

CIRJE-J-47

解説 X-12-ARIMA2000

東京大学大学院経済学研究科

国友直人

2001年3月

このディスカッション・ペーパーは、内部での討論に資するための未定稿の段階にある論文草稿である。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。

On X-12-ARIMA2000

Abstract :

This is a tentative Japanese translation of the first part (Chapter 1 ~ Chapter 5) of the X-12-ARIMA Manual with two short appendices on the seasonal adjustment programs. Because the X-12-ARIMA program developed by the time series research group of the U.S. Census Bureau uses the statistical time series analysis extensively, it may be helpful for practitioners in Japan to use or understand the X-12-ARIMA program although the evaluation of the X-12-ARIMA program is still under discussion. We shall translate the remaining parts of the X-12-ARIMA Manual in the near future hopefully.

解説 X-12-ARIMA2000 *

国友直人[†]

2001年1月
(暫定版)

本解説の主旨

この解説稿は米国センサス局 (U.S.Census Bureau) 発行の X-12-ARIMA Reference Manual (2000年5月16日に発行された Version 0.2.7) 1章～5章の翻訳版に季節調整法と季節調整プログラム利用に関する短い付録をつけたものである。原マニュアルの6章と変数表はとりあえず省略したので命令文における変数の指定等の細部については原マニュアルを参照されたい。

センサス局時系列研究スタッフにより開発された季節調整法 X-12-ARIMA には統計の時系列解析をはじめ数理統計的手法が多く取り入れられているので、こうした分野についての専門的予備知識が欠けている場合にはその利用は必ずしも容易でない。季節調整法としての X-12-ARIMA 法の妥当性の評価についてはなお様々な意見があり得るが、官庁統計家など実務関係者にとりまづはその正確な理解とより適切な利用に資することが必要と判断してこの解説稿を作成した。出来る限り原文の意図の正確な理解を期したつもりであるが、誤解等についてのコメントを歓迎する。

*オリジナル・ファイル X-12-ARIMA Reference Manual (Version 0.2.7 のオリジナル TEX ファイル) を提供してくれた米国センサス局のデビット・フィンドレー博士 (Dr. David Findley) とブライアン・モンセル博士 (Dr. Brian Monsell) のご好意に特に感謝する。またこの解説原稿を整理してくれた東京大学経済学部学生の一場知之君に感謝する。

[†]東京大学経済学部

目次

1. はじめに	4
2. X-12-ARIMA の起動	6
2.1 入力 (Input)	7
2.2 出力 (Output)	7
2.3 入力エラー	8
2.4 出力ファイル名の指定	8
2.5 複数系列への X-12-ARIMA 起動法	8
2.6 ログ・ファイル (Log Files)	11
2.7 フラッグ (Flags)	11
2.8 プログラム制約 (Program limits)	14
3. X-12-ARIMA における RegARIMA モデル	14
3.1 モデルの一般型	15
3.2 入力データと変換	16
3.3 回帰変数の特定化	16
3.4 ARIMA モデルの識別と特定化	19
3.5 モデルの推定と統計的推測	20
3.6 外れ値を含む診断	21
3.7 予測 (Forecasting)	22
4. RegARIMA モデル推定の関連事項	23
4.1 母数推定の初期値設定と収束	23
4.2 MA 作用素の可逆性	24
4.3 AR 作用素の定常性	24
4.4 AR・MA 要素の共通因子と過剰階差問題	25
4.5 モデル選択基準	26
5. スペック・ファイルと書式	27
5.1 印刷と保存	31
5.2 日付	33
5.3 入力ファイルの書式	33
引用文献	36
付録	40

A. X-12-ARIMA(2000) とは？	40
B. 季節調整法小史	40

1. はじめに

季節調整プログラム X-12-ARIMA はシスキン・ヤング・マスグレーブ (1967) により開発されたセンサス局 II-X11 法の改良版である。より分かりやすく X-11 法を改良するとともにユーザーが様々なオプションをもとに季節性や曜日効果を検出し調整する新たな診断方法を提供している。さらにプログラムでは様々な新たな統計的方法を用いて季節調整の問題を解決し、様々な経済データを適度に調整することを可能にした。こうした方法の応用例については Findley=Hood (1999) が説明している。

この新しい方法は主として我々が RegARIMA モデルと呼んでいる統計的時系列分析用モデルをプログラムの中で広範に用いることにより実現されている。この RegARIMA モデルは ARIMA (自己回帰和分移動平均) 過程にしたがう誤差項を持つ回帰モデルである。より正確に云えば、この RegARIMA モデルでは時系列 (あるいはその対数変換値) は回帰変数の線形結合で表され、共分散構造は ARIMA モデルで決められると仮定されている。もし回帰変数を全く用いなければ、期待値ゼロの仮定の下では RegARIMA モデルは ARIMA モデルに帰着される。プログラム X-12-ARIMA の中にはフローやストック系列の取引日効果や休日効果を直接的に推定する回帰変数が組み込まれている。さらに時系列におけるある種の突然の攪乱やレベルの突然の変化がある時には X11 プログラムを使って季節調整を行なう前に回帰変数を用いて処理する。またプログラムで提供されていない問題に対処するためにはモデルのフィットに際してユーザーが自ら回帰変数を定義して利用することも可能である。X-12 プログラムにおける RegARIMA モデル部分はセンサス局の時系列研究グループが開発した RegARIMA プログラムにもとづいている。

時系列の分析において回帰変数の利用が必要となることがあってもなくても、予測に際して RegARIMA モデルを利用することが X-12-ARIMA におけるもっとも重要な改善点であり、この方法により最新時点 (あるいは初期時点) の季節調整値を改善することが可能となっている。時系列の末端部分の扱いにおいて X11 法で用いられているトレンド推定や非対称の季節移動平均過程にまつわる問題の扱いを改善している。こうした重要な技術的改善の試みとしてはカナダ統計局によって開発され利用されている X-11-ARIMA 法が既に知られている。また、こうした方法に関する理論的あるいは実利的利点に関してはこれまでに多くの文献、例えば Geweke (1978), Dagun (1988), Bobbitt=Otto (1990) あるいはそれらが引用している論文もある。

X-12-ARIMA プログラムはパーソナル・コンピュータ (386 あるいはそれ以上の数学演算装置を備えていることが必要) で計算するプログラムとして利用可能であり、DOS (3.0 版以上)、サン・ユニックス・ワークステーション (Sun 4 UNIX)、汎用大型計算機 (VAX/VMS) により実行可能である。他の計算機システムで X-12-ARIMA を実行させるプログラムを作るにはフォートラン (FORTRAN) ・ソース・コードを利用することもできる。様々な計算機で実行させることが可能なプログラムの最新テスト版を含むファイル・解説・実例などはインターネット URL

<http://www.census.gov/srd/www/x12a/>

に保管されているので利用者は直接にダウン・ロードすることができる。あるいは利用者は ftp を用いて—pub/ts/x12a—を—ftp.census.gov—からコピーできる。(その際、ログインは “—anonymous—” として自分のメール住所をパスワードとするだけでよい。) また限定された範囲内であるが、通常の郵便や電子郵便による問い合わせに対してセンサス局からのプログラム支援も提供されている。もし特定の入力ファイルを実行させた時に問題が生じたなら入力ファイルと出力ファイルを提供すると問題の所在を識別される可能性が高くなる。

このプログラムの季節調整部分ではシスキン・ヤング・マスグレーブ (1967) 及び Dagun (1988) で詳しく説明されている X11 季節調整法を用いている。した

がって X-12-ARIMA では X11 法と X11 プログラムにおけるすべての季節調整方法を実行することができ、X11 プログラムと同一の季節移動平均やトレンド移動平均を実行することができるとともに、X11 法における曜日効果や休日効果の除去方法も実行することもできる。

季節調整部分では幾つかの追加的なオプションを付け加える改善がなされたが、次のような新しいオプションがある。

- (a) 期間変更の安定性 (sliding spans) 法、この分析方法については Findley=Monsell=Shulman=Pugh (1990) が例示している。
- (b) ある系列についての季節調整値の改訂の履歴を計算する能力。
- (c) 新しいヘンダーソン型フィルター法により任意の奇数項ヘンダーソン型フィルターを利用者が選択できるようになったこと。
- (d) 季節フィルターの新しいオプション。
- (e) 季節調整値の不規則要素に対する新しい異常値処理オプション。
- (f) カレンダー上の曜日の要因表。
- (g) 擬加法的季節調整法。

X-12-ARIMA では季節性の伴う経済時系列に対して RegARIMA モデルを用いて統計的分析を行なっている。この目的の為に幾つかの回帰変数を提供しているが、例えばトレンド係数、全平均、固定季節効果、曜日効果、休日効果、一時的効果 (加法的外れ値)、水準の変化、一時的変化、傾斜変化などがある。さらに利用者が回帰変数を定義してモデル分析に利用することも容易に可能である。こうしたプログラムは RegARIMA モデル分析を行なうために作られたのでより一般的な統計計算パッケージ用ではないことに注意しておく。特に X-12-ARIMA は時系列の高性能プロットが可能なグラフィック・ソフトとともに用いられるべきであろう。

この X-12-ARIMA プログラムによりモデル分析や季節調整が可能な観測値 (データ) は量的変数のみであり、二値変数やカテゴリー変数を扱うには適当ではない。観測値は一定の間隔を持った時系列であることが必要であり、欠則値があることは許されない。またプログラム X-12-ARIMA は一変数時系列モデルのみを扱うことができ、異なる時系列間の関係を推定することはできない。

プログラム X-12-ARIMA では季節 ARIMA モデルに対して標準的な記号 $(p d q) \times (P D Q)_s$ を用いる。ここで $(p d q)$ はそれぞれ非季節的自己回帰 (AR) 作用素、階差、移動平均 (MA) 作用素を表している。 $(P D Q)_s$ の方は季節自己回帰作用素、季節階差、季節移動平均作用素を表している。

標準的な統計的時系列モデルをデータにフィットしたい利用者にはプログラム X-12-ARIMA により識別・推定・診断という三つのモデル分析を行なうことができる。RegARIMA モデルを特定化して回帰変数を導入することによって誤差項に ARIMA モデル $(p d q) \times (P D Q)_s$ を持つ回帰モデルを分析に用いることもできる。どのような回帰変数を用いるかはモデル分析を行なう利用者の時系列についての知識にもとづくことになる。回帰モデルの誤差項についての ARIMA モデルの識別はプログラムで提供される標本自己相関や標本偏自己相関などを利用した標準的な方法を用いて行なわれる。RegARIMA モデルが決定されるとプログラムでは繰り返し一般化最小二乗アルゴリズム (IGLS) を用いた最尤法によりモデルの母数を推定する。さらにモデル診断においては推定した残差によりその適切さを調べることができる。プログラム X-12-ARIMA ではモデルの妥当性を調べる残差についての標準的診断法を提供するとともに、加法的外れ値、水準

の変化などについてのより複雑な方法を実行することができる。さらに推定されたモデルを用いて点予測値、予測の標準誤差、予測区間なども計算することができる。

こうしたモデル分析の方法に加えて自動モデル選択の機能や利用者が修正 AIC (赤池情報量基準 AIC をデータの長さを考慮して修正した量) を用いて回帰変数 (例えば曜日効果や復活祭の変数など) を分析に用いるか否かを定めることができる。またモデルの尤度統計量 (修正 AIC など) や予測値などの履歴を保存して、競合する統計的モデルを比較することに役立たせることもできる。

これから 5 章をつかってプログラム X-12-ARIMA の能力について詳細に説明する。2 章ではプログラム X-12-ARIMA の実行法と利用者が変更することも可能なプログラム制約について説明する。3 章ではプログラム X-12-ARIMA による RegARIMA モデルと RegARIMA モデルを用いた分析や予測などの技術的要点について議論する。4 章では統計的モデル分析を行なう利用者が知っておくべき時系列モデルの推定や統計的推測についての鍵となる事柄を説明する。5 章では入力ファイル (スペック・ファイルと呼ぶ) を説明するが、特にスペック・ファイルの記述法や関連する問題にも言及しておく。こうして 2 章から 5 章までを使ってプログラム X-12-ARIMA の機能や利用法の概略を説明する。次に 6 章はプログラム X-12-ARIMA の入力スペック・ファイルを作成するときに必要なスペック・コマンドを説明する。スペック・ファイルに書かれる可能性のある各コマンド (命令) を詳細に説明する。利用者は 6 章に書かれている説明により命令文を用いて X-12-ARIMA の実行を制御したり、プログラムで用意されているオプションの選択することができるよう。

2. プログラム X-12-ARIMA の実行

プログラム X-12-ARIMA をインストールする方法は計算機に依存している。この問題についての情報は本マニュアルに説明している他にインターネット ftp.census.gov 上で公開している `/pub/ts/x12a/ReadMe` の中のファイルからも得ることができる。DOS システム上のパソコン上にインストールされていればプログラム X-12-ARIMA を実行するには

```
> path\x12a path\filename
```

とすればよい。この命令文では `path\filename.spc` が入力ファイル (スペック・ファイル) となる。この時プログラム X-12-ARIMA は出力ファイルとして

```
path\filename.out
```

を作り出す。ここで `path` の指定は使用しているディレクトリー (カレント・ディレクトリー) 上にプログラム X-12-ARIMA が存在する必要がある、入力ファイルについても同様である。

ここでファイル名のみを利用者が決めることができることに注意しておく。プログラムは実行時に指定された入力ファイル名を用いてファイルを作り出す。例えばスペック・ファイルが `filename.spc` のときには出力ファイルは `filename.out` という名で書かれ、エラー・メッセージは `filename.err` として書かれることになる。したがって、スペック・ファイル `xuu1.spc` がパソコン DOS 上の現ディレクトリーに置かれていれば

```
> x12a xuu1
```

として改行キー <return> (あるいは <enter>) を押せばプログラムが実行され、ファイル `xuu1.out` と `xuu1.err` がディレクトリー上に作られることになる。

入出力のプログラムは以下で簡単に言及しておくが後でより詳細に説明することとなる。UNIX システム上でプログラムを実行させるにはこれまでのスラッシュを適切に変更するだけでよい。またOSとしてDOS、UNIX、VAX/VMSの為の簡単なマニュアルも利用可能であり、プログラム X-12-ARIMA を実行するより詳しい説明がなされている。

2.1 入力 (Input)

プログラム X-12-ARIMA を実行するには入力ファイル (スペック・ファイルと呼ぶ) を作成する必要がある。このファイルはアスキー形式 (あるいはテキスト形式) でありプログラムを実行する上で必要な時系列、実行したい分析、必要な出力についての情報が要求される。プログラムはスペック・ファイルが拡張子 .spc の形式で書かれていることを想定している。したがって、例えば入力ファイルは *path\filename* となっていればよい。

スペック・ファイル以外でプログラム X-12-ARIMA の入力ファイルとなりうるのは時系列データを含むオプション・ファイルであり、利用者が定義する回帰変数と事前調整要素と自動モデル選択に関するモデル型などがそれにあたる。こうしたファイル・コマンドはスペック・ファイルの中で適当な形式で書かれていなければならない。こうした付加的な入力ファイルが利用可能になったのでスペック・ファイル上で利用可能なデータ値を含めたり自動モデル選択の為のモデルを自動的に設定することができる。

2.2 出力 (Output)

標準的には出力ファイルは *path\filename.out* として書かれる。個々のスペック・ファイルの内容に対応して出力ファイルが作られるが、その場合には 5 . 1 節に言及される出力印刷 (print) 変数が使われる。コマンドの保存 (save) では例えば残差系列を保存し、グラフ・プログラムでデータをプロットするなど更なる分析の為に他の出力ファイルを作り出す為に使うことができる。ここで保存 (save) を利用するときにはプログラム上では 5 . 1 節に書かれてある標準的な変数名を用いた出力ファイルが作られることに注意しておく。このときもし同一の変数名を用いたファイルが存在している場合には書き換えられることになる。そこでプログラムの利用者が出力を保存する場合には注意を払う必要がある。このことについては 5 . 1 節の説明をよく読む必要がある。

2.3 入力・エラー (Input errors)

プログラムが入力エラーを発見するとその報告の為に適当なエラー・メッセージが書き出される。こうしたエラー・メッセージは *path\filename.err* という名前のファイルの中に保存される。プログラムが実行上でエラーの部分を持定化できればエラーが起きた場所 (^) をスペック・ファイルとともに出力する。そのエラーが全体に影響を及ぼす場合にはメッセージ ERROR を個々のエラー・メッセージの前に出力するが、その時にはどの様にプログラムを修正したらよいかなどのメッセージを付けることがある。メッセージ WARNING はエラーとはみなされないが注意を喚起する為のものである。

