

## 科学研究費助成事業 研究成果報告書

平成 29 年 4 月 26 日現在

機関番号：12601

研究種目：基盤研究(C) (一般)

研究期間：2014～2016

課題番号：26380267

研究課題名(和文) 正規有限混合分布モデルの統計的推測理論の開発

研究課題名(英文) Development of Inference Methods for Finite Normal Mixture Models

研究代表者

下津 克己 (SHIMOTSU, Katsumi)

東京大学・大学院経済学研究科(経済学部)・教授

研究者番号：50547510

交付決定額(研究期間全体)：(直接経費) 3,600,000円

研究成果の概要(和文)：本研究は、混合正規回帰モデルならびに多変量混合正規モデルにおける統計的推測の理論を発展させた。混合正規回帰モデル・多変量混合正規モデルの分析においては、モデルを構成する要素の個数をデータから決定することが非常に重要となる。しかしながら、これらのモデルは、その尤度関数が特殊な形状をしているため、要素数に関する統計的推測が非常に困難であることが知られていた。

本研究では、これらのモデルの尤度関数を分析する新しいアプローチを開発し、これらのモデルの要素の個数を検定する統計的手法を構築した。さらに、コンピューター・シミュレーションによって、本研究が提唱した手法が、実用上有用であることを確認した。

研究成果の概要(英文)：This research project developed new inference methods for finite mixture of normal regression models (switching regression models) and multivariate normal mixture models. These mixture models have been used in numerous empirical applications in economics. The number of components is an important element of these mixture models. However, because the log-likelihood function of a normal mixture model has a non-standard structure, there exists no procedure to statistically test the number of components in these models. We develop a new approach to analyzing the log-likelihood function of these two models, and construct likelihood-based tests for the number of components. Further, we confirmed by computer simulations that the proposed tests show good finite sample performance.

研究分野：計量経済学

キーワード：正規有限混合モデル 計量経済学 統計的推測 漸近理論

### 1. 研究開始当初の背景

混合正規回帰モデル(スイッチング回帰モデル)ならびに多変量混合正規モデルは、観測されない多様性を柔軟にモデルにすることが可能なため、経済学・生物学・ファイナンスなどの多様な分野の応用において非常に幅広く使われている。以下では、混合正規回帰モデルと多変量混合正規モデルをまとめて、混合正規モデルと呼ぶ。経済学では、混合正規モデルは、観測可能な変数でコントロールできない経済主体の異質性をモデルするために用いられる。例えば、多数の国のマクロ経済データを用いて経済成長のモデルを推定する場合には、各国独自の観測できない要因(例えば、歴史的要因)が経済成長に影響を与えている可能性がある。この要因を無視して、各国を同一として扱った場合には、推定量にはバイアスが生じる。混合正規モデルは、このような観測できない要因を適切にモデルすることを可能とする。

混合正規モデルにおいては、モデルを構成する要素の個数は、観測されない異質性の取り得る値の数を決定するため、モデルの特定化において重要な要素である。実際の応用では、経済理論が要素数に関する情報を提供することは稀であるため、混合正規モデルの要素数を決定するための統計手法を開発することは重要な課題である。

多くの統計的仮説検定においては、尤度比検定が有効な手法である。しかしながら、混合正規モデルは数学的に特殊な構造を持つため、その要素数を決定するために尤度比検定を用いた場合には、尤度比検定等計量の漸近分布は通常のカイ2乗分布に従わず、要素数に関する尤度比検定統計量の漸近分布は未解決の問題である。そのため、実際の応用では、混合正規モデルの要素数を統計的に推測できない状況にあり、混合正規モデルの要素数を検定する手法を開発することは、目下の重要な課題となっている。

