

平成22年 4月 20日現在

研究種目：若手研究(B)
 研究期間：2007 ～ 2009
 課題番号：19700265
 研究課題名（和文） ハイパーパラメータ選択のための情報量規準の漸近的特性
 研究課題名（英文） Asymptotic properties of information criteria for selecting hyper parameter
 研究代表者
 柳原 宏和 (Yanagihara Hirokazu)
 広島大学・大学院理学研究科・准教授
 研究者番号：70342615

研究成果の概要（和文）：leave-one-out CV 規準から leave-k-out CV 規準までを重みを付けて足すことにより CV 規準のバイアスを $O(n^{2k})$ に減らしたバイアス補正 CV 規準と、ジャックナイフ法でバイアスを評価することでバイアスを $O(n^{-3})$ に減らしたバイアス補正 CV 規準を提案した。さらに、正規性を仮定した多変量リッジ回帰におけるリッジパラメータの選択ための C_p 規準において、バイアスを完全に除去した C_p 規準を提案した。

研究成果の概要（英文）：We proposed two bias-corrected CV criteria, one is reducing the bias of CV to $O(n^{2k})$ by the weighed sum of leave-k-out CV criteria, and the other is reducing the bias of CV to $O(n^{-3})$ by applying the jackknife method to an evaluation of the bias. Moreover, we proposed an unbiased C_p criterion for multivariate ridge regression under the normal assumption.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2007 年度	1,100,000	0	1,100,000
2008 年度	1,100,000	330,000	1,430,000
2009 年度	1,000,000	300,000	1,300,000
年度			
年度			
総計	3,200,000	630,000	3,830,000

研究分野：統計科学

科研費の分科・細目：情報学・統計科学

キーワード：モデル選択, 情報量規準, 罰則付き尤度, バイアス補正

1. 研究開始当初の背景

価値観や生活様式の多様化により、社会現象が複雑化してきている今日において、複雑化した現象を表現するために、B-スプラインやニューラルネットワークなどの多変量ノンパラメトリックモデルのような従来の線形モデルよりも柔軟な統計モデルでの解析が求められている。ノンパラメトリックモデル

の最適化を尤度関数の最大化により行うと、強度にデータに依存した結果になるため、実解析上あまり好ましくない。そのため、罰則付き対数尤度関数の最大化により最適化を行うことが通常よく用いられる手法であるが、罰則項を制御するハイパーパラメータにより結果が大きく変化してしまうため、その選択問題は非常に重要な問題であると言え

る。ハイパーパラメータの選択問題において、赤池情報量規準 (Akaike's Information Criterion; AIC, Akaike, 1974, IEEE Transactions on Automatic Control) に代表される情報量規準の最小化により最適なハイパーパラメータを選ぶ方法が最も頻繁に用いられる手法である。情報量規準はカルバック=ライブラー情報量等の真のモデルと候補のモデルに関する予測ダイバージェンスに基づくリスク関数の推定量であるため、第一にそのバイアスが問題となり、必要であればそのバイアスを補正することを考えなくてはならない。情報量規準はモデルの当てはめの良さを測る標本ダイバージェンスにモデルの複雑さに関する罰則項を足すことで定義される。もし候補のモデルのパラメータ数が増えれば、標本ダイバージェンスは小さくなるが、その分モデルの複雑さは大きくなる。情報量規準が小さくなるモデルとは、モデルのあてはまりの良さと複雑さがちょうど釣り合ったモデルであると言える。このとき、バイアスが大きくなることは、モデルの複雑さが大きくまたは小さくなることに対応し、過大なバイアスのため、必要以上に単純なモデル、または複雑なモデルが選びやすくなってしまいう危険性がある。そのため、最終的なモデルの決定において深刻な間違いを起こす可能性がある。一方、クロスバリデーション (Cross-Validation; CV, Stone, 1974, Journal of the Royal Statistical Society Series B) 規準は、AIC などその他の情報量規準よりもバイアスが小さい規準量であることが知られている。しかしながら、バイアスが小さいといえども、標本数が小さいときやモデルの未知パラメータ数が大きいときに、無視できないバイアスが存在する可能性がある。より良いモデルを選ぶためには、よりバイアスの小さい情報量規準が必要となる。

2. 研究の目的

AIC に代表される情報量規準はカルバック=ライブラー情報量等の予測ダイバージェンスに基づくリスク関数の推定量であるため、そのバイアスが問題となり、必要に応じて補正を考えなくてはならない。最も簡単でよく用いられるバイアス補正法は、テーラー展開により得たバイアスの漸近展開式を用いる方法であるが、そのような補正法では、理論的にはバイアスを補正しているが数値的にはほとんどバイアスを補正していないという状況がしばしば起こる。漸近展開により求めた補正項は真の分布の高次キュムラントに依存していることが多く、実際に補正に用いるためにはそのキュムラントの推定量を使用する必要がある。しかしながら、このような高次キュムラントの推定量は標本数が少なくないときですら大きなバイアスを持ち、その結果、バイアス補正項がうまく

