

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成24年 6月 8日現在

機関番号：62603

研究種目：基盤研究(C)

研究期間：2009～2011

課題番号：21500287

研究課題名（和文） 多様な解像度を持つ多変量時系列データからの現象モデリング

研究課題名（英文） Statistical modeling based on multiple time series with various time resolution

研究代表者

川崎 能典 (KAWASAKI YOSHINORI)

統計数理研究所・モデリング研究系・准教授

研究者番号：70249910

研究成果の概要（和文）：本研究では、データの採取条件によって観測頻度（時間間隔）が異なり、観測間隔が不規則にしか得られないような、異なる時間解像度からなる時系列データ群に関する統計的モデリングの研究を行った。具体的には、異なるサンプリング間隔を持つ時系列データの相互依存関係を利用する補間法の研究を行った。具体的応用問題として、より信頼が置けるが観測頻度が低い年次系列（典型例としてはGDP速報）を、速報性を重視するために精度に限界はあるものの観測頻度がより高い系列（補助系列、典型的にはGDP速報）との整合性を考慮して新たに四半期系列を推計する問題を研究した。

研究成果の概要（英文）：In this research project, we developed a method of statistical modeling to work with multiple time series observed with various resolutions or different observation intervals. Especially, we are interested in some interpolation method making use of the inter-relationship between the time series with different observation frequencies. In application, we focused on the de-convolution of annual time series (the final estimate of GDP for example) with the help of closely related quarterly time series (the quick estimates of quarterly GDP for example).

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2009年度	700,000	210,000	910,000
2010年度	500,000	150,000	650,000
2011年度	500,000	150,000	650,000
総計	1,700,000	510,000	2,210,000

研究分野：総合領域

科研費の分科・細目：情報学，統計科学

キーワード：時系列解析

1. 研究開始当初の背景

計算機の高速度化とストレージの廉価化は、科学技術の分野のみならず、実社会においても計測とデータの保存コスト低下に貢献している。このような観点から、少なくとも逆問題の重要性を理解している学術分野においては、大規模・高次元データに対応する必

要性が共通認識となりつつある。その一方、データの規模や次元だけではなく、異質性が拡大しつつあることには、比較的注意が払われていないと思われる。

統計科学とその周辺で近年発達してきたいくつかの手法は、本研究が問題にする「時間解像度の異質性」にうまく対処できる可能性がある。具体的には、混合データサンプリ

ング回帰、関数データ解析が有望である。これらの手法は、分布ラグの多項式近似、データの関数近似、基底展開による近似、正則化といった要素に共通性が見いだせる。現在のところ、この2つの研究分野は個別に発展を遂げているという印象だが、相違点を明らかにした上で、総合的なデータ解析技術を構築していくことは、統計科学の発展にとって重要であろう。

より俯瞰的な立場からは、状態空間モデルによる補間、ウェーブレット解析、点過程モデル、連続時間モデルによる近似なども関連性が見いだされる可能性がある。

2. 研究の目的

本研究の目的は、データの採取条件によって観測頻度(時間間隔)が異なり、観測間隔が不規則にしか得られないような、異なる時間解像度からなる時系列データ群に関する、システム接近型の統計的モデリングを研究することである。換言すれば、等間隔かつ観測時点に変量間で共通であることを前提とする多変量時系列モデルでは対処できない問題への解決法を、実例を通じつつさまざまな角度から研究する。

具体的には、異なるサンプリング間隔を持つ時系列データの相互依存関係を利用する補間法の研究を行う。その応用として、集計量の分解問題が挙げられる。特に、より信頼がおけるが観測頻度が低い系列(典型例としてはGDP 確報)を、速報性を重視するために精度に限界はあるものの観測頻度がより高い系列(補助系列、典型的にはGDP 速報)との整合性を考慮して新たに四半期系列を推計する問題はベンチマーク問題と呼ばれている。

