

季節調整プログラム DECOMP と その後の展開

統計数理研究所 北 川 源 四 郎

(1997年7月 受付)

1. はじめに

経済時系列の季節調整法としては Shiskin et al. (1967) が開発したセンサス局法 X-11 が広く使われてきたが、両側フィルターを用いるため改定変動量が大いことやパラメータの設定が恣意的に行われるなどの問題点が指摘され、1980年前後には多くの統計学者により改定量を減少させるための改良 (Dagum (1975)) や ARIMA モデル (Hilmer and Tiao (1982)), ベイズモデル (Akaike (1980a, 1980b), Akaike and Ishiguro (1980)) や状態空間モデル (Kitagawa (1981), Kitagawa and Gersch (1984), Harvey (1989)) などに基づく新しい季節調整法の開発が試みられた。このような季節調整法の再検討の流れの形成にはセンサス局自身が催した 1981 年の国際シンポジウムやアメリカ統計学会との数年にわたる研究プロジェクトなどが寄与していた。

ちょうどこのような時期に、Akaike (1980a) はベイズ型情報量規準 ABIC に基づくベイズモデルの構成法を提案し、新しい季節調整法を発表した。本稿の状態空間モデルに基づく方法は、この Akaike のベイズモデルが状態空間モデルの形で表現できることを利用したもので、これによってカルマンフィルタなどの効率的な計算法が利用できるようになるとともに、定常 AR 成分を含むより複雑な分解も可能となった。

DECOMP はセンサス局における Gersch 教授との共同研究の結果に基づき開発された状態空間モデルによる季節調整のためのプログラムで、時系列解析プログラムパッケージ TIMSAC-84 (Akaike et al. (1985)) に収められている。本稿では、プログラム DECOMP のモデルおよび計算法を簡単に解説し、その問題点および解決の方法を示すとともに、非ガウス型モデル、非線形モデル、離散分布モデルの開発などその後の発展の様子を紹介する。

2. DECOMP のモデル

DECOMP では時系列 y_n を

$$(2.1) \quad y_n = t_n + s_n + p_n + td_n + w_n$$

のようにトレンド y_n , 季節成分 s_n , 定常変動成分 p_n , 曜日効果項 td_n および偶然変動成分 w_n の和として表現する。それぞれの成分は以下のようなモデルで表現できるものと仮定する。

$$(2.2) \quad \begin{aligned} \Delta^k t_n &= v_{1n}, & p_n &= \sum_{j=1}^m a_j p_{n-j} + v_{3n} \\ \Delta_q^l s_n &= v_{2n}, & td_n &= \sum_{i=1}^7 \beta_i d_{in} \end{aligned}$$

ただし、 Δ と Δ_q はそれぞれ階差および周期が q の季節階差を表し、 B を $Bx_n=x_{n-1}$ によって定義されるシフトオペレータとすると、 $\Delta=1-B$ 、 $\Delta_q=1-B^q$ と表すことができる。また、ノイズ項については $v_{1n}\sim N(0, \sigma_1^2)$ 、 $v_{2n}\sim N(0, \sigma_2^2)$ 、 $v_{3n}\sim N(0, \sigma_3^2)$ 、 $w_n\sim N(0, \sigma^2)$ と仮定する。これらの分散は以下の時系列の分解において、信号対ノイズ比の役割をはたす重要なパラメータである。 d_{in} は n 番目の観測期間中に含まれる i 番目の曜日の数である。DECOMPの特徴はARモデルで表現できる定常変動成分を極めて自然にモデルに取り込むことができることである。

DECOMPではトレンドモデルの階差 k としては1, 2, 3のいずれかが利用できるようになっているが、 $k=3$ を使うことはほとんどないといってよい。季節階差の次数 ℓ としては1と2が利用できるが、季節成分の時間変化に一定の傾向が見られるとき以外は通常 $\ell=1$ が用いられる。ARモデルの次数は0以上で上限はプログラムの設定によるが、通常は0と2程度で十分なことが多く、あまり大きな次数を試す必要はないようである。曜日効果項については、これを入れたモデルと入れないモデルの2種類だけが計算できる。実際の計算では $\beta_1+\dots+\beta_7=0$ という制約が加えられているので自由なパラメータは6個であり、曜日効果項は

$$(2.3) \quad td_n = \sum_{i=1}^6 \beta_i (d_{in} - d_{7n})$$

と表される。DECOMPでは、 $k=2$ 、 $\ell=1$ 、 $m=0$ 、曜日調整なしがデフォルトとなっている。サブルーチンのレベルでは、このほかに外生変数に対する回帰項を入れることも可能になっているが、メインプログラムではこの機能は使われていない。

(2.1)、(2.2)のモデルは $d=k+(p-1)\ell+m+\delta$ 次元の状態空間モデル

$$(2.4) \quad x_n = F_n x_{n-1} + G_n v_n$$

$$(2.5) \quad y_n = H_n x_n + w_n$$

で表すことができる。ただし、 δ は曜日効果項を含む場合は6、含まない場合は0である。 $p=12$ 、 $k=2$ 、 $\ell=1$ 、 $m=2$ 、 $\delta=0$ の場合には状態ベクトルは15次元となり

$$(2.6) \quad x_n = [t_n, t_{n-1} \mid s_n, s_{n-1}, \dots, s_{n-10} \mid d_n, d_{n-1}]^T$$

$$F_n = \begin{bmatrix} 2 & -1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ \hline 0 & 0 & -1 & -1 & \cdots & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & \cdots & 0 & a_1 & a_2 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 0 \end{bmatrix}, \quad G_n = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ \hline 0 & 1 & 0 \\ 0 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 0 \\ \hline 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$H_n = [1 \ 0 \mid 1 \ 0 \ \cdots \ 0 \mid 1 \ 0]$$

と設定される。このように、DECOMPで用いる状態空間モデルの係数行列 F_n 、 G_n 、 H_n は実際には時間的に変化しない場合が多い。ただし、曜日効果項を考慮し $\delta=6$ とした場合には、 $H_n = [1 \ 0 \mid 1 \ 0 \ \cdots \ 0 \mid 1 \ 0 \mid d_{1n}^* \ \cdots \ d_{6n}^*]$ 、($d_{in}^* = d_{in} - d_{7n}$)となり、 H_n だけは時間変化する。