推定されないということが起きる。一方、CV 規準は、AIC などのその他の情報量規準よりもバイアスが小さい規準量であることが知られているが、小さいといえども、無視できないバイアスが存在する可能性がある。本研究の目的は、漸近展開式を利用してバイアスを補正する従来の方法ではない方法により CV 規準のバイアス補正を行い、できるだけバイアスの小さい規準量を提案することにある。

3. 研究の方法

本研究の目的は、バイアス項の漸近展開式を導出し、その展開式の初項または第二項での未知パラメータを推定量で置き換えたバイアス補正項を加える従来のバイアス補正法を用いることなく、ハイパーパラメータ選択のための情報量規準のバイアスを補正することである。しかしながら、実際に用いないといえども、バイアス項がどのような形になっているか調べることは重要である。実際に、罰則無しの対数尤度関数により未知パラメータの推定し、カルバック=ライブラー情報量によりモデルを評価した場合での CV 規準のバイアスの漸近展開式は、Yanagihara, Tonda & Matsumoto (2006, Journal of Multivariate Analysis) により求められているが、未知パラメータを罰則付き対数尤度関数により推定し、モデルの評価にカルバック=ライブラー情報量を用いた場合での CV 規準のバイアスの漸近展開式はまだ求められていない。また、実際のバイアス補正には以下のような方法を用いる。従来の CV 規準は $n-1$ 個の Calibration-Sample により予測モデルを構築し、残りの 1 個の Validation-Sample により予測モデルのあてはまりを測るものであるが、 $n-k$ 個の Calibration-Sample により予測モデルを構築し残りの k 個の Validation-Sample により予測モデルのあてはまりを測っている CV 規準、leave-k-out CV 規準もある。もちろん、従来の CV 規準は leave-one-out CV 規準である。これらの leave-one-out CV 規準から leave-k-out CV 規準までの重み付き和により、CV 規準のバイアス補正を行う。実際のバイアスをうまく補正するためには、この重みに関する係数をうまく選ぶ必要がある。これらの係数は真の分布の未知パラメータに依存しない形になることが予想される。このような係数の具体的な形を求め、本研究の目的の一つであるバイアス補正問題を解決する。また、1 回目の繰り返しブートストラップで得たバイアス補正項から k 回目の繰り返しブートストラップで得たバイアス補正項までの重み付き和によりバイアス補正を行う、繰り返しブートストラップ法によるバイアス補正法 (Hall & Martin, 1988, Biometrika) でも同様なバイアス補正を行うことが出来る。しかしながら、

繰り返しブートストラップ法では、その繰り返し回数が増えれば増えるほど推定値が不安定になる。一方、本研究でのバイアス補正法は繰り返し回数の増加による推定値の不安定さは繰り返しブートストラップ法に比べ小さい。さらに、実際のバイアス補正のオーダーは繰り返しブートストラップ法によるバイアス補正よりも小さくなることが予想される。これら二つの補正法の違いや関連性もバイアス補正に用いる重みに関する係数を比較することにより明らかにする。

4. 研究成果

情報量規準のバイアス項がどのような形で展開されているかを明らかにするため、未知パラメータの推定に用いる関数とモデルの評価に用いる関数が異なる場合での CV 規準のバイアスの展開式を n^{-1} 項まで求めた。同様の計算で一般化情報量規準 (Generalised Information Criterion; GIC, Konishi & Kitagawa, 1996, Biometrika) のバイアスの展開式も n^{-1} の項まで求められた。バイアス補正はバイアスの漸近展開式を利用する方法ではなく、leave-one-out CV 規準から leave-k-out CV 規準までを重みを付けて足すことにより行った。バイアスをうまく補正するためには、この重みに関する係数をうまく選ぶ必要があり、これらの重みに関する係数の一般形を具体的に求めた。この規準量のバイアスは $O(n^{-2k})$ となる。また、提案したバイアス補正法と、繰り返しブートストラップによるバイアス補正法は非常に似たものとなり、新しいバイアス補正情報量規準での係数は繰り返しブートストラップ法によりバイアス補正を施した規準量での係数に収束することがわかった。また、バイアス項をジャックナイフ法で推定することにより、新たなバイアス補正 CV 規準を提案した。この新たな CV 規準には、leave-one-out CV 規準から leave-k-out CV 規準までを重みを付けてバイアスを補正したバイアス補正 CV 規準と同様に、具体的な展開式を求めなくてもバイアス補正規準を計算することができる利点がある。バイアスのオーダーは CV 規準のバイアスが $O(n^{-2})$ であるのに対して $O(n^{-3})$ となり、重み付き和による補正よりも低次になるが、簡単に求めることができるといった利点を持つ。一方、CV 規準は計算時間がかかりすぎるという欠点がある。計算時間の短縮のために、モデルや規準量を限定した下でバイアス補正を試みた。まず、正規性を仮定した多変量リッジ回帰におけるリッジパラメータの選択のための C_p 規準において、バイアス補正を行った MC_p 規準を導出した。この規準量はバイアスを補正しているだけでなく、従来