ベンチマーク問題については、日本のGDP 統計でこれまで利用されてきたプロラタ法と呼ばれる比例配分法に関し、速報系列の年集計値と確報系列の作成時に推計される年次系列の差が小さくない場合には断層問題が生じている可能性がしばしば指摘されている。実際、[雑誌論文]①に整理した通り、従来に利用されていた方法では例えば第1四半期の変化率の推定値などにベンチマークの効果がしわ寄せされる可能性が高かったことが伺われる。

3. 研究の方法

簡単化のため一般性をやや犠牲にして、年次データの四半期分割のイメージで説明する。第一の方法は、新たに生成する四半期系列(推定系列)の総和が実際に観測される年次系列に等しい、という制約の下で、推定系列と参照系列の比の時間変化の総和を最小

化することである。これはしばしば比例デントン法と呼ばれる。

第二の方法は、参照系列に対する線形回帰として推定系列を位置づけ、制約つき線形回帰モデルの予測誤差最小化問題として定式化することである。この枠組みはしばしばチャウ・リン法として言及される。

これらの方法論に基づく若干の一般化を考察することができる。

4. 研究成果

(1) 方法論的成果

まず、プロラタ法の問題点と経済時系列の断層問題に関する議論の整理を行った。

その上で、比例デントン法に関して、線形制約の下での二次計画問題として論点の整理を行った。国友・川崎(2011) ([雑誌論文]①)の補題1、定理1を参照。オリジナルの比例デントン法での評価基準を一般化できること、直近データのベンチマークに関心が集中している場合の加重2乗和基準の採用、参照系列が(実務的にはそう信じられているにもかかわらず)必ずしも正の系列である必要のない点を示した。

チャウ・リン法とその一般化に関しては、推定誤差系列の確率ベクトルについてその分散共分散行列の逆行列 W が明示的に与えられたと仮定した場合に、この問題がラグランジュ形式の条件付き最小化問題に帰着することを示した。国友・川崎(2011)の定理2を参照。 W の選択は、推定誤差系列がどのような確率過程に従うと仮定するかという論点と密接に結びついており、チャウ・リン法はそのひとつの定式化に埋め込まれることを明らかにした。

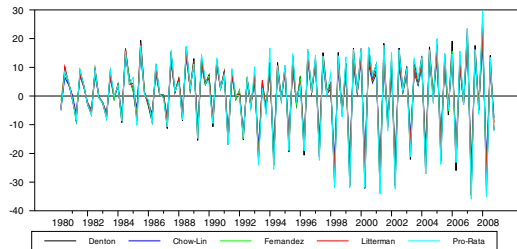
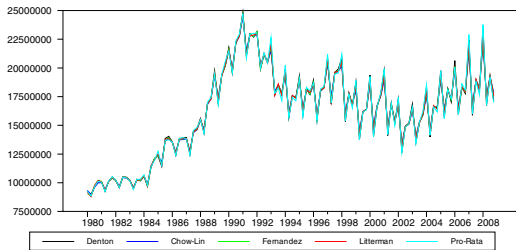
(2) 応用・実証分析的知見

日本の設備投資系列(1980年から2008年)と出荷系列(2001年から2008年)の原系列を実例として、年次系列から補助系列としての四半期系列を参照しつつ、四半期データを生成する諸手法を適用して比較した。ここで取り上げる方法は、プロラタ法、デントン法、チャウ・リン法、フェルナンデス法、リッターマン法である。チャウ・リン法、フェルナンデス法、リッターマン法の違いは、分配後の誤差系列にどのような時系列構造を仮定するかの違いであり、チャウ・リン法はAR(1)、フェルナンデス法は1次のランダムウォーク、リッターマン法はARIMA(1, 1, 0)である。

今回に行ったデータ分析からは次のようなことがわかった。

設備投資に関してはどの手法を採用しても結果に大きな違いがない。ただ、ベンチマーク系列と補助系列の比を観察すると、デントン法は滑らかな動きであり、チャウ・リン法、フェルナンデス法、リッターマン法はやや変動が大きい。

