

科学研究費助成事業（学術研究助成基金助成金）研究成果報告書

平成 25 年 5 月 24 日現在

機関番号：14301

研究種目：若手研究（B）

研究期間：2011～2012

課題番号：23730215

研究課題名（和文） セミパラメトリックモデルのための集中情報量規準の開発

研究課題名（英文） Focused information criterion for semiparametric models

研究代表者

末石 直也（SUEISHI NAOYA）

京都大学・経済学研究科（研究院）・講師

研究者番号：40596251

研究成果の概要（和文）：モデル選択の目的は、観測されたデータを基に「最適な」モデルを選ぶことである。しかしながら、最適なモデルはモデルの使用目的に応じて異なる。本研究では、モーメント制約によってモデルが記述されているときに、経験尤度推定量を用いて、興味のあるパラメータを正確に推定するためのモデル選択の方法を考察した。また、推定量の平均 2 乗誤差を最小にすること目的とした、モーメント制約モデルのためのモデルアベレージングの方法を提案した。

研究成果の概要（英文）：The goal of model selection is to select a “best” model in a data-driven way. However, the best model generally differs for different intended use of the model. This study develops an empirical likelihood-based model selection method for moment restriction models that is designed to obtain a good estimate for a specific parameter of interest. This study also investigates a model averaging method for moment restriction models that minimizes the mean squared error of the estimator.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
交付決定額	2,000,000	600,000	2,600,000

研究分野：社会科学

科研費の分科・細目：経済学・経済統計学

キーワード：集中情報量規準、モデル選択、モデルアベレージング、経験尤度法

1. 研究開始当初の背景

モデル選択の目的は、潜在的に複数のモデルの候補が存在し、かつ、経済理論などの先験的な知識からはどのモデルを使うことが適切であるか決定できないときに、観測されたデータを基に、最適なモデルを選ぶことにある。

赤池情報量規準やバイズ情報量規準などの既存の情報量規準は、大域的に当てはまりの良いモデル、あるいは、正しいモデルを選ぶことを目的としている。しかしながら、応用上は、データの分布全体としての当てはまりの良さよりも、分布を特徴づける特定のパラメータのみに関心があることが多い。例え

ば、計量経済学で基本となる線形モデルの推定においては、正しいモデルを見つけることよりも、特定の変数間の因果的効果を正確に測定することのほうが重要であることが多い。そのような状況で、当てはまりの良いモデル、あるいは、正しいモデルが、関心のあるパラメータを正確に推定する上で有用なモデルであるとは限らない。通常、モデル選択はそれ自身が目的ではない。情報量規準によってモデルが選ばれた後に、選ばれたモデルを用いて行われる統計的推測にこそ、本来の興味はある。従って、モデルの使用目的を考慮に入れた情報量規準を構築することが望ましい。

このような問題意識の下、Claeskens and Hjort (2003, JASA) (以下、CH (2003)) は集中情報量規準 (FIC; focused information criterion) と呼ばれる情報量規準を提案している。彼らはパラメトリックモデルのモデル選択方法を考察しており、候補となっているそれぞれのモデルを用いて、モデル間で共通の興味のあるパラメータ (focus parameter) を最尤推定した場合、推定量の極限分布の平均 2 乗誤差を最小にするようなモデル選択の方法を提案している。

CH (2003) はモデル選択の新たな方向性を示しているが、計量経済学へ応用するには、残された問題がある。彼らの FIC は最尤推定量を基にした情報量規準であるため、モデルはパラメトリックモデルでなければならない。しかし、計量経済学では、データの分布を完全に定式化するよりも、モーメント制約モデルなどのセミパラメトリックモデルを使うことが一般的である。そのため、FIC が適用可能な範囲は限定的である。

パラメトリックモデルのモデル選択の研究の歴史は長く、統計学において多くの方法が提案されている。それに対し、モーメント制約モデルのためのモデル選択の研究は始まったばかりで、発展途上の段階にある。既存研究では、Andrews and Lu (2001) が GMM の J 統計量を基にした情報量規準を提案し、Hong, Preston, and Shum (2003) が Andrews and Lu の結果を経験尤度 (EL; empirical likelihood) 法へと拡張している。また、本研究の研究代表者も別の研究においてモーメント制約モデルのための AIC を導出しているが、これらの研究では、興味のあるパラメータに応じて最適なモデルを選択するという問題は考察されていない。

