

公開講演会要旨

非定常時系列モデルと最適化

統計数理研究所 北 川 源 四 郎

(平成元年 11 月 1 日, 統計数理研究所 講堂)

1. はじめに: パラメトリックモデルとモデル評価

統計的データ解析では、与えられたデータに対し適当な統計的モデルを構成し、それにもとづいて有益な情報を抽出しようとする。よい情報を得るためにはよいモデルが必要で、そのためには統計的モデルを評価する基準にもとづく最適化が必要である。赤池によって提唱されたエントロピー最大化原理に従えば、統計的モデルのよさは KL 情報量 (エントロピー) によって評価される。しかし、実際のデータ解析においては通常、真の分布はわからないので KL 情報量はそのままの形では役に立たない。対数尤度は (定数項を無視して) KL 情報量を推定したものであると解釈できる。従ってモデルが未知のパラメータを持つ場合には、対数尤度を評価基準として最適化を行なうことによって、よいモデルが得られる。これが最尤法であり、統計的モデルの統一的な推定法を与える。

例として、非定常時系列のトレンドを多項式によって推定する問題を考えてみよう。多項式回帰モデルには係数と残差分散がパラメータとして含まれているが、これらのパラメータは最尤法によって簡単に推定できる。しかし、統計的モデリングの観点からすると、これは最適化の第 1 のステップにすぎない。多項式回帰の場合、いくら精度よくパラメータ推定を行なっても、多項式の次数が不適當であれば、まったく無意味なトレンドの推定値が得られることがあることはよく知られている。赤池情報量規準 AIC はパラメータが最尤法で定められたモデルの KL 情報量を (定数項を無視して) 推定したものである。従って AIC を最小化することによって次数の決定ができ、これが統計的モデリングにおける最適化の第 2 ステップを構成する。

AIC の導入により次数選択の問題もほぼ機械的に行なえるようになったので、統計的モデリングとは与えられたモデル族の中で AIC 最小のモデルを機械的に捜せば事足りるという印象を与えるかもしれない。しかし実際はまったく逆で、モデル間の客観的な比較が可能となったことによって、よりよいモデルを利用することがたいせつであることがはっきりと認識されるようになった。これによってモデル選択という最適化の第 3 ステップが生じるが (図 1)、このステップは前もって決まったモデル族の中で AIC を最小化するだけで解決するのではなく、対象の特性および解析の目的を考慮して、よりよいモデルの開発が要求される。

本稿では、非定常時系列解析におけるトレンド推定の問題を例としながら、統計的モデリングにおいてこのような意味での最適化、すなわちモデルの改良が行なわれる様子を考えることにする。

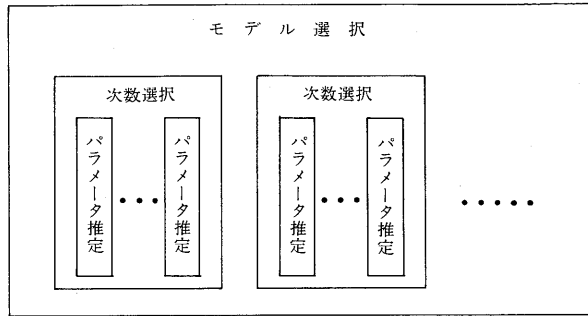


図1. 統計的モデリングにおける3段階のモデリング.

2. ベイズモデルと状態空間表現の利用

現実のデータのトレンドは、多項式のような簡単なパラメトリックモデルではうまく表現できないものが多い。このような場合には、大規模なパラメトリックモデルが用いられるようになってきている。トレンド推定では

$$(2.1) \quad \text{観測値 } (y_n) = \text{トレンド}(t_n) + \text{白色雑音 } (w_n)$$

という表現を考えるが、従来のモデリングではトレンド t_n に多項式などの硬いモデルを想定した。しかしデータが少ない場合、AICによる次数選択では低い次数が選択され、トレンドの動きがうまく表現できないことがある。この場合、単に次数を高くしても変動が大きくなるだけで、よい結果は得られないことはよく知られている。この問題を解決するために、最近ではトレンド成分に直接、多項式のような関数形を仮定せず、トレンド成分の時間変化の仕方にモデルを導入する方法が用いられている。例えば、トレンド成分は滑らかに変化する成分であると考え、トレンド項のモデル(成分モデル)として k 階差分が0に近いとするモデル、すなわち