図の上段はレベルで、下段は変化率で結果を表示している。なお、本報告書所収の図はすべて国友・川崎(2011)に収められている。



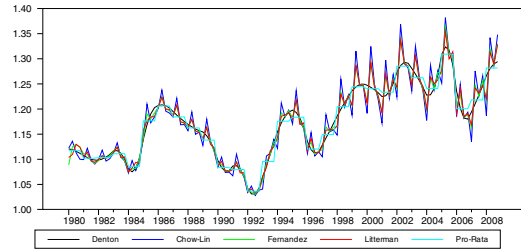
チャウ・リンモデルの自己回帰パラメータ推定値は 0.99、リッターマンモデルでの自己回帰パラメータ推定値は 1.00 であった。従って、チャウ・リン法、フェルナンデス法が図で殆ど重なっていることは納得がいくが、なぜリッターマン法までもが同じ軌跡辿るのか、すぐには理解が難しいように思われる。2 階差モデルなのだが、イノベーションの分散が非常に小さいということかもしれない。

ベンチマークを行う際、参照する系列に季節性が含まれていれば、必然的にベンチマーク系列も季節性を反映する。前期比伸び率を計算したい場合など、足下の動向・変化を把握したいとなれば、ベンチマーク系列を季節調整することになる。一方、参照系列を季節調整してしまい、季節調整系列に対してベンチマークをとる方法も考えられる。この方法は、季節調整系列だけではなくベンチマーク原系列とも言うべき系列も公表しなければならない場合には問題であるが、季節調整がベンチマークの前後にそれぞれ入った場合の結果を比較しておくことには一定の意義がある。

今回取り上げた時系列の中では、出荷系列には明瞭な季節成分は含まれないことから、以下設備投資系列だけを取り上げる。参照系列の季節調整は、対数加法型でトレンド 2 階差分の標準季節調整モデルが支持されたので、これにより行った。季節調整参照系列に対

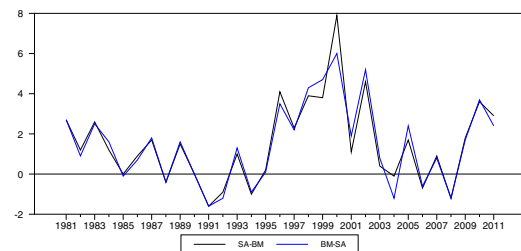
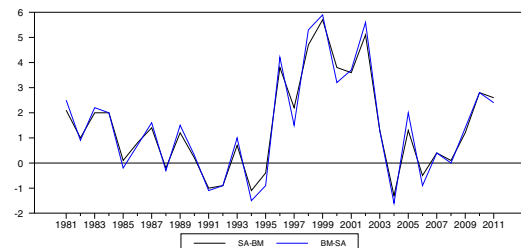
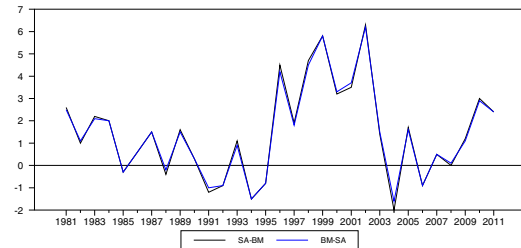
し、ベンチマーク法をさまざまに変えて結果を比較した。

レベルで観察すると、どの方法も結果が重なっているように見えるが、変化率で見ると総じてプロラタ法の結果で変動が大きいことが観察される。同様のことが、ベンチマークとインデックス(参照系列)の比率を通じても観察される。



次に関心のあることは、先に季節調整をした場合(SA-BM)と、ベンチマーク後に季節調整した場合(BM-SA)とでどのような違いが生じるかを、比較することである。性急な一般化は避けなければならないが、チャウ・リン法やプロラタ法に比べると、デントン法は季節調整に関して頑健であるという印象を与える。

原四半期のベンチマーク値の季節調整値と季節調整値のベンチマーク値を比較すると、前者の方が若干ではあるがより安定する結果が得られた。季節調整法を一種の平滑化と理解すると当然な結果と解釈できる。設備投資系列についてはその差は小さい。