プログラム X-12-ARIMA はまずスペック・ファイル全体を読み、入力エラーがあれば幾つかのエラーとして検出する。このとき利用者は一度プログラムを動かすだけで幾つかのエラーを修正することが可能となる。しかしながら、エラーがある場合にはしばしば最初かあるいは最初の幾つかのエラーの情報だけしか実は意味がなく、はじめの部分のエラーが系列 (series) の定義等により他のエラー

を引き起こすことも少なくない。プログラムは重大なエラーを検出したときには実行を停止する。メッセージ WARNING の場合にはプログラムは停止されないが、利用者は入出力ファイルを調べ望ましい結果が得られているか否か慎重に検討することが望まれる。

2.4 出力ファイルの設定

前にも述べたようにスペックファイルを *filename.spc* とするとプログラム X-12-ARIMA では出力ファイルは *filename.out*、エラー・ファイルは *filename.err* 等となる。与えられた原系列を用いて様々な調整を行ったり RegARIMA モデルのオプションを利用したりする目的の為にしばしば入力ファイルで用いた名前ではないファイル名を使いたいことがある。出力ファイルとして他の名前を使うには

```
> path\x12a path\filename path\outname (2.1)
```

とすればよい。この場合にはスペック・ファイルは *filename.spc* であるが出力ファイルは *outname.out*、エラー・ファイルは *outname.err* として保存される。すなわちすべての出力ファイルは通常の入出力ファイル名ではなく利用者が指定するパス名とファイル名に保存されることになる。

2.5 複数系列への X-12-ARIMA の適用

実用的にはプログラム X-12-ARIMA の一度の実行で複数の原系列を処理する必要が生じることがある。複数系列に対して一度に X-12-ARIMA を実行する方法としては二つのやり方がある。

- (a) 複数スペック形式 (multi-spec mode) : ここでは個々の系列についてそれぞれ入力スペック・ファイルを用意する形式をとる。
- (b) 単一スペック形式 (single spec mode) : 各系列に対して一つの入力スペック・ファイルからのオプションを用いて実行する形式をとる。

いずれの形式でプログラムを実行する場合にもまずメタ・ファイル (*metafile*) を作る必要がある。このファイルはアスキー・ファイル形式で X-12-ARIMA が扱うことの出来るファイル名を含んでいる必要がある。X-12-ARIMA が扱う 2 種類のメタ・ファイルとしては入力メタ・ファイルとデータ・メタ・ファイルがある。

メタ・ファイルを実行中にスペック・ファイルの一つでエラーが発生するとプログラムは適当なエラー・メッセージを出力する。その際、エラーが発生したスペック・ファイルの計算は停止するが、他のスペック・ファイルはそのまま継続して実行される。すべての入力ファイルの実行で発生したエラーは X-12-ARIMA のログ・ファイル (2.7 節参照) に書き込まれる。

2.5.1 複数スペック・ファイルの実行

複数スペック形式で X-12-ARIMA を実行するにはまず入力メタ・ファイル (*input metafile*) を作っておく必要がある。このファイルはアスキー形式で X-12-ARIMA で読み込むファイル名を含んでいる必要がある。入力メタ・ファイルは 1 行ごとに 2 つまでの内容、すなわちある系列に対する入力スペック・ファイル (必要ならパス名を含む) とその系列に対する出力ファイル名を含む必要がある。(ただし後者を書きこむのはオプションである。) もし出力ファイルが利用者により指定されなければ、出力ファイルのパス名とファイル名は入力ファイル名にしたがって

指定される。また入力ファイルは入力メタ・ファイル上に書かれた順にしたがって実行される。

ここで例を用いて以上で述べたことを説明すると3つのスペック・ファイル xuu1.spc, xuu2.spc, xuu3.spc を実行させるには次のようなメタ・ファイルを作ればよい。

```
xuu1  
  
xuu2  
  
xuu3
```

ここではすべてのスペック・ファイルは現ディレクトリ 上にあると仮定している。もしこれらのスペック・ファイルが DOS 上のディレクトリー c:\export\specs に保存されている時にはメタ・ファイルは

```
c:\export\specs\xuu1  
c:\export\specs\xuu2  
c:\export\specs\xuu3
```

となる。入力メタ・ファイルを用いて X-12-ARIMA を実行するには

```
> x12a -m metafile
```

とすればよい。ここでメタ・ファイル (metafile.mta) はメタ・ファイル名であり -m は信号 (フラッグ) で X-12-ARIMA の実行にあたってメタ・ファイルの存在を示している。例えば上に定義されたメタ・ファイルが exports.mta として保存されていればスペック・ファイルを実行するには

```
> x12a -m exports
```

と打鍵して改行キーを押せばよい。

ここで上の例のように入力メタ・ファイル名が与えられるときにはファイル名のみ指定されているので、入力メタ・ファイルの拡張子は .mta とする必要がある。また必要であればパス情報も入力メタ・ファイルに付けておくことが望ましい。

X-12-ARIMA で出力ファイルとして使われるファイル名はメタ・ファイルの中のスペック・ファイルから取られるのであってメタ・ファイル名を利用するわけではない。上の例ではメタ・ファイル exports.mta 上の個別の入力スペック・ファイルに対応して出力ファイル名として xuu1.out, xuu2.out, xuu3.out が作られるが、集計した出力ファイルとして exports.out が作られるわけではない。上の例で別の名前での出力ファイルを作らせるには次のように入力メタ・ファイルの各行に希望する出力ファイル名を加えておくだけでよい。

```
c:\export\specs\xuu1 c:\export\output\xuu1  
c:\export\specs\xuu2 c:\export\output\xuu2  
c:\export\specs\xuu3 c:\export\output\xuu3
```

2.5.2 単一スペック形式による X-12-ARIMA の実行

X-12-ARIMA を用いて複数の時系列の各系列に対して同一の計算を実行させるにはデータ・メタファイル (*data metafile*) を作る必要がある。データ・メタ・ファイルでは各行は2つの内容まで含むことができ、各系列に対するデータ・ファイルの中のファイル名 (必要ならパス名) と出力ファイル名 (必ずしも必要ではない)

を書くことになる。もし出力ファイル名が利用者により指定されなければデータ・ファイルのパス名とファイル名が出力ファイルには使われる。

注意点：データ・メタ・ファイルでは個別のデータ・ファイルには拡張子を用いない。もしデータ・ファイルが当該ディレクトリーになければ拡張子がパス名とファイル名とともに指定されていなくてはならない。

データ・ファイルはデータ・メタ・ファイルの中に表われる順番にしたがい処理される。個々のデータ・ファイルについてのオプションは実行時において識別される各入力スペック・ファイルにしたがい処理される。このことはデータ・メタ・ファイルのすべてのデータ・ファイルが同一の形式となっていなければならないことを意味している。この場合には X-12-ARIMA でサポートしている幾つかの形式は避けて使わない必要があるが、こうした点について詳しくは6章のスペック・コマンドの系列 (series) で説明する。

例えば3つのデータ・ファイル xuu1.dat, xuu2.dat, xuu3.dat を一度に処理する為にはデータ・メタ・ファイルは次のようになる。

```
xuu1.dat
```

```
xuu2.dat
```

```
xuu3.dat
```

ここではすべてのデータ・ファイルは現ディレクトリーの中にあると想定されていることに注意しておこう。ここで c:\export\data の中に保存されているファイルを実行する場合にはメタ・ファイルは次のようになる。

```
c:\export\data\xuu1.dat
```

```
c:\export\data\xuu2.dat
```

```
c:\export\data\xuu3.dat
```

データ・メタ・ファイルを用いて X-12-ARIMA を実行するには命令文

```
> x12a specfile -d metafile
```

を用いる。ここで metafile.dta はデータ・メタ・ファイルであり、-d は X-12-ARIMA の利用にデータ・メタ・ファイルを用いるという信号 (フラグ) であり、specfile.spc はデータ・メタ・ファイル上の各系列で用いられるスペック・ファイルである。例えば上の例で用いられた3つの系列からなるデータ・メタ・ファイルの名が exports.dta であれば

```
> x12a default -d exports
```

と打鍵して改行キーを押せば default.spc を入力スペック・ファイルとして X-12-ARIMA が実行される。

この例でデータ・メタ・ファイル名が与えられる時にはファイル名のみが定められその拡張子は定められてはいないので、入力メタ・ファイルには拡張子 .dta が必要となる。さらに必要ならばパス名もデータ・ファイル名に付け加えることになる。

X-12-ARIMA で出力ファイルの作成に使われるファイル名はメタ・ファイル自身ではなく、メタ・ファイルの中のデータ・ファイルからとられる。上の例ではメタ・ファイル exports.dta の中の各データ・ファイルに対応した出力ファイル xuu1.out, xuu2.out, xuu3.out が作られるが、メタ・ファイル名を用いた exports.out ではない。この例で別の出力ファイル名を指定するには次のようにデータ・メタ・ファイルの各行に指定する出力ファイル名を書けばよい。

```
c:\export\data\xuu1.dat c:\export\output\xuu1
```

```
c:\export\data\xuu2.dat c:\export\output\xuu2
```

```
c:\export\data\xuu3.dat c:\export\output\xuu3
```

2.6 ログ・ファイル(Log Files)

プログラム X-12-ARIMA を実行するごとにログ・ファイル(log file) が作り出され、各系列あるいは各スペック・ファイルにたいしてモデル分析の結果や季節調整の診断結果が保存される。X-12-ARIMA が複数あるいは単独のスペック・ファイルにより前節までに説明したように実行するとログ・ファイルはメタ・ファイル(複数スペック形式)あるいはデータ・メタ・ファイル(単独スペック形式)と同一の名前とディレクトリーに拡張子.log をつけて保存される。例えば

```
> x12a -m exports
```

とすると、メタ・ファイル exports.mta の中のスペック・ファイルが実行され、ファイル exports.log の中に利用者が選んだ診断結果が保存される。ここで一系列のみが処理される場合には出力ディレクトリーと出力ファイル名はログ・ファイルを作る拡張子 .log が用いられる。

利用者はログ・ファイルに保存すべき診断結果を指定することができるが、例えばスペック命令である系列(series), 集計(composite), 変換(transform), x11, x11 回帰(x11 regression), 回帰(regression), 自動モデル(automdl), 推定値(estimate), 診断(check), 期間変更の安定性(slidingspans), 履歴(history) などではログ・ファイルや出力ファイルを指定することができる。これら個々のスペック命令とそこでの保存可能なログ・ファイルの詳細は6章で説明する。前節でも述べたようにメタ・ファイルのスペック・ファイルの実行時にエラーが発生した場合にはログ・ファイルに入力ファイルとともに書き込まれる。

2.7 フラッグ(Flags)

前節の例ではプログラムを実行する為には命令文において信号(フラッグ) `-m` と `-d` が必要であった。命令文の中で入出力オプションを指定するには他のやり方もあり得る。命令文の一般的な形式は

```
> path\x12a arg1 arg2 ... argN
```

となる。ここで x12a に続く変数は状況により信号(フラッグ・コマンド)かファイル名となる。表 2.1 には X-12-ARIMA で用いられるフラッグ・コマンドが要約されているが、本節ではこれから少し詳しく各フラッグの意味を説明することとする。このフラッグ命令文の中では順番は問題とはならないが、幾つかのフラッグではその後で適当なファイル名を指定することが必要となる。既に前節ではフラッグ `-m` と `-d` を説明しておいたが、この両方のフラッグを一緒に実行することはできない。

フラッグ `-i` は次の変数が入力スペック・ファイルのパス名とファイル名であることを示している。このフラッグは入力スペック・ファイルが最初の変数である場合には必要ない。したがって二つの命令文 `x12a test` と `x12a -i test` は同等である。二つのフラッグ命令 `-i` と `-m` は同時に実行できない。

フラッグ `-i` と同様にフラッグ `-o` は次の変数が出力のためのパス名とファイル名であることを示す。前に説明した拡張子(.out と .err) やコマンドの保存(save)に関する拡張子などもファイル名に付け加えることができる。このフラッグは入力スペック・ファイルが次に出力ファイル名が2行目(式 2.1 のように)ならば必要ない。したがって次に挙げる命令文はすべて同等となる。

```
x12a test test2
x12a -i test -o test2
x12a -o test2 -i test
```

ここで `x12a -i test test2` とすると、最初の変数がスペック・ファイルでなくフラグ `-i` なのでエラーとなる。二つのフラグ `-o` と `-m` は同時に実行することはできない。

フラグ `-s` は出力の中のある季節調整と RegARIMA モデルの診断結果を主出力ファイルと分けたファイルとして保存する為に必要となる。この結果は出力の時系列表だけでなく出力ファイルの中の表も含んでいるが、必ずしも各時系列表には使われていない形式でも保存できる。フラグ `-s` が宣言されると X-12-ARIMA は自動的に診断の重要な結果を別のファイルに保存するので診断結果を要約する際に利用することができる。このファイルは季節調整診断ファイル (*seasonal adjustment diagnostics file*) と呼ばれるが、主出力ファイルと同一のパス名とファイル名に拡張子 `.xdg` を付けたものとなる。例えばコマンド

```
> x12a test -s
```

に対する季節調整診断ファイルは `test.xdg` に保存される。またコマンド

```
> x12a test -s -o testout
```

に対する季節調整診断ファイルは `testout.xdg` となる。

季節調整診断ファイル以外にもフラグ `-s` を用いるとプログラムは RegARIMA モデル分析の結果をモデル診断ファイル (*model diagnostics file*) の中に保存する。このモデル診断ファイルは主出力ファイルと同一のパス名とファイル名を持ち拡張子は `.mdg` となる。

したがって命令文

```
> x12a test -s
```

に対するモデル診断ファイルは `test.mdg` に保存され、命令文

```
> x12a test -s -o testout
```

に対するモデル診断ファイルは `testout.mdg` に保存される。

季節調整診断ファイルについてはインターネット上で `ftp.census.gov` に一つのプログラムが公開されていて、それを使うと季節調整診断ファイルを読んだり要約したりすることができる。そのプログラムはプログラム言語 Icon により書かれている。(Griswold=Griswold (1997) を参照されたい)。

フラグ `-g` は次の変数名がディレクトリーのパス名であって別のグラフィクス・プログラムの入力としてその出力を用いることを意味している。この出力には次のようなファイルがある。

- (1) グラフにする診断ファイルでスペック・ファイル (`.spc`) の中のオプションにより作られるもの、
- (2) ファイル名を含むグラフ・メタ・ファイル、
- (3) 実行される時系列や季節調整の型の情報を含む季節調整・診断・ファイル、
- (4) 時系列に当てはめる RegARIMA モデルの情報を含むモデル診断ファイル。

グラフィック・メタ・ファイルの拡張子は `.gmt`、季節調整診断ファイルの拡張子は `.xdg`、またモデル診断ファイルの拡張子は `.mdg` である。こうしたファイルのファイル名はプログラムの主出力ファイル名が用いられる。例えば利用者が

```
> x12a test -g c:\sagraph
```

と入力すると、グラフィック・メタ・ファイルは c:\sagraph\test.gmt に保存され、季節調整診断ファイルは c:\sagraph\test.xdg に保存され、モデル診断ファイルは c:\sagraph\test.mdg にそれぞれ保存される。命令文

```
> x12a test -g c:\sagraph -o testout
```

に対してはグラフ・メタ・ファイルは c:\sagraph\testout.gmt に保存され、季節調整診断ファイルは c:\sagraph\testout.xdg に、さらにモデル診断ファイルは c:\sagraph\testout.mdg にそれぞれ保存される。この二つの例ではともに季節調整のグラフィックに必要なファイルはディレクトリー c:\sagraph に保存される。ここでの注意点としてはフラッグ -g が入るディレクトリーが出力ファイルとは別に作られている必要があることであるが、同一ディレクトリーの副ディレクトリーであってもよいことである。

グラフィック形式の出力を作るグラフ・ソフトとして SAS/GRAPH(SAS) を使うプログラム X-12-Graph がある。(これについては Hood (1998a)、Hood (1998b) を見よ。) グラフィック形式で X-12-ARIMA の実行により保存されるファイル一覧とそれらのファイルを記述するグラフィック・メタ・ファイルで用いられるコードについては表 2-2 を見ればよい。