3. DECOMPの計算法

季節調整のためには、モデルの選択、パラメータの推定と最終的に得られたモデルによる時

系列の分解のための計算が必要である。パラメータの推定、モデルの選択のためには対数尤度を利用しているが、その計算には下記(3.3)式のように時系列 y_n の一期先の予測値および予測誤差分散が必要である。一方、時系列の分解のためには、すべてのデータにもとづく各状態 x_n の事後分布すなわち平滑化分布を求める必要がある。DECOMP は状態空間表現を利用し、これらの計算を効率よく行っている。

3.1 カルマンフィルタによる計算

$x_{n|n-1}$ と $V_{n|n-1}$ は時刻 $n-1$ までの観測値 $Y_{n-1} \equiv \{y_1, \dots, y_{n-1}\}$ にもとづいて求められた x_n の一期先予測分布 $p(x_n|Y_{n-1})$ の平均と分散共分散行列、 $x_{n|n}$ と $V_{n|n}$ はフィルタ分布 $p(x_n|Y_n)$ の平均と分散共分散行列とする。このとき、よく知られているように、これらは以下のカルマンフィルタにより逐次的に効率よく計算することができる (Anderson and Moore (1979), 北川 (1993))。

一期先予測

$$(3.1) \quad \begin{aligned} x_{n|n-1} &= F_n x_{n-1|n-1} \\ V_{n|n-1} &= F_n V_{n-1|n-1} F_n^T + G_n Q_n G_n^T \end{aligned}$$

フィルタ

$$(3.2) \quad \begin{aligned} K_n &= V_{n|n-1} H_n^T (H_n V_{n|n-1} H_n^T + R_n)^{-1} \\ x_{n|n} &= x_{n|n-1} + K_n (y_n - H_n x_{n|n-1}) \\ V_{n|n} &= (I - K_n H_n) V_{n|n-1} \end{aligned}$$

この逐次的なアルゴリズムでは初期値、正確には時刻 0 における状態の平均 $x_{0|0}$ および分散共分散行列 $V_{0|0}$ が必要である。この初期値の決め方については 5.2 節で説明する。

上記のカルマンフィルタの結果を用いると、季節調整モデルの対数尤度は以下のようにして簡単に求められる。

$$(3.3) \quad \ell(\theta) = -\frac{1}{2} \left\{ N \log 2\pi + \sum_{n=1}^N \log r_n + \sum_{n=1}^N \frac{\varepsilon_n^2}{r_n} \right\}$$

ただし、 $\varepsilon_n = y_n - H_n x_{n|n-1}$ 、 $r_n = H_n V_{n|n-1} H_n^T + \sigma^2$ である。

DECOMP のモデルはデータの長さが N の場合、デフォルトモデルでも $2N$ 個、AR モデルを含む場合には $3N$ 個の状態変数を持つ。したがって、通常の最小二乗法では $O(N^3)$ のオーダーの計算量が必要となる。実際には分散 σ^2 や自己回帰係数 a_j の最尤推定を非線形最適化の方法によって行うので、上記の尤度計算を多数回繰り返す必要がある。カルマンフィルタを用いると逐次的な計算により一回のフィルタリングを $O(d^3 N)$ 以下の計算量で行うことができる。

また、カルマンフィルタを用いると観測値の一部が欠測の場合にも何ら問題なく状態推定を行うことができる (北川 (1993))。さらに X-11 では両端での推定に片側フィルタを用いるために大きな改定が生じることが知られているが、カルマンフィルタでは利用できる情報をすべて用いた最適な状態推定を行うのでそのような問題が生じる余地がない。

最終的な時系列の分解のためには固定区間平滑化が行われる。これは一期先予測とフィルタの結果を利用して、後ろ向きの逐次計算

$$(3.4) \quad \begin{aligned} A_n &= V_{n|n} F_{n+1}^{-1} V_{n+1|n}^{-1} \\ x_{n|N} &= x_{n|n} + A_n (x_{n+1|N} - x_{n+1|n}) \\ V_{n|N} &= V_{n|n} + A_n (V_{n+1|N} - V_{n+1|n}) A_n^T \end{aligned}$$

を行うことにより求めることができる。

3.2 平方根フィルタ

DECOMP のプログラムは実際には前節のカルマンフィルタを使っているわけではなく、情報行列平方根フィルタ (Information Square-Root Filter, Kitagawa (1981)) と呼ばれるアルゴリズムを用いている。このアルゴリズムでは、分散共分散行列 $V_{n|n-1}$ と $V_{n|n}$ の代わりにその逆行列の平方根分解、すなわち $V_{n|n-1} = R_n^{-T} R_n^{-1}$, $V_{n|n} = S_n^{-T} S_n^{-1}$ を考え、以下のようにこれらの間の逐次的な計算を行う。

一期先予測

$$(3.5) \quad \left[\begin{array}{c|c|c} W_n & T_n & a_n \\ \hline 0 & R_n^{-1} & b_n \end{array} \right] = P \left[\begin{array}{c|c|c} U_{n-1}^{-1} & 0 & 0 \\ \hline S_{n-1}^{-1} F_{n-1}^{-1} G_n & S_{n-1}^{-1} F_{n-1}^{-1} & C_{n-1} \end{array} \right]$$

フィルタ

$$(3.6) \quad \left[\begin{array}{c|c} S_n^{-1} & c_n \\ \hline 0 & e_n \end{array} \right] = P' \left[\begin{array}{c|c} R_n^{-1} & b_n \\ \hline V_n^{-1} H_n & V_n^{-1} y_n \end{array} \right]$$

ただし、状態ベクトルおよびシステムノイズの次元をそれぞれ d および q とするとき、左辺の行列はそれぞれ $(d+q) \times (d+q+1)$ および $(d+1) \times (d+1)$ で P および P' は上三角行列に縮約するための適当な Householder 変換を表す。情報行列平方根フィルタではこの Householder 変換を繰り返し適用することにより、カルマンフィルタに対応する逐次的な計算アルゴリズムを構成している。

この逐次計算では状態の推定値は直接は計算していないが、必要な場合には

$$(3.7) \quad \begin{aligned} x_{n|n-1} &= R_n b_n \\ x_{n|n} &= S_n c_n \end{aligned}$$

により簡単に求められる。

また、平滑化分布の平均は

$$(3.8) \quad \begin{aligned} v_{n-1|N} &= W_n^{-1} (a_n - T_n x_{n|N}) \\ x_{n-1|N} &= F_n^{-1} (x_{n|N} - G_n v_{n-1|N}) \end{aligned}$$

