

令和 元年 5月 31日現在

機関番号：15301

研究種目：若手研究(B)

研究期間：2015～2018

課題番号：15K17038

研究課題名(和文)高次元長期記憶時系列に対する統計的分析法とその応用に関する研究

研究課題名(英文)Statistical analysis of large dimensional long-memory time series and its applications

研究代表者

生川 雅紀(Narukawa, Masaki)

岡山大学・社会文化科学研究科・准教授

研究者番号：30588489

交付決定額(研究期間全体)：(直接経費) 2,200,000円

研究成果の概要(和文)：長期記憶性を有する高次元時系列におけるセミパラメトリックな統計的分析法を拡張された因子モデルを用いて構築した。具体的には、第1段階では共通要素を主成分分析法から推定し、第2段階として各記憶パラメータを局所法で推測する。2段階セミパラメトリック法を考案し、推定量の漸近的な性質や有限標本での挙動を調べた。また、非定常長期記憶過程に対する最大効率テイパーを取り入れた多変量セミパラメトリック法を提案し、為替レートデータ分析への応用を行った。

研究成果の学術的意義や社会的意義

次元の大きい長期記憶性を持つ時系列に対しては、既存分析手法の適用が困難な上に関連研究の蓄積が乏しい中、本研究で考案している時間領域の因子モデルと主成分分析法に周波数領域のセミパラメトリック推測を組み合わせた2段階アプローチは、二つの領域の手法を駆使することで高次元長期記憶時系列における統計的推測法を提供するという大きな意義があり、時系列データの統計的分析に新たな視点と方向性をもたらすであろう。

研究成果の概要(英文)：This research developed semiparametric statistical analysis of large dimensional time series with long-range dependence by using extended factor models. Specifically, we proposed the two-step semiparametric approach in which the common components are estimated by principal components analysis as the first step and the estimators of the memory parameters are obtained by the local Whittle method as the second step. We also investigated the asymptotic properties and the finite sample performance. Moreover, we constructed multivariate local Whittle estimators by incorporating the maximal efficient taper, and provided an empirical application of the proposed method to exchange rate data.

研究分野：計量経済学

キーワード：長期記憶 高次元時系列 セミパラメトリック法 因子モデル Taper

様式 C - 19、F - 19 - 1、Z - 19、CK - 19 (共通)

1. 研究開始当初の背景

時系列とは時の経過に沿って観測されるデータであり、時間的な相関がその構造を特徴付け、その持続性により、定常過程は短期記憶過程と長期記憶過程へ、さらに非定常過程も非定常長期記憶過程と単位根過程に分類される。長期記憶過程とは、短期記憶過程と単位根過程の狭間に位置する概念とも考えられ、マクロ経済やファイナンスデータを始めとして多様な分野のデータが長期記憶性を示すことが知られている。また、観測個体数や変数が大きい「高次元時系列」データが近年容易に利用できるようになって来ており、複数の経済時系列が相互に関連しつつ変動していることを踏まえると、クロスセクション・多変量方向の相関を考慮しないことが致命的となる一方で、数変量程度を前提とする従来の多変量時系列に対する手法では高次元データを扱う上での困難は避けられない。そのため、このような高次元の長期記憶時系列における統計的分析法の構築が急務となっていると言えよう。

長期記憶過程に従う時系列では記憶パラメータと呼ばれる、時間方向の相関構造を規定するパラメータが統計的推測の主要な関心対象となり、モデル定式化の誤りを回避し、記憶パラメータの頑健で柔軟な推測を目的とするセミパラメトリックなモデリングに基づいた、周波数領域における推定や検定法が中心的なアプローチの一つとなっている。多くの研究蓄積があり、スペクトル密度の使用する周波数に応じて局所法と大域法に大別できる。とくに、長期記憶性とその原点近傍の挙動に特徴的に表れることに基づいて推測を行う前者に関しては、効率的とされる局所 Whittle 法を多変量時系列へと拡張した分析手法がいくつか提案されているが、数変量程度を想定しており、観測個体数が大きい多変量データの場合には夥しいパラメータ数に伴う数値計算上の問題が生じることから実行可能とは言い難く、高次元多変量・クロスセクション時系列を対象とした進展はほとんど見られない状況と考えられる。