このフラッグ -g を用いて作られた季節調整診断ファイルは外部のグラフ・ソフト SAS/GRAPH に必要な季節調整の重要な情報のみを保存するものであり、フラッグ -s オプションを使うとファイルの一部分のみ保存される。ここで両方のオプションを同時に使うと季節調整診断ファイルすべて(主出力ファイルではなく)は -g オプションにより指定されたディレクトリーに保存される。モデル診断ファイルが作られればそのファイルはグラフ・ディレクトリーに保存される。利用者が季節調整診断ファイルを使うとある種の警告がスクリーン上に表われ、またモデル診断ファイルの場合もグラフィック・ディレクトリー警告が表われる。

フラッグ -n, -w, -p はどれもプログラムの主出力を制御している。オプション -n により利用者は主出力ファイルの表数を制限することができる。プログラム X-12-ARIMA は主出力としてかなりの数の表を作り出すが、X-12-ARIMA では利用者がその出力表を柔軟に決められることができるので、時には幾つかの表に絞った方が便利となる。このことを簡単に実行するには例えばフラッグ -n を使えばよく、X-12-ARIMA デフォルトとしての表は主出力ファイルには出なくなり、利用者がスペック・ファイルに書いた表のみが出力されることとなる。

フラッグ -w は主出力ファイルの表を横幅のある(132文字)に指定する為に用いられる。このフラッグがなければデフォルトは80文字の形式である。表の正確な文字数は系列の大きさや系列(series)での精度の指定に依存して決まる。フラッグ -p により主出力ファイルにおいて改ページやヘッダーを省略することができる。このオプションを用いなければ改ページが各出力表の先頭に挿入され、実行タイトル、系列名、ページ数も挿入される。

フラッグ -c は集計季節調整を入力メタ・ファイル(-m)に制限するときのみ使われる。集計季節調整では X-12-ARIMA では通常は構成要素の時系列とともに合成系列(集計量と呼ばれる)の季節調整を行なうが、通常は足し算で集計される。(その詳細は6章のコマンドの集計(composite)を参照されたい。)この命令を使うときには各系列についてスペック・ファイルが必要となる。フラッグ -c を使うときには各系列に対する入力ファイルの中に指定されている季節調整やモデルのオプションは無視され、各系列は集計系列を作る為だけに使われる。このオプションでは集計系列に対する RegARIMA モデルを識別するために利用される。

最後にフラッグ -v により X-12-ARIMA を実行するに際しては利用者が入力確認モードを使ってエラーが一つあるいは複数の入力スペック・ファイルにあるか否かを確かめることができる。この方法で利用者は全系列に X-12-ARIMA を実行させることなくエラーのあるプログラム・オプションを調べることができる。

このフラグ $-v$ は他のフラグ $-s, -c, -n, -w, -p$ などと同時に用いることは出来ない。

2.8 プログラム制限

プログラム X-12-ARIMA には系列の最大数、回帰変数の最大数等々の制限がある。こうした制限の数値は大多数の応用上には十分大きいですが、計算機への負荷があまりに大きくなったり膨大なメモリーを使いすぎたりした結果、プログラムの実行時間が長くなるのを防ぐ為に設定されている。もし必要ならばこうした制限は修正することが可能であるが、そのためにはフォートランで書かれた原プログラムを変換するなどして新たな制限を設定する必要がある。こうした修正を行なう為のパラメータ設定はファイル `model.prm` と `srslen.prm` の中に書かれている。プログラムの制約を修正するときにはオリジナル・ファイルを特にバックアップしておくことを推奨する。表 2-3 には利用者が修正するプログラムの制限に対応したパラメータ変数はファイル `model.prm` と `srslen.prm` の中にある。

3. RegARIMA モデル分析

3.1 節ではプログラム X-12-ARIMA で扱う一般的な統計的モデルを説明する。3.2 節から 3.7 節では RegARIMA モデル分析と予測の様々な段階、すなわちデータ入力、変換、回帰、説明変数選択、ARIMA モデルの識別と選択、モデルの推定と統計的推測、外れ値を含む診断、予測、などにおける X-12-ARIMA の能力について説明する。こうした課題の実行を制御する入力（スペック・ファイルの書き方）についても以下の節で説明する。スペック・ファイルのより細かなことについては 6 章で説明する。

RegARIMA モデルを作る際には高性能の時系列プロットを用いることを強く勧めたい。プロットすることにより季節性のパターンをはじめ外れ値の可能性や確率的非定常性といった貴重な情報が得られる。さらに系列の変換や様々な階差変換などの影響を調べることに有力な情報を提供する。X-12-ARIMA 自体はそうしたプロット能力を持っていないので他のソフトウェアを用いる必要がある。

3.1 モデルの一般型

季節性をもつ時系列についてもボックス・ジェンキンス (1976) により議論されている ARIMA モデルはしばしば利用されている。時系列 z_t に対する乗法型季節 ARIMA モデルの一般型は

$$\phi(B)\Phi(B^s)(1-B)^d(1-B^s)^D z_t = \theta(B)\Theta(B^s)a_t \quad (1)$$

と表現することができる。ここで B は遅れ作用素 (backshift operator $Bz_t = z_{t-1}$)、 s は季節周期、 $\phi(B) = (1 - \phi_1 B - \dots - \phi_p B^p)$ は非季節自己回帰 (AR) 作用素、 $\Phi(B^s) = (1 - \Phi_1 B^s - \dots - \Phi_p B^{ps})$ は季節自己回帰作用素、 $\theta(B) = (1 - \theta_1 B - \dots - \theta_q B^q)$ は非季節移動平均 (MA) 作用素、 $\Theta(B^s) = (1 - \Theta_1 B^s - \dots - \Theta_Q B^{Qs})$ は季節移動平均作用素、 a_t 項は独立・同一分布にしたがい、期待値ゼロで分散一定 σ^2 のホワイト・ノイズである。さらに $(1-B)^d(1-B^s)^D$ は次数 d の非季節階差と次数 D の季節階差を表している。もし $d = D = 0$ (階差なし) であれば、通常は (1) 式の z_t をその平均からの差、すなわち $z_t - \mu$ (ここで $\mu = E[z_t]$ とする) で置き換える。

この ARIMA モデルを拡張して時間に依存する期待値を回帰モデルを組み込むことで有益な統計的モデルを得ることができる。より具体的には時系列 y_t を

$$y_t = \sum_i \beta_i x_{it} + z_t \quad (2)$$

と表現する。ここで y_t は (被説明) 時系列であり、 x_{it} は y_t とともに観察可能な回帰変数、 β_i は回帰母数、 $z_t = y_t - \sum \beta_i x_{it}$ であり、回帰の誤差項は (1) 式で与えられる ARIMA モデルにしたがうと仮定される。時系列 z_t について ARIMA モデルを利用するので標準的な回帰分析を時系列に用いることによる基本的問題が生じる。標準的回帰分析では回帰誤差項 (すなわち (2) 式の z_t) は時間について無相関であることが仮定される。ところが時系列データでは (2) 式の誤差項は通常は自己相関がある上にしばしば階差操作を行なう必要もある。そうした状況で確率変数列 z_t が互いに無相関であることを仮定すると、通常は適切でない結果をもたらしがちである。

プログラム X-12-ARIMA において用いられる RegARIMA モデルは (1) 式と (2) 式で定められる。この二つの式を組み合わせると統計的モデルは

$$\phi(B)\Phi(B^s)(1-B)^d(1-B^s)^D \left(y_t - \sum_i \beta_i x_{it} \right) = \theta(B)\Theta(B^s)a_t. \quad (3)$$

と表現されることになる。この RegARIMA モデルは (1) 式で与えられる純粋な ARIMA モデルに回帰関数 ($\sum \beta_i x_{it}$) を加えるか、あるいは (2) 式で与えられる回帰モデルの誤差項 z_t に (1) 式で与えられる ARIMA モデルを加えることにより得ることができる。いずれにしても RegARIMA モデルではまず時系列 y_t から回帰効果を引き去り、期待値ゼロの系列 z_t を得た後に、誤差系列 z_t を階差操作することにより定常時系列 w_t を求めることになる。この確率変数列 w_t は定常 ARMA モデル $\phi(B)\Phi(B^s)w_t = \theta(B)\Theta(B^s)a_t$ にしたがうと仮定される。したがって (3) 式で与えられる RegARIMA モデルを別の形で表現すると

$$(1-B)^d(1-B^s)^D y_t = \sum_i \beta_i (1-B)^d(1-B^s)^D x_{it} + w_t. \quad (4)$$

となる。ここで w_t は定常 ARMA モデルにしたがっている確率変数列である。(4) 式の形では RegARIMA モデルにおける回帰変数 x_{it} や原系列 y_t に対して階差作用素 $(1-B)^d(1-B^s)^D$ を用いる ARIMA モデルを強調した形式となっている。

ここで (3) 式で表現された RegARIMA モデルでは回帰変数 x_{it} は同時点の被説明変数系列のみに影響すると仮定していることに注意しておく。すなわちモデル (3) では $\beta x_{i,t-1}$ など遅れをもつ効果は明示的には考慮されていないのである。むろん遅れのある回帰効果を X-12-ARIMA に含めることは可能であるが、それには利用者がラグ (遅れを持つ) 変数を定義する必要がある。

プログラム X-12-ARIMA では RegARIMA モデルの ARIMA 部分を定式化する際に様々な柔軟性を備えているが、例えば次の事などが挙げられる。(i) 二つ以上の乗法的 ARIMA 因子、(ii) AR 多項式と MA 多項式でのラグの一部分の省略、(iii) モデルを推定した時に各 AR 項と MA 項の母数を利用者が特定の値に設定できること、(iv) トレンドの定数項を導入し、階差系列 $((1-B)^d(1-B^s)^D y_t)$ に対してゼロでない平均を設定する。こうした RegARIMA モデルの定式化についての詳細は 6 章で議論し、かつ例も示される。

時系列 ARIMA モデル分析自体に関してはボックス・ジェンキンス (1976) の古典的研究に詳しく説明されているが、他の統計的時系列分析の教科書、例えば Abraham=Ledolter (1983), Vandaele (1983) などでも説明されている。また Bell (1999) は特に RegARIMA モデルについて議論しているが、ちょうどプログラム X-12-ARIMA の議論に適している。

3.2 データ入力と変換

分析が必要な原系列の観察値は系列 (series) を使ってプログラムに読み込まれる。データは系列 (series) に含まれるか、あるいは別のファイルとして置くかいずれかの形をとる。分析に使うデータの長さを制御するには期間 (span) やモデル期間 (modelspan) を使うが、原系列の最初や最後の幾つかのデータを分析に含めないことも可能である。系列 (series) を使って初期時点、(必要なら) 季節周期、時系列のタイトルを指定することもできる。

コマンドの変換 transform によりデータの非線型変換及び事前要素による修正を行なうことができる。利用できる非線型変換として対数変換や平方根変換などボックス・コックス (Box-Cox) 変換やロジット変換 (0 と 1 の間の時系列に役立つ) などが含まれている。事前調整を定義すると月の長さ (length of month) (あるいは四半期系列の場合には四半期の長さ (length of quarter) により観測値を月次系列を分割したり、月次 (四半期) 平均でスケール変換することも可能である。さらに、2 月についてはうるう年要素により調整することもできる。また一連の利用者が定義することにより原系列に割ったり引いたりする事前調整を行なうこともできる。

3.3 回帰変数の指定

RgARIMA モデルを定式化するには回帰変数 ((2) 式の x_{it}) 及び回帰の誤差項 z_t に対する (1) 式で与えられる ARIMA モデルを定める必要がある。最初の問題は回帰 (regression) を使って行ない、後者はアリマ (arima) (3.4 節で議論する) を用いて行なう。どのような回帰変数を選択するかは利用者がモデル分析しようとする時系列についての事前知識を使って定めることになる。経済季節時系列のモデル分析を行なう際にしばしば利用されている回帰変数は X-12-ARIMA の中に組み込まれているので簡単に利用できる。その方法についてはこれから説明するが、回帰変数リストは本章の表 3.1 にまとめられている。これらの変数の定義と用法については 6 章の回帰 (regression) の項目で説明される。さらに利用者が入力するデータ変数 (利用者の定義する回帰変数) を使うこともできる。統計モデルの推定 (3.5 節) に際しては X-12-ARIMA は各説明変数の統計的有意性を評価する t -統計量や曜日効果など) 特別の効果全体に対する回帰係数の有意性を χ^2 -統計量を使って調べている。

もっとも基本的な回帰変数としては定数項がある。もし ARIMA モデルが階差操作を含まなければ定数項は通常の切片をあらわす回帰変数であり、他の変数を含まなければ定常時系列の期待値 (平均) を表現する。もし ARIMA モデルが階差操作を含んでいれば (4) 式で与えられる ARIMA モデルにもとづく階差に対して定数 1 からなる変数をプログラム X-12-ARIMA は作り出す。この変数に対する母数はトレンド定数と呼ばれるが、これは階差の次数と同じ多項式トレンドを作るからである。例えば非季節階差が正 ($d > 0$) で季節階差がゼロ ($D = 0$) とすると、もとの多項式トレンドは変数 t^d に比例する。ここで注意しておく必要があるのはより低次の多項式 t^j ($0 \leq j < d$) は階差 $(1 - B)^d$ をとるとゼロになり係数を推定できないので回帰変数に含まれないことである。むしろ、トレンド定数が回帰に含まれるか否かにかかわらずモデル (3)(あるいは (4)) では階差操作をとおして低次の多項式トレンド項を含んでいる。ここで季節階差操作 ($D > 0$) を含む時には原系列でのトレンド定数項はより複雑になり、次数 $d + D$ の多項式を含むことを意味する。トレンド定数を含まなければ次数 $d + D - 1$ の多項式トレンドを含んでいることを意味している。

月次系列においては固定的季節効果 (Fixed seasonal effects) は各月に対応する 12 個のダミー変数によって表すことができる。ただし、こうしたダミー変数は和は 1 になるので全平均効果と区別できない。このことから 2 つの問題が生じ

る。階差なしのモデルでは定数項との多重共線性、12個の変数を階差を取ると和がゼロになるので階差モデルでは積率行列の非正則性が生じる。この場合のよく知られた対処法としては、12個のダミー変数の代わりに11個のダミー変数を使うことでモデルの母数を変換すればよい。あるいは別の方法としては11個の変数を使って固定的月次季節性をフーリエ（三角関数）表現の中に組み込むことも考えられる。このような母数の変換に用いられる変数は表3-1に与えられている。プログラムX-12-ARIMAではこうしたオプションも利用できるし、特定の周波数のみに三角関数を利用することも可能である。四半期系列や他の季節周期系列に対してはX-12-ARIMAにより必要な変数をつく出すこともできる。ただし、こうした変数はすべて階差をとるとゼロになる場合には季節階差を含むモデルで利用することはできない。

曜日効果 (*Trading-day effects*) は時系列の変動が異なる年の同月の曜日の構成に影響される時に生じる。曜日効果としては7個の変数でおのこの各曜日の数(月曜の数) - (火曜の数), ..., (土曜の数) - (日曜の数)をあらわす変数、月の長さ(1om)の変数やその季節性を除いた変数、あるいはうるう年変数(1pyear)などである。X-12-ARIMAでは6つの曜日変数はtdnolpyearと呼ばれている。7個目の回帰変数を追加して用いるよりもより簡単で賢く乗法的にうるう年効果を扱う方法としては原系列に対して変換 $\bar{m}_{Feb} Y_t / m_t$ (ここで Y_t は変換前の原系列、 m_t は t 月の長さ(28か29)、 $\bar{m}_{Feb} = 28.25$ は2月の平均長さ)として2月の値 Y_t のみをスケール変換してしまうことであろう。もしRegARIMAモデルが季節効果を含んでいる場合には2月の周期を除いてすでに月の長さの効果は含まれているので、曜日効果はうるう年効果のみを扱えばよいことになる。この場合には変数tdnolpyearをモデルに組み込むだけでよい。X-12-ARIMAではどちらのやり方も可能であり、さらにオプション(td)があり、月の長さの効果を自動的に扱ってくれるが、この方法については回帰(regression)の説明を参照されたい。

