

Equivalence between Adaptive-Lasso and Generalized Ridge Estimates in Linear Regression with Orthogonal Explanatory Variables after Optimizing Regularization Parameters

広島大・理 大石 峰暉 柳原 宏和

\mathbf{y} を n 次元目的変数ベクトル, \mathbf{X} を中心化 ($\mathbf{X}'\mathbf{1}_n = \mathbf{0}_k$) された $n \times k$ 説明変数行列とした以下のような線形回帰モデルを扱う.

$$\mathbf{y} = \mu\mathbf{1}_n + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}.$$

ここで, $\mathbf{1}_n$ はすべての成分が1である n 次元ベクトル, $\mathbf{0}_k$ はすべての成分が0である k 次元ベクトル, μ は位置母数, $\boldsymbol{\beta}$ は k 次元回帰係数ベクトル, $\boldsymbol{\varepsilon}$ は各成分が独立に平均0, 分散 σ^2 の分布に従う n 次元誤差ベクトルである. 本発表では説明変数が直交する場合を扱うため, \mathbf{X} に以下を仮定する.

$$\mathbf{X}'\mathbf{X} = \mathbf{D} = \text{diag}(d_1, \dots, d_k), \quad d_1 \geq \dots \geq d_k > 0.$$

さらに, $\text{rank}(\mathbf{X}) = k < n - 1$ とする. 回帰係数の推定法として, Adaptive-Lasso (AL) 型 ℓ_1 ペナルティと Generalized Ridge (GR) 型 ℓ_2 ペナルティを用いた罰則付き推定法を考える. これらのペナルティを用いて得られた推定値である, AL-推定値 $\hat{\boldsymbol{\beta}}_{\boldsymbol{\theta}}^{\text{AL}}$, GR-推定値 $\hat{\boldsymbol{\beta}}_{\boldsymbol{\theta}}^{\text{GR}}$ の第 j 成分 ($j = 1, \dots, k$) はそれぞれ以下のように与えられる.

$$\hat{\beta}_{\boldsymbol{\theta},j}^{\text{AL}} = \begin{cases} \frac{-\theta w_j \text{sign } z_j + 2d_j^{1/2} z_j}{2d_j} & (\theta < 2d_j^{1/2}|z_j|/w_j), \\ 0 & (\theta \geq 2d_j^{1/2}|z_j|/w_j) \end{cases}, \quad \hat{\beta}_{\boldsymbol{\theta},j}^{\text{GR}} = \frac{d_j^{1/2} z_j}{d_j + \theta_j}.$$

ここで, θ , $\boldsymbol{\theta} = (\theta_1, \dots, \theta_k)'$ はそれぞれのペナルティに対する正則化パラメータ, z_j は $\mathbf{D}^{-1/2}\mathbf{X}'\mathbf{y}$ の第 j 成分, w_j は AL のペナルティの重みである. AL-推定値は0を含むことから AL ではスパースな推定値が得られ, 一方で GR-推定値は0を含まないことから GR ではスパースな推定値は得られないことがわかる.

これらの推定値はいずれも正則化パラメータの値によって変化するため, 正則化パラメータの最適化が重要な問題となる. その最適化法の1つとしてモデル選択規準最小化法がある. GR の場合, Nagai *et al.* (2012) などによる GC_p 規準最小化法や, Yanagihara (2013) などによる GCV 規準最小化法を用いた正則化パラメータの最適化法が提案されており, これらの手法によって最適化された正則化パラメータは ∞ を含む形で得られる. GR-推定値の形から $\hat{\theta}_j = \infty$ のとき, $\hat{\beta}_{\boldsymbol{\theta},j}^{\text{GR}} = 0$ となることがわかる. すなわち, 正則化パラメータを最適化することで, GR でスパースな推定値を得ることができる.

本発表では, GC_p 規準や GCV 規準などの最小化法によって最適化された正則化パラメータの下での AL-推定値と GR-推定値が同等であることを示す. さらに, 多くのモデル選択規準は分散の推定量と一般化自由度を用いて構成されていることから, それらの2変数関数としてモデル選択規準を一般化し, その最小化法によって最適化された正則化パラメータの下でも上記の2つの推定値が同等であることを示す.

引用文献

- [1] Nagai, I. Yanagihara, H. & Satoh, K. (2012). Optimization of ridge parameters in multivariate generalized ridge regression by plug-in methods. *Hiroshima Math. J.*, **42**, 301–324.
- [2] Yanagihara, H. (2013). Explicit solution to the minimization problem of generalized cross-validation criterion for selecting ridge parameters in generalized ridge regression. *TR-No. 13-07*, Hiroshima Statistical Research Group, Hiroshima University.