により計算できる (Kitagawa (1981))。

3.3 パラメータの最尤推定

DECOMP のモデルには一般に、分散パラメータ $\sigma^2, \tau_1^2, \tau_2^2, \tau_3^2$ と係数パラメータ $a_1, \dots, a_m, \beta_1, \dots, \beta_6$ が含まれる。もちろん、実際には m や δ の値により、これらの一部だけが使われる場合が多い。これらのパラメータのうち、観測ノイズの分散 σ^2 および曜日効果の係数 β_1, \dots, β_6 はフィルタの内部で自動的に推定される。一方、システムノイズの分散および AR 係数は、前小節の方法により計算される対数尤度を目的関数とし、BFGS 公式を用いた準ニュートン法により最尤推定値を求めている (北川 (1993))。

4. DECOMP による計算例

図 1(a) は合衆国センサス局により公表されている INEWUO (Industry Division, newspaper, unfilled orders) データに対するトレンド、季節成分および観測ノイズからなる DECOMP の標

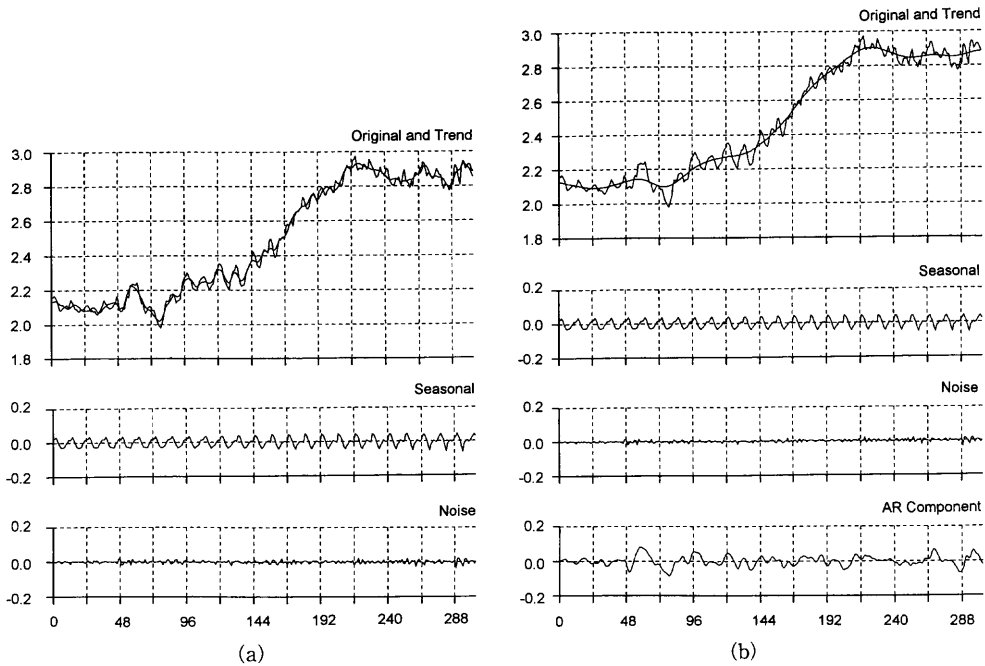


図1. INEWU データ：(a)標準的モデルによる季節調整，(b)定常 AR 項を含むモデルによる季節調整。

準モデルによる季節調整の結果である。推定されたトレンドにはうねりが見られ、季節成分は徐々に変化している。

一方、図1(b)は同じデータに対して定常 AR 項を入れて季節調整を行った結果である。この場合にはトレンドは極めて滑らかとなり、だいたいの傾向をよく捉えているように見える。標準モデルで見られたトレンドのうねりは AR モデルによって表現されている。このように標準型の季節調整と AR モデルを含む季節調整では推定されたトレンドが大きく異なることがある。ただし、その場合でも季節成分は現在の例のようにそれほど変わらない場合が多い。したがって、季節成分だけを取り除くという従来の季節調整の目的に関しては AR 項の導入はそれほど影響を及ぼさないと見える。しかし、将来の予測を考える場合には二つのモデルではその予測能力に大きな違いが見られる（北川（1986））。

図2(a)はセンサ局により公表された BFRNWS (Business Division, whole sale furniture sales) データの標準的なモデルによる季節調整の結果を示す。トレンドは非常に滑らかでデータの傾向をよく捉えているように見える。季節成分はほとんど固定的である。これに対して2(b)は曜日効果項を含む季節調整の結果を示す。トレンドは標準モデルの場合より少し変動が大きいが季節成分はほぼ同じである。曜日効果項はそれほど大きくはないが、ノイズ項の大きさと同程度で、これに対応してノイズ項は標準モデルの場合とかなり異なっている。季節成分+曜日効果の部分は現データの小さな変動までをよく再現しており、このデータの変動に対しては曜日効果が極めて重要な役割を果たしていることがわかる。