2. 研究の目的

本研究の主要な目的は以下の 2 つである。

(1) CH (2003) のアイデアを基に、モーメント制約モデルのための FIC を導出する。モーメント制約モデルの推定方法として、ノンパラメトリックな最尤推定量としての解釈が可能な EL 推定量を採用する。focus parameter の EL 推定量の極限分布の平均 2 乗誤差を最小にするようなモデル選択の方法を提案する。

(2) (1) で得られた結果を拡張し、モーメント制約モデルのモデルアベレージング (model averaging) の方法を考察する。

モデルアベレージングはモデル選択の代替的な方法である。モデル選択後のモデルの推定は、選ばれた単独のモデルに基づいて行われる。それに対し、モデルアベレージングでは、候補となっているすべてのモデルを用

いて推定を行い、それぞれのモデルから得られた推定量の加重平均を求めることで推定量を構築する。本研究では、EL 推定量に基づく focus parameter のアベレージング推定量を提案し、極限分布の平均 2 乗誤差を最小にする加重の決め方を考察する。

3. 研究の方法

本研究で考察するモーメント制約モデルは CH (2003) のモデルとは異なるが、FIC 導出のためのいくつかの基本的なアイデアは踏襲することができる。

まず、候補となっているモデルの集合の中に、推定するパラメータの数が最も多い最大のモデル (これをフルモデルと呼ぶ) が含まれると仮定し、フルモデルは正しく定式化されているものとする。その他のモデルは、フルモデルにネストし、これらの定式化は誤っているものとする。また、すべてのモデルには共通のパラメータが含まれているものとする。

このような設定の下で、focus parameter の EL 推定量の平均 2 乗誤差を最小にするようなモデルを選ぶ。一般に、推定するパラメータの数が少ないほど、定式化の誤りから生じるバイアスが大きくなるが、推定量の分散は小さくなる。パラメータの数が多きときには、逆の関係が成り立つ。従って、バイアスと分散のトレードオフを考慮してモデルを選択することが望ましい。しかし、フルモデル以外のモデルの定式化が大域的に誤っている (global misspecification) 場合、つまり、定式化の誤りが漸近的に消失しない場合には、推定量のバイアスのオーダーの方が分散のオーダーより大きくなってしまふ。フルモデルにはバイアスが存在しないため、このことは、フルモデルが漸近的には常に最もよいモデルになることを意味する。このため、漸近的な枠組みでバイアスと分散のトレードオフを考慮することができない。

そこで、CH (2003) に倣い、ピットマン・ドリフトのような局所的な定式化の誤り (local misspecification) を仮定する。つまり、フルモデル以外のすべてのモデルは定式化を誤っているが、真のデータ生成過程とのかい離がサンプルサイズの平方根のオーダーで消えていくことを仮定する。このフレームワークの下で、focus parameter の EL 推定量のバイアスの 2 乗の項と分散の項のオーダーが同じになり、漸近的な枠組みでのバイアスと分散の比較が可能になる。

4. 研究成果

(1) FIC の導出

まず、「研究の方法」で述べた局所的な定式の誤りの下で、EL 推定量の極限分布を導出した。GMM 推定量において、Newey (1985) が

局所的な定式化の誤りは極限分布の分散には影響を与えず、バイアスにしか影響を与えないことを示しているが、同様の結果を EL 推定量についても示した。

この結果を基にして、focus parameter の EL 推定量の極限分布の平均 2 乗誤差を求めることができ、平均 2 乗誤差を推定したものが FIC となる。しかしながら、局所的な定式化の誤りの下では、バイアスの一致推定量を得ることができないという問題がある。そこで、CH (2003) と同様に、平均 2 乗誤差の漸近的に不偏な推定量を導出し、これを FIC とした。