$$(2.2) \quad \Delta^k t_n = v_{1n}$$

が考えられる。 v_{1n} は平均0、分散 τ^2 の正規分布に従う白色雑音で、 $\tau^2=0$ の場合には、このトレンド成分モデルの実現値は $k-1$ 次の多項式となるが、一般に $\tau^2>0$ の場合には、局所的には多項式に近い性質を持ちながら大局的には極めて柔軟な関数を表すことができる。

(2.1) と (2.2) の二つのモデルは ($k=2$ の場合)

$$x_n = \begin{bmatrix} t_n \\ t_{n-1} \end{bmatrix} \quad F = \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix} \quad G = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad H = [1 \quad 0]$$

とおくことによって状態空間表現

$$(2.3) \quad \begin{aligned} x_n &= Fx_{n-1} + Gv_n \\ y_n &= Hx_n + w_n \end{aligned}$$

の形で表現することができるので、カルマンフィルタを利用すれば条件付き分布 $p(y_n|y_1, \dots, y_{n-1}, \theta)$ が自動的に計算できる。

一般に時系列モデルの場合には観測値は独立ではないので、尤度の計算のためには観測値 y_1, \dots, y_N の同時分布を考える必要があるが、条件付き分布を用いると

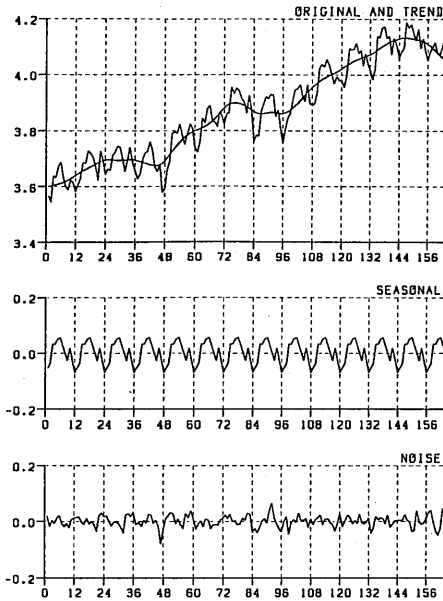


図2. RAUTODLRS データ (アメリカの自動車の小売り販売高) の対数の標準的季節調整.

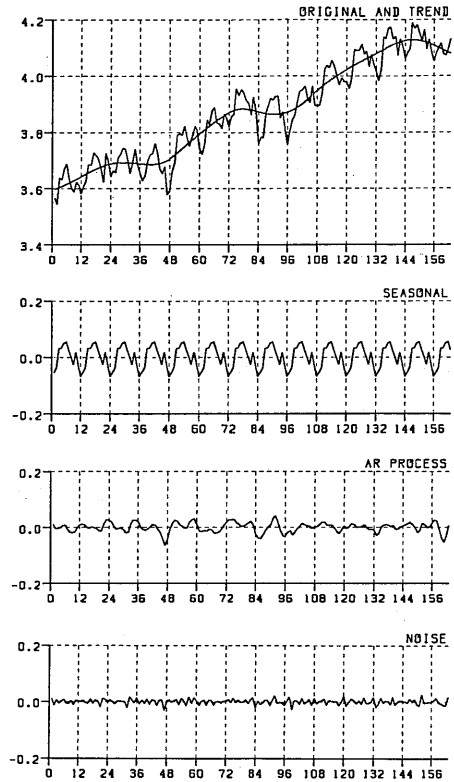


図3. RAUTODLRS データの AR 項を含む季節調整.