季節調整済系列でベンチマークした場合と、ベンチマーク系列を季節調整した場合での、前年同期比の比較を行ったのが上の図である。上段からデントン法、チャウ・リン法、プロラタ法。対象は設備投資系列である。青と黒のラインの重なりが多いほど、季節調整のタイミングに関してベンチマーク法が頑健であることを示唆している。

既存のベンチマーク法では実際に観察される時系列におけるトレンド項や季節性の変動パターンを組み込んだ考察はあまりなされていないようである。我々が行った限定的な実験によれば、原系列のベンチマーク値に季節調整を行う方が原系列の季節調整値にベンチマーク値を求めるよりも平滑化されるという結果が得られている。

(3) その他の成果

為替のティックデータに基づき、取引の生起度をモデル化するための点過程モデルの開発を行った。（〔学会発表〕③）実際の取引とクォートが区別できるようなデータが利用可能になったことも相まって、クォートと取引という二つの不等間隔データの相互依存関係を分析するツールが出来上がった。これにより、非同期観測のデータ間の相互依存関係を分析することが可能になった。

空間的にも時間的にもデータが不等間隔に配置されている問題として、琵琶湖の水質に関する長期多変量時系列データに基づくトレンド解析を行った。（〔雑誌論文〕④）

統計モデリングにおいて応答が非線形の場合の要因探索に関する研究も行った。（〔学会発表〕③）

観測打ち切りのある金融市場のデータ（株式、商品先物など）に対する時系列分析法を研究し、多様な時間解像度の予測ホライズンに対応する取引リスクの推定に応用し、研究発表を行った。（〔学会発表〕①）

正則化を利用して回帰モデルにおける説明変数の自動グループ化を行う方法（〔雑誌論文〕②）や、多変量非線形時系列間の因果解析を検定する方法の基礎として、関数データ解析に類似の発想に基づく、非線形な関係を含む時系列モデルの定式化検定に関する研究でも成果があった。（〔雑誌論文〕③）

5. 主な発表論文等

（研究代表者、研究分担者及び連携研究者には下線）

〔雑誌論文〕（計4件）

- ① 国友直人、川崎能典、ベンチマーク問題と経済時系列：GDP速報とGDP速報を巡って、経済学論集（東京大学経済学部）、査読無、77巻、2011、2-19.

- ② Ueki, M., Kawasaki, Y., Automatic grouping using smooth-threshold estimating equations, Electronic Journal of Statistics, 査読有, Vol. 5, 2011, 309-328.

- ③ Nishiyama, Y., Hitomi, K., Kawasaki, Y., Jeong, K., A Consistent nonparametric test for nonlinear causality - Specification in time series regression, Journal of Econometrics, 査読有, Vol. 165, 2011, 112-127.

- ④ Kawasaki, Y., Kawai, K., Okubo, T., Kanefuji, K., Long term trend analysis of water quality in Lake Biwa, Proc. of MODSIM09, 査読有, 2009, 3172-3178.

〔学会発表〕（計3件）

- ① 青木義充、川崎能典、打ち切りを考慮した株価変動の時系列モデル、2011年度統計関連学会連合大会、2011年9月7日、九州大学（福岡市）。
- ② Kawasaki, Y., Ueki, M., Variable selection in discrete choice models, Joint Statistical Meeting 2010, 2010年8月7日, Vancouver, Canada.
- ③ 川崎能典、高頻度金融データの日内季節調整について、ファイナンス計量分析の新展開と日本の金融市場、2010年3月31日、東京大学経済学部、東京。

〔その他〕

ソフトウェアの配布：〔雑誌論文〕③で提案した方法について、GAUSS言語による解析プログラムを配布中。要求次第著者から個別に配布。報告書提出時点での照会件数は、海外から5件、国内から0件。

6. 研究組織

(1) 研究代表者

川崎 能典 (KAWASAKI YOSHINORI)
統計数理研究所・モデリング研究系
・准教授
研究者番号：70249910