次に、モンテカルロ実験により、FIC のパフォーマンスを評価した。データ生成過程は、線形の操作変数モデルを採用した。線形モデルを用いた実証研究においてはしばしば、興味のあるパラメータは内生説明変数の係数だけで、外生変数については欠落変数バイアスを回避するためだけにモデルに入れられることがある。このような状況では、バイアスが十分に小さければ、外生変数をモデルから除外したほうが、推定量の平均 2 乗誤差を小さくできる可能性がある。そこで、内生説明変数の係数を focus parameter として、モデル選択を行った。実験の結果、バイアスが小さいときには、フルモデルを使うのに比べて、FIC によってモデル選択を行ったほうが、推定量の平均 2 乗誤差を小さくすることができることが確認された。また、Hong, Preston, and Shum (2003) で提案されている情報量規準と比較しても、FIC は概ね良好なパフォーマンスを示した。

CH (2003) 以降、Claeskens, Croux, Kerckhoven (2006) や Hjort and Claeskens (2006) など、様々なモデルの FIC が導出されてきた。しかし、これらの論文で考察されているモデルは尤度をベースとしたモデルであり、その意味では、CH (2003) の特殊ケースに過ぎないともいえる。それに対して、本研究はセミパラメトリックモデルが対象であり、FIC の適用範囲を本質的に拡張するものである。

モンテカルロ実験のデザインとして採用したように、線形の操作変数モデルは本研究の FIC の主要な適用例のひとつである。近年、操作変数の選択方法に関する研究は盛んに行われているが、モデル選択の研究は数少ない。本研究は、従来とは異なる視点から、操作変数モデルの分析方法を示すものでもある。

(2) アベレージング推定量の提案

(1) の結果の拡張として、モーメント制約モデルの focus parameter のアベレージング推定量を提案した。

局所的な定式化の誤りの下でのモデルアベレージングの方法は、Hjort and Claeskens

(2003) においても考察されているが、本研究とは目的が異なっている。Hjort and Claeskens (2003) はアベレージング推定量の信頼区間を求めることに主眼を置いているのに対し、本研究では、推定量の平均 2 乗誤差を最小化することを目的としている。

本研究では、アベレージングのウェイトが非確率的であると仮定した下で、アベレージング推定量の極限分布の平均 2 乗誤差を求め、これを最小にするウェイトの選び方を考察した。ただし、(1) と同様に、推定量のバイアスを一致推定することはできない。そのため、推定された平均 2 乗誤差を最小するようにウェイトを選ぶと、ウェイトは定数には確率収束せず、確率的なウェイトに分布収束することとなる。よって、最適なウェイトを一致推定することはできないという問題は残されている。

アベレージング推定量の有限標本の性質を検証するため、(1) と同様のデザインの下でモンテカルロ実験を行った。その結果、本研究の FIC を含む、過去に提案された情報量規準で選ばれた単独のモデルに基づいた推定量に比べ、アベレージング推定量はかなり低い平均 2 乗誤差を示した。多くの先行研究と同様に、アベレージングには推定量のリスクを軽減する効果があることが確認された。

本研究の関連研究として、Liu (2012) は局所的な定式化の誤りの下で、線形モデルのアベレージング方法を提案している。ただし、Liu (2012) は最小 2 乗法を用いたアベレージング推定量を提案しており、内生性がある場合のアベレージング方法は考察されていない。従って、本研究の成果は Liu (2012) を補完する結果にもなっている。

モーメント制約モデルのアベレージング方法に関する研究は、萌芽的な段階にある。本研究以外では、Martins and Gabriel (2012) が GMM 推定量を用いた異なるアベレージングの方法を提案しているが、残された課題は多い。本研究はこの分野における先鞭をつけるものである。

(1) と (2) で得られた研究成果は論文にまとめられ (論文タイトル: Empirical Likelihood-Based Focused Information Criterion and Model Averaging)、現在、国際学術誌に投稿中である。論文は研究代表者のホームページでも公開している。