2. 研究の目的

このような背景の下、本研究では観測個体数や変数が大きい高次元な長期記憶時系列におけるセミパラメトリックな統計的推測法の体系的な構築とその考案された手法を実際の経済データ分析へと応用することを目的とする。根幹となる高次元長期記憶時系列に対する統計的分析法に関しては、パラメータ数を減少させられる因子構造を取り入れたモデルを拡張し適用することによって2段階で構成される推測法、具体的には、各系列の時間方向への確率的依存関係は1つの記憶パラメータによって規定され、それ以外は特定の定式化をしないセミパラメトリックなモデリングを、クロスセクション・多変量方向への相関構造については共通要素で表現されていると想定し、第1段階として、因子モデルを用いた主成分分析法により各系列を独立な単変量へと分解し、第2段階では、それらに含まれる各記憶パラメータをセミパラメトリックな手法によって推測するアプローチを確立することを目指す。

さらに、現実の経済・ファイナンスデータ分析への応用を念頭に置いた枠組みへの拡張を目的として、記憶パラメータが実数と虚数を表すことに着目し、2段階セミパラメトリック法における推定量の理論的な性質を利用した高次元長期記憶時系列における実数と虚数と検定や周期的な長期記憶性へも展開させ、経済・ファイナンスデータを対象とした分析から新たな知見を引き出すことを試みる。以上が本研究課題の当初の目的である。

3. 研究の方法

既存の多変量時系列を前提とした方法で問題となるのは、次元の上昇に応じて著しいパラメータ数の増大が生じることであり、この観点から高次元多変量として扱うのではなく、クロスセクション(多変量)方向の相関を因子モデルで表現し、各系列間の相関構造を共通要素と捉えることで抽出して取り除き、高次元長期記憶時系列を単変量時系列へと分割する手法を検討する。このような段階を踏むことで、パラメータ数に起因する実装上の問題を克服することが可能となり、単変量化された各長期記憶時系列に対する記憶パラメータの推測においては既存の単変量セミパラメトリック法が利用できる、2段階での推測アプローチが構築される。そのためには、まず、近年の研究進展に伴い時間領域のパネルデータ分析における主要なツールとなりつつある因子モデルを長期記憶時系列が取り扱えるよう発展させる必要がある。高次元時系列の各系列が因子負荷・共通因子の積となる因子構造を持つ共通要素と誤差項によって構成されており、クロスセクション(多変量)方向への相関は共通要素で描写されているとするが、本研究では、誤差項は定常長期記憶過程に従うとし、その方向に関しては無相関(独立)である厳密な因子モデルと設定する。このモデル化における共通要素を主成分分析法によって推定し除去することで各系列の誤差項に対応する残差が求められ、これが第1段階となる。

次に、単変量へと分解された各系列における記憶パラメータが関心対象となるため、一般に周波数領域では Whittle 尤度に基づくセミパラメトリック法が効率的なことを踏まえ、その局所法と大域法を候補として取り上げ、各記憶パラメータをセミパラメトリックに推測することを第2段階とした2段階アプローチへの展開に取り組む。第1段階の主成分分析法を用いた共通要素の推定も上記以上のモデル特定化や分布等の仮定を要しないため、以上の方法は2段階セミパラメトリック法と見なせる。適切に機能するか吟味するために数値シミュレーションによるパフォーマンスの検証を行うとともに、採用したセミパラメトリック推定量を中心に漸近的性質の導出を試みる。2段階アプローチであることに留意すると、第1段階の共通要素推定量の一致性のみならず収束速度を導出する必要があり、とくに後者は因子モデルの構造から時

