

統計モデルの接空間を使った LARS 法

北海道大・情報科学 廣瀬 善大

1 背景

スパース推定・スパースモデリングの研究はこの 20 年ほどの間に大きく進展した。その代表的な手法のひとつが Least Absolute Shrinkage and Selection Operator (LASSO) [5] である。LASSO は線形回帰問題に対するパラメータ選択と変数選択を同時に行う手法であり、最適化問題として定式化される。この最適化問題を効率的に解くためのアルゴリズムに Least Angle Regression (LARS) [3] がある。LARS は反応変数と説明変数の相関を利用したアルゴリズムであり、誤差として正規分布を仮定することで、ユークリッド幾何による記述が可能となっている。

ユークリッド幾何の一般化のひとつとして情報幾何が知られている [1]。情報幾何は、リーマン幾何的手法を統計学の問題に持ち込むことを可能にする。たとえば、正規分布がユークリッド幾何として記述されることは、情報幾何の一例として扱うことができる。

2 提案手法

本発表では、LARS を利用して、一般化線形モデルに対するパラメータ推定手法を提案する。一般化線形モデルは、情報幾何の観点からは多様体になっており、ユークリッド幾何では扱うことができない。この多様体は双対平坦と呼ばれる性質をもつことが知られており、双対平坦性にもとづく LARS の拡張が提案されている [4, 2]。提案手法では、接空間がユークリッド幾何で扱えることを利用し、多様体を直接に使うことはせず接空間において LARS を実行する。それにより計算効率が非常に高い手法となる。数値実験により、一般化線形モデルに対する LASSO との比較を行い、接空間による統計モデルの簡略化をしても推定の性能が劣らないことを確認する。

参考文献

- [1] Amari (2016). *Information Geometry and Its Applications*, Springer.
- [2] Augugliaro, Mineo, and Wit (2013). dgLARS: a Differential Geometric Approach to Sparse Generalized Linear Models, *Journal of the Royal Statistical Society. Series B*, **75**, 471–498.
- [3] Efron, Hastie, Johnstone and Tibshirani (2004). Least Angle Regression, *Annals of Statistics*, **32**, 407–499.
- [4] Hirose and Komaki (2010). An Extension of Least Angle Regression Based on the Information Geometry of Dually Flat Spaces, *Journal of Computational and Graphical Statistics*, **19**, 1007–1023.
- [5] Tibshirani (1996). Regression Shrinkage and Selection via the Lasso, *Journal of the Royal Statistical Society. Series B*, **58**, 267–288.