(3) 大域的な定式化の誤りの下での exponential tilting (ET) 推定量の挙動の分析

これは、研究 (1) からの派生的な研究の成果である

(1) において、局所的な定式化の誤りの下での EL 推定量の極限分布を導出しているが、

大域的な定式化の誤りの下では、EL 推定量の漸近的性質は大きく異なる。Schennach (2007, AS) は、モーメント制約モデルに大域的な定式化の誤りが存在し、かつ、モーメント関数が有界でないとき、EL 推定量はいかなるパラメータの値にもサンプルサイズの平方根のオーダーでは確率収束できない (root-n-consistency が無い) ことを示している。

ET 推定量は EL 推定量と同様に、モデルと真の分布との間の Kullback-Leibler ダイバージェンスを最小にする推定量になっている。Schennach (2007) では、ET 推定量についても、一定の条件の下で、root-n-consistency が無いことを示唆しているが、擬似真値 (pseudo-true value) に確率収束する可能性については排除していなかった。本研究では、Schennach (2007) の設定の下では ET 推定量の擬似真値は最小化問題の双対問題からは識別されないことを示し、従って、ET 推定量は擬似真値に確率収束しえないことを明らかにした。

White (1982) や Vuong (1989) を始めとして、定式化の誤りの下でのパラメトリックモデルの推測に関する研究は、過去に数多く行われている。パラメトリックモデルでは、Kullback-Leibler ダイバージェンスによって擬似真値が自然に定義されるので、仮にモデルが誤っていても、意味のある推測を行うことが可能である。一方、モーメント制約モデルに対しても、Kullback-Leibler ダイバージェンスを定義することが可能であることから、近年では、モーメント制約モデルについても定式化の誤りの下での推測に関する研究が盛んに行われている。本研究の結果は、Kullback-Leibler ダイバージェンスを定義する上でのパラメトリックモデルとモーメント制約モデルの本質的差異を明らかにするものであり、定式化の誤りの下でのモーメント制約モデルの分析方法について注意を喚起するものである。

本研究成果は、学術誌 *Economics Letters* に掲載されている。

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 1 件)

- ① Sueishi, N. (2013): “Identification Problem of the Exponential Tilting Estimator under Misspecification,” *Economics Letters*, vol. 118, pp. 509-511. 査読有

DOI:10.1016/j.econlet.2012.12.024

京都大学学術情報リポジトリで公開

<http://repository.kulib.kyoto-u.ac.jp/d>

[space/handle/2433/171235](http://space.handle/2433/171235)

[学会発表] (計 6 件)

- ① 末石直也, 吉村有博 (発表者), “Variable and Order Selection Criterion for Series Estimation in Partially Linear Models,” 日本統計学会春季大会, 2013 年 3 月 3 日, 学習院大学
- ② 末石直也, “A New Interpretation of Empirical Likelihood for Time Series Models and Its Application to Model Selection Testing,” Hitotsubashi-Sogang Conference on Econometrics, 2012 年 11 月 17 日, Sogang University (韓国)
- ③ 末石直也, “Model Selection Criterion for Infinite Dimensional Instrumental Variable Models,” Institute of Mathematical Statistics Asia Pacific Rim Meeting (国際学会), 2012 年 7 月 3 日, Tsukuba International Congress Center
- ④ 末石直也, “Information Criteria for Moment Restriction Models,” Recent Development in Statistics, Empirical Finance and Econometrics (国際研究集会), 2011 年 11 月 30 日, 京都大学
- ⑤ 末石直也, “Information Criteria for Moment Restriction Models,” 統計関連学会連合大会, 2011 年 9 月 5 日, 九州大学
- ⑥ 末石直也, “Empirical Likelihood-Based Focused Information Criterion,” Summer Workshop on Economic Theory, 2011 年 8 月 10 日, 小樽商科大学

[その他]

研究成果は下記ホームページでも公開している。

<http://www.econ.kyoto-u.ac.jp/~sueishi/>

6. 研究組織

(1) 研究代表者

末石 直也 (SUEISHI NAOYA)

京都大学・経済学研究科 (研究院)・講師

研究者番号: 40596251