ここまでの説明では時系列は(観測できない)ある日次系列を月別に集計して作られていると仮定していた。このような系列は月次のフロー(*flow*)系列と呼ばれている。これに対して系列がある日次系列の月末での値となっている時にはストック(*stock*)系列と呼ばれるが、この場合には他の回帰変数を用いる方が適当であろう。月末ストック系列の曜日効果は月末値の曜日を7つのダミー変数を使ってモデル分析することが考えられよう。ここでダミー変数の和は常に1なので積率行列が非正則性となるのを防ぐ為に6個のダミー変数だけ用いればよいことに注意しておく。(表3-1を参照のこと。)X-12-ARIMAでは月末でない他に日(例えば月初め日)による計算される月次ストック・データについても適当な回帰変数を利用できる。

四半期フロー系列に対してもX-12-ARIMAでは月次系列と同一の曜日効果のオプションを用意している。ただし、四半期データにおける曜日効果は比較的まれにしか検出されない。四半期におけるカレンダーの構成は月次系列に比較して比率の意味では時間の経過とともにそれほど変化しない。月次と四半期以外の季節周期を持つフロー系列及び月次以外のストック系列に対しては曜日効果の説明変数は利用できない。

さらにX-12-ARIMAでは月次系列と四半期系列における曜日効果変動を簡単に説明する一つの回帰変数(平日・週末対比変数)

$$T_t = (\text{平日日数}) - \frac{5}{2}(\text{土曜} \cdot \text{日曜数})$$

を使うことができる。

このモデルでは月曜から金曜までの全ての曜日は同一の効果を持ち、土曜と日曜は同一の効果を持つと仮定している。X-12-ARIMAではこのモデルは次の二つのやり方で推定できる。まず利用者がオプションtdと同様に月の長さの効果をモデル分析するためにプログラムを書くのであればオプションtd1coefを規定

すればよい。また月の長さ効果を導入するオプション `tdnolpyear` とともにオプション `tdinolpyear` を使うこともできる。

月次フロー系列における 休日効果 (*Holiday effects*) は系列で計測する経済活動が休日の日付とともに変動することから生じるが、例えば (i) 経済活動が休日の日付の周辺で増加したり減少したりすること、あるいは (ii) 各年に変化する休日の日付により 2 ヶ月以上にわたり異なる効果がありうること、などが挙げられる。クリスマスのように固定された日付の休日は固定季節効果と識別可能ではない。復活祭効果 (*Easter effects*) は米国の経済時系列ではもっともよく観察されている休日効果である。というのは復活祭の日付は 3 月 22 日から 4 月 25 日まで変動するからである。レーバー日 (*Labor Day*) や感謝祭日 (*Thanksgiving*) の効果がある可能性もあるが、これらの休日効果はそれほど一般的には観察されない。X-12-ARIMA における復活祭日やレーバー日の基本モデルではある特定の整数を w として経済活動の水準がその日の w 日前から変化し、休日当日までその水準にとどまると仮定している。感謝祭日のモデルではその日の前後固定された日数だけ経済活動の水準が変化し、12 月 24 日までその水準にしたがうと仮定されている。こうした休日効果を扱う回帰変数では所与の月 t とその効果が持続する月数により構成される。(実際、表 3-1 で示されているようにこうした回帰変数は長期の全月平均からの乖離幅として基準化されている。) 原理的には同じ復活祭効果の変数は四半期フロー系列にたいしても適用可能であるが、レーバー日と感謝祭日効果の方は四半期系列には利用可能でない。X-12-ARIMA ではストック変数については休日効果に対処する変数は用意していない。

プログラム X-12-ARIMA には系列の水準の (一時的ないし恒久的) 急激な変化を扱う 4 種類の回帰変数が用意されている。これらの変数は加法的外れ値 (*additive outliers*) (AO), 水準変化 (*level shifts*) (LS), 一時的変化 (*temporary changes*) (TC), 傾斜的水準変化 (*ramps*) とそれぞれ呼ばれている。変数 AO では時系列の特定の観察値のみの変化を表現し、変数 LS ではある時点から後の全ての時系列に一定値の増加・減少することを表現し、変数 TC では系列の水準が特定時点で変化した後指数的に急速にもとの水準に戻ることを表現し、変数 ramp では一定の期間に線形的に増加あるいは減少することを表現している。こうした効果をモデル分析する為に利用できる回帰変数は表 3-1 にまとめられている。(回帰変数 LS は -1 と 0 をとるように定義し、はじめにゼロをとり後に 1 をとる変数としていないのは、全体の回帰関数の平均水準の予測値を時系列の最新時点の水準と矛盾しないようにしたことによるものである。同様にして変数 ramp も定義している。)

コマンドの回帰 (*regression*) を使うと事前知識として時系列に既知の時刻に変化があると考えられる場合に変数 AO, 変数 LS, 変数 TC, 変数 ramps などを用いて対処することができる。しかしながら、しばしば多くの季節変動では水準の変化が起こったか否かを識別することが困難であろう。そこで潜在的な外れ値の場所や性質を決めることが外れ値 (*outlier*) による外れ値の検出方法の目的となる。この問題については 3.6 節か 6 章の外れ値 (*outlier*) を参照されたい。この方法では変数 AO, 変数 TC, 変数 LS (変数 ramp は除かれる) の効果を検出し、検出した効果をそのまま回帰変数として自動的に付け加えることができる。

ここで説明した変数 AO, LS, TC, ramp はボックス・チャオ (1975) がインターベンション (*interventions*) と呼んでいた時系列分析の簡単な形である。X-12-ARIMA はボックス・チャオが動学的インターベンション効果として議論している分析をすべて扱えるわけではないが、ここで言及した変数 AO, 変数 LS, 変数 TC, 変数 ramp を適当に選んでつないでいけばかなり複雑な動学的インターベンション分析を行なうことができ、しかも各ステップで生じる追加的な母数の数は 1 か 2 の増加で押さえられる。同様にして利用者が設定する変数を含む (確率過程としての時系列) RegARIMA モデルとボックス・ジェンキンス (1976) の 10 章と 11 章で議論した動学的トランスファー関数モデルと解釈することもできよう。したがっ

て RegARIMA モデルはしばしばより一般的な動学的トランスファー関数モデルとして用いることができる。他方、トランスファー・モデルの方は一般的には確率的説明変数が未知なので予測については特別の扱いが必要という問題がある。(この問題についてはボックス・ジェンキンス(1976)の11.5節を参照されたい。)

3.4 ARIMA モデルの識別と特定化

RegARIMA モデルにおける ARIMA 部分は (1) 式における次数とその構造、すなわち 3 つ組 $(p \ d \ q)$, $(P \ D \ Q)$, s により決定される。もし回帰変数がモデルに含まなければ通常の ARIMA モデルについての次数決定(すなわち ARIMA モデルの識別と呼ばれている)を通常によく知られた方法で行なえばよい。すなわち、時系列 y_t とその階差についての標本自己相関関数 (ACF) や標本偏自己相関関数 (PACF) に基づく方法である。RegARIMA モデルでは回帰効果が存在することから ACF や PACF を見かけ上ゆがめるので、それを修正する必要がある。典型的には階差次数は時系列 y_t とその階差系列を検討してモデルを識別することができる。次に階差データの回帰変数の階差系列への回帰残差を得ることができる。こうして得られた残差の ACF と PACF により RegARIMA モデルにおける誤差項の自己回帰項 (AR) と移動平均項 (MA) の次数の識別を行なえば良い。こうした RegARIMA モデルの識別は Bell=Hillmer(1983) や Bell(1999) に議論され例示されている方法である。

RegARIMA を用いて分析方法を行なう鍵となるのはコマンドの識別 (identify) である。これを月次時系列の例で説明してみよう。階差次数を決めることはまず時系列 y_t , $(1 - B)y_t$, $(1 - B^{12})y_t$, $(1 - B)(1 - B^{12})y_t$ の自己相関関数 ACF を利用する。これは識別 (identify) を使えば一度に実行することができる。階差次数が決まれば識別 (identify) と回帰 (regression) を使って (i) 時系列 y_t の階差系列を回帰説明変数の階差系列に回帰, (ii) 回帰残差の ACF と PACF を使って自己回帰成分と移動平均成分を識別する、などを行なう。例えば 1 次の非季節階差と 1 次の季節階差を指定 ($d = 1$ かつ $D = 1$) して、識別 (identify) と回帰 (regression) を使って ARIMA モデル

$$(1 - B)(1 - B^{12})y_t = \sum_i \beta_i(1 - B)(1 - B^{12})x_{it} + w_t \quad (5)$$

を最小二乗法 (OLS) で当てはめれば (5) 式の回帰残差の ACF と PACF が計算される。

ここでもし候補となる階差 (非季節階差と季節階差ともに) の最大値をあらかじめ指定して良ければ必ずしもプログラム X-12-ARIMA を二度も実行させる必要はない。例えばこの最大階差の次数を $d = 1$ と $D = 1$ と想定してみよう。このとき識別 (identify) と回帰 (regression) を用いると (i) (5) 式に OLS により回帰を行ない母数 β_i の推定値を求め, (ii) 推定された (原系列からの) 回帰残差系列 $\tilde{z}_t = y_t - \sum_i \beta_i x_{it}$ を計算し, さらに (iii) \tilde{z}_t , $(1 - B)\tilde{z}_t$, $(1 - B^{12})\tilde{z}_t$, $(1 - B)(1 - B^{12})\tilde{z}_t$ の ACF と PACF を計算してくれる。

これまでの説明について例外が一つあるので注意しておく。回帰 (regression) で定数項を設定すると、(5) 式にもとづく最小二乗回帰では定数項は含まれるが、データから回帰効果は取り除かれない。したがって、 $\tilde{\beta}_1 x_{1t}$ がトレンド項の場合には $\tilde{z}_t = y_t - \sum_{i \geq 2} \beta_i x_{it}$ となる。このように扱う理由を (5) 式により説明すると、3.3 節で述べたように (5) 式のトレンド定数項は二次多項式を意味するが、階差操作 $(1 - B)(1 - B^{12})$ により定数項と一次トレンド項を $t^0 \equiv 1$ と t について含む形にしている。定数項と一次トレンド項の係数を推定することができないので

系列 y_t から完全な形で多項式効果を取り除くことは出来ないからである。そこで X-12-ARIMA では階差のない回帰残差 z_t を推定する際には 2 次多項式項 t^2 を取り除くのではなく推定されるトレンド定数を無視することで解決を図っている。(4) 式で表されるより一般 ARIMA モデルにおいても同様に扱われている。識別 (identify) におけるトレンド定数項を入れる効果は (4) 式の $i \geq 2$ に対する係数 β_i の回帰による推定値に影響が生じるこののみである。

3.5 モデル推定と統計的推測

回帰 (regression) とアリマ (arima) により RegARIMA モデルを定式化することができる。次に推定値 (estimate) により厳密な最尤法、あるいは条件付最尤法 (ボックス・ジェンキンス (1976) の 209 項-212 項の説明を見ればよいがしばしば条件付最小二乗法とも呼ばれている) の一種により時系列モデルの母数を推定できる。利用者は完全な厳密尤度、自己回帰 (AR) 部分には条件付尤度で移動平均 (MA) 部分は厳密尤度、あるいは両方を条件付推定、といったいずれの最大化を指定することができる。自己回帰 (AR) 母数については完全な厳密最尤推定値と条件付尤度推定値の差は一般的に小さい上にどちらの方法が良いかについての説明も可能なことがある。移動平均 (MA) 母数については厳密尤度と条件付尤度の違いはより本質的であり、厳密尤度の方が推奨される。移動平均母数の条件付最尤推定はオプションとして与えられているが、主として他のソフトウェアとの比較や厳密尤度の最大化に際して収束問題が発生した時に初期値を設定する為に利用することができる。(この収束問題については 4.1 節を参照されたい。) 特別の指定が無ければ AR 母数と MA 母数の推定では厳密最尤法を用いる。

いずれの推定方法を採用するにせよ純粋な ARIMA モデルの尤度関数は二乗和関数に帰着されるので非線形最小二乗 (More=Garbow=Hillstrom (1980) で説明されている MINPACK) プログラムにより最小化される。RegARIMA モデルの全体についての尤度関数の最大化は Burman=Otto=Bell (1987) が用いた繰り返し一般化最小二乗 (IGLS) アルゴリズムを用いている。このアルゴリズムは次の二つのステップから構成されている。すなわち

(i) 所与の AR と MA 母数にたいして回帰係数を一般化最小二乗 (GLS) 法で推定するが、その際には ARIMA モデルで決まる共分散構造を利用する。

(ii) 回帰モデルの母数 β_i の値を所与として最尤法により ARIMA モデルを回帰残差 $z_t = y_t - \sum \beta_i x_{it}$ に適用する。

このように IGLS ではこの二つのステップを収束が達成されるまで繰り返す。(ここでコマンド推定値 (estimate) のオプションを使うと希望するなら繰り返し推定の途中結果を表示することもできる。) 厳密 (あるいは条件付) 尤度関数の評価法については Box=Jenkins (1976, 7 章), Ljung=Box (1979), Hillmer=Tiao (1979), Wilson (1983) などを参考にしたが、アルゴリズムの詳細は別の技術メモに書かれている。また 4 章では全ての利用者が母数推定について注意すべき事項を述べている。

RegARIMA モデルの統計的推測については ARIMA モデルの推定及び RegARIMA モデルの推定についての漸近理論 (Box=Jenkins (1976) の 7 章, Brockwell=Davis (1987) の 8 章, Pierce (1971), Bell (1999)) を利用することができよう。漸近理論によれば適当な仮定の下では母数の推定量は近似的に正規分布にしたがい、その平均は真の母数であり、その共分散行列は推定可能である。ここで適当な仮定とは例えば正しい統計的モデルの下で自己回帰係数は定常性条件を満たし、移動平均係数は反転条件を満たし、時系列データが漸近理論が適用できるほど十分長く存在していることなどである。こうした理論的結果にもとづき X-12-ARIMA では ARMA モデルの母数推定値や回帰の母数推定値などの標準偏差や場合により ARMA 母数と回帰母数全体の推定値の相関 (共分散) 行列を計算している。

(漸近的には回帰推定値と ARMA 推定値は無相関となる。) こうした結果により 3.3 節で述べたように母数推定に関する標準的な正規分布にもとづく議論に利用することもできるし t -統計量や χ^2 -統計量を使って個別的回帰母数の有意性や回帰効果を検定するために幾つかの母数をまとめた有意性の検定に利用することもできる。これに加えて X-12-ARIMA では対数尤度関数の最大値も出力するので様々な統計モデルを推定した時に各種の尤度比検定などを行なうことが可能である。

X-12-ARIMA では誤差分散の推定量 σ^2 として $\hat{\sigma}^2 = SS/(n-d-s \cdot D)$ を用いている。ここで SS とは残差平方和であり、 $n-d-s \cdot D$ は階差操作の後の実効的な観測数である。(尤度関数が AR 母数について条件付きで計算されている場合には $n-d-s \cdot D$ は $n-p-d-s \cdot P-s \cdot D$ で置き換える必要がある。) ここでは推定された母数について自由度を調整していないことに注意しておく。このためにもプログラム X-12-ARIMA により純粋な回帰モデル(つまり ARIMA(0 0 0) モデル)を推定すると推定値 $\hat{\sigma}^2$ は通常回帰分析における分散の不偏推定値とは異なる。したがって推定の標準偏差、 t -統計量、 χ^2 -統計量なども標準的な回帰分析プログラムとは若干のずれが生じることになっている。

統計的推測についてはもう一つの別の方法として尤度にもとづくモデル選択基準を用いることが X-12-ARIMA で可能であり、AIC(赤池情報量基準)、AICC(F-修正 AIC)、Hannan-Quinn の基準、シュワルツの BIC などが計算される。これらの統計量を用いるときにはそれぞれより小さな値をとるモデルを選択するという基準となる。こうした基準よりもむしろ伝統的な基準である t -統計量、 χ^2 -統計量、尤度比検定統計量などと比較すると非入れ子型 (*nonnested*) モデル(すなわち一つのモデルの母数を取り除くなどすることにより他のモデルが導かれないモデル選択の状況で、例えば AR(1) モデルと MA(1) モデルの比較が例である)における選択が可能となりうることを挙げることができよう。ただし、モデル選択基準を用いるときにはかなり注意を払わねばならない。4.5 節では RegARIMA モデル分析で生じる状況で標準的な尤度比検定が適当でないばかりかこうしたモデル選択基準も適当でないについて言及する。