### 2. 研究の目的

本研究の目的は、混合正規モデルの要素数に関する統計的推測の手法を開発すること、ならびにそのパフォーマンスをシミュレーションで検証することである。

### 3. 研究の方法

混合正規モデルの要素の個数に関する統計的推測、すなわち要素の数が  $m$  個である帰無仮説を  $m+1$  個である対立仮説に対して検定することが困難であるのは、以下の理由による。要素の数が  $m+1$  個のモデルが、「要素の数が  $m$  個である」という帰無仮説を満たすのは、(1) 1つの要素がゼロの実現確率を持つ場合、(2) 2つの要素のパラメータの値が等しい場合、の2つの場合である。両方の場合において、通常の尤度比検定の正規条件は満たされない。混合正規モデルでは、ケース(1)においては、スコアの分散が無限大になる

(Li et al., 2009)。このため、ケース(1)の分析は著しく困難となる。

ケース(2)では、帰無仮説の下では、特定のパラメータに関してのスコアがゼロとなるため、フィッシャー情報行列が特異となる。Kasahara and Shimotsu (2012)は、通常の(正規分布でない)有限混合モデルでは、あるパラメータの変換を行うと、特定のパラメータに関してのスコアがゼロとなるため、フィッシャー情報行列が特異となる方向を特定できること、さらに、そのようなパラメータに関しては、対数尤度関数の2次の導関数がスコアの役割を果たすことを発見した。この発見により、通常の有限混合モデルでは、対数尤度関数に4次のテイラー展開を適用することで、対数尤度関数を漸近的にパラメータの多項式の2次関数として近似できる。この結果を利用して、Kasahara and Shimotsu (2012)は、通常の有限混合モデルの要素数の検定統計量の漸近分布を導出した。

混合正規モデルにおいては、the loss of strong identifiability (Chen, 1995)と呼ばれる、パラメータに関する高次の導関数が互いに線形従属する現象が発生し、フィッシャー情報行列は通常の混合モデルの場合よりも複雑な特異性を持つことが知られている。このため、Kasahara and Shimotsu (2012)のパラメータ変換は適用不可能である。

本研究の鍵となる発見は、混合正規モデルにおいてケース(2)の分析を可能にするための新しいパラメータの変換である。本研究は、あるパラメータの変換を行うと、特定のパラメータに関してのスコアがゼロとなるため、フィッシャー情報行列が特異となる方向を特定できること、さらに、そのようなパラメータに関しては、対数尤度関数の3次・4次の導関数がスコアの役割を果たすことを発見した。この発見により、混合正規モデルでは、対数尤度関数に8次のテイラー展開を適用することで、対数尤度関数が漸近的にパラメータの多項式の2次関数として近似できる。この変換を用いると、各要素の実現確率が0ならびに1の値をとらない仮定の下で、要素の数が  $m$  個である帰無仮説を  $m+1$  個である対立仮説に対して検定する尤度比検定統計量の漸近分布の導出が可能となる。

実際の検定においては、Li et al. (2009)が提唱したEM testが、実行の容易性等の点で優れている。本研究では、Li et al. (2009)のEM approachを拡張して、要素の数が  $m$  個である帰無仮説を  $m+1$  個である対立仮説に対して検定するmodified EM testを構築する。

### 4. 研究成果

本研究は、以下の3つの主要な研究成果をあげた。(1) 正規混合モデルの対数尤度関数をパラメータの多項式の2次関数によって漸近的に近似する理論の構築。(2) 正規混合モデルの要素数に関する新しい検定統計量で

ある modified EM test の構築とその漸近分布の導出。(3) modified EM test の有限標本下でのパフォーマンスのシミュレーションによる検証。

(1) 正規混合モデルの対数尤度関数の新しい漸近近似理論の構築

本研究は、正規混合モデルの対数尤度関数をパラメータの多項式の2次関数によって漸近的に近似する、新しい理論を構築した。正規混合モデルの対数尤度関数の漸近近似は、近年まで未解決の問題であった。本研究は、あるパラメータの変換を行うと、特定のパラメータに関してのスコアがゼロとなるため、フィッシャー情報行列が特異となる方向を特定できること、さらに、そのようなパラメータに関しては、対数尤度関数の3次・4次の導関数がスコアの役割を果たすことを発見した。この発見により、正規混合モデルでは、対数尤度関数に8次のテイラー展開を適用することが適切であることが判明し、その結果、対数尤度関数が漸近的にパラメータの多項式の2次関数として近似できることが示された。