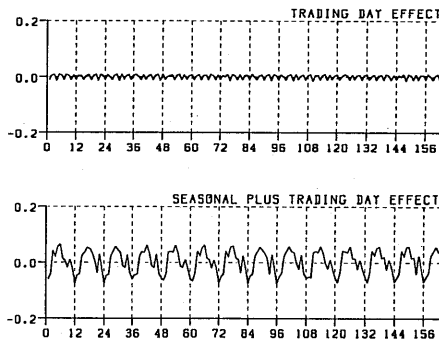


図4. RAUTODLRS データの曜日効果の推定値と曜日効果と季節成分の和.

という分解を考えるとよい.

図3には(3.2)の場合と同様に, 状態空間表現を利用してこの分解を行なった結果が示してある. 標準的モデルの場合には $AIC=580.1$ であったが, この AR 成分を含むモデルの場合 $AIC=550.3$ で, あてはまりがよいことがわかる. 北川(1986)にはこの二つのモデルによって長期予測を行なった場合の相違が示してある.

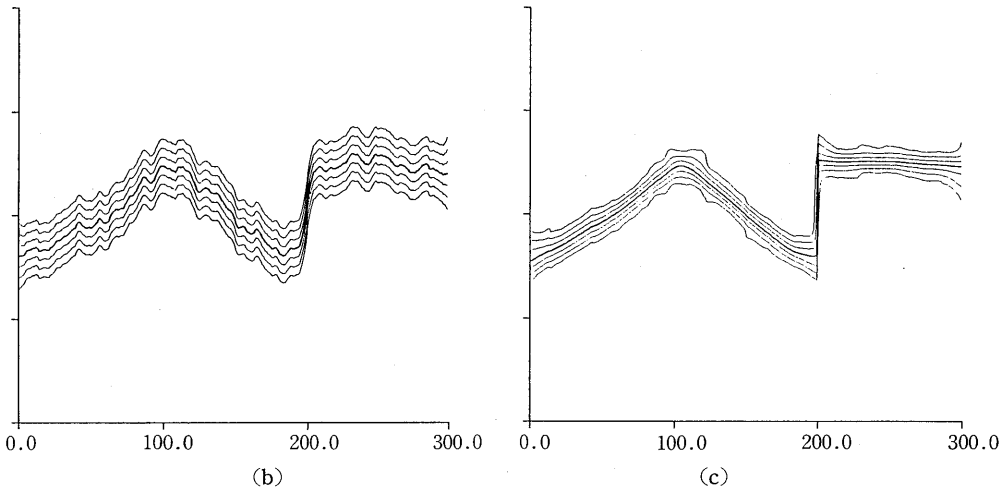
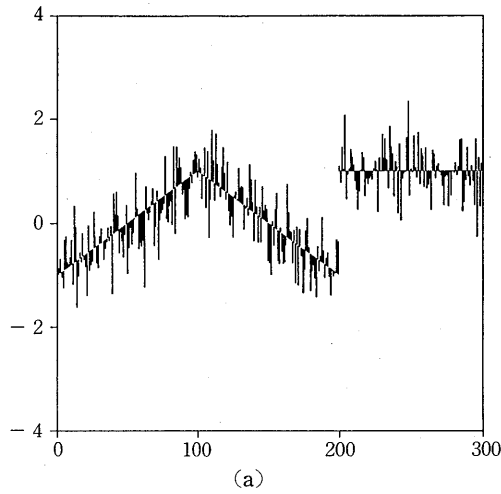


図5. (a) データ, (b) ガウス型モデルによるトレンドの推定値, (c) 非ガウス型モデルによるトレンドの推定値.