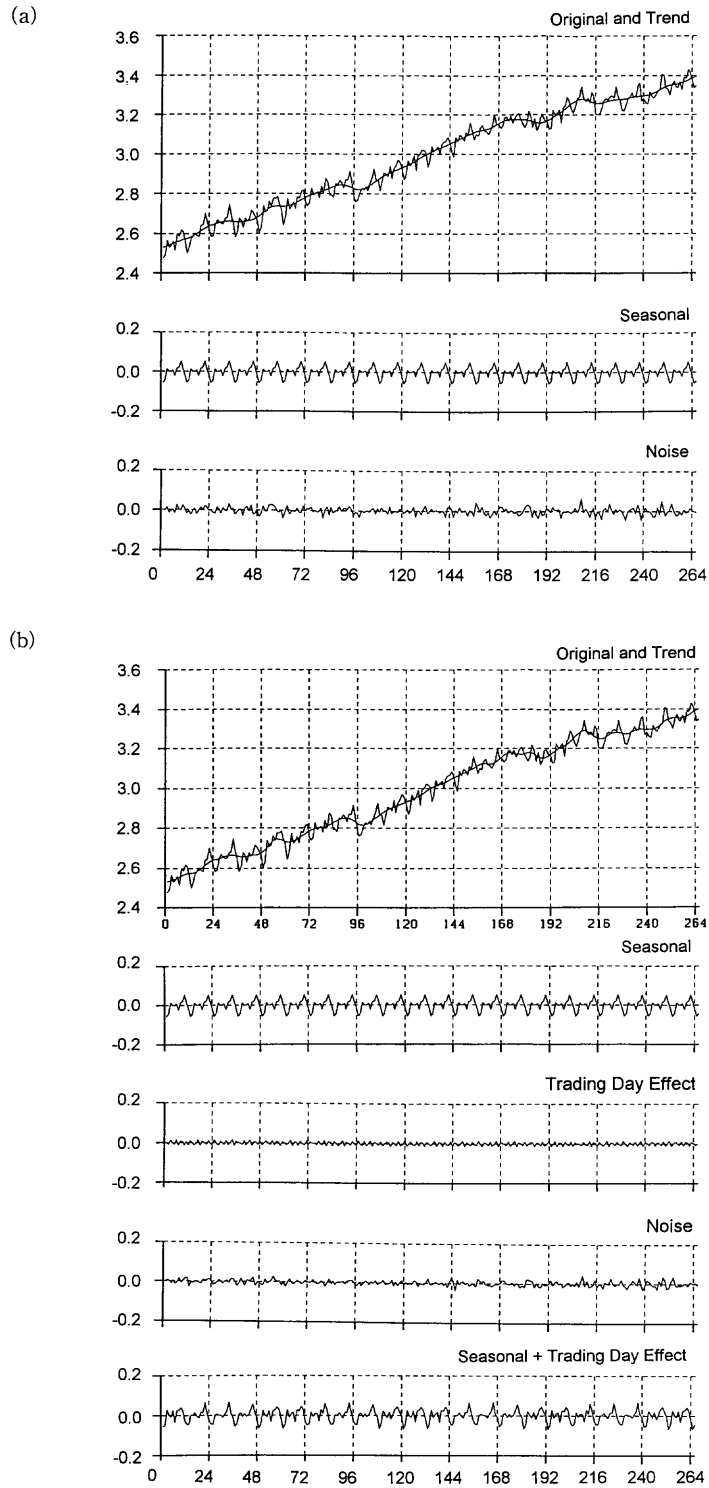


図2. BFRNWS データ：(a)標準的モデルによる季節調整，(b)曜日効果項を含むモデルによる季節調整。

5. DECOMP の問題点と改良の可能性

5.1 平方根フィルタの利用について

3.1 節に示したように DECOMP では通常のカルマンフィルタではなく、情報行列平方根フィルタを用いている。これは季節調整モデルのように非定常成分を含み非常に曖昧な事前分布から出発した場合に生じる数値的不安定性を解消することを目的としたものである。情報行列（共分散行列の逆行列）の初期値を 0 とすることにより初期状態に関する情報が全くない場合（分散無限大）にも対応できる。さらに、最小二乗法におけるハウスホルダー法に相当する平方根アルゴリズムの採用により、通常のカルマンフィルタと比較して 2 倍の計算精度が実現されている。

このように DECOMP では非定常成分の推定を前提として、やや特殊な計算法が採用されている。DECOMP を単に季節調整ソフトウェアとして利用する場合にはこれは大きな長所であるが、このプログラムを修正して独自のモデルを開発しようとする場合には、カルマンフィルタと比較してなじみの少ないアルゴリズムであるため一部の人以外には変更が困難かもしれない。この点を考慮するとむしろ汎用のカルマンフィルタを用いた方がよいかもしれない。北川 (1993) にはカルマンフィルタによる季節調整プログラムが与えられている。

5.2 初期分布の設定法について

状態空間モデルでは状態ベクトルの初期分布を与える必要があるが、DECOMP では、時間を逆転した後向きデータに対してフィルタを適用し、時刻 0 における状態の平均および分散共分散行列を求めている。これはすべてのデータを用いて初期分布を決めていることに相当する。ただし、実際には時間反転のデータを用いなくても固定点平滑化を行えば、同じ分布を求めることは可能である。ただし、いずれにせよこれらの方法では初期分布の推定のためだけに通常のフィルタと同程度の計算量が必要であり、全体では 2 倍の計算量が必要になる。

初期分布の決め方としては、 σ_0^2 を大きな値、 I_d を d 次元単位行列として、 $\sigma_0^2 I_d$ を分散共分散行列として用いる方法や最尤法などで推定する方法などが提案されている。ただし、前者の方法では σ_0^2 の大きさをどのように定めるかが問題であり、後者の方法では、得られる平均は現在のものとほぼ同じだが、必要な計算時間ははるかに長くかかるものと考えられる。

DECOMP の開発後、Ansley and Kohn (1985)、De Jong (1988)、Bell and Hillmer (1991) などによって、非定常成分に対する初期分布の決め方に関する研究が行われた。Bell and Hillmer の方法に従えば、状態ベクトルの次元を d とするとき

$$(5.1) \quad \begin{aligned} x_{d,d} &= A_d^{-1} Y_d \\ V_{d,d} &= A_d^{-1} (\Omega + \sigma^2 I_d) A_d^{-T} \end{aligned}$$

を初期値として時刻 d から計算を開始すればよい。ただし、

$$(5.2) \quad \begin{aligned} Y_d &= [y_1, \dots, y_d] \\ A_d &= [F^{-(d-1)T} H_1^T, \dots, F^{-T} H_d^T] \\ \Omega_{ij} = \Omega_{ji} &= \begin{cases} \sum_{k=1}^{d-i} H_i F^{-k} Q F^{\lambda - (k+i-j)T} H_j, & 1 \leq j \leq i \leq d-1 \text{ のとき} \\ 0, & 1 \leq j \leq i = d \text{ のとき} \end{cases} \end{aligned}$$

である。

5.3 季節成分のモデル

DECOMP で自動的に推定された季節成分はしばしば固定的になるという指摘がある。これ

は最尤法で推定された分散がほとんど0になったことによる。季節成分が固定的になることの是非については今後の検討が必要であるが、最も安易な対策としては非線形最適化において対応する分散がある範囲以下にならないようにしておくことが考えられる。

より本格的な対策としては季節成分のモデルをより定常なモデルとし

$$(5.3) \quad \begin{aligned} (1-\alpha B)(1+B+\cdots+B^{11})S_n &= v_{2n} \\ (1-\alpha B^{12})S_n &= v_{2n} \\ (1+\alpha B+\cdots+\alpha^{11})S_n &= v_{2n} \end{aligned}$$