(2) 正規混合モデルの要素数に関する新しい検定統計量である modified EM test の構築とその漸近分布の導出

前述した、正規混合モデルの対数尤度関数のパラメータの多項式の2次関数による漸近近似に基づき、さらに、Li et al. (2009) によって提唱された EM test を拡張することにより、modified EM test を構築した。modified EM test は、Li et al. (2009) の EM test と比較すると、罰則項を必要としないという利点を持つ。

modified EM test の漸近分布は、 $m$  個の確率変数の最大値で表され、さらに各々の確率変数は、正規分布に従う確率変数を多項式からなる錐に投影することにより得られることが示された。

(3) modified EM test の有限標本下でのパフォーマンスのシミュレーションによる検証

実際の応用で想定される、様々な種類の正規混合モデルを用いた、広範なコンピュータ・シミュレーションを行い、modified EM test のパフォーマンスを検証した。その結果、

回帰モデルでない正規混合モデルでは、modified EM test の有限標本下での帰無分布は、その漸近分布でよく近似されること、正規混合回帰モデルでは、modified EM test の有限標本下での帰無分布の近似にはブートストラップを用いたほうがよいこと、多変量正規混合モデルでは、modified EM test の漸近分布をシミュレーションで計算することは計算量が多くなるため、帰無分布の近似にはブートストラップを用いたほうがよいこと、対立仮説が正しい場合は、

modified EM test は、帰無仮説を棄却する能力があること、modified EM test を行うのに必要な計算時間は過大ではなく、modified EM test は十分実用に供することのできるものであることが確認された。

参考文献

Chen, J. (1995). "Optimal Rate of Convergence for Finite Mixture Models." *Annals of Statistics*, 23, 221-233.

Kasahara, H. and Shimotsu, K. (2012). "Testing the Number of Components in Finite Mixture Models." CIRJE Discussion Paper, University of Tokyo, F-867.

Li, P., Chen, J., and Marriott, P. (2009). "Non-finite Fisher Information and Homogeneity: An EM Approach." *Biometrika*, 96, 411-426.

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2件)

Hiroyuki Kasahara, Katsumi Shimotsu. Testing the Order of Multivariate Normal Mixture Models. CIRJE Discussion Paper, University of Tokyo, F-1044. 査読無. 2017, 1-33.

Hiroyuki Kasahara, Katsumi Shimotsu. Testing the Number of Components in Normal Mixture Regression Models, *Journal of the American Statistical Association*, 110. 査読有. 2015, 1632-1645.

[学会発表](計6件)

下津克己、The 26th Annual Meeting of the Midwest Econometrics Group、Testing the Number of Regimes in Markov Regime Switching Models、2016年10月22日、University of Illinois at Urbana-Champaign、シャンペーン(米国)

下津克己、2016 Asia Meeting of the Econometric Society、Testing the Number of Regimes in Markov Regime Switching Models、2016年8月12日、同志社大学(京都府京都市)

下津克己、The 3rd Shandong Econometrics Conference、Testing the Number of Regimes in Markov Regime Switching Models、2016年6月17日、山東大学、済南(中国)

下津克己、日本経済学会 2015 年度秋季大会、Testing the Number of Components in Normal Mixture Regression Models、2015 年 10 月 11 日、上智大学（東京都千代田区）

下津克己、2015 Joint Statistical Meeting、Testing the Number of Components in Normal Mixture Regression Models、2015 年 8 月 9 日、Washington State Convention Center、シアトル（米国）

下津克己、International Conference on “Frontiers of Theoretical Econometrics”、Testing the Number of Components in Normal Mixture Regression Models、2015 年 8 月 1 日、University of Konstanz、コンスタンツ（ドイツ）

## 6 . 研究組織

### (1)研究代表者

下津 克己 (SHIMOTSU, Katsumi)  
東京大学・大学院経済学研究科・教授  
研究者番号：50547510