

Title	NONNEGATIVE GARROTE IN BOOSTING AND REGRESSION TREES
Author(s)	元垣内, 広毅
Citation	大阪大学, 2012, 博士論文
Version Type	
URL	https://hdl.handle.net/11094/59082
rights	
Note	著者からインターネット公開の許諾が得られていないため、論文の要旨のみを公開しています。全文のご利用をご希望の場合は、 〈a href="https://www.library.osaka-u.ac.jp/thesis/#closed"〉 大阪大学の博士論文について 〈/a〉 をご参照ください。

Osaka University Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

Osaka University

氏名	元垣内 広毅
博士の専攻分野の名称	博士(工学)
学位記番号	第 25249 号
学位授与年月日	平成24年3月22日
学位授与の要件	学位規則第4条第1項該当 基礎工学研究科システム創成専攻
学位論文名	NONNEGATIVE GARROTE IN BOOSTING AND REGRESSION TREES (ブースティング及び回帰樹木における非負圧縮推定量)
論文審査委員	(主査) 教授 白旗 慎吾 (副査) 教授 狩野 裕 教授 内田 雅之

論文内容の要旨

一般に、回帰解析の目標には、応答の適合予測と回帰関係の適切な解釈を与える二つの立場がある。本研究では、これらの二つの目標を同時に満たす回帰係数の縮小推定量として、非負圧縮推定量をとりあげ、近年の計算機志向型の接近法との融合を試みた。

第一に、弱学習器を組み合わせることによって、応答の予測性能を向上させる方法論として、ブースティング法に焦点をあて、非負圧縮推定量にブースティング法を導入する2種の方法を提案し、文献事例及びシミュレーションによって、ブースティング法の導入が、パラメータの推定量の安定性の向上においても有用であることを示した。

第二に、データの層化と回帰解析を同時に遂行しながら、データの背後に潜む回帰関係を捉えることを可能とする樹木構造接近法として、多変量適応型回帰スプラインが知られている。本研究では、多変量適応型回帰スプラインにおける基底関数の剪定を非負圧縮推定量により遂行する方法を提案し、文献事例及びシミュレーションによってその性能を評価した。その結果、提案法は、回帰関係の解釈が容易な単一の樹木構造の形式をとるとともに、より安定的に基底関数の刈り込みを行い、さらに、多くの状況下において、通常多変量適応型回帰スプラインを上回る予測性能を示した。

第三に、回帰樹木の生成にランダム性を注入し、ランダム性を有する多数の樹木によるアンサンブル学習から、強力な予測力を保持する方法論として、バギング樹木及びランダム・フォレストが知られている。他方、多数の樹木で構成されるこれらの方法論は、その複雑なモデルからの回帰関係の解釈が困難であるという欠点を有する。そこで、本研究では、バギング樹木とランダム・フォレストに非負圧縮推定量を導入し、応答の予測力のさらなる向上に加え、多数のランダムな樹木の順序付けと絞り込みを行うことで、代表的な樹木からの回帰関係の解釈を行う方法を提案した。文献事例及びシミュレーションを通して、その性能を評価した結果、提案法は、代表的な樹木による回帰関係の解釈が可能であるとともに、高い予測力を有することが示唆された。

論文審査の結果の要旨

回帰分析の目的には、応答の予測と回帰関係の適切な解釈を与える2つの立場がある。本研究では、これらの2つの目標を同時に満たす回帰係数の推定量として、縮小推定量である非負圧縮推定量をとりあげ、近年の計算機志向型統計解析の接近法との融合を試みている。

まず第1に、弱学習器を組み合わせることによって、応答の予測性能を向上させる方法論として、ブースティング法に焦点をあて、非負圧縮推定量にブースティング法を導入する2種の方法を提案し、文献事例及びシミュレーションによって、ブースティング法の導入が、パラメータの推定量の安定性の向上においても有用であることを示した。

第2に、データの層化と回帰分析を同時に遂行しながら、データの背後に潜む回帰関係を捉えることを可能とする樹木構造接近法として、多変量適応型回帰スプラインが知られている。多変量適応型回帰スプラインにおける基底関数の剪定を非負圧縮推定量により行う方法を提案し、文献事例及びシミュレーションによってその性能を評価した。その結果、提案法は、回帰関係の解釈が容易な単一の樹木構造の形式をとるとともに、より安定的に基底関数の刈り込みを行い、さらに、多くの状況下において、通常多変量適応型回帰スプラインを上回る予測性能を示した。

第3に、回帰樹木の生成にランダム性を注入し、ランダム性を有する多数の樹木によるアンサンブル学習から、強力な予測力を保持する方法論として、バギング樹木及びランダム・フォレストが知られている。他方、多数の樹木で構成されるこれらの方法論は、その複雑なモデルからの回帰関係の解釈が困難であるという欠点を有する。そこで、本研究では、バギング樹木とランダム・フォレストに非負圧縮推定量を導入し、応答の予測力のさらなる向上に加え、多数のランダムな樹木の順序付けと絞り込みを行うことで、代表的な樹木からの回帰関係の解釈を行う方法を提案した。そして、文献事例及びシミュレーションを通して、その性能を評価した結果、提案法は、代表的な樹木による回帰関係の解釈が可能であるとともに、高い予測力を有することが示唆された。

このように本論文は回帰分析における新たな方法論を提案し、実データおよびシミュレーションによりその良さを示しており、博士(工学)の学位論文として価値のあるものと認める。