のいずれかを用いることが考えられる。

5.4 トレンドのモデルと予測

DECOMPではトレンドのモデルとして1次から3次までの階差モデルを採用している。トレンドの推定はこのモデルで十分なことが多いが、例えば1次と2次の中間のモデル

$$(5.4) \quad (1-B)(1-\alpha B)y_n = v_{1n}$$

を考えることは可能である。このようなモデルを用いることにより、長期予測に関してはより良い結果が得られる可能性がある。

さらに、予測においてはトレンド成分は次数に対応して一定値、直線または2次曲線を伸ばすだけであり、このトレンド成分の予測誤差が全体の予測誤差の大半を占める場合が多い。トレンド成分の誤差の経験分布を用いることにより、より現実的な予測ができる可能性があり、今後の研究の課題である。

また、Harvey型のトレンドモデル

$$(5.5) \quad \begin{bmatrix} t_n \\ \Delta t_n \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} t_{n-1} \\ \Delta t_{n-1} \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_n \\ v_{1,n} \end{bmatrix}$$

を用いるとレベルシフト型のノイズを含む一般化されたトレンドモデルが得られる。ここで、 $\varepsilon_n=0$ あるいは ε_n の分散を0とおくとDECOMPのトレンドモデルが得られる。後述の非ガウス型のモデルなどでは、この二つのタイプのノイズを同時に考慮することによりトレンドの傾きやレベルが急激に変化する構造変化を自動的に処理することができるようになる。

5.5 定常成分のモデル

DECOMPでは定常変動成分に対してARモデルを仮定している。解析の目的が、例えば数年程度の循環的変動のように特定の周波数成分を捉えることにある場合には、このARモデルの係数の推定において固有根の偏角が対応する角度の範囲に入るように制限を加えてやればよい。例えば、次数 $2k$ の場合には

$$(5.6) \quad B^{2k} + a_1 B^{2k-1} + \cdots + a_{2k} = \prod_{i=1}^k (B^2 - 2r_i \cos \theta_i B + r_i^2)$$

とし、 θ_i を適当に定めた $\theta_{\min} < \theta < \theta_{\max}$ の範囲で探せばよい。

一方、定常変動成分を入れたモデルでは、トレンド t_n と定常変動成分 p_n の取り合いにより、ARモデルの次数を変えたときにそれぞれの成分の推定値が大きく変化することがある。この不安定性を避けるためには、ARモデルを用いる場合には $w_n=0$ と仮定することが考えられる。これは、ノイズ項 w_n が白色ではなく、AR過程であるとみなすことに相当する。

5.6 曜日・祝日調整のモデル

DECOMP の曜日調整では、すべての曜日それぞれが別々の係数を持つ最も一般のモデルを用いている。しかし、たとえば通常の営業日をすべて同じように扱うモデルを導入することは簡単である。そのためには曜日効果項を

$$(5.7) \quad td_n = \beta_1(d_{1n} - d_{7n}) + \beta_2(d_{2n} + \dots + d_{6n} - 5d_{7n})$$

とすればよい。ただし、 β_1 は日曜日の係数、 β_2 は月曜-金曜に共通の係数である。

祝日の影響については、ある程度は季節変動の中に組み込まれ自動的に処理されている。ただし、季節変動としては祝日が日曜、土曜などと重なる影響までは調整できないので、厳密には祝日項を導入した方がよいと考えられる。この場合には、 \bar{d}_{in} を祝日以外の i 番目の曜日の数、 \bar{d}_{0n} は第 n 期中の祝日の総数、 $\bar{\beta}_0, \dots, \bar{\beta}_6$ をパラメータとして、

$$(5.8) \quad td_n = \bar{\beta}_0 \bar{d}_{0n} + \sum_{i=1}^6 \bar{\beta}_i (\bar{d}_{in} - \bar{d}_{7n})$$

と表される。

5.7 改定変動量

DECOMP による季節調整は利用できるすべてのデータを用いてベイズ推定を行っている。したがって、新しいデータが利用できるようになるとそれに伴って過去の推定値も修正されることとなる。これは最適あるいは準最適な推定値の性質として当然のことと考えるべきであるが、季節調整の分野では改定変動と呼ばれ、実務家にとっては好ましくないことと考えられているようである。

X-11 における改定の主要な原因は、端点において片側フィルタを用いて最適でない推定を行ったことにあるので、DECOMP の場合にもこのような要請に応じることが適当かどうか自体が問題である。しかしながら、もし必要であれば、固定区間平滑化の代わりにフィルタや固定ラグ平滑化を使えば一定期間後はまったく改定がないようにすることは極めて簡単である。また、なるべく改定が少ない範囲でよい最終推定値を求めるためには改定量にペナルティを課したフィルタを構成することも考えられる。

5.8 スペクトルの溝について

DECOMP による分解の残差系列にスペクトルの溝が見られることがあるという指摘があるが(木村(1996))、この問題については川崎・佐藤(1997)の議論にゆずりたい。

6. ユーザインターフェイスの開発

DECOMP は計算結果としてパラメータや AIC などの統計量とともに各成分の図形をプリントあるいはディスプレイに出力するようになっている。プログラムは FORTRAN で大型計算機用として開発されたものであるが、CALCOMP 系のグラフィックライブラリを用いるとワークステーションやパーソナルコンピュータでも計算や図の出力ができる。ただし、それらのソフトウェアはそれほど普及しているわけではない。

しかしながら、最近、以下のようなユーザインターフェイスを加えたソフトウェアが開発されており、WS や PC でもかなり手軽に DECOMP を利用できるようになっている。

6.1 TIMSAC for Windows

統計数理研究所の田村義保氏は TIMSAC を WINDOWS 上で利用できるソフトウェア

TIMSAC for Windows を開発している。これは TIMSAC シリーズの多くのプログラムを含む大規模なものであるが、TIMSAC-84 のプログラムのひとつとして、DECOMP も Windows 上で計算できるようになっている。ただし、プログラムは C 言語によりカルマンフィルタを用いて書かれている。

6.2 S-DECOMP

統計数理研究所の佐藤整尚氏は統計解析システム S 上で DECOMP を使えるようなソフトウェア S-DECOMP を開発している。現在では S は WS でも PC でも利用できるもので、S-DECOMP はどちらでも利用できることになる。

6.3 Web DECOMP

同じく佐藤整尚氏は、ホームページ上(URL は <http://www.ism.ac.jp/~sato>)で DECOMP による季節調整を行うシステム Web DECOMP を開発している。

このシステムを用いると、インターネットへの接続が可能であればソフトウェアをインストールすることなく、DECOMP での季節調整を手軽に実行することが可能である。詳しくは佐藤 (1997) を参照されたい。