3.6 外れ値を含む診断

RegARIMA モデルの診断はモデル推定から得られる残差の分析から行なわれるが、いずれも真の残差(3)式の a_t がホワイト・ノイズと見なせるか、すなわち $N(0, \sigma^2)$ にしたがう独立・同一分布の確率変数か否かを調べることにその目的となる。(ここで分布の正規性は標本からの推定の漸近論や統計的推測には必ずしも必要でないが、予測における予測区間の正当化には極めて重要となることに注意しておく。) チェック (check) を使えば当てはめた統計モデルから残差にもとづく様々な診断のための統計量が作られる。またこのコマンドでは自己相関関数を調べる為に残差の ACF と PACF やその標準偏差推定値をともに求め、リュンゲン・ボックス (Ljung=Box(1978)) の Q-統計量も計算する。さらに X-12-ARIMA では残差の記述統計量や基準化残差のヒストグラムも計算するが、求めた残差は他のソフトウェアによるグラフィックスを含むさらなる分析の為にファイルに書くこともできる。

時系列モデルにおける診断の一つの重要な要素として外れ値の検出がある。外れ値 (outlier) を用いることで自動的に加法的な外れ値を変数 AO で、一時的変化としての外れ値を変数 TC で、さらに水準の変化を変数 LS により検出することができる。こうした外れ値のタイプ (AO, TC, LS) やそれに関連する回帰変数については 3.3 節で既に説明した。X-12-ARIMA における外れ値の検出法は基本的には Chang=Tiao(1983) にもとづいている (Chang=Tiao=Chen(1988) を見よ) がさらに Bell(1983, 1999) や Otto=Bell(1990) により拡張された方法にもよっている。その一般的な方法は繰り返し GLS 回帰と同様であるが、利用する回帰変数と

しては変数 AO, LS, TC などを用いて外れ値検出に関するすべての期間について行なう、すなわち n の長さの時系列に対して変数 AO, LS, TC の検出を実行する。実際には $3n$ 個の変数を使うのではなく少なめの変数を用いる事になるが、その理由については6章の外れ値 (outlier) の詳細 (DETAILS) に議論されている。

簡単に要約するとこの方法では各時点において計算した t 統計量が有意となるとモデルに対応する変数 AO, LS, TC を回帰変数に組み込む操作を行なう。あまりに過度の計算負荷を避けるために AR 項と MA 項の母数は各時点で外れ値に対する t 統計量を計算する際には固定しておく。X-12-ARIMA ではこうした一般的外れ値検出について二つのやり方が可能である。外れ値推定 1 (addone 法と呼ぶ) では各外れ値をモデルに加えるごとに全体のモデルを再推定する方法であり、外れ値推定 2 (addall 法と呼ぶ) では検出された幾つかの外れ値を加えるごとにモデルを再推定する手続きをとる。こうした二つの方法については6章の外れ値 (outlier) で説明されているがより詳しい形で Findley=Monsell=Bell=Otto=Chen(1997) の付録 B に説明されている。

外れ値の検出を行なっている際には残差の標準偏差についての頑健な推定値 $1.48 \times$ 残差メディアン絶対偏差 (Hampel et al. (1986)105 項による) により推定を行なっている。外れ値は有意な外れ値を全ての (あるいは一部の) 観測時点について探すので通常の正規分布に基づく臨界点 (すなわちほぼ 2.0) では外れ値の有意性検出には低すぎる。プログラムでは 3.8 を標準的臨界点としているが利用者が変えることも可能である。

モデルが複数の水準の変化を含む場合は外れ値の検出や回帰 (regression) を使うときに起こり得るが、X-12-ARIMA ではオプションとして 2 個のシフト、3 個のシフト等々が一時的水準の変化 (*temporary level shift*) として隠れていたという帰無仮説を検定する t 統計量を作り出す。もし一つの変化が別の変化と相殺していれば 2 個の連続する水準の変化が一時的水準変化を隠していることがあり、このときには二つに回帰母数の和はゼロとなる。同様のもし 3 つの回帰母数の和がゼロならば三個の水準変化が相殺することもあり得る。(こうした連続する水準変化の個数には上限がある。) こうした t 統計量は水準変化の母数推定値の和をその標準偏差推定値で割って計算される。水準の変化について有意でない t 統計量 (例えば絶対値が 2 よりも小さい) 水準の連続的变化が一時的水準変化を隠している仮説は棄却される。こうした検定により利用者はモデルにおける水準の変化のインパクトを評価することが可能となる。もちろんこうした統計量の一つか複数が有意となれば利用者は一時的レベル変化の変数の代わりに適当な水準の変化をあらわす回帰変数により再度定式化すればよい。(こうしたことは利用者が変数を定義することでも、あるいは水準の変化に対応する変数の母係数の和がゼロになるように設定しても実行可能である。) 水準の変化 (これは系列の水準が恒久的に変化することに対応する) で記述するか、あるいは一時的な水準の変化で定式化するかはデータ期間の末端部分において系列の水準変化を伴う予測と関連する重要な問題である。

3.7 予測

X-12-ARIMA により RegARIMA の母数が推定されれば予測 (forecast) により点予測値、予測誤差、予測区間を求めることができる。点予測値は真のモデルを用いていると仮定し、現在・過去の時系列 y_t の値を所与として将来の線形予測の中で平均二乗誤差 MMSE を最小化することで求められる。すなわちここでは RegARIMA モデルが正しく、正しい回帰変数を用い、追加的な外れ値や水準の変化が予測期間に起こらず、ARIMA モデルが正しく、母数推定値が真の値に一致していることを意味している。こうした仮定は標準的であるが当然ながら実際の応用では現実的ではないかもしれない。実際に期待できることとしては RegARIMA モデルが真の状況にかなり近い近似となっていて、真のモデルがわからなくともその結果が近似的に適切となっていることであろう。ここでは 2 つのやり方で予

測誤差を計算することができる。まずすべての母数の値を既知と仮定して予測値を計算することが考えられる。さらに、回帰母数を推定することから追加的な予測誤差が発生するが AR 項と MA 項の母数は既知とする方法も考えられよう。かなり長い時系列ではボックス・ジェンキンス (1976, 267 項-269 項) によれば AR 項と MA 項の推定誤差は小さいので予測の標準偏差を計算するに際してはその誤差を無視できるとの正当化がありえよう。

もし時系列が変数変換されているときには予測結果はまず変換したスケールで求め、次に原系列の予測に変換しなおすことが考えられる。例えば原系列 Y_t に対数変換 $y_t = \log(Y_t)$ を施してモデル (3) を当てはめたとすると、まず系列 y_t に対して予測を行ない、求めた点予測値と予測区間を原系列のスケールに変換しなおすことが考えられる。ただし、標準的な仮定の下で予測と述べたのは変換後の変数 y_t についての MMSE の点予測値であって原系列ではないことに注意しておく。同様の予測方法を X-12-ARIMA が用意している変数変換についても行なうことができる。もし事前調整がなされていれば点予測値と区間予測値はもとのスケールに変換されることになる。

さて利用者が定義した回帰の説明変数が含まれていると X-12-ARIMA では予測区間におけるその値を与えることが必要となる。既成の回帰変数については必要な将来の値はプログラム上で作り出す。利用者が自ら事前調整要素を定義する場合には予測期間についてもその値を定めることが必要となる。

4. RegARIMA モデルの推定に関連する事項

プログラム X-12-ARIMA で使っている IGLS アルゴリズムと非線形最小二乗法情報は RegARIMA モデルの最尤推定値を見つける為には信頼できる方法であるが時々推定上に問題が発生することがある。モデルと推定する上で生じる可能性のある幾つかの問題について以下では発生した時の対処の方法とともに議論する。さらに X-12-ARIMA におけるモデル選択基準の利用に関して重要な注意点も説明する。

4.1 母数の初期値と収束問題への対応

利用者は繰り返しによる最尤推定の最適化において AR 母数と MA 母数の初期値を設定することができる。しかしながらこの設定はめったに必要としないし、一般的には設定するようにも推奨もできない。何も指定しなければ X-12-ARIMA では AR 母数と MA 母数の全てについて 0.1 が設定される。(回帰モデル部分については GLS 回帰で推定するので初期値は必要でない。) X-12-ARIMA で自動設定されている値は多くの場合には適当であるように判断されている。より適切な初期値を設定したからと言って(例えば条件付最尤推定により推定した値を使うなどして)初期値をわざわざ推定することを正当化するような収束スピードの向上などは期待できそうもない。ただし、例外的としては例えば時系列に幾つかの観測値を追加するなど初期値としてかなり正確な値がわかっているときには設定することでメリットが生じよう。しかしながら初期値を設定することを可能にした真の理由は推定が困難な場合に収束問題に対処するためにあることに言及しておこう。

X-12-ARIMA における繰り返し推定により収束しない場合には幾つかの対処法がある。プログラムが繰り返しの最大数に達して収束を止めた場合には(この場合には警告のメッセージが出て最後の繰り返しで得られた値が出力される)最後の繰り返しで得られた値を初期値と設定してプログラムを再び実行すると収束する可能性がある。より簡単ではあるが計算時間はかかる方法としては繰り返し制約数を増加させプログラムを実行することである。もしプログラムが収束以前あるいは最大の繰り返し数以前にトラブルが発生させた場合にはまず条件付尤度

によりモデルを推定し、そこで得られた推定値を初期値として利用して厳密な最尤推定を行なうことができる。他方、我々の経験からは収束問題はしばしば推定するモデルがかなり複雑な場合（例えば高次のモデル等）や適切でない条件のもとで起こりがちである。そのような場合には推定結果を検討し、より単純なモデルを選択することが適切な対処と言えよう。4.2節から4.4節にかけては生じうる推定問題とその場合の適切なモデルの選択についての幾つかの状況について説明する。

4.2 MA 項の反転可能性

移動平均 MA 項の多項式 $\theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$ は方程式 $\theta(B) = 0$ のすべての根 G_1, \dots, G_q が単位円の外にあるとき ($|G_j| > 1$ 、反転可能と呼ばれる)。例えば Brockwell-Davis (1991, 123 項-125 項) が説明しているように ARIMA モデルにおける MA 作用素の一つ以上が反転可能でないと同一の自己相関構造を持ち同一の条件付尤度を持つ反転不能な一つ以上の MA 作用素が存在し得る。したがってデータだけでは反転可能なモデルと反転不能なモデルを区別することは出来ないため、反転可能なモデルを選択した方がよい。この問題は予測において特に重要であり、反転不能なモデルを用いると大域的に正しくない予測を行なうことになってしまう。一つの例外としては、MA 多項式の根が単位円上に存在する場合 ($|G_j| = 1$) であるが、これは反転可能領域の境界であり MA モデルの厳密な最尤推定を行なうなど適切に対処すれば予測における問題は生じない。

X-12-ARIMA による推定では尤度関数の繰り返し非線形最適化において MA 母数について反転可能性の制約条件を課している。したがって正確に云うと X-12-ARIMA により推定されたモデルは反転可能である。もし最尤推定値 (MLE) が境界上にあれば (すなわち MA 作用素の根が単位円上に厳密にある場合) X-12-ARIMA では非線形最適化により内側から反転可能領域に近づき収束基準、あるいは繰り返し最大数により境界領域に近づく方法をとっている。これにより X-12-ARIMA では結果として反転境界上のモデルを推定できている。このような状況では母数空間の制約の境界上での尤度の最大化は困難な最適化となるので繰り返し収束には時間がかかる。より重要な点として反転可能境界上の推定値の収束にかかわる状況はもとのモデルが適切でないことが多く、利用者へ結果 (及び繰り返し推定の細かな出力) を調べることを警告しているともとれる。4.4 節で不適切なモデル化のもっとも重要な原因となる AR 因子と MA 因子が共通因子を持つ状況や過剰階差問題とその適切な対処法について説明する。

推定するモデルが季節階差や季節 MA 項を含む時に反転不能な結果がもたらされることが少なくない。(すなわち、母数 θ の MLE が 1 のときに $1 - \theta B^s$ は反転可能ではない。) また季節 ARIMA モデルは季節的経済時系列ではよく用いられるので利用者は反転可能性問題について注意し、問題が生じたときには 4.4 節で説明している対処法を理解しておく必要がある。

4.3 AR 項の定常性

自己回帰 AR 多項式 $\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ において $\phi(B) = 0$ の根がすべて単位円外にあれば定常時系列であるが、そうでなければ非定常となる。ここでより正確に述べると (1) 式において $\phi(B)\Phi(B^s)w_t = \theta(B)\Theta(B^s)a_t$ にしたがう確率過程 $w_t = (1 - B)^d(1 - B^s)^D z_t$ は AR 多項式のゼロ根すべてが単位円外にあるとき定常過程となる。AR 項を含む厳密な尤度関数はすべての AR 作用素が定常性の条件を満たすものと仮定して導かれる。すなわち AR 項を含む厳密な尤度関数の評価、推定、予測など他の分析はすべてこの AR 項についての条件が満たされていることを仮定して行なわれる。したがって AR 項を含む尤度関数の厳密な評価ではプログラム X-12-ARIMA では推定上で母数の制約条件を課して

いる。次節で述べるように AR 項と MA 項に共通因子が存在する状況でなければ X-12-ARIMA においては尤度が $-\infty$ に近くなる定常性の境界に近づく可能性は低い。

尤度を AR 母数について条件付で定義すれば X-12-ARIMA でも定常性を仮定したり制約にする必要は無い。この場合にはモデルの推定や予測などは母数の定常性領域に制約されずに実行できる。しかしながら 3.5 節で説明した統計的推測の議論は成立しない。定常条件が無ければ AR 母数の推測については例えば Fuller(1976) の 8.5 節にあるような様々な込入った議論を用いる必要が生じる。

4.4 AR 項と MA 項の共通因子と過剰識別

自己回帰移動平均 (ARMA) モデルを推定する際には AR 項と MA 項が相殺する可能性が生じる。(1) 式 (あるいは (3) 式) において $p > 0$ かつ $q > 0$ 、あるいは $P > 0$ かつ $Q > 0$ のいずれかが満たされる時に ARMA モデルと呼ばれている。理論的には $p > 0$ かつ $Q > 0$ 、あるいは $P > 0$ かつ $q > 0$ の場合も ARMA モデルとなるがそのような場合にはこの AR 項と MA 項の相殺問題はほぼ生じない。この問題のもっとも簡単な例としては ARMA(1,1) モデル $(1 - \phi B)z_t = (1 - \theta B)a_t$ において $\phi = \theta$ となる状況である。両辺の $(1 - \phi B)$ 要素をとれば $z_t = a_t$ という簡単なモデルである。したがってこの場合には尤度関数は母数空間上の $\phi = \theta$ 周辺でほぼ一定となる。こうした状況では ARMA(1,1) モデルの MLE がほぼ制約 $\hat{\phi} = \hat{\theta}$ を満たしているので非線形推定の収束が困難となる。また同様の問題は AR 多項式と MA 多項式が共通のゼロ根を持つときに生じることになる。この問題についてのより詳細な議論についてはボックス・ジェンキンス (1976, 248 項-250 項) を参照されたい。

X-12-ARIMA プログラムによるこの混合 ARMA モデルの推定で収束に問題が発生するときにはこの AR 項と MA 項の相殺効果が原因となることが少なくない。どちらにしても AR 多項式と MA 多項式のゼロ根を調べることにより簡単にこの問題を検討できる (コマンド推定値 (estimate) の中の print=roots を使えばよい)。仮に共通ゼロ根が見つければモデルを修正して共通因子を相殺すればよいので AR 多項式と MA 多項式の次数を落として再推定すれば良い。この共通因子は厳密な意味では必ずしもなく、AR 項のゼロ根と MA 項のゼロ根がほぼ共通であっても問題が生じるので対処する必要がある。

ところで推定された MA 多項式は階差作用素と相殺する可能性がある。特に非季節階差をとって推定した MA モデルが $(1 - B)$ 因子を含んでいたり季節階差をとって推定した季節 MA モデルが $(1 - B^s)$ 因子を含む場合などに発生することがある。例えばモデル $(1 - B)(1 - B^s)z_t = (1 - \theta B)(1 - \Theta B^s)a_t$ では θ か Θ が 1 に近く推定されていれば問題が生じる。この種の共通因子は過剰階差問題と呼ばれているが、これは時系列を定常化するために必要以上に階差操作を行なったことから生じていると考えられる。階差操作と MA 項に過剰階差が生じてモデルを単純化したときには利用者は階差作用素により前には消されていた非確率的な時間の関数をモデルの回帰項を加えておく必要が生じることに注意しておこう。このことは非季節階差が $\hat{\theta} = 1$ 付近で MA 要素 $(1 - \theta B)$ と相殺する時にはより簡略化したモデルにはトレンド定数項 (あるいは元のモデルが 1 次の階差であれば全平均) を含めるべきということを意味する。もし季節階差が $\hat{\theta} = 1$ 付近で季節 MA 項 $(1 - \Theta B^s)$ とが相殺される時にはより簡略化したモデルにはトレンド定数と固定的季節要素を含める必要があることになる。過剰階差問題については Abraham=Box(1978) や Bell(1987) が議論している。

