

機関番号：14303

研究種目：基盤研究(C)

研究期間：2007～2010

課題番号：19530182

研究課題名(和文) 無限次元局外母数がセミパラメトリック推定量に与える影響

研究課題名(英文) Effects of infinite dimensional nuisance parameters on semiparametric estimators

研究代表者

人見 光太郎 (HITOMI KOHTARO)

京都工芸繊維大学・工芸科学研究科・准教授

研究者番号：00283680

研究成果の概要(和文)：有限次元のベクトルと無限次元の関数という2種類のパラメータを持つセミパラメトリック統計モデルにおいて、無限次元のパラメータが既知の場合の方が、無限次元のパラメータが未知で推定する必要のある場合よりも有限次元パラメータベクトルの推定量の分散が小さくなるという逆説的な現象がおこることを理論的に予測し、その分散の逆転がおこる必要十分条件と一つのわかりやすい十分条件を導き出した。

研究成果の概要(英文)：This paper considers a puzzling phenomenon that is observed in some semiparametric estimation problems. In some cases, using estimated values of the nuisance parameters provides a more efficient estimator for the parameters of interest than does using the true values. This phenomenon takes place even in cases of semi-nonparametric models in which the nuisance parameters are infinite dimensional and can not be estimated at the parametric rate. We examine the structure and present the necessary and sufficient condition for the occurrence of this puzzle. We also provide a simple sufficient condition. It shows that the puzzle occurs when the term accounting for the effect of estimation of nuisance parameters is included in the tangent space.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2007年度	1,000,000	300,000	1,300,000
2008年度	700,000	210,000	910,000
2009年度	500,000	150,000	650,000
2010年度	700,000	210,000	910,000
年度			
総計	2,900,000	870,000	3,770,000

研究分野：計量経済学

科研費の分科・細目：経済学・経済統計学

キーワード：計量経済学、セミパラ推定、無限次元母数、分散の逆転

1. 研究開始当初の背景

通常、計量経済モデルは多くのパラメータを持つが、研究者が興味のあるのはその一部のパラメータであるのが普通で、それ以外は興味の対象ではない局外母数である。 β を研究者が興味があるパラメータとして、 γ を局外母数とする。もし、 γ の値がわかっていた場合はモデルに対する知識が増えるために β

の推定はより容易になる。この γ の値がわかっていた場合の β の推定量を β_1 としよう。 γ の値がわからずに、推定した γ を使った β の推定量を β_2 としよう。

通常は、より多くの知識があるために β_1 の分散の方が、 β_2 の分散より小さくなる。しかし、Pierce (1982)は古典的な有限次元のパラメータのケースで γ を推定する必要があ

る β_2 の分散の方が小さくなる場合があるということを見つけた。この現象をここでは推定量の分散の逆転と呼ぶことにしよう。その後、同じ現象がミッシングデータに関するモデルで多く見つかると、このパズルがどのような場合に起きるかということについての理解が進んだ。例えば、Robins, Mark and Newey (1992)、Imbens (1992)、Robins, Rotnitzky and Zhao (1994)。

しかし、近年多くのセミパラメトリックモデルが作られるようになると、局外母数が密度関数、条件付き期待値関数などの無限次元のパラメータの場合にも、推定量の分散の逆転現象が起ることが見つかってきた。例えば、Hirano, Imbens and Ridder (2003)、Magnac and Maurin (forthcoming)等がこの特定のセミパラメトリックモデルでこの現象が起ることを報告している。局外母数が無限次元パラメータの場合には、この推定量の分散の逆転はより不思議な現象となる。なぜなら、有限次元のパラメータの場合には推定量は $N^{-1/2}$ (N は標本数) の収束速度で真のパラメータに近づいていくのに対し、関数などの無限次元のパラメータの場合には $N^{-1/2}$ よりも遅い速度でしか真のパラメータに収束しない。これは無限次元パラメータを推定する必要がある場合には、計量経済モデルに対する知識がよりゆっくりとしか蓄積しないことを意味するからである。

2. 研究の目的

本研究は、この無限次元の局外母数が未知で推定する必要がある場合の方が、無限次元の局外母数が既知の場合よりも興味がある有限次元パラメータの推定量の分散が小さくなるという不思議な現象を解明する。無限次元パラメータの収束速度は有限次元パラメータの収束速度よりも遅いため、このような逆説的な現象を起らないと考えられていたが、そのような現象が発生するいくつかのセミパラメトリック推定量が発見されたために、なぜそのような現象が発生するのかというメカニズムを解明する。