7. 非線形・非ガウス型の季節調整モデル

7.1 一般の状態空間モデルと非ガウス型フィルタ

一般化状態空間モデル (Kitagawa (1996))

$$(7.1) \quad \begin{aligned} x_n &\sim Q(x|x_{n-1}) \\ y_n &\sim R(y|x_n) \end{aligned}$$

は、線形・ガウス型の状態空間モデルを一般化し、モデルが非線形あるいは分布が非ガウスの場合にも適用できるようにしたものである。一般化状態空間モデルの利用により、線形・ガウスという従来の枠にとらわれずに自由なモデリングが可能となる。

このような一般の状態空間モデルに対してはカルマンフィルタではよい状態推定値は得られないが、近年、さまざまなタイプの非線形・非ガウス型フィルタが開発されている (Kitagawa (1987, 1994, 1996))。以下に季節調整に利用される状態推定アルゴリズムを示す。

7.1.1 ガウス和フィルタ

ガウス和フィルタは元来、非線形システムのフィルタとして開発されたものであるが (Alspach and Sorenson (1972))、線形・非ガウス型状態空間モデルのフィルタとしても利用できる。線形状態空間モデル (2.4), (2.5) において、システムノイズあるいは観測ノイズの分布が混合ガウス分布で

$$(7.2) \quad q(v) = \sum_{i=1}^{K_v} \alpha_i \varphi_i^q(v), \quad r(w) = \sum_{j=1}^{K_w} \beta_j \varphi_j^r(w)$$

と表されるものとする。ここで、 α_i と β_j は ($\alpha_1 + \dots + \alpha_{K_v} = 1, \beta_1 + \dots + \beta_{K_w} = 1$) をみたす重み係数(混合比)、 $\varphi_i^q(v)$ と $\varphi_j^r(w)$ は適当な正規密度関数である。初期分布も混合ガウス分布に従うとき、状態 x_n の予測分布 $p(x_n|Y_{n-1})$ および、フィルタの分布 $p(x_n|Y_n)$ も混合ガウス分布で

$$(7.3) \quad \begin{aligned} p(x_n|Y_{n-1}) &= \sum_{k=1}^{L_n} \gamma_{kn} \varphi_k(x_n|Y_{n-1}) \\ p(x_n|Y_n) &= \sum_{\ell=1}^{M_n} \delta_{\ell n} \varphi_{\ell}(x_n|Y_n) \end{aligned}$$

と表すことができる。ガウス和フィルタの特長はこの表現において、混合比 γ_{kn} , δ_{en} および各正規分布の平均および分散共分散行列などがカルマンフィルタを利用して計算できることである。すなわち、カルマンフィルタを(たとえば5-20個位)並列に走らせることにより、非ガウス分布に対応できるフィルタが得られる。Kitagawa (1994) では混合ガウス近似を用いた平滑化アルゴリズムも与えられている。

7.1.2 モンテカルロ・フィルタ

モンテカルロ・フィルタは任意の確率分布を多数(例えば1万個)の粒子で近似することに基づいている。予測分布とフィルタ分布を近似する粒子をそれぞれ $p_n^{(j)}$, $f_n^{(j)}$ ($j=1, \dots, M$, M は粒子数)と表すことにすると、これらはシステムノイズに従う乱数 $v_n^{(j)}$ と観測値 y_n を用いて簡単に計算できる(Kitagawa (1996))。この方法は、乱数を用いた近似を行うためにある程度の誤差変動は避けられないが、極めて広範な非線形・非ガウス型のモデルに対して適用できることが特長である。

実際、モンテカルロフィルタの利用により、非線形・非ガウス型モデルによる季節調整、離散変数の季節調整あるいは自己組織型のモデルなどが比較的簡単に実現できる。

7.2 さまざまな季節調整モデル

7.2.1 非ガウス型モデル

線形の季節調整モデル(2.1), (2.2)においてシステムノイズあるいは観測ノイズの分布に対して正規性を仮定しないものが非ガウス型季節調整モデルである。

Kitagawa (1994) ではシステムノイズの分布として、二つの正規分布の混合

$$(7.4) \quad q(v) = (1 - \alpha)\varphi_0(v) + \alpha\varphi_1(v)$$

を利用することにより、経済時系列のレベルシフトや季節パターンの急激な変化などの構造変化を自動的に検出できることが示されている。ただし、 α は構造変化の確率に対応する混合比、 φ_0 と φ_1 はそれぞれ通常の変化および構造変化時の急激な変化に対応する正規密度関数である。

図3は日本の民間企業在庫データの解析例である。データの中ほどに大きな変化が見られるが、標準的な季節調整ではトレンドは滑らかに変動し、急激な変化は捉えられない。また、

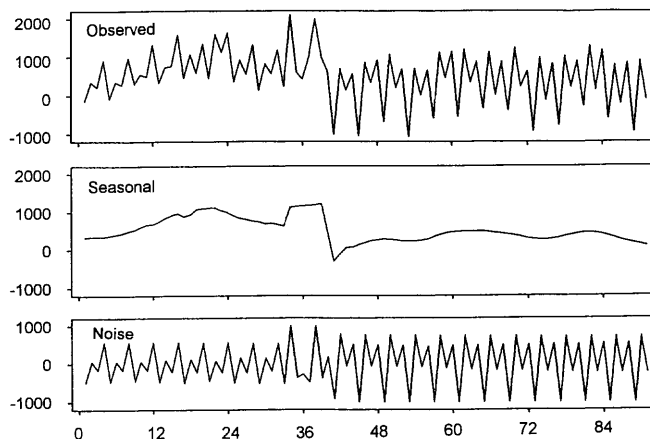


図3. 民間企業在庫データ：構造変化を含むデータと混合ガウスモデルをシステムノイズに用いたモデルによる季節調整。

季節成分はデータ区間の最初から最後にかけて徐々に変化する。これに対して、非ガウス型のシステムノイズを用いて計算した図3はこれとかなり異なった結果を示している。推定されたトレンドは中ほどで2回大きくジャンプし、大きな構造変化があったことを示唆している。また、季節成分の途中で急激に変化している。このように非ガウス型のノイズを用いることにより急激な変化を捉えることができる。この方法はその変化時点を前もって与えておかなくても自動的に検出できることに特長がある。