間・クロスセクション方向双方の標本サイズに依存する予想でき、第2段階はセミパラメトリック推定量であるため、収束速度が時間方向の標本サイズより遅くなると考えられる。これらを援用すれば、2段階セミパラメトリック法における理論的性質の確立が見込める。さらに、因子モデル表現の誤差項が周期性を有する長期記憶を描写できる Gegenbauer 過程に従うと想定することで周期的長期記憶時系列への発展性についても検討し、また、記憶パラメータの漸近分布から、差分を取った高次元時系列の各系列に対して、単位根を含む実数積分に関する仮説検定をパネル単位根検定の手法の援用も考慮しつつ構築することが期待される。

4. 研究成果

(1) 高次元定常長期記憶時系列に対する統計的推測法の構築

本研究の中核となる高次元時系列の統計的分析法について、因子構造を取り入れた2段階セミパラメトリック法を考案できた。研究方法に基づき、既存の因子モデルを用いた高次元時系列分析法における仮定を綿密に検討することで、定常な高次元長期記憶時系列を表現できるよう拡張したモデル化を行い、主成分分析法を用いて共通要素を抜き取れることを確認し、各系列の残差から局所法による記憶パラメータの推定量が得られる一連の手順を確立した。各系列における記憶パラメータの推定パフォーマンスを数値シミュレーションによって詳細に調べ、第2段階で採用した局所 Whittle 法が十分に機能すること、第1段階の主成分分析法による共通要素の推定が定常長期記憶過程の場合にも適用可能であることが分かった。これらを裏付ける理論的な性質の導出については、基盤となる因子構造に課される仮定を既存研究よりも大きく緩めて整備した下で、定常長期記憶過程を想定した因子モデルにおける因子負荷・共通因子の主成分分析による推定量の一致性と収束速度を導いた。予想どおりに時間・クロスセクション方向双方の標本サイズが増大して行く状況ではそれらの条件に依存することが確認できたのみならず、各系列に含まれる記憶パラメータの最大値が時間方向に関しては大きく影響することも示された。以上で明らかになった共通要素推定量の特性を踏まえ、第2段階の局所 Whittle 法による記憶パラメータ推定量の漸近的性質について調べた。時間・クロスセクション方向における各標本サイズの増加オーダー関係に条件を課す必要があるものの、共通要素推定量によってクロスセクション(多変量)方向の相関が除去でき、影響を無視できうることを見出した。また、それを援用することで記憶パラメータ推定量の一致性や単変量の場合と同様の漸近分布を有することが導出でき、この理論的な結果から実数積分検定も行えうることが示せた。

(2) 因子モデル拡張による2段階セミパラメトリック法の一般化

2段階アプローチによる推測法における理論的性質に必要な条件の整備やその導出過程で、当初のモデル化では現実のデータ分析へと応用する際に制約的となる側面が浮かび上がったため、研究計画を一部変更し、この問題に対応できるよう因子モデルの更なる拡張に取り組んだ。(1)で提案している2段階セミパラメトリック法では、基盤となっている因子モデルの共通要素が含む共通因子に長記憶性が存在する可能性を暗に排除している点が制約となりうる。モデルの構造から各系列に共通の支配的な要因と捉えられる共通因子が時間方向の強い相関を持たないのはやや不自然に見受けられ、実用上の分析対象が狭まる恐れもあるため、共通因子自体も定常長期記憶過程に従い、かつ、誤差項より強い時間的相関となる因子モデルへと大幅に拡張することを試みた。この場合、共通因子が記憶パラメータを有するので、第2段階で共通因子推定量・残差の双方に含まれる記憶パラメータを局所 Whittle 法で推定するアプローチへと発展させた。さらに、共通因子の記憶パラメータが共通の値を取るとするモデル化から実数積分関係を描写できることも見出せ、その際に複数存在しうる共通因子を多変量時系列と見なし多変量局所 Whittle 法を修正して共通記憶パラメータを推定する方法を考案するに至った。以上のようにして一般化された2段階セミパラメトリック法の各推定量が十分に良好なパフォーマンスを有することが数値シミュレーションから示され、拡張因子モデルに対しても主成分分析法や記憶パラメータの局所 Whittle 法が問題なく機能することが確認できた。また、漸近的な性質に関しては、(1)で得られた成果を足掛かりとして、共通因子に長期記憶性を導入するために関連する仮定を再吟味して緩和させ、その影響を慎重に見極めつつ一般化した枠組みの下での主成分分析法による推定量や各記憶パラメータ推定量の漸近的性質の導出を行った。結果的には当初の計画以上に高次元時系列の統計的分析法として一般化でき、併せて理論的な性質の確立にも目途がついたため、より実用性の高い体系的な方法論とすることができた。

