

Bayesian generalized fused lasso modeling via NEG distribution

中央大学大学院 嶋村 海人
久留米大学 バイオ統計センター 植木 優夫
中央大学 小西 貞則

1. はじめに Fused lasso (Tibshirani *et al.*, 2005) は、回帰係数の L_1 ノルムに加え、連続した回帰係数の差の L_1 ノルムを罰則として課すスパースモデリングである。この罰則により、回帰係数の推定値はスパース化だけでなく平滑化も同時に行うことができる。このため、生命科学におけるゲノムデータ解析や画像処理など様々な分野への応用を目的として、fused lasso の特徴を有するスパース正則化法が提唱され、これらを総称して generalized fused lasso と呼ばれている。Bayesian fused lasso は、fused lasso をベイズアプローチによって捉え、推定に MCMC を適用したものである。本報告では、generalized fused lasso をベイズの観点から考察し、NEG (normal-exponential-Gamma) 分布に基づくベイズ型スパースモデリングを提唱し、数値実験を用いて有効性を検証する。

2. 提案手法 ガウス型線形回帰モデルにおける正則化最尤法による推定をベイズの観点から捉えると、与えられたデータ \mathbf{y} に対して回帰係数 $\boldsymbol{\beta}$ は事後分布 $\pi(\boldsymbol{\beta}|\mathbf{y}) \propto f(\mathbf{y}|\boldsymbol{\beta}, \sigma^2)\pi(\boldsymbol{\beta}|\sigma^2)\pi(\sigma^2)$ のモードとして推定される。また、 L_1 ノルム罰則はベイズの枠組みで捉えると、ラプラス事前分布に対応する。Kyung *et al.* (2010) では、 $\boldsymbol{\beta}$ の事前分布に2つのラプラス分布の積を仮定することで、fused lasso をベイズの枠組みで定式化することによって得られた Bayesian fused lasso を提案した。ここで、fused lasso は次の目的関数を最小化する手法である。

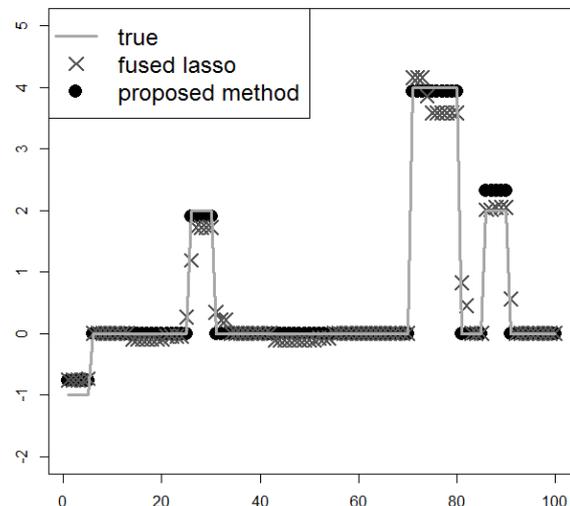
$$\log f(\mathbf{y}|\boldsymbol{\beta}, \sigma^2) - \gamma_1 \sum_{j=1}^p |\beta_j| - \gamma_2 \sum_{j=2}^p |\beta_j - \beta_{j-1}|.$$

本研究では、Bayesian fused lasso におけるラプラス分布を、次の式で与えられる NEG 分布に置き換えることで新たなベイズモデリングを提案する。

$$\text{NEG}(\beta_j|\lambda, \gamma) = \kappa \exp\left(\frac{\beta_j^2}{4\gamma^2}\right) D_{-2\lambda-1}\left(\frac{|\beta_j|}{\gamma}\right).$$

ただし、 $D_{-2\lambda-1}$ は放物柱関数 (parabolic cylinder function) である。この fused lasso 型ベイズスパースモデリングでは、NEG 分布の階層表現を利用して、ギブスサンプリングによるベイズ推定が可能となる。

3. 数値実験・まとめ Friedman *et al.* (2007) の数値実験に基づき、手法の有効性を検証する。この数値実験は染色体異常の検出などを想定したものである。右の図には、真のモデル (実線) に対して、従来手法である fused lasso (×) と提案手法 (●) を比較した結果を与えた。図より、提案手法が従来手法よりも真の構造をよく捉えていることがわかる。NEG 分布を利用した提案手法は、従来手法と比べて柔軟に正則化項を設定できることから、より汎用性の高いスパースなモデルの構築に有用であると考えられる。



参考文献

- [1] Friedman, J., Hastie, T., Hofling, H. and Tibshirani, R. (2007). *Ann. Appl. Stat.*, **1**, 302–332.
- [2] Kyung, M., Gill, J., Ghosh, M. and Casalla, G. (2010). *Baysn. Anal.*, **5**, 369–412.
- [3] Tibshirani, R., Saunders, M., Rosset, S., Zhu, J. and Knight, K. (2005). *J. Roy. Stat. B.*, **67**, 91–108.