

氏名	駒 木 文 保
学位（専攻分野）	博士（学術）
学 位 記 番 号	総研大甲第2号
学位授与の日付	平成4年 3月16日
学位授与の要件	数物科学研究科 統計科学専攻 学位規則第4条第1項該当
学位論文題目	State-Space Modelling of Time Series Sampled from Continuous Processes with Pulses
論文審査委員	主 査 教 授 北 川 源四郎 教 授 赤 池 弘 次 教 授 田 辺 國 士 教 授 松 縄 規 助教授 尾 形 良 彦

論文内容の要旨

本論文は、これまでほとんど未解決であった、パルスをもつ時系列データのためのモデルと解析方法を提案し、問題点を解決したものである。

線形ガウス時系列モデルは、様々な分野で有効に利用されるようになっている。これには、モデルのパラメータ値を最尤法によって数値的に推定することが実用化されたこと、モデルの客観的な選択がAICによって可能になったこと、が大きい。しかし、線形ガウス時系列モデルでは適切に取り扱うことのできない時系列データも多くあらわれる。典型的な例として、データがパルス状のパターンを示す場合がある。たとえば、血中のホルモン濃度、神経の電位、河川の流量、などがあげられる。これらの中で、河川流量のデータに対して雨量のデータが知られている場合のように、入力と見なせるデータが得られている場合には、入出力システムとして定式化することにより解析することができる。しかし、入力を得られない場合についての解析方法についての研究は、きわめて限られていた。この様な時系列データに対して有効な統計モデルおよび解析手法を提案し、関連する問題を解決することが本論文のねらいである。

本論文は4つの章から構成されている。第1章では、研究の動機、目的、論文の構成について述べている。

第2章では、これまでに他の研究者によって行われた、パルスをもつ生体ホルモン時系列の統計的な解析に関する代表的なモデルと方法の紹介を行い、それらの特徴と限界について述べた。哺乳類の黄体形成ホルモン(Luteinizing Hormone (LH))の血液中のレベルの時系列データの場合には、急激に濃度が上昇し、その後ゆっくりと指数的に下降する、というパルスの特徴的なパターンがあることが知られている。このようなデータは、通常の線形ガウスモデルでは適切に取り扱うことができない。そのため、生理学の分野では様々なきわめて ad hoc な方法が使用されてきた。統計学の分野では、このような生体ホルモン時系列データを方法論的な視点から取り扱ったものに O'Sullivan and O'Sullivan (1988)、Diggle and Zeger (1989) がある。しかし、これらの方法も次のような重大な欠点を抱えている。前者は、ピークの検出を目的とした方法である。しかし、モデルの推定に組合せ論的複雑さがともなう、また時系列モデル

ではないのでシミュレーションや予測に利用できない、などの問題点がある。後者は、フィードバックを考慮した新しいタイプの時系列モデルを導入した。しかし、1階のマルコフモデルなので、パルスが始まりからピークに達するまでに2観測時点以上かかる場合は、データの一部の値に人工的に手を加えてから解析しなければならない、パルスの形に比べ観測時点が粗いので解析方法として無理がある、などの問題点がある。そのため、このようなパルスを持つ時系列データをそのまま人工的な手を加えることなく解析でき、かつモデルが組合せ論的複雑さをともなわない方法の開発が現実問題として切実に必要とされてきた。このことを踏まえて本論文がまとめられている。

第3章が本論文の主要部分である。観測時刻間を細分割し観測数より多い時点を設定してから状態空間時系列モデルを利用してベイズ的方法を適用することにより、前章で考察した問題点を解決した。ベイズ的方法に基づく時系列解析(Akaike(1980))は、時系列のトレンドの推測などに応用されている。また非ガウスモデルによる解析方法(Kitagawa(1987))も最近発展してきており、トレンドが不連続点を含む場合の推測などに有効に利用されている。ここでは、パルスをもつ時系列データを取り扱うことのできる新しい状態空間モデルを提案し、例として、Diggle and Zeger の取り扱ったデータと同じ8頭の牛の黄体形成ホルモンの時系列データを解析した。

モデルは、各時刻における状態が増加モードと減少モードのふたつのモード($m(t)=0, 1$)と実数値を取るレベル($x(t)$)からなり、増加モードのときはレベルがゆらぎながらほぼ直線状に増加し、減少モードのときは、レベルがゆらぎながら指数関数的に減少するというものである。ふたつのモードは、レベルの値に依存して確率的に入れかわる。得られたデータはレベルに観測ノイズが加わったものとする。このモデルの状態の時系列は定常分布を持つことを理論的に示した。これにより初期分布として定常分布を使った尤度計算が可能になった。

本論文で提案した状態空間モデルを用いることにより、尤度に基づく推測を行って、AICによるモデルの比較やベイズ的推測を行うことができる。LHデータの解析に際しては、状態空間上の確率分布をヒストグラム状の関数で近似して対数尤度を計算し(Kitagawa(1987))、勾配を数値微分をもちいて計算し準二

