

Model estimation and selection schemes for efficient and accurate forecasting

吉田, 航

<https://hdl.handle.net/2324/7363909>

出版情報 : Kyushu University, 2024, 博士 (数理学) , 課程博士
バージョン :
権利関係 :



氏 名 : 吉田 航

論 文 名 : Model estimation and selection schemes for efficient and accurate forecasting

(効率的かつ高精度な予測のためのモデル推定と選択)

区 分 : 甲

論 文 内 容 の 要 旨

本論文では、統計モデルを用いた予測の実用性を高めるために、高次元データ、モデル選択、計算の高速化といった現実的な課題に対処する手法を提案する。

本論文は2部構成になっている。第1部では、線形ガウス状態空間モデルを用いた多変量予測に関する内容である。このモデルは、時間変動を柔軟に表現できることから広く研究されている一方で、学習に大きな計算コストを要するという課題がある。特にデータの次元が増えると推定すべきパラメータが膨大になり、従来の方法では計算負荷が実用的な範囲を超えてしまう。この問題に対処するため、本論文ではパラメータベクトルの推定値を各成分（次元）ごとに独立に計算する手法を提案し、計算時間を大幅に削減した。この高速化手法の導出は、予測誤差の分布に基づいている。

さらに、モデル選択の課題にも取り組む。モデルの候補が多い場合、従来の方法ではそれらを個別に評価する必要があり、計算コストが急増する。特に時系列分析では、この問題がさらに複雑化する。例えば、回帰係数の時間変化を許した回帰モデルである動的回帰モデルを考える。動的回帰のモデル選択では、通常回帰モデルのように「どの変数をモデルに含めるか」を決めるだけでなく、「各回帰係数は時変性を有するか否か」をも考慮する必要がある。本論文では、このような時系列特有のモデル選択問題に着目し、スパース推定やベイズ分析に基づく効率的な手法を新たに提案する。

第2部では、線形回帰モデルを用いた予測について議論する。本研究では特に、データの収集コストが高い、取得頻度が低いなどの理由で、サンプルサイズが小さい場合の予測に注目する。このような状況では、推定やモデル選択の不安定性が予測精度を低下させる要因となる。例えば、小サンプル下での線形回帰では縮小推定が頻繁に用いられるが、適切なモデルを安定的に得ることは難しい。そのため、この不安定性を低減し、小サンプル下でも有効な予測を実現する必要がある。

そこで我々は、ブートストラップによるリサンプリング法を用いて複数の予測結果を集約することで、不安定性の解消を試みた。ここで、ブートストラップでは通常、最小二乗法に基づき推定された分布を用いて実行される。しかし、実験の結果、この方法では予測精度が改善されないことが判明した。その原因を調べたところ、リサンプリングで使用する分布の選択が予測精度に大きく影響することがわかった。これを受けて、我々はリサンプリング分布と予測精度の関係に着目し、最小二乗法ではなく交差検証法によって得られた分布に基づいてリサンプリングする方法を提案した。その結果、予測精度を向上させるのに適した分布を探索的に選択できるようになり、小サンプル下でも安定した予測を実現できることを確認した。