルを仮定して異常値の探索をすると、異常値は検出されなかった。また、どちらのケースでも、REGARIMAモデルにおいて標準曜日調整（スペック・ファイルでの表記ではTDNOLPYEAR）とうるう年調整（同じくLPYEAR）の説明変数を使用して、標準曜日調整（カレンダーの土曜日、日曜日だけを曜日調整する。）とうるう年調整を行った。また、PICKMDL（version 0.2のAUTOMDLに相当する。）の場合の候補となるARIMAモデルについては、X-12-ARIMAでdefaultとして用意されている5つのモデルを用いた場合と、先に述べたように筆者が従来から用いてきた $(0 \ 1 \ 0)(0 \ 1 \ 0)$ から $(2 \ 1 \ 2)(2 \ 1 \ 2)$ までの81とおりのARIMAモデルを用いた場合の両方をやってみたが、選ばれたモデルはどちらも $(2 \ 1$

表 4 日本の大口径電力使用量のデータにversion 0.3のPICKMDLとAUTOMDLを適用した結果

	PICKMDL	AUTOMDL
選ばれたモデル	(2 1 2)(0 1 1)	(0 1 1)(0 1 1)
モデルの当てはまりについての統計量		
AIC	2858.6587	2867.1242
BIC	2901.7304	2900.2562
MAPR	0.32	0.32

2)(0 1 1)で同じであった。これに対して、version 0.3のAUTOMDLを用いた場合は、モデル (0 1 1)(0 11) が選ばれている。これは表3で87.3%のケースでモデル (0 1 1)(01 1) が選ばれているのを見たが、日本のデータでもモデル (0 1 1)(0 1 1) が選ばれることが多いのではないかと推測される結果であった。

この結果は表4のとおりである。選ばれたモデルについては、PICKMDLでは (2 1 2)(0 1 1) であり、AUTOMDLでは (0 1 1)(0 1 1) という結果になっている。

表4のモデルの当てはまりについての統計量を見てみると、AICで見ると、PICKMDLでは2858.6587であるのに対し、AUTOMDLでは2867.1242となっていて、PICKMDLの方が数値が小さく、より当てはまりのよいモデルを選んでいるということになる。しかし、BICで見ると、PICKMDLでは2901.7304であるのに対し、AUTOMDLでは2900.2562となっていて、AUTOMDLの方が数値が小さく、より当てはまりのよいモデルを選んでいるということになる。このようにAICを用いるか、BICを用いるかによって、モデルの当てはまりのよさの評価が異なる結果になっている。Version 0.3のAUTOMDLはBICを用いて最

適なモデルを選んでいるので、BICを基準として見れば、PICKMDLが選んだ(2 1 2)(0 1 1)よりも、AUTOMDLが選んだ(0 1 1)(0 1 1)の方がモデルのデータへの当てはまりがよいということになる。

なお、表4では、季節調整値の安定性の指標であるMAPR (Mean Absolute Percent Revision) も示したが、この指標ではどちらも0.32と同じ値になっている。この値が同じということは、実用上は、どちらのモデルを選んでも、季節調整値としての安定性(新しいデータが付け加わったときの、最近時点の季節調整値がどの程度安定しているか)では同等であるということになる。

5. 季節調整値の年合計値の調整についての新たな機能

X-12-ARIMAのversion 0.3では、このほかに、季節調整値における年合計値(暦年または年度)を原数値の年合計値に一致させるということがオプションでできる機能が加わった。従来から、月次または四半期の時系列データを季節調整した場合、季節調整値では、1月から12月までの1年間のデータを合計しても、原数値の同じ期間の1年分の合計値とはわずかであるが差があるのが普通であった。そして、このような差が生じるのは、季節調整という複雑な計算を時系列データに対して行う結果としてはやむをえないと考えられてきた。X-12-ARIMAのversion 0.3で、この点について改善を行い、季節調整値における年合計値を原数値の年合計値に一致させるという機能が付け加わったことは、季節調整方法の発達の中でも1つの進歩と言ってよいであろう。

この機能を用いるのは、プログラムを書き表すスバック・ファイルの中で、FORCEというコマンドを用いる。この調整方法は、カナダ統計局が開発したもので、 λ と ρ の2つのパラメーターを用いる回帰式を利用する方法である。ここで λ は、回帰方程式のウェイト・マトリックスを決めるのに使われ、 ρ は1次のARモデルの係数の値を示す。この方

法のより詳しい説明は、X-12-ARIMA Reference Manual version 0.3のpp. 100-102で説明されている。

Monsell (2007) p. 4では、この機能について、次のような注意事項が述べられている。

原数値と季節調整の年合計を一致させることは、ある目的にとっては便利なものである。しかし、季節変動要因が乗法的 (multiplicative, 季節変動要因が傾向・循環変動に対して掛け算の形で結びついている) な場合や、季節変動要因が時間の経過につれて変化していく場合には、季節調整の品質 (quality) を低下させるものであることに留意する必要がある。

この季節調整値の年合計値を原数値の年合計値と一致させる新しい機能を、筆者は4. (3)と同様に日本の大口電力使用量のデータ (データ期間は同じ) に適用してみた。この機能を使うためのforceというコマンドを、X-12-ARIMAでX-11を用いて季節調整を行うスペック・ファイルにおいて、

```
x11 {
    }
force {start=january type=regress }
```