3. 研究の方法

局外母数が有限次元の場合は、推定量の分散の逆転が起るのは、2段階推定の1段階目で効率的に(最尤法で)局外母数 γ を推定するが、2段階目の興味があるパラメータ β の推定には最尤法は使わずに、他のモーメント条件が使われている。そして、推定した γ を使う β_2 の influence function が真の γ を使う β_1 の influence function からその influence function の γ のスコアへの射影を取り除いたものとなっているために β_2 の分散の方が小さくなっている。これを無限次元のパラメータの場合に拡張するには、

そのセミパラメトリックモデルに含まれる様々な有限次元の γ によるパラメトリックサブモデルを考え、各々のパラメトリックサブモデルでの γ のスコアで張られる接平面 (tangent space) と β_2 の influence function の関係に読み替えてやることで可能となるという予想のもとで以下のような予備研究を行ってから理論的分析を行った。

- (1) 文献検索と通じて、推定量の分散の逆転が起っているセミパラメトリック推定量の例をより多く見つけ、それらがこの研究で予想しているような接平面と influence function の関係を持っているかどうかを確かめる。
- (2) いくつかのセミパラメトリック推定量について、コンピュータによるモンテカルロシミュレーションを使って無限次元局外母数を推定した場合と推定しない場合の分散の大きさの比較を行なう。

4. 研究成果

当初の研究目的から得られた成果とそこから派生した成果を述べる。

- (1) これは当初の研究目的から得られた成果で雑誌論文②にまとめた成果である。有限次元のベクトルと無限次元の関数という2種類のパラメータを持つセミパラメトリック統計モデルにおいて、無限次元のパラメータがわかっている場合の方が、無限次元のパラメータが未知で推定する必要がある場合よりも有限次元パラメータベクトルの推定量の分散が小さくなるという逆説的な現象がおこることを理論的に予測し、その分散の逆転がおこる必要十分条件と一つのわかりやすい十分条件を導き出した。

基本的な考え方は以下のようなものである。局外母数が有限次元の場合は、推定量の分散の逆転が起るのは、2段階推定の1段階目で効率的に(最尤法で)局外母数を推定するが、2段階目の興味があるパラメータの推定には最尤法は使わずに、他のモーメント条件が使われている。そして、推定した局外母数を使う推定量の influence function が真の局外母数推定量の influence function からその influence function の局外母数のスコアへの射影を取り除いたものとなっているために推定した局外母数を使った場合の分散の方が小さくなっている。

これを無限次元のパラメータの場合に拡張し、そのセミパラメトリックモデルに含まれる様々な有限次元の局外母数によるパラメトリックサブモデルを考え、各々のパラメトリックサブモデルでの局外母数のスコアで張られる接平面

(tangent space)と興味あるパラメータの influence function の関係に読み替えてやることで無限次元の局外母数の場合に拡張した。

そこから導かれる一つの十分条件は局外母数の推定が与える効果がパラメトリックサブモデルでの局外母数のスコアで張られる接平面 (tangent space)に含まれていることである。

- (2) 当初の研究目的から派生した研究成果で雑誌論文①にまとめたものであるが、Granger の因果性の統計的な検定方法を開発した。

経済時系列での代表的な因果関係の概念である Granger の因果性はある確率変数 $X(t)$ の予測値が X の過去の値 ($X(t-1)$, $X(t-2)$, ...) だけでなく他の確率変数 $\{Y(t)\}$ のも使って過去の値にも条件付けた方が予測値の分散が小さくなることと定義される。そして、通常は Granger の因果性での予測値は予測誤差を最小にする過去の値の最良線形写像で定義される。しかし、これでは非線形の影響を除けないため、非線形な関係を許すように定義を拡張し、そのような強い意味での Granger の因果性が存在するかどうか、どのような非線形な関係でも検出できるような統計的な検定を提案した。また、期待値(1次モーメント)だけではなく、近年金融データで注目されている2次モーメント、あるいはより一般的に k 次のモーメントにまで同時に Granger の意味で因果性があるかどうかの統計的な検定を提案した。

この検定の一つの面白い特徴は平滑化を行うカーネルを使った検定にも関わらず、 $N^{1/2}$ (N はサンプル数) で近づいてくるような局所対立仮説に対してもパワーを持つことである。