[曜日調整]

デパートの毎月の売上高などのデータでは、ひと月の中に含まれる日曜日や土曜日などの数の違いが大きな影響を及ぼすことがある。この場合、第 n 番目の月に含まれる各曜日の数を $d_{ni}, i=1, \dots, 7$ と表すと

$$(3.5) \quad td_n = \beta_1 d_{n1} + \dots + \beta_7 d_{n7}$$

によって曜日効果を表すことができる。従って

$$(3.6) \quad \text{観測値}(y_n) = \text{トレンド}(t_n) + \text{季節成分}(s_n) + \text{曜日効果}(td_n) + \text{ノイズ}(w_n)$$

という分解を行えば、曜日効果を考慮にいたした季節調整ができることになる。必要ならばAR項を加えてもよい。

図4には、図2, 3と同じデータにこのタイプの季節調整を行なった結果が示してある。季節成分と曜日効果の和がトレンドのまわりの変動をよく表していることがわかる。このモデルの場合、 $AIC=503.1$ で標準的モデルのAICとくらべてかなりよいことが確認できる。

4. 分布の形を変えてみる：非ガウス型モデル

これまでトレンドの推定のためにいろいろなモデルを考えてきたが、実際のデータ解析ではトレンドにジャンプがあったり、異常値が含まれていて、これらの取扱いに困ることがある。この場合には、(2.3)において、 v_n あるいは w_n にコーシー分布などの非ガウス分布を用いた非ガウス型のモデルを考えると便利である。このような非ガウス型モデルに対しては、線形ガウス性を利用したカルマンフィルタや固定区間平滑化のアルゴリズムは利用できないが、類似の非ガウス型フィルタや平滑化の公式を導くことができるので、数値近似等を利用した計算が可能である (Kitagawa (1987), 北川 (1989))。

図5にはトレンドにジャンプと折れ曲がりがあるデータと、ガウス型および非ガウス型モデルによるトレンドの推定値が示してある。ガウス型モデルではトレンドが波打ち、またジャンプが明確には捉えられていないが、非ガウス型モデルの利用によって、滑らかでかつジャンプのあるトレンドをうまく再現できていることがわかる。この方法は非ガウス型の季節調整に拡張できる (Kitagawa (1989))。

5. ほかの情報を使ってみる

いかに巧妙なモデルを考えても、決まった情報を用いていたのでは限度がある。このような場合、ほかに使える情報を捜すことによって、著しい改善が得られることがある。例えば、地下水位の観測データの解析においては、単純にトレンドを推定してもほとんど有益な情報は得られない。しかし、気圧、地球潮汐、降雨の効果を表現する適当なモデルを構成し、水位データを

$$(5.1) \quad \text{水位} = \text{トレンド} + \text{気圧効果} + \text{潮汐効果} + \text{降雨効果} + \text{ノイズ}$$

と分解してみると、トレンド項に地震の影響がはっきり現れてくる (松本他 (1989))。

参 考 文 献

[情報量規準, ベイズモデル, 回帰モデル]

Akaike, H. (1980). Likelihood and Bayes procedure, *Bayesian Statistics* (eds. J.M. Bernardo, M.H. De Groot, D.V. Lindley and A.F.M. Smith), 143-166, University Press, Valencia, Spain.

坂元慶行, 石黒真木夫, 北川源四郎 (1983). 情報量統計学, 共立出版.

[季節調整]

Gersch, W. and Kitagawa, G. (1983). The prediction of time series with trends and seasonalities, *Journal of Business and Economic Statistics*, **1**, 253-264.

Kitagawa, G. and Gersch, W. (1984). A smoothness priors-state space approach to the modeling of time series with trend and seasonality, *J. Amer. Statist. Assoc.*, **79**, 378-389.

北川源四郎 (1986). 時系列の分解, 統計数理, **34**, 255-271.

[非ガウス型モデル]

Kitagawa, G. (1987). Non-Gaussian state space modeling of nonstationary time series, *J. Amer.*

Statist. Assoc., **76**, 1032-1063.

北川源四郎 (1989). 非ガウス型時系列モデリング, オペレーションズリサーチ, **34**, 541-546.

Kitagawa, G. (1989). Non-Gaussian seasonal adjustment, *Computers Math. Appl.*, **18**, 503-514.

[地下水位データの解析]

松本則夫, 高橋 誠, 北川源四郎 (1989). 地震にともなう地下水位変動の定量的な検出法の開発, 地質調査所月報, **40**, 613-623.