一方、観測ノイズに対して同様に混合ガウス分布

$$(7.5) \quad r(w) = (1 - \beta)\varphi_0(w) + \beta\varphi_1(w)$$

を用いると時系列の異常値の処理を自動的に行うことができる (Kitagawa (1994))。この論文では、人工的に異常値を加えたデータを解析し、標準的な季節調整と非ガウスモデルによる季節調整を比較し、以下のようなことが示されている。標準的な季節調整法では異常値の影響を強く受け、推定されたトレンドは大きく変動している。また、異常値がない場合の結果と比較するとその他の部分のトレンドも変動が大きくなっていることがわかる。これに対して、観測ノイズに混合ガウス分布を用いた季節調整では異常値の影響をほとんど受けずに、異常値はすべて観測ノイズとして表現される。またこの論文では、混合ガウス型分布による季節調整法は、異常値を含まないデータに適用してもほとんど悪影響がないことも示されている。

7.2.2 非線形モデル

Kitagawa and Nagahara (1995) にはモンテカルロフィルタを用いると自由に非線形モデルにもとづく季節調整ができることが示されている。例えば、 w_n を対数正規分布として積型の季節調整モデル

$$(7.6) \quad y_n = T_n \times S_n \times W_n$$

を(2.2)式とともに直接推定することができる。また、ノイズ項だけが加法的な混合モデル

$$(7.7) \quad y_n = T_n \times S_n + W_n$$

も推定することができる。さらに、Kitagawa (1995) には、より複雑な非線形季節調整モデル

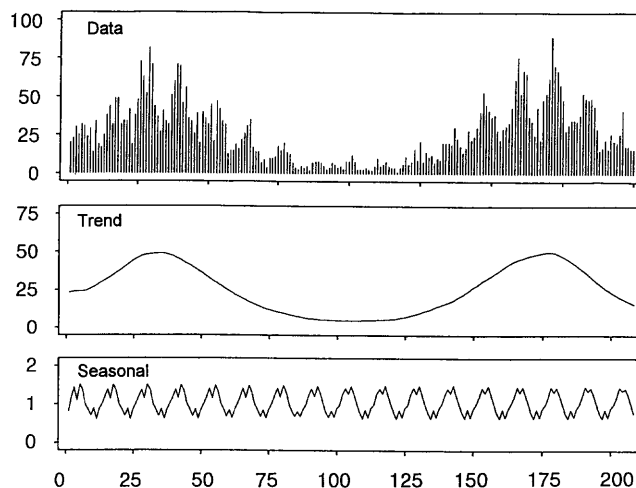


図4. 平均値が季節変動する非定常ポアソン過程の解析例。人工的に生成したデータと離散分布モデルによる季節調整。

がとりあげられている。

7.2.3 離散・カウントデータのモデル

経済データあるいは医学統計のデータのなかには、計量ではなく計数によって得られたデータが多く、特にその数が少ない場合には、連続変量とみなす近似は不適當になる。モンテカルロフィルタでは、このような場合、観測値にポアソン分布などの離散分布を想定し、その平均などがトレンドと季節変動をもつ以下のようなモデルをあてはめることができる。

$$\begin{aligned}
 (7.8) \quad y_n &\sim \frac{e^{-\lambda_n} \lambda_n^{y_n}}{y_n} \\
 \lambda_n &= \exp\{t_n + S_n\} \\
 \Delta^k t_n &= v_{1n} \\
 \Delta^q S_n &= v_{2n}
 \end{aligned}$$

図4は平均値がトレンドと季節変動をもつポアソン過程から人工的に発生させたデータの解析結果である。上から順に、生成されたデータ、推定されたトレンドと季節成分を示す。非線形・非ガウス型モデルを用いるとこの例の計数データのように直接ポアソン分布等の離散分布をあてはめることができる。さらにこの方法は、この例の場合よりもさらに平均が小さく0が観測されるような場合にも問題なく適用できる。

7.3 自己組織型モデルとパラメータの自動調整

モデル型季節調整法には通常いくつかのパラメータが含まれる。これまでに開発された多くの季節調整法では最尤法などの方法により推定値を求めている。したがって、1回の計算時間は短くてもそれを数十回あるいは数百回計算する必要がある。さらに、いろいろな次数やオプションを試す場合にはその計算を何回も繰り返すことになる。計算機が高速となった現在では、カルマンフィルタが利用できる標準の季節調整モデルの場合はこれでも実用上ほとんど問題無いが、計算量が大きな非線形・非ガウス型モデルの場合にはやはり大きな問題である。

Kitagawa (1995) では状態ベクトルのなかに未知のパラメータを含めることにより、季節調整のための分解とパラメータ推定を同時に行う方法を提案している。この方法をパラメータが3つだけの最も簡単な DECOMP のデフォルトの季節調整モデルに適用するには、状態ベクトルを

$$(7.9) \quad z_n = (t_n, t_{n-1}, S_n, \dots, S_{n-10}, \log \tau_{1n}^2, \log \tau_{2n}^2, \log \sigma_n^2)^T$$

と拡大し、モンテカルロフィルタで z_n を推定すればよい。

(7.9)から明らかのように z_n の事後分布を求めることにより、トレンドや季節成分と同時に $\tau_{1n}^2, \tau_{2n}^2, \sigma_n^2$ などのパラメータも推定することができる。この方法の特徴は、従来の DECOMP などが最尤法でパラメータ推定を行った後に平滑化により成分分解を行うのに対して、パラメータ推定と状態推定を同時に行うことである。ただし、 $\log \tau_{1n}^2, \log \tau_{2n}^2, \log \sigma_n^2$ に対するモデルが必要であるが、通常はランダムウォーク型のモデル

$$\begin{aligned}
 (7.10) \quad \log \tau_{1n}^2 &= \log \tau_{1, n-1}^2 + \varepsilon_{1n}, & \varepsilon_{1n} &\sim N(0, \xi_1^2) \\
 \log \tau_{2n}^2 &= \log \tau_{2, n-1}^2 + \varepsilon_{2n}, & \varepsilon_{2n} &\sim N(0, \xi_2^2) \\
 \log \sigma_n^2 &= \log \sigma_{n-1}^2 + \varepsilon_{3n}, & \varepsilon_{3n} &\sim N(0, \xi_3^2)
 \end{aligned}$$

を仮定する。ここで、 $\xi_i^2 > 0$ とすることによりパラメータが時間変化する場合にも対応できる。数値的には、 $\xi_i^2 = 0$ とした場合には粒子の数が減少する問題があり、むしろ $\xi_i^2 > 0$ の場合のほうが安定した結果が得られる。