過剰階差を含むモデルでは推定結果が収束するときには、相殺している階差操作と MA 要素を取り除き、適当な回帰変数を付け加えてモデルを推定するのは常によいとは限らないが、それはこの共通因子の問題が必ず推定問題や他の問題

を (MA 母数について厳密な尤度関数を利用すると仮定しても) もたらずとは限らないからである。特に予測結果は過剰階差を含むモデルとそれを修正したモデルでは同一の結果をもたらす、回帰や ARMA 母数の推定値のその標準偏差はほぼ同一となる。(むろん次節で述べるように対数尤度の値や対応するモデル選択基準は異なる。) このことはここでの状況が AR 項と MA 項の共通因子の問題とはかなり異なることを意味している。AR 項と MA 項の共通因子問題では推定の収束問題を生じさせるために発生すれば必ずモデルから取り除き、モデルを再推定することが必要である。

4.5 モデル選択基準の利用

X-12-ARIMA プログラムは次に挙げるモデル選択基準を提供している。すなわち、AIC (赤池情報量基準 Akaike(1973), Findley(1985),(1999) を参照), AICC (Hurvich=Tsai(1989) を参照), Hannan=Quinn(1979), BIC (Schwarz(1978)) などを提供している。ホワイト・ノイズの分散を含め推定する母数の数を n_p としよう。さらに階差や季節階差を適用した後のデータ数を N 、推定により得られた厳密な尤度関数の最大値を L_N とするとこれらのモデル選択基準は次の式で与えられる:

$$\begin{aligned} AIC_N &= -2L_N + 2n_p \\ AICC_N &= -2L_N + 2n_p \left\{ \frac{N}{N - \frac{n_p+1}{N}} \right\} \\ Hannan - Quinn_N &= -2L_N + 2n_p \log \log N \\ BIC_N &= -2L_N + n_p \log N. \end{aligned}$$

各基準の中では与えられた時系列に対してその値がもっとも小さくなるモデルが選択されることになる。こうしたモデル選択基準を用いる場合には注意すべき幾つかの事項がある。第一にこうしたモデル選択が可能なのは同一の次数の階差モデルに対して正当化が可能なことである。したがってモデル選択基準を階差操作の次数選択に使うことはできないことになる。第二にはモデル選択を可能とするには AR 母数と MA 母数について厳密な最尤法を使わねばならないことであるが、X-12-ARIMA ではこの推定方法が自動選択される。このことから X-12-ARIMA では条件付尤度関数を使った場合にはモデル選択基準の値は出力されない。

さらにモデル選択基準は異なる外れ値 (変数 AO, LS, TS, ramp) の回帰変数を含む RegARIMA モデルには利用すべきでないことである。特にどの外れ値をモデルに含めるか否かを定める際には用いるべきでない。この点についての理論的問題としてモデル選択基準を導出するにあたってはモデルの未知母数は将来データが得られるにつれてより正確に推定できることが仮定されているが、外れ値の母数に関してはこのことは成り立たないことから生じる。

なお同一の外れ値変数を含むモデルでモデルを比較するときには経験的か直感的には外れ値変数による回帰推定では外れ値は最尤推定値に対しては同一の方向で影響を与える (つまりその影響は対数尤度の差を考える限り相殺する) ことが云えるようである。例えば AO 変数を時刻 t に使うと時刻 t の観測値が欠落しているかのように扱ってもほぼ同様の効果があり (Gomez=Maravall=Pena(1999) を参照)、本質的には比較する複数のモデルにおいて当該観測値を欠測値と扱ったかのような状況になることが知られている。これに対して、変数 AO を回帰変数とすると実際には異なるデータ・セットからモデルを推定するような状況となり、対数尤度関数やその対数尤度から計算される他の統計量を比較することについての正当化の根拠はほとんどない。(この点は二つのモデルを比較する際に季

節階差と非季節階差の次数を同一に取る必要があることと本質的に同じである。) 尤度が階差データから計算されるときには階差の次数が異なるとデータは本質的に異なると見なければならぬ。))

実際に異なる外れ値変数を含むモデルを比較してみると通常は多くのこうした変数を含むモデルが選択される傾向にあるが、予測はより悪くなる可能性がある。さらに選択された外れ値のかなりが1年後、あるいは二つのデータが加わると外れ値として識別されないことも少なくない。すなわちモデル識別がかなり不安定となる可能性がある。

変数変換や事前調整はモデル選択基準に影響を与える。例えば変換 (transform) の中に変数として含まれない関数 (function) による変数があったり $\lambda \neq 1$ であるべき (power) 変換があると変換コマンドは原系列に関する尤度関数に影響を及ぼし、その結果モデル選択基準も影響される。このことから X-12-ARIMA では原時系列に対する対数尤度を計算する際には変数変換調整を計算し、その変数変換調整済み対数尤度によりモデル選択基準が計算される。このように与えられた時系列に対して異なる変数変換、あるいは事前調整などの要素が処理されている。したがって X-12-ARIMA にかける入力前に時系列が変数変換されていると (つまり例えば原系列の対数変換値が X-12-ARIMA 読み込まれたりすると) 変数変換を X-12-ARIMA の中で行なったときとは異なるモデル選択基準の値となることに注意しておく。同様のことが事前調整要素についても当てはまることに注意しておく。したがって、モデル選択基準を用いる場合には同一の入力系列に対する異なる統計モデルの比較に限定すべきである。このことから系列 (series) の中で期間 (span) を用いて系列の異なる部分に対してモデル選択基準を用いることができないことも明らかであろう。

5. スペック・ファイルと文章構成法

X-12-ARIMA の主たる入力はスペック・ファイルと呼ばれている特別な入力ファイルで作られる。このファイルは一連のコマンド・スペック (specs) によりデータ、利用したい季節調整オプション、出力、利用する時系列等々の情報を X-12-ARIMA に与える。ここで幾つかのスペック命令文としては次のようなものがある。

系列 (series) : このスペック命令は集計値の調整を除き常に必要である。このスペックにより時系列データ、開始時点、季節周期、分析期間、系列のタイトルが指定される。

集計 (composite) : このスペック命令により直接的あるいは間接的な集計値の調整が指定され、スペック命令である系列 (series) の代わりに用いられる。

変換 (transform) : このスペック・コマンド (命令) ではデータ系列の変換や事前調整を指定する。

x11 : このスペック命令は季節調整オプションを指定し、季節調整の形式、季節成分とトレンド成分へのフィルター、復活祭調整オプション、季節調整の診断オプション等を指定する。

x11 回帰 (x11 regression) : このスペック命令は非標準的回帰オプション、例えば回帰変数の指定や不規則要素に対する頑健な回帰を行なうために指定する。

自動選択 (automdl) : このスペック命令は自動的モデル選択の方法を指定する。

アリマ (arima) : RegARIMA モデルの ARIMA モデル部分を指定する。

回帰 (regression) : このスペック命令は RegARIMA モデルの回帰部分で利用する説明変数を指定し、識別 (identify) により取り除く回帰効果を指定する。

推定値 (estimate) : このスペック命令では回帰 (regression) とアリマ (arima) により特定化されたモデルの推定と尤度の評価、それから推定におけるオプションを指定する。

チェック (check) : このスペック命令は推定されたモデルを診断する際に有用な統計量を指定する。

予測 (forecast) : このスペック命令は推定したモデルに基づく予測を行なうことを指定する。

外れ値 (outlier) : このスペック命令は加法的な外れ値や水準の変化の自動検出を指定する。さらに、一時的な水準変化の検定が利用可能である。

識別 (identify) : このスペック命令により決められたデータの階差次数を用いて ARIMA モデル識別の為に回帰効果を含んだ (回帰 (regression) により別に定める) 自己回帰関数と偏自己回帰関数を計算する。

期間変更の安定性 (slidingspans) : このスペック命令により季節調整値の安定性について期間変更の安定性分析を行なう。

履歴 (history) : このスペック命令により季節調整の改定値の履歴や関連する RegARIMA モデルの統計量を求めることができる。

各スペック命令はその名前によりスペック・ファイルとして定義され、それに続いて変数 (*arguments*) と指定された値 (*values*) を含んだ {} との形をとる。変数と指定された値は変数 (*argument*) = 値 (*value*) あるいは複数の数値が必要な時には変数 (*argument*) = (値₁, 値₂, ...) という形になる。ここで様々な値 (*value*) の形があり、タイトル、変数名、鍵言葉、数値、それに日付等がある。こうした例は 6 章の個々のスペック命令文についての説明で定義されたり例示されることになる。幾つかのスペック命令においてしばしば表われるので 5.1 ではスペック命令である 印刷 (*print*) と保存 (*save*) における変数 (*arguments*) について、5.2 節では日付 (*date*) の変数について説明する。

以下で見るように系列 (*series*) あるいは集計 (*composite*) を除いては決められた変数は必要としない。ほとんどの変数はあらかじめ指定された値を持っていてその値は各スペック命令の解説に書かれている。特に指定が無ければ利用する変数すべてにあらかじめ指定した値が使われる。

一般的にはある一つのスペック・ファイルにおいて多くのスペック命令は必要としない。実際、ほとんどすべての X-12-ARIMA の計算 (複合的計算は除かれるが) は一つだけスペック・ファイルとして必要な命令の系列 (*series*) があれば実行される。このスペック命令ではデータ (*data*) あるいはファイル (*file*) 変数が必要である。ただし、例外としてはデータ・メタ・ファイルを使う場合があるが、この場合についてより詳しくは 2.5.2 節を参照されたい。したがって X-12-ARIMA にとり最低限のスペック・ファイルとしては一系列 (*series*) *data*=(データの値 (*data values*))—となる。季節調整を実行するにはスペック・コマンドの *x11* が必要であり、もしそうでなければスペック命令である *x11* 回帰 (*x11 regression*), 期間変更の安定性 (*slidingspans*), 履歴 (*history*) (推定値 (*estimates*)) を含めて季節調整履歴を調べる) などの中で少なくとも一つが必要となる。その場合には X-12-ARIMA ではスペック・コマンド *x11* が規定値として指定されていると扱うのでスペック・ファイルの中に—*x11*—と書かれていることと同等になっている。

モデルの識別を実行するには識別 (identify) が必要であり、モデルの推定には通常はスペック命令アリマ (arima), 回帰 (regression), 推定値 (estimate) が必要となる。推定値 (estimate) がなくてもスペック命令として外れ値 (outlier), 自動モデル (automdl), チェック (check), 予測 (forecast), x11, 期間変更の安定性 (slidingspans), 履歴 (history) 等が一つでもあればモデルの推定は行なわれる。その場合には X-12-ARIMA ではコマンド推定値 (estimate) があたかも存在するかのごとくに扱い、スペック・ファイルに—estimate—と書かれていると同等となる。またスペック命令のアリマ (arima) がなければ推定は ARIMA(0 0 0) モデル (すなわちホワイト・ノイズ) の指定として推定が実行されることになる。この場合はスペック・ファイルに—arima—と書かれていることと同等である。

一つの例外を除きスペック・ファイルに書かれている指示の順番とスペック命令のカッコ内の変数の順番は問題ではなく同一の結果をもたらす。ここでの例外は系列 (series) や集計 (composite) であるが、これらのコマンドは最初のスペック命令として書かれなければならないことである。スペック・ファイルの形式は自由であるが空白、タブ、空白行などを用いるとスペック・ファイルが読みやすくなる。コメントを付ける事もできる。入力表示にコメントやその他のルールの使い方については 5.3 節で議論する。

重要事項 スペック・ファイルの最後には必ず改行を指定すべきである。そうでないと当該行が読まれないことになるが、これはフォートラン言語の制約である。

次に挙げる例 5.1 によりデフォルト X-11 を実行させる単純なスペック・ファイルを示しておく。このファイルを実行した出力ファイルのスペクトル診断によると曜日効果の存在を示唆しているが、この結果についてのメッセージが出力には書かれている。RegARIMA モデルを用いれば季節調整に先立ち曜日効果と予測により系列を延ばすことと同一の事を行なうことができる。例 5.2 と例 5.3 では最終の季節性と曜日効果を取り除く (例 5.4) 前にモデルの ARIMA モデル部分を識別するスペック・ファイルを示している。ここでももちろん、X-11 のスペック・ファイルとともに 6 章で説明する X-11 の曜日効果の除去法を利用する事もできよう。

例 5.1 : デフォルト X-11 用の X-12-ARIMA スペック・ファイル

```
series{title = "Monthly Retail Sales of Household Appliance Stores"
data = ( 530  529  526  532  568  785  543  510  554  523  540  599
        574  619  619  600  652  877  597  540  594  572  592  590
        632  644  621  604  613  828  578  533  582  605  660  677
        682  684  700  705  747 1065  692  654  719  690  706  759
        769  730  740  765  791 1114  695  680  788  778  780  805
        852  823  831  836  913 1265  726  711  823  780  844  870
        865  915  920  935 1030 1361  859  852  954  895  993 1109
        1094 1173 1120 1159 1189 1539 1022  987 1024 1005 1054 1098
        1191 1191 1161 1201 1294 1782 1154 1059 1178 1126 1120 1233
        1260 1311 1302 1365 1395 1899 1123 1087 1210 1157 1159 1260
        1357 1265 1231 1287 1452 2186 1309 1242 1388 1400 1397 1527
        1654 1650 1555 1560 1836 2762 1541 1480 1619 1455 1510 1698
        1651 1749 1783 1863 2074 3051 1836 1690 1856 1796 1904 1927
        1978 2055 1976 2204 2423 3502 1977 1767 1935 1900 2073 2143
        2299 2247 2162 2274 2529 3731 2184 1901 2058 1974 2018 2091
        2239 2253 2157 2190 2397 3659 2170 2086 2297 2251 2311 2520)
start = 1972.jul}
x11{}
```

時系列データに対してモデル分析を行なうときには少なくとも X-12-ARIMA を 2 度実行してみることを薦める。最初の実行はモデルの ARIMA モデル部分を

識別することの為であり、二度目の実行は RegARIMA モデルを推定し、可能な限りモデルを使って予測する為である。最初の実行の為のスペック・ファイルでは系列 (series) と識別 (identify) を使うと共に変換 (transform) と回帰 (regression) 等を用いる。二度目の実行ファイルで系列 (series), アリマ (arima), 推定値 (estimate) 等を使い、さらに必要なら変換 (transform), 回帰 (regression), 外れ値 (outlier), チェック (check), 予測 (forecast) 等を用いる。プログラム X-12-ARIMA を二度実行させるときには二つの異なるスペック・ファイルが必要であるが、実際には最初のスペック・ファイルを修正して二度目の実行ファイルを作れば良い。もし診断により推定されたモデルを修正することが必要であればスペック・ファイルをさらに訂正して X-12-ARIMA をもう一度実行すればよい。

例 5.2 : RegARIMA モデル識別用の X-12-ARIMA スペック・ファイル

```
series{title = "Monthly Retail Sales of Household Appliance Stores"
data = ( 530 529 526 532 568 785 543 510 554 523 540 599
        574 619 619 600 652 877 597 540 594 572 592 590
        .
        .
        .
        2239 2253 2157 2190 2397 3659 2170 2086 2297 2251 2311 2520)
start = 1972.jul}
transform{function = log}
regression{variables = td}      # Comment: Series has trading-day effects
identify{diff=(0, 1) sdiff = (0, 1)}
```

モデルの識別を実行する為の典型的なスペック・ファイルは例 5.2 の形式のようになろう。この例では系列 (series), 変換 (transform), 回帰 (regression), 識別 (identify) などのスペック命令を使っている。まずデータが系列 (series) の中に定義され、系列の対数変換を行ない (変数変換 (transform) による)、系列に影響すると知られていたり可能性がある変数を回帰変数をコマンド回帰 (regression) により定める。ここで変数 (variables) = td は 6 個の曜日効果変数 (td6) を含むと共にうるう年効果を調整することにも言及しておこう。(3.3 節と 6 章の回帰 (regression) の説明を参照されたい。) コマンドの識別 (identify) により変換された変数の階差 (月の長さも調整して) に回帰変数 (6 個の曜日効果の変数) により回帰を行なう。ここで季節階差と非季節階差の最高次 $(1 - B)(1 - B^{12})$ を適用して回帰を行なっている。それからスペック・コマンドの識別 (identify) により原系列の回帰残差系列を計算し、季節階差と非季節階差のすべての組み合わせに対して標本自己相関関数と標本偏自己相関関数の表とプリンター・プロットを出力している。