の C_p 規準よりも分散が小さいものになっており、実際にはリスク関数の最小分散不偏推定量となっている。数値実験の結果、ほとんどの場合、 MC_p 規準でリッジパラメータを最適化した方が、 C_p 規準により最適化するよりも予測 MSE が小さくなることがわかった。しかしながら、リッジ回帰において、情報量規準を最小にする最適なリッジパラメータは陽な形で求めることができず、最適解を求めるためには計算機による繰り返し計算が必要となる。一方、Hoerl and Kennard (1970, Technometrics) では、リッジパラメータを説明変数の個数まで追加した一般化リッジ回帰も提案されている。この一般化リッジ回帰を多変量に拡張した。リッジ回帰では LSE を一様に縮小しているのに対し、一般化リッジ回帰では個々の変数ごとに縮小する度合いが異なるため、リッジ回帰より柔軟な縮小推定が行えると期待できる。多変量一般化リッジ回帰においても、リッジパラメータ選択のための MC_p 規準を提案した。この MC_p 規準もリッジ回帰の場合と同様に、バイアスを補正しているだけでなく、分散も改善しており、リスク関数の最小分散不偏推定量となっている。最適化を必要とするリッジパラメータが増加しているにもかかわらず、 C_p 規準や MC_p 規準を最小にするリッジパラメータは陽な形で求めることができることがわかった。また、 C_p 規準で最適化したリッジパラメータは、分散共分散行列を既知としたときの経験ベイズ法により得られた推定量に分散共分散の不偏推定量を代入したものと一致することもわかった。さらに、数値実験により、ほとんどの場合、 MC_p 規準でリッジパラメータを最適化した方が、 C_p 規準により最適化するよりも予測 MSE が小さくなることがわかった。一般化リッジ回帰で言えば、予測 MSE を最小にするリッジパラメータが未知パラメータを含んだ形で陽にかけるため、それらの未知パラメータを推定量に置き換えることで、情報量規準の最小化とは異なった、リッジパラメータの最適化法が考えられる。それらの最適化法は、単変量の場合において、提案されていた。これら結果を多変量に拡張し、新たなリッジパラメータの最適化法を提案した。具体的には、予測 MSE を最小にするリッジパラメータの未知回帰係数行列と分散共分散行列に LSE と不偏推定量を代入する方法、LSE ではなく、上記のリッジパラメータを考慮した一般化リッジ回帰推定量を代入したもの、その代入を無限回繰り返し続けたものの 3 通りである。数値実験の結果より、未知パラメータ行列が小さい場合

は無限回繰り返す方法が、未知パラメータ行列が大きい場合は、従来のLSEを代入する方法が予測MSEを小さくすることがわかった。平均的にみれば、 MC_p で最適化する方法が最も良いこともわかった。

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5件)

- ① Yanagihara, H. and Satoh, K., An unbiased C_p criterion for multivariate ridge regression, Journal of Multivariate Analysis, 査読有り, Vol. 101, 2010, 1226-1238.
- ② 柳原宏和・永井 勇・佐藤健一, 多変量一般化リッジ回帰におけるリッジパラメータ最適化のためのバイアス補正 C_p 規準, 応用統計学, 査読有り, 38 巻, 2009, 151-172.

[学会発表] (計6件)

- ① 永井 勇・柳原宏和・佐藤健一, 多変量一般化リッジ回帰モデルにおけるリッジパラメータ最適化法の比較, 2009 年度統計関連学会連合大会, 2009 年 9 月 9 日, 同志社大学.
- ② Yanagihara, H., Bias corrections of cross-validation criterion, 2009 年度統計関連学会連合大会, 2009 年 9 月 7 日, 同志社大学 (招待講演)
- ③ Yanagihara, H., Simple bias-corrected cross-validation criterion, The 2009 Joint Statistical Meetings, August 3, 2009., Walter E. Washington Convention Center USA (Poster Session).
- ④ 永井 勇・柳原宏和・佐藤健一, 多変量一般化リッジ回帰モデルにおけるリッジパラメータの選択法, 2008 年度統計関連学会連合大会, 2008 年 9 月 8 日, 慶応大学.
- ⑤ Satoh, K. and Yanagihara, H., Modified C_p in multivariate ridge regression, East Asia Regional Biometric Conference 2007, December 10, 2007, University of Tokyo (Poster Session).
- ⑥ 柳原宏和, クロスバリデーション規準の簡単なバイアス補正法, 2007 年度統計関係学会 連合大会, 2007 年 9 月 8 日, 神戸大学.

[その他]

ホームページ等

<http://www.math.sci.hiroshima-u.ac.jp/~yanagi/>

6. 研究組織

(1)研究代表者

柳原 宏和 (Yanagihara Hirokazu)

広島大学・大学院理学研究科・准教授

研究者番号：70342615

(2)研究分担者

()

研究者番号：

(3)連携研究者

()

研究者番号：