8. ま と め

季節調整プログラム DECOMP のモデル, 計算法および計算例を紹介するとともに, 現在までに指摘されたいくつかの問題点に関して改良の可能性を検討した。さらに最近急速に発達した非線形・非ガウス型の状態空間モデルを用いることにより, 構造変化や異常値への対応が自動的に行えること, 非線形の季節調整モデルや離散分布のモデルなどが自由に取り扱えることを示した。

なお, 査読者から季節成分の分散が小さすぎる場合があることに関連して, 異なるモデルの比較のための評価規準に関しての議論を加えることが望ましいという指摘があったが, この問題を議論することは本稿の範囲を超えるのでここではこの問題についてはとりあげなかった。この問題に関連してベイズモデルの予測能力の評価のための情報量規準が Kitagawa (1997) に与えられている。

謝 辞

日頃, 季節調整に関する議論に加わっていただいている統計数理研究所の樋口知之氏, 川崎能典氏, 佐藤整尚氏およびコメントをいただいた査読者および編集委員に感謝いたします。なお, この研究の一部は文部省科学研究補助金の国際学術研究 08045018, 基盤研究 C(2)09680318 および統計数理研究所共同研究 9-共研 A-41 として行われた。

参 考 文 献

- Akaike, H. (1980a). Likelihood and the Bayes procedure in Bayesian statistics, *Bayesian Statistics* (eds. J. N. Bernardo, M. H. DeGroot, D. V. Lindley and A. F. M. Smith), 141-166, University Press, Valencia, Spain.
- Akaike, H. (1980b). Seasonal adjustment by a Bayesian modeling, *J. Time Ser. Anal.*, **1**, 1-13.
- Akaike, H. and Ishiguro, M. (1980). BAYSEA, a Bayesian seasonal adjustment program, *Comput. Sci. Monographs*, No. 13, The Institute of Statistical Mathematics, Tokyo.
- Akaike, H., Ozaki, T., Ishiguro, M., Ogata, Y., Kitagawa, G., Tamura, Y.-H., Arahata, E., Katsura, K. and Tamura, Y. (1985). TIMSAC-84 Part 1, *Comput. Sci. Monographs*, No. 22, The Institute of Statistical Mathematics, Tokyo.
- Alspach, D. L. and Sorenson, H. W. (1972). Nonlinear Bayesian estimation using Gaussian sum approximations, *IEEE Trans. Automat. Control*, **AC-17**, 439-448.
- Anderson, B. D. O. and Moore, J. B. (1979). *Optimal Filtering*, Prentice-Hall, New Jersey.
- Ansley, C. F. and Kohn, R. (1985). Estimation, filtering, and smoothing in state space models with incompletely specified initial conditions, *Ann. Statist.*, **13**, 1286-1316.
- Bell, W. and Hillmer, S. (1991). Initializing the Kalman filter for nonstationary time series models, *J. Time Ser. Anal.*, **12**, 283-300.
- Dagum, E. B. (1975). Seasonal factors forecasts from ARIMA models, *Proceedings of the 40th Session of the International Statistical Institute, Warsaw*, 206-219.
- De Jong, P. (1988). The likelihood for a state space model, *Biometrika*, **75**, 165-169.
- Harvey, A. (1989). *Forecasting Structural Time Series Models and the Kalman Filter*, Cambridge University Press, Cambridge.
- Hilmer, S. and Tiao, G. (1982). An ARIMA-based approach to seasonal adjustment, *J. Amer. Statist. Assoc.*, **77**, 63-70.
- 川崎能典, 佐藤整尚 (1997). 季節調整の「最適性」について, *統計数理*, **45**(2), 245-263.
- 木村武 (1996). 季節調整法の評価に関する実証分析, *日本統計学会誌*, **26**(3), 269-286.
- Kitagawa, G. (1987). Non-Gaussian state space modeling of nonstationary time series (with discussion), *J. Amer. Statist. Assoc.*, **82**, 1032-1063.
- Kitagawa, G. (1981). A nonstationary time series model and its fitting by a recursive problems, *J. Time*

- Ser. Anal.*, **2**, 103-116.
- 北川源四郎 (1986). 時系列の分解—プログラム DECOMP の紹介—, *統計数理*, **34**(2), 255-271.
- 北川源四郎 (1993). 『時系列解析プログラミング』岩波書店, 東京.
- Kitagawa, G. (1994). The two-filter formula for smoothing and an implementation of the Gaussian-sum smoother, *Ann. Inst. Statist. Math.*, **46**(4), 605-623.
- Kitagawa, G. (1995). Self-organizing state space model, Research Memo., No. 578, The Institute of Statistical Mathematics, Tokyo.
- Kitagawa, G. (1996). Monte Carlo filter and smoother for non-Gaussian nonlinear state space models, *Journal of Computational and Graphical Statistics*, **5**(1), 1-25.
- Kitagawa, G. (1997). Information criteria for the predictive evaluation of Bayesian models, *Comm. Statist. Theory Methods*, **26**(9), 2223-2246.
- Kitagawa, G. and Gersch, W. (1984). A smoothness priors-state space modeling of time series with trend and seasonality, *J. Amer. Statist. Assoc.*, **79**, 378-389.
- Kitagawa, G. and Nagahara, Y. (1995). Monte Carlo smoothing method for seasonal adjustment, Research Memo., No. 579, The Institute of Statistical Mathematics, Tokyo.
- 佐藤整尚 (1997). Web DECOMP の紹介—WWW 上で行う季節調整システム—, *統計数理*, **45**(2), 233-243.
- Shiskin, J., Young, A.H. and Musgrave, J. C. (1967). The X-11 variant of the census method II seasonal adjustment program, Tech. Paper 15, Bureau of the Census, Washington, D. C.

Seasonal Adjustment Program DECOMP and Development of Various Models Thereafter

Genshiro Kitagawa

(The Institute of Statistical Mathematics)

The basic models and computational method of the seasonal adjustment program DECOMP are shown. The models are expressed in linear Gaussian state space model and the unknown state vectors are estimated by information square root filter and smoother algorithms. Two numerical examples are shown to exemplify the method.

Based on the experiences since the development of DECOMP, various possibilities of improving the method are shown. Recent developments related to the user interfaces such as the Web DECOMP are briefly shown. Finally, the use of nonlinear non-Gaussian seasonal adjustment model is discussed. It is shown that by using a proper non-Gaussian model, the structural changes and outliers can be automatically treated. The use of nonlinear seasonal adjustment model and seasonal adjustment of discrete process are also discussed.