というように、x11のコマンドの下に書き加えることにより、この機能を実行することができる。ここで、forceのかけこ内に書いてあるstart=januaryは、年合計値を計算する時の最初の月が1月であることを示す。すなわち、1月から12月までのデータを合計した暦年のデータで、季節調整値の合計が原数値の合計と同じ大きさに調整するという指示である。もし、日本の年度 (4月から翌年3月までの1年間) で年合計値を調整する場合であれば、start=aprilとなる。

次のtype=regressは、先に述べたカナダ統計局が開発した回帰式を用いる調整方法を用いるということである。typeで指定できる調整方

法には、この外にDenton法と呼ばれる方法もtype=Dentonとして指定できる。この方法は、季節調整値の年合計値と原数値の年合計値の差を、季節調整の結果をなるべくゆがめないように、毎月のデータに配分する方法であるが、Monsell氏はあまり推奨していない方法である。

上に述べた、暦年の年合計値で調整し、方法としては回帰式を用いる方法で、大口電力使用量のデータに適用した結果を表5に示した。

表5 大口電力使用量のデータにforceを適用しない場合と適用した場合の季節調整値の比較

年(暦年)	原数値 数 値	季節調整値			
		forceなし		forceあり	
		数 値	誤差率	数 値	誤差率
2000	266,707	266,216	-0.18	266,707	0.00
2001	259,858	260,119	0.10	259,858	0.00
2002	258,833	258,897	0.02	258,833	0.00
2003	260,847	261,069	0.09	260,847	0.00
2004	268,389	267,809	-0.22	268,389	0.00
2005	271,784	271,874	0.03	271,784	0.00
2006	283,658	283,812	0.05	283,658	0.00
2007	296,027	296,117	0.03	296,027	0.00

(備考) 1. 数値の単位は百万kwh

2. 誤差率は原数値に対してどの程度相違しているかの割合(%)

この表では1990年から2007年までの結果のうち、2000年から2007年までの結果だけを示した。これを見ると、forceコマンドを用いないで年合計値の調整をしなかった場合には、季節調整値の年合計値は原数値の年合計値に対して0.02から0.22%までの誤差率で誤差があることがわかるが、forceコマンドを用いて年合計値を調整した場合は、6桁のデー

タで見える場合、誤差率0.00%で完全に一致している。

6. おわりに

本稿では、X-12-ARIMAのversion 0.3で付け加わった新しい機能について、ARIMAモデルの自動選択機能を中心に見てきた。この自動選択機能は、スペイン銀行のVíctor GómezとAgustín Maravallが開発した季節調整プログラムであるTRAMO-SEATSの中のTRAMOで採用されている方法を基本的に取り入れたものである。ただし、取り入れるに当たっては、ある程度の修正が行われているので、元のTRAMOの方法が完全にそのまま取り入れられているわけではない。

この新しいversionのX-12-ARIMAは、このARIMAモデルに自動選択機能やその他の改良点加わって、ユーザーは従来よりも高度な手法が便利に使えるようになった。今後、こうした機能を日本経済のデータについて実際に適用されて、その結果について検討を行った上で、季節調整の手法がさらに改善されていくことが期待される。

【参考文献】

日本語文献

- [1] 東 晋司, 2003年「季節調整プログラムX-12-ARIMAとTRAMO-SEATSの分解構造——試験的な実証分析を添えて——」ESRI Discussion Paper Series No. 63, 内閣府経済社会総合研究所。
- [2] 奥本佳伸, 2000年『季節調整法の比較研究——センサス局法X-12-ARIMAの我が国経済統計への適用』（経済企画庁経済研究所「経済分析 政策研究の視点シリーズ」17）大蔵省印刷局。
- [3] 山本 拓, 1988年『経済の時系列分析』創文社。

英語文献

- [1] Box, George E.P. and Gwilym M. Jenkins, 1970, *Time Series Analysis: Forecasting and*

Control, Holden-Day, San-Francisco.

- [2] Dent, Ayonda M., Catherine C. Harvill Hood, Kathleen M. McDonald-Johnson and Roxanne M. Feldpausch, 2007, "Comparing the Automatic ARIMA Model Selection Procedures of X-12-ARIMA Version 0.2 and 0.3," (<http://www.census.gov/srd/www/sapaper/sapaper.html>).
- [3] Gómez, V. and Maravall, A., 1997, "Programs TRAMO and SEATS: Instructions for the User, Beta Version," Banco de Espana, Madrid.
- [4] Monsell, Brian C., 2007, "Release Notes for Version 0.3 of X-12-ARIMA," U.S. Census Bureau.
- [5] Schwarz, G. 1978, "Estimating the Dimensions of a Model," *Annals of Statistics*, 6, 461-464.
- [6] Time Series Staff, Statistical Research Division, 2002, *X-12-ARIMA Reference Manual, Version 0.2.10*, U.S. Census Bureau.
- [7] Time Series Staff, Statistical Research Division, 2007, *X-12-ARIMA Reference Manual, Version 0.3*, U.S. Census Bureau.

(2009年12月2日受理)