最初の出力結果を検討し ARIMA モデル部分の識別 (例えば $(0 \ 1 \ 1)(0 \ 1 \ 1)_{12}$ であったとすると) を行なう識別 (identify) の後にアリマ (arima) と推定値 (estimate) をスペック・ファイルに付け加えることが考えられる。結果として得られるスペック・ファイルは例 5.3 で与えられる。(ただし、データのすべては載せていない。)

例 5.3 : RegARIMA モデル推定用の X-12-ARIMA スペック・ファイル

```
series{title = "Monthly Retail Sales of Household Appliance Stores"
data = ( 530 529 526 532 568 785 543 510 554 523 540 599
        574 619 619 600 652 877 597 540 594 572 592 590
        .
        .
        .
        2239 2253 2157 2190 2397 3659 2170 2086 2297 2251 2311 2520)
```

```

start = 1972.jul}
transform{function = log}
regression{variables = td}      # Comment: Series has trading-day effects
# identify{diff=(0, 1) sdiff = (0, 1)}
arima{model = (0,1,1)(0,1,1)}
estimate{print = iterations}

```

このスペック・ファイルには命令として系列 (series), 変換 (transform), 回帰 (regression), アリマ (arima), and 推定値 (estimate) が含まれている。ここでコマンドの回帰 (regression)、アリマ (arima)、推定値 (estimate) により次のモデルを指定し推定している :

$$(1 - B)(1 - B^{12})\left(y_t - \sum_{i=1}^6 \beta_i T_{it}\right) = (1 - \theta B)(1 - \Theta B^{12})a_t,$$

ただし T_{it} は 6 個の曜日回帰変数である。系列 y_t は原データ系列からうる年効果を調整して対数変換を行なってからモデル分析が行なわれる。残差分析、外れ値の検出、あるいは予測などを行ないたいければそれに応じたスペック・コマンドを付け加えればよい。

求めた統計的モデルが適当である事を仮定するとスペック命令の x11 と予測 (forecast) を入力スペック・ファイルに加える事により予測による外挿を利用する季節調整を行なう事ができる。そのようなスペック・ファイルとして例 5.4 を挙げておく。なおデータは一部分のみでは省略した。

例 5.4 : 季節調整用の X-12-ARIMA スペック・ファイル

```

series{title = "Monthly Retail Sales of Household Appliance Stores"
data = ( 530  529  526  532  568  785  543  510  554  523  540  599
        574  619  619  600  652  877  597  540  594  572  592  590
        .
        .
        .
        2239 2253 2157 2190 2397 3659 2170 2086 2297 2251 2311 2520)
start = 1972.jul}
transform{function = log}
regression{variables = td}      # Comment: Series has trading-day effects
# identify{diff=(0, 1) sdiff = (0, 1)}
arima{model = (0,1,1)(0,1,1)}
estimate{print = iterations}
forecast{maxlead = 60}
x11{seasonalma = s3x9}

```

このスペック・ファイルにより 3×9 季節フィルター (x11) を曜日効果を調整した系列に適用して季節調整を行なうことができる。事前調整した系列は季節調整に先立ち計算した 60 期間の予測値 (予測 (forecast)) により外挿する。主要な出力ファイルにより得られた季節調整の良さについての診断ができる。また 6 章に説明する適当なスペック命令を含める事による追加的な診断結果を入れる事もできる。

5.1 プリントと保存

X-12-ARIMA からの出力の制御はスペック命令である印刷 (print) と保存 (save) の変数を使って行なう。印刷 (print) の変数は主要な出力ファイルへの出力を制

御し、保存 (save) の変数はファイルに書き出す出力の表を制御する。説明の都合上、これら印刷 (print) と保存 (save) により制御される出力を表 (Tables) と呼ぶことにする。むしろ、この種の表の中には (例えば ACF のプロットなど) 通常は表の形とはみなされない出力も含まれている。スペック命令である印刷 (print) と保存 (save) により制御されている表は印刷 (print) と保存 (save) の変数として説明されているように、指定されたプリント形式と拡張子とともに一覧となっている。保存 (save) を用いた出力表は精度が確保される形式で書かれているが、他のソフトウェアでさらに分析することができるようにラベル情報は最小限としている。保存された表は単一のタブをそれぞれ含む一貫した形式になっている。

スペック命令である印刷 (print) がなかったり (—print=default—)

あるいは—print=()—となっていればスペックであらかじめ指定している出力が主な出力ファイルに書き出される。主たる出力ファイルへ出力することを止めなければ print=none とする必要がある。(ただし、このスペック print=none によっても幾つかのラベル情報などがスクリーンに書き出されることに注意しておく。) 他方、主な出力ファイルに書かれるすべての利用可能な出力表とプロットを得るには print=all と設定することが必要である。また主な出力ファイルに書かれる (プロットを除く) 出力ファイルすべてを得るには print=alltables と指定すればよい。画面上に書き出される出力表の幾つかのみ得るには print=brief と指定すればよい。スペック・コマンドの default, brief, none などの出力水準に対して、さらに個々の表を加えるにはその名前を変数として付け加えればよい。これらのコマンドは必ずと言うわけではないが+の後にかかれる事が多い。例えば推定値 (estimate) の中で print = (+iterations + residuals) は

```
print = (default + iterations + residuals)
```

と同等であるが、これはデフォルト default の出力に加えて、あらかじめ指定された繰り返し推定と推定された残差の結果の出力を意味している。コマンド print=(none estimates) を使うと出力は母数推定値のみとなる。デフォルトでは個々の表は出力されないが、出力変数名として- (すなわち print=brief -acf や print=(all -iterations)) とするとすべて出力される。利用者がどれかの出力表をファイルに保存したい時には適当なスペック・ファイルの中に保存 (save) 変数を書いておかなければならない。例えば推定値 (estimate) の中に save=(mdl estimates) とする等である。保存可能な表は印刷 (print) と保存 (save) など中にリストにある長い (long) 方の名前をスペック命令の記述に用いるか、保存する表の拡張子と同一の短い (short) 3文字名を用いることにより得る事ができる。例えば推定値 (estimate) の中の regcmatrix は rcm と同一の指定である。ただし、既に出てきた鍵言葉である none, all, alltables, default, brief などは保存 (save) の中では使えない。また保存する表の名前は+か-を前につけてはならない。すべての表が保存できるとは限らないし、またいずれのスペック・ファイルでも表を保存する事ができるわけでもない。

保存 (save) により各ファイルに指定の表を書き込む事ができるわけであるが、保存されたファイルは出力と同一のディレクトリーに置かれ、主な出力ファイルと同一のファイル名と3文字の拡張子名を持つことになる。もし同じ名前のファイルが既に存在していれば新しく書き換えられる。ここで用いられる拡張子は印刷 (print) と保存 (save) の説明の中で一覧となってリストアップされている。例えば DOS 計算機上で入力ファイルとしてディレクトリー C:\T SERIES の中の SALES.SPC を使って X-12-ARIMA を実行させるとしてみよう。推定値 (estimate) が save = (mdl estimates) を含んでいる時には結果として保存されるモデルの表と母数の推定値はファイル C:\T SERIES\SALES.MDL 及び C:\T SERIES\SALES.EST に書かれる。X-12-ARIMA によって使われる拡張子は明らかな矛盾を避ける (すなわち .dat, .exe, .com, .for, .spc など) が、利用者は X-12-ARIMA により保存され

ているファイルが意図しないうちに書き換えられてしまうことを注意深く避ける必要がある。保存したファイルのリストが出力の最初に書かれてるが、*がついているファイルは存在していたファイルが書き換えられたことを意味する。またスペック・ファイルにエラーがあったり、その他の理由でプログラムが実行を停止した場合には幾つかあるいはすべての保存ファイルが書かれていない可能性もある。

5.2 日付

日付変数は幾つかのスペックに表われるが、その値は同一の形式に規定されている。月次データの日付は *year.month* であるが、この形式は他の周期（例えば *year.quarter*）等に一般化されている。年を表示するには4桁すべてを含めることが必要である。したがって 67 は AD(あるいは CE)67 年であり AD1967 年ではない。

月次データでは月は整数 1 から 12、あるいは 3 文字の略月名 (*jan, feb, mar, apr, may, jun, jul, aug, sep, oct, nov, and dec*) を用いる。したがって 1967.12 と 1967.*dec* は同一月を意味する。四半期データやその他の季節周期のデータでは整数のみが使われるので例えば 1967.1 や 1967.4 となる。

日付は系列の初期時点を定義したり時系列の一部を分析する際に使われる。また外れ値の回帰変数を指定する時にも利用される。例えば 1978 年 4 月に加法的な外れ値の為に回帰変数を指定し、1982 年 9 月に水準の変化が始まることを指定するのは次のように書く。

```
regression { variables=(ao1978.apr ls1982.sep) }
```

日付けの季節性はスペック命令の系列 (*series*) 中のデータの季節性に一致している必要がある。すなわち、*ao1972.jan* は月次データには適当であるが、四半期データには使用が認められていない。

5.3 入力文での一般ルール

入力文字 (Allowable input characters)

引用で使われる文字を除き入力として認められる文字は英文字、数値、空白、タブ、新行 (*newline*) 文字、及びその他: = . , { } () [] + - # である。プログラムではスペック・ファイルにおける他の ASCII 文字は無視するが、それはフラッグや警告文が作られることでわかる。引用文の中では次のような文字がさらに使える: ! % ' * / : ; i j ? @ \ - | ~ 等である。二重の引用も複数の引用文で区切られた文の中では認められる。

大括弧 (Braces), 丸括弧 (parentheses), 角括弧 (brackets)

文字 { }, (), [] はそれぞれ異なる機能があり、互いに代替はできない。文字 { } は変数を含むスペックに使われ、() は変数への複数の数値リストを含むときに利用され、[] が利用されるのは (i) 例えば復活祭の回帰変数など特殊な変数を定義する値、すなわち回帰 (*regression*) { variables = (td Easter[14]) } 等、あるいは (ii) 一部のラグの欠けた ARIMA モデルに表われるラグ、たとえば *arima* { model = (0 1 [1,3]) } などで利用される。

活字の大きさ (Case sensitivity)

スペック名・変数・日付・鍵言葉 (例えば *none* と *all*) 等、事前に定められている回帰変数名 (*td* や *seasonal* など) は活字の大きさは問題とならない。したがって X-12-ARIMA では TD と *td* は同一の変数であり共にスペック命令の回帰 (*regression*) や *x11* 回帰 (*x11regression*) では変数 (*variables*) と認識される。

コメント (Comments)

記号#の後に来る1行は#が引用文の中のない限りコメントとなる。スペックに一部分だけにコメントが必要なら、大括弧・丸括弧・角括弧で区切りをつけないなければならない。

等号 (Equals sign)

変数に値を指定するには等号=が用いられる。すなわち `print = none` あるいは `title = "Monthly Retail Sales of Household Appliance Stores"` などである。

スペック・ファイルでの文長

スペック・ファイルでの行は132文字に制限されている。132文字を超える欄は無視される。したがって特にデータがスペック・ファイルの中で132文字を超える行を含んでいると入力データが切断されることになることに注意が必要である。ただし、1行132文字の制約は(スペック・ファイルでない)コマンドファイル(file)によって別のファイルから読み込まれるデータには適用されない。他のファイルの法はフォートラン(FORTRAN)での入力長の制約に依存しているのでシステムによって異なりうる。

重変数値

複数の数値からなる重変数値は例えば `variables=(td seasonal const)` などのように括弧でまとめなければならない。一変数は一つの値をとるか複数の数値を一つとしてとるが、一つの場合は括弧は付けても付けなくても良い。つまり `variables=td`, `variables=(td)`, `variables = (td seasonal)`, `start=1967.4`, `start=(1967.4)` などすべて可能である。

空白リスト (Null list)

空白の変数、例えば `outlier{}` などを使う事ができる。空白の変数リストの変数はあらかじめ指定されたデフォルトの値をとることになる。

数値 (Numerical values)

数値は指数記号を含め自由な形式で使うことができる。例えば `400`, `400.0`, `400.` と `4.e+2` はすべて同一な値である。

順番 (Ordering)

スペック・コマンドの順番についての唯一の制約は系列(series)あるいは集計(composite)が始めのスペックとなることである。スペックの中では変数の順番には制約がない。むろん、複数の値が指定される時には順番が依存することが明らかなる場合もあり、例えばデータの変数、(系列(series), 変換(transform), 回帰(regression)など), モデル(model)の変数(アリマ(arima)などARIMAモデルの特定化、期間(span)変数の日付(スペック系列(series)と外れ値(outlier)など)である。

区切り (Separators)

空白、タブ、空白行を文や行を区別するために使う事ができる。重変数値のリストの中ではカンマを `data=(0, 1, 2, 3, 4, 5)` などと用いる事ができる。欠落値をあらかじめ指定するデフォルト値で置き換える場合(数値が必要となる場合)には必ずカンマを用いなくてはならない。例えばスペック命令の期間(span)では二つの値が要求される。コマンド文が仮に `span=(1967.4,)` となっている時には1967.4の後にカンマが存在するので期間(span)の二つ目の数値が欠けていることを意味する。この時にはあらかじめ定められた値(最後の観測の日付)が使われることになる。

タイトルとファイル名

時系列名などのタイトルは少なくとも1つの許容される入力文字を含み、空白があっても('title' or "title")などと書かれている必要がある。タイトルでは大文字と小文字の区別は維持される。タイトルに文字#が表われるときにはタイトルの一部分と見なされコメントとは見なされない。タイトルは1行79文字以内

でなければいけない。またパス名を含みファイル名もタイトルについてのこれらの規則が適用される。

参考文献

- Abraham, B., and G.E.P. Box (1978), "Deterministic and Forecast-Adaptive Time-Dependent Models," *Applied Statistics*, 27, 120-130.
- Abraham, B., and J. Ledolter (1983), *Statistical Methods for Forecasting*, New York: John Wiley and Sons.
- 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.
- Baxter, M.A. (1994) "A Guide to Seasonal Adjustment of Monthly Data With X-11 (Third Edition)," Central Statistical Office, United Kingdom.
- Bell, W.R. (1983), "A Computer Program for Detecting Outliers in Time Series," *Proceedings of the American Statistical Association, Business and Economic Statistics Section*, 634-639.
- Bell, W.R. (1987), "A Note on Overdifferencing and the Equivalence of Seasonal Time Series Models With Monthly Means and Models With (0,1,1)₁₂ Seasonal Parts When $\Theta = 1$," *Journal of Business and Economic Statistics*, 5, 383-387.
- Bell, W.R. (1999), "An Overview of *regarima* Modeling," forthcoming Research Report, Statistical Research Division, U. S. Census Bureau.
- Bell, W.R., and S.C. Hillmer (1983), "Modeling Time Series with Calendar Variation," *Journal of the American Statistical Association*, 78, 526-534.
- Bobbitt, L. and Otto, M. C. (1990), "Effects of Forecasts on the Revisions of Seasonally Adjusted Values Using the X-11 Seasonal Adjustment Procedure," *Proceedings of the American Statistical Association, Business and Economic Statistics Section*, 449-453.
- Box, G.E.P., and D.R. Cox (1964), "An Analysis of Transformations," *Journal of the Royal Statistical Society, B*, 26, 211-252.
- Box, G.E.P., and G.M. Jenkins (1976), *Time Series Analysis: Forecasting and Control*, San Francisco: Holden Day.
- Box, G.E.P., and G.C. Tiao (1975), "Intervention Analysis with Applications to Economic and Environmental Problems," *Journal of the American Statistical Association*, 70, 70-79.
- Brockwell, P.J., and R.A. Davis (1991), *Time Series: Theory and Methods*, New York: Springer-Verlag.
- Burman, J.P., M.C. Otto, and W. R. Bell (1987), "An Iterative GLS Approach to Maximum Likelihood Estimation of Regression Models with ARIMA Errors," Research Report No. 87/34, U. S. Census Bureau.
- Chang, I., and G.C. Tiao (1983), "Estimation of Time Series Parameters in the Presence of Outliers," Technical Report No. 8, Statistics Research Center, University of Chicago.
- Chang, I., G.C. Tiao, and C. Chen (1988), "Estimation of Time Series Parameters in the Presence of Outliers," *Technometrics*, 30, 193-204.
- Chen, B., and D.F. Findley (1998) "Comparison of X-11 and *regarima* Easter Holiday Adjustments," Statistical Research Division, U. S. Census Bureau, manuscript in preparation.