- (3) これも当初の研究目的から派生した成果で雑誌論文④にまとめたものである。個人情報秘匿等の目的のために所得統計などの多くのデータはそのまま公表されるのではなく、階級別の頻度、階級別平均の形でしか公表されない。そのような階級別の平均、頻度しか得られない場合には通常の最尤法は使うことができず、今までは階級別の頻度のみを使った擬似的な最尤法がパラメータの推定に使われてきた。本研究ではデータから得られる階級別頻度、階級別平均が発生する同時密度関数を明示的に導出し、それを使ってこのデータから最も効率的に分布のパラメータを推定するための最尤推定量を導

出した。

また、一方で階級別頻度、階級別平均を使って、パラメータを推定する GMM 推定量を導出し、その GMM 推定量の漸近分布が階級頻度、階級平均を使った西遊推定量の漸近分布に等しいことを示した。

また階級頻度、階級平均をもとにした尤度関数を定義できたので、その尤度に基づくモデル選択の手法も導出することができた。

- (4) 当初の研究目的から派生した研究成果で雑誌論文③にまとめたものである。発展途上国では初等教育であっても退学率が高い。その原因を調べるためにエジプトの小学生の年代の子供のパネルデータを使って、子供の成績の上昇で計った学校の質と退学率を導入した家計の行動モデルを推定した。その結果、子供の潜在能力と達成度を説明変数によってコントロールすると質の高い小学校の方が退学率が低くなることを示すことができた。これは家計の経済合理性と整合的な結果であることを見いだした。

- (5) これは当初の研究目的より派生して学会発表①にまとめたものである。この研究で行ったことは、確率変数に高次のモーメントが存在するかどうかの統計的な検定を作り出した。基本的な考え方は2次モーメントの存在に関する検定を作ると一般の k 次モーメントの検定は確率変数 x を $y = |x|^{k/2}$ と変換してやることによって y の2次モーメントの検定にできるので2次モーメントの検定さえつくればよい。2次モーメントの存在と特性関数が原点で2回微分可能であるのと同値であるという関係を使って、標本特性関数の原点での2階微係数の推定量を使った検定を導出した。

- (6) これは当初の研究目的より派生して学会発表②にまとめたものである。離散的な確率過程 $\{X(t)\}$ が ARMA 過程に従う場合、AR 部分の係数の絶対値が1を超える場合には確率過程は急速に発散し、1を超えない場合には定常確率過程となるため、AR 係数の値が1を超えるかどうかの検定が重要な問題となり単位根検定といわれている。通常はすでに計測されたサンプル数が決まっているデータを用いて検定を

行うが、本論文ではデータが次々に入ってくる逐次的な設定での単位根検定を導出した。検定統計量は局所対立仮説のもとで漸近的に正規分布に従うことを示した。手法的には離散的な時系列データを用いた逐次検定を連続時間の確率過程で拡散近似して漸近的な分布を得る一般的な手法を提案し、その応用として離散時間での逐次的な単位根検定と移民のある分岐過程のパラメータの検定の漸近分布を導出した。

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 Y. Nishiyama, K. Hitomi, Y. Kawasaki, Kiho Jeong, “A Consistent Nonparametric Test for Nonlinear Causality”, Journal of Econometrics, 印刷中, 査読有
- ② K. Hitomi, Y. Nishiyama, and Ryo Okui, “A Puzzling Phenomenon in Semiparametric Estimation Problems with Infinite-Dimensional Nuisance Parameters,” Econometric Theory (2008), 24:1717-1728, 査読有
- ③ Eric A. Hanushek, Victor Lavy and K. Hitomi “Do Students Care about School Quality? Determinants of Dropout Behavior in Developing Countries,” Journal of Human Capital (2008), vol. 2, No. 1, pp.69-105, 査読有
- ④ K. Hitomi, Qing-Feng Liu, Yoshihiko Nishiyama and Naoya Sueishi, “Efficient Estimation and Model Selection for Grouped Data with Local Moments,” Journal of the Japan Statistical Society (2008), vol. 38, No. 1, pp.131-143. 査読有

[学会発表] (計 2 件)

- ① K. Hitomi, Y. Nishiyama, K. Nagai, “A Nonparametric test for the existence of moments”, 2009 年度日本統計学会春季大会発表論文 (2009 年 3 月 6 日)、統計数理研究所
- ② 永井圭二、人見光太郎、西山慶彦、「単位根と分岐過程の criticality の逐次検定」、2008 年度日本統計学会春季大会発表論文(2008 年 3 月 1 日)、成城大学

6. 研究組織

(1) 研究代表者

人見 光太郎 (HITOMI KOHTARO)
京都工芸繊維大学・工学科学研究科・
准教授
研究者番号：00283680

(2) 研究分担者
()

研究者番号：

(3) 連携研究者
()

研究者番号：