(3) 多変量非定常長期記憶過程におけるセミパラメトリック推定法

2段階アプローチを進めて行く途上で既存研究から非定常長期記憶過程における新たな多変量局所 Whittle 法を構築できる可能性を見出し、本研究課題の高次元長期記憶時系列と上述してきたように密接に関連するため、その確立に取り掛かった。非定常性や過剰差分を扱う上でピリオドグラムに生じるバイアスを減少させる手法である Taper を取り入れると、通常、記憶パラメータ推定量の効率性が落ちることが知られているが、本研究では最大効率 Taper と呼ばれる、理論上は漸近分散の膨張が引き起こされず、有限標本においても大幅にその膨張を抑えられる方法を多変量ピリオドグラムに組み込むことで、効率性低下を回避した多変量局所 Whittle 法を提案するに至った。差分系列に基づく上に最大効率 Taper はシフト不変な性質を持つため、幅広い潜在的に非定常な時系列を対象とでき、差分に応じた多項式トレンドが多変

量を構成する各系列に含まれていても影響しない頑健な記憶パラメータベクトルの推定量が得られる。また、記憶パラメータ推定量の一致性や漸近正規性といった漸近的な性質に加え、数値シミュレーションの実施から安定したパフォーマンスを発揮することを確認した。主要な代替レートデータの対数値を多変量時系列と見なして各系列の記憶パラメータをセミパラメトリック推測する経済データ分析への応用も行い、これらの結果を研究論文としてまとめた。

以上の成果は、高次元長期記憶時系列の統計的分析に関する研究蓄積が乏しい現状において、(1)と(2)はその分析手法の方向性を提供しうるものであり、さらに定常過程から一般的な単位根過程を含む非定常長期記憶過程への発展を目指す上で(3)は共通因子の推測に有益な手法となりうると言えるだろう。その一方で、理論的な性質の確立が難航したことや当初の計画にはない想定外の展開があったために実データへの応用が不十分となった点や内容が膨らみ取りまとめ切れていない側面が生じたことも否めず、今後の課題にしたいと考えている。

5. 主な発表論文等

〔雑誌論文〕(計2件)

Narukawa, M. Efficient tapered semiparametric estimation of multivariate fractional time series. The Economic Association of Okayama University Discussion Paper, 査読無, 1-99, 2018, 1-35.

Narukawa, M. Semiparametric Whittle estimation of a cyclical long-memory time series based on generalised exponential models. Journal of Nonparametric Statistics, 査読有, Vol.28, 2016, 272-295. DOI:10.1080/10485252.2016.1163350

〔学会発表〕(計3件)

生川雅紀. Efficient tapered semiparametric estimation of multivariate fractional processes. The 5th Institute of Mathematical Statistics Asia Pacific Rim Meeting, 2018年6月29日, Singapore.

生川雅紀. 効率的Taperを用いた多変量局所Whittle法について 統計関連学会連合大会, 2016年9月5日, 金沢大学.

生川雅紀. Semiparametric Whittle estimation of a cyclical long-memory time series based on GEXP models. Recent Progress in Time Series and Related Fields, 2015年12月11日, 東北大学.