- Cholette, P.A. (1978) "A Comparison and Assessment of Various Adjustment Methods of Sub-Annual Series to Yearly Benchmarks," Research Paper, Seasonal Adjustment and Time Series Staff, Statistics Canada.
- Cholette, P.A. (1979) "A Comparison of Various Trend Cycle Estimators," Research Paper, Seasonal Adjustment and Time Series Staff, Statistics Canada.
- Dagum, E.B. (1988) "X-11-ARIMA/88 Seasonal Adjustment Method - Foundations and Users Manual," Statistics Canada.
- Doherty, M. (1991) "Surrogate Henderson Filters in X-11," Technical Report, Statistics New Zealand, Wellington, New Zealand.
- Duffet-Smith, P (1981) "Practical Astronomy With Your Calculator, 2nd Edition," Cambridge University Press.
- Findley, D.F. (1985) "On the Unbiasedness Property of AIC for Exact or Approximating Linear Stochastic Time Series Models," *Journal of Time Series Analysis*, **6**, 229–252.
- Findley, D.F. (1999) "Akaike's Information Criterion II" in the *Encyclopedia of Statistical Science, Update Volume 3*, eds. S. Kotz, C. B. Read, and D. L. Banks. New York: John Wiley and Sons, 2–6.
- Findley, D.F. and C.C. Hood (1999). "X-12-ARIMA and Its Application to Some Italian Indicator Series," to appear in *Seasonal Adjustment Procedures – Experiences and Perspectives*, Istituto Nazionale di Statistica (ISTAT), Rome, 231-251.
- Findley, D.F., B.C. Monsell, H.B. Shulman, and M.G. Pugh (1990), "Sliding Spans Diagnostics for Seasonal and Related Adjustments," *Journal of the American Statistical Association*, **85**, 345-355.
- 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).
- Fuller, W.A. (1976), *Introduction to Statistical Time Series*, New York: John Wiley.
- Geweke, J. (1978), "Revision of Seasonally Adjusted Time Series," SSRI Report No. 7822, University of Wisconsin, Department of Statistics.
- Gomez, V., A. Maravall, and D. Peña (1999). "Missing Observations in ARIMA Models: Skipping Approach versus Additive Outlier Approach," *Journal of Econometrics*, **88**, 341-363.
- Griswold, R.E., and M.T. Griswold (1997) *The ICON Programming Language, 3rd Edition*, San Jose: Peer-to-Peer Communications.
- Hampel, F.R., E.M. Ronchetti, P.J. Rousseeuw, and W.A. Stahel (1986), *Robust Statistics: The Approach Based on Influence Functions*, New York: John Wiley.
- Hannan, E.J., and B.G. Quinn (1979), "The Determination of the Order of an Autoregression," *Journal of the Royal Statistical Society*, **B**, **41**, 190-195.
- Hillmer, S.C., and G.C. Tiao (1979), "Likelihood Function of Stationary Multiple Autoregressive Moving Average Models," *Journal of the American Statistical Association*, **74**, 652-660.
- Hood, C.C. (1998a) "X-12-Graph: A SAS/GRAPH Program for X-12-ARIMA Output, User's Guide for X-12-Graph Interactive for PC/Windows, Version 1.0",

- U. S. Census Bureau, Washington, DC.
- Hood, C.C. (1998b) "X-12-Graph: A SAS/GRAPH Program for X-12-ARIMA Output, User's Guide for X-12-Graph Batch for PC/Windows, Version 1.0", U. S. Census Bureau, Washington, DC.
- Huot, G. (1975) "Quadratic Minimization Adjustment of Monthly or Quarterly Series to Annual Totals," Research Paper, Seasonal Adjustment and Time Series Staff, Statistics Canada.
- Hurvich, C.M., and C. Tsai (1989), "Regression and Time Series Model Selection in Small Samples," *Biometrika*, 76, 297-307.
- Leser, C.E.V. (1963), "Estimation of Quasi-Linear Trend and Seasonal Variation," *Journal of the American Statistical Association*, 58, 1033-1043.
- Ljung, G.M. (1993), "On Outlier Detection in Time Series," *Journal of the Royal Statistical Society*, B, 55, 559-567.
- Ljung, G.M., and G.E.P. Box (1978), "On a Measure of Lack of Fit in Time Series Models," *Biometrika*, 65, 297-304.
- Ljung, G.M., and G.E.P. Box (1979), "The Likelihood Function of Stationary Autoregressive-Moving Average Models," *Biometrika*, 66, 265-270.
- Lothian, J. and M. Morry (1978), "A Test of Quality Control Statistics for the X-11-ARIMA Seasonal Adjustment Program," Research Paper, Seasonal Adjustment and Time Series Staff, Statistics Canada.
- Monsell, B.C. (1989) "Supplement to Technical Paper No. 15 - The Uses and Features of X-11.2 and X-11Q.2," Statistical Research Division, U. S. Census Bureau.
- Montes, M. J. (1997), "Frequency of the Date of Easter over one 400 year Gregorian Cycle" [Online]. Available: <http://www.smart.net/~mmontes/freq2.html> [1999, December 8].
- Montes, M. J. (1998), "Calculation of the Ecclesiastical Calendar" [Online]. Available: <http://www.smart.net/~mmontes/ec-cal.html> [1999, December 8].
- More, J.J., B.S. Garbow, and K.E. Hillstrom (1980), User Guide for MINPACK-1, Report ANL-80-74, Argonne National Laboratory, Argonne, Illinois.
- Otto, M.C., and W.R. Bell (1990), "Two Issues in Time Series Outlier Detection Using Indicator Variables," *Proceedings of the American Statistical Association, Business and Economic Statistics Section*, 182-187.
- Otto, M.C., and W.R. Bell (1993), "Detecting Temporary Changes in Level in Time Series," *Proceedings of the American Statistical Association, Business and Economic Statistics Section*, 170-174.
- Pierce, D.A. (1971), "Least Squares Estimation in the Regression Model With Autoregressive-Moving Average Errors," *Biometrika*, 58, 299-312.
- SAS Institute Inc. (1990), *SAS/GRAPH Software: Reference, Version 6, First Edition, Volume 1*, Cary, NC: SAS Institute.
- Schwarz, G. (1978), "Estimating the Dimension of a Model," *Annals of Statistics*, 6, 461-464.
- Shiskin, J., A.H. Young, and J.C. Musgrave (1967) "The X-11 Variant of the Census Method II Seasonal Adjustment Program," Technical Paper No. 15, U.S.

Department of Commerce, U. S. Census Bureau.

Soukup, R.J. and Findley, D. F. (2000) "Detection and Modeling of Trading Day Effects," to appear as an SRD Technical Report, U. S. Census Bureau.

Thomson, P., and T. Ozaki (1992) "Transformation and Seasonal Adjustment," submitted for publication.

Vandaele, W. (1983), *Applied Time Series and Box-Jenkins Models*, New York: Academic Press.

Wilson, G.T. (1983), "The Estimation for Time Series Models, Part I. Yet Another Algorithm for the Exact Likelihood of ARMA Models," Technical Report No. 2528, Mathematics Research Center, University of Wisconsin-Madison.

付録

この付録ではプログラム X-12-ARIMA を実際に利用する上で特に注意する必要があると思われる二つの基本的事柄について簡単に解説する。

A. X-12-ARIMA(2000) とは？

米国センサス局においては1980年代末頃から同局時系列研究グループにより新たな季節調整法の開発が X-12 開発計画として進行していた。1996年になり時系列研究グループはその頃から一般にも利用が広がりつつあったインターネットによりプログラムの公開を開始した。ここで公開と云う意味は実はセンサス局の公式なソフトウェアとして配布したということではなく、あくまで実験版 (Experimental Version) プログラムを公開したということであり、その形式は2000年12月の時点でも変化していないと思われる。1996年にインターネットで無料配布を開始したのは正確には X-12-ARIMA(β -Version) と呼ばれている版であり、 β -Version と云う名前の通りその後たびたびプログラムの誤り (バグ) 等の修正が行なわれている。(その具体的な修正箇所・改善箇所についての情報はセンサス局の X-12-ARIMA についてのインターネット URL 上にファイル Corrections(訂正箇所)として公開されている。)

その後、X-12-ARIMA プログラムは1998年になり X-12-ARIMA として利用可能になったが、ここではこのプログラムを X-12-ARIMA(1998) と呼んでおくことにする。このときに β -Version と云う呼び名がなくなったことからプログラム X-12-ARIMA の基本的な部分については計算プログラム上の問題はほぼなくなったと理解してよいと判断されよう。2000年12月時点で利用可能な版は2000年5月に修正した Version 0.2.7 であるのでここではこれを X-12-ARIMA(2000) と呼ぶことにするが、この解説稿のもとになっているマニュアルはこのプログラムに対応して2000年5月に作られた版である。

本解説の第1章にもあるようにプログラム X-12-ARIMA を動かすにはまず二通りのやり方のどちらかで最新版の計算プログラムと関係ファイルをセンサス局の URL からダウンロードする必要がある。我々は2000年11月にインターネットを利用して卓上計算機に転送したファイルをセンサス局のホームページに書かれている指示にしたがいプログラムを解凍し利用している。本解説の第1章～第5章のもととなる原稿はこの X-12-ARIMA(2000) のマニュアル原稿を直接に書いたセンサス局の X-12-ARIMA 開発の当事者から送られたものである。

なお、センサス局 X-12-ARIMA(β -Version) に関する注意点や疑問点、あるいは季節調整法を巡る論争点などについては国友直人 (1997) ”季節調整法 X-12-ARIMA の特長と問題点” (東京大学経済学部 Discussion Paper No. J-97-1, 誤植訂正版 (<http://www.e.u-tokyo.ac.jp> 中の国友の URL 上でダウンロードが可能)), 統計数理 (1997) ”季節調整法特集” (統計数理研究所) などに掲載された論文や討論などを参考にされたい。X-12-ARIMA プログラムの開発当事者による技術的部分の説明としては引用文献中の Findley et.al. (1998) が比較的詳しいので参考になろう。ただし、この論文の最終版は同名の論文 (1996) の改訂版であるが、X-12-ARIMA プログラムに関する元々の技術的説明も少し変化していることに注意しておく。

B. 季節調整法小史

経済時系列における季節的変動の分析は19世紀まで遡ることができる。スタンレー・ジェボンズを始めとする何人かの著名な経済学者が経済変数の変動におい

て季節性がかなり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議論していた。

近代的な統計学の成立とともに時系列データの分析方法も発展を遂げていく中で統計的時系列分析 (statistical time series analysis) と呼ばれる分野も成立し発展してきたが、その中では季節変動を巡る問題は常に重要な意味を持ち続けて来たようである。特に戦間期から 1950 年代頃にかけて統計的時系列分析では記述統計的手法やスペクトル分析 (spectral analysis) におけるピリオドグラム (periodogram) 解析などの手法が開発されたのでその応用などが試みられている。

こうした中で特に官庁における季節調整法に関して注目すべきこととしては NBER (National Bureau of Economic Research) で経済時系列を研究していたジュリアス・シスキン (Julius Shiskin) が米国政府統計局 (センサス局) に研究者として招かれ、当時に利用可能であった統計的分析手法を使って季節調整プログラムの開発に着手した。当時に利用可能であった季節性調整法としては連環比率法や BLS (Bureau of Labor Statistics) 調整法などがあったが、シスキンは当時ようやく利用可能となりつつあった電子計算機を利用して汎用となる季節調整プログラムの開発に乗り出したわけである。こうした中で 1954 年当時にセンサス局の季節調整プログラムとして初めて実用化されたプログラムはセンサス局 I 法と呼ばれている。この方法では季節調整の基本的方法としてよく知られるようになった移動平均法 (Moving average method) を利用して季節性を取り除き、原系列から季節調整値を計算しようとするものであった。1950 年代から 1960 年代にかけて国際的にも官庁統計では季節調整法を開発しようとする機運が高まり、ヨーロッパの当局やわが国の当局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。わが国では当時の通商産業省が MITI 法と呼ばれる季節調整プログラムを開発したのに続き、当時の経済企画庁も EPA 法という季節調整法を開発している。これら二つの季節調整プログラムはセンサス局 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。なお、この EPA 法は 1978 年に運用を中止し、MITI 法は改良された MITI 法-III として 1998 年まで運用を続けていた。ヨーロッパでもドイツのブンデス・バンク法や英国中央銀行の季節調整法などが当時から開発されいたが、近年ではヨーロッパ統計局を中心に研究が行なわれている。

一方、米国センサス局ではセンサス I 法を開発した後もシスキンを中心として精力的に改良を重ね、次々に新しく改良された方法を研究し開発を続けていた。これらの一連の方法は実験用 (Experimental methods) という意味でセンサス局 II 法の X シリーズという名前になった。数々の試行錯誤を重ねプログラム修正を重ねた末に 1965 年 (解説マニュアルは内部資料 Shiskin et al. (1967) 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。このプログラムは当時、一般に利用可能となってきたメイン・フレームの電子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を受けて利用できる形をとった。こうした中で 1970 年ごろの時期からわ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中央官庁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。ただし、センサス局 X-11 法はその内部にかなり複雑なオプションを含んでいるので X-11 法を採用した各官庁がどの様に季節調整法を利用しているかについては当時も完全に一般に公開されていたわけではない。

さて、センサス局では季節調整法としては X-11 法を開発してからしばらくの間は特に重要な研究・開発はされなかったようである。他方、統計的時系列分析の分野では 1970 年代にはボックス・ジェンキンスが提唱した ARMA (自己回帰移動平均) モデルにもとづく予測の方法が特にその実用性の観点からかなりの注目を浴びるようになっていた。こうした機運の中でカナダ・センサス局の E. Dagun を中心とするグループはセンサス局 X-11-ARIMA と呼ばれる季節調整法プログラムを 1975 年頃開発した。このプログラムが 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末端処理の問題を改善しようとしたことにある。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官庁統計ではそれほど用いられなかったようである。

時代が 1980 年代半ばになるとよく知られているように計算機を取り巻く環境が劇的に変化し始めて来た。こうした中で米国センサス局においても 1980 年代末頃から再び季節調整法を検討しようとする機運が高まって来た。より具体的にはセンサス局の研究セクションの中に時系列解析研究グループがデビット・フィンドレー (David Findley) を中心に形成され、X-12 開発の計画が具体化しはじめていった。時系列研究グループが開発した経緯を文献から判断すると従来から存在している X-11 法をそれに対してかけられてきた幾つかの批判をその枠内で解決しようとしたと見る事ができよう。基本的にはカナダ・センサス局が開発した X-11-ARIMA を基礎として、それに更に新たに幾つかの機能を付け加えることが研究開発の内容である。より具体的には本解説の第 1 章～第 5 章に述べたように従来の季節調整法 X-11 法に様々な機能を付け加えることにより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ようにしたことが大きな改訂点である。また、それとともに 1990 年代になって実現し始めたインターネットによる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公開したが、この方法は官庁統計のあり方に対する一つの画期的な新しい出来事であろう。

ここで季節調整法についてはセンサス局法以外にもこれまでに様々な方法が開発されていることにも最後に言及しておこう。特に 1970 年代末にシカゴ大学の A.Zellner 教授の季節調整法比較プロジェクトがあり、その中で報告し注目を浴びたわが国の統計数理研究所の赤池弘次 (X-12-ARIMA マニュアルに登場する H.Akaike と同一)・石黒真木夫が開発したプログラム BAYSEA が専門家の間ではよく知られている。この季節調整法はセンサス局法が基礎とする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を直接的に活用すると云う統計学的な新しいアイデアにより季節性を取り除くユニークな方法である。この季節調整プログラム BAYSEA はその後同研究所の北川源四郎により改良され、1987 年に状態空間モデルに基づく季節調整プログラム DECOMP が研究・開発されている (北川源四郎 (1993) "時系列プログラミング" (岩波) を参照)。さらに、同研究所の佐藤整尚により 1998 年に Web-decomp が開発され、インターネットの Web 上 (<http://www.ims.ac.jp/sato>) で具体的な統計計算を容易に実行できるという新しい方法により公開されている。