ニュートン法によりパラメータの最尤推定値をもとめた。この様にして得た推定値を利用して、非ガウススムージングを行うことによりホルモンのレベルやパルスのピークの位置などについての事後分布を計算し、これらについて適切な推定が行えることを示した。また、8頭の牛それぞれについての最尤推定値を利用したシミュレーションにより得られたデータは、それぞれ元のデータときわめて似た特徴を示した。これは、この方法が安定に働いていることを示している。この増加モードでレベルが直線状に増加するモデルと指数関数的に上昇するモデルとの AIC による比較を行い前者の方があてはまりが良いことを確認した。また、モデルがいろいろなパラメータのもとで、さまざまな頻度のパルスを持つパターンを生成することを確認し、パラメータと生成されるデータとの関係を考察した。さらに、元の牛のデータより複雑な、シミュレーションにより生成されたデータの推測を行ない、適切に情報を取り出せることを示した。

第4章では、本研究を通じて得られた知見および考察が述べられている。また、基礎的な概念、手法や副次的な理論の証明が付録にまとめられている。

最後に、本論文の主要な貢献を要約する。新しく提案した状態空間モデルによる時系列解析には次のような顕著な利点がある。

- ・時点を細かくとることにより粗い観測点からでもホルモンのレベルやパルスのピークの位置などについて、事後分布を求めることにより適切な推定ができる。
- ・計算量は観測値の数に比例する。
- ・データに人工的な手を加える必要がない。
- ・観測値のサンプリング間隔が一定でない場合にも適用できる。
- ・モデルをシミュレーションや予測に利用することが可能である。
- ・尤度に基づいた方法なので、モデル比較に AIC が適用できる。

これらの点は、従来のモデルおよび関連する手法に比べて優れている。

参考文献

- Akaike, H. (1980) Likelihood and the Bayes Procedure. in Bayesian Statistics, eds. J.M. Bernardo and M.H. De Groot et al., Valencia Uni-

versity Press, pp.143-166.

Diggle, P. and Zeger, S. (1989) A Non-Gaussian Model for Time Series with Pulses. J. Amer. Statist. Assoc. 84, 354-359.

Kitagawa, G. (1987) Non-Gaussian State-Space Modeling of Nonstationary Time Series. J. Amer. Statist. Assoc. 82, 1032-1041.

O'Sullivan, F. and O'Sullivan, J. (1988) Deconvolution of Episodic Hormone Data: An Analysis of the Role of Season on the Onset of Puberty in Cows. Biometrics 44, 339-353.

論文の審査結果の要旨

審査委員会は上記論文について、数物科学研究科における課程博士の授与に係わる論文審査等の手続き等に関する規程にもとづき、公開の論文発表会を開催し審査を行なった結果、以下の理由により合格と判断する。

総合評価

本論文は、従来ほとんど未解決であったパルス状の変動をともなう時系列の解析のための重要なモデルを提案し、さらにその推定法、計算法を導いたものである。本論文で提案されたモデルは、従来の解析法の問題点を的確に捉えた上で最近急速に発展しつつあるベイズモデルの考え方にもとづいて考案されたものであり、さまざまな長所を持っている。さらに、最新の非ガウス型フィルタを用いることによりパラメータ推定や状態推定のための実用的な計算法を実現することに成功している。

一般的な解析法が完成された観がある線形ガウス型時系列モデルに対し、非線形あるいは非ガウス型時系列のモデリングはいまだ個別的な試みの段階にあるが、本論文でなされたパルス状の変化を含む時系列のモデリングの実用化は時系列解析における重要な貢献であると評価できる。以上の研究内容は、著者が統計学および時系列解析における最近の成果、動向を熟知した上で行なったことを示すものであり、オリジナリティのある仕事として高く評価することができる。

本論文の主要な貢献点

本論文で提案された時系列解析法には次のような顕著な利点がある。(a) 観測の間隔よりも精密な精度でパルスのピーク的位置などに関する推測が可能である。(b) 観測を等間隔で行なう必要がない。(c) データに人工的な加工を行なう必要がない。(d) モデルの推定に必要な計算量は観測数に比例するだけであるので大量のデータの処理が可能である。(e) 推定されたモデルはシミュレーションや予測などへの応用が可能である。(f) 本論文で解析の対象となったホルモンのデータだけでなく、神経データ、河川流量などの未知のパルス状の入力の影響を受ける多くの時系列への適用が期待される。

本論文のモデルにもとづく解析法は、汎用性、実用性、高精度の三つの特長

を備え、従来の解析法と比較して格段に優れた方法である。

論文の公表について

本論文の主要部分は国際的に権威ある統計科学の学術誌である *Biometrika* に投稿中であり、編集者からの手紙によれば、部分的な小修正の後に受理されることがほぼ確実である。