2018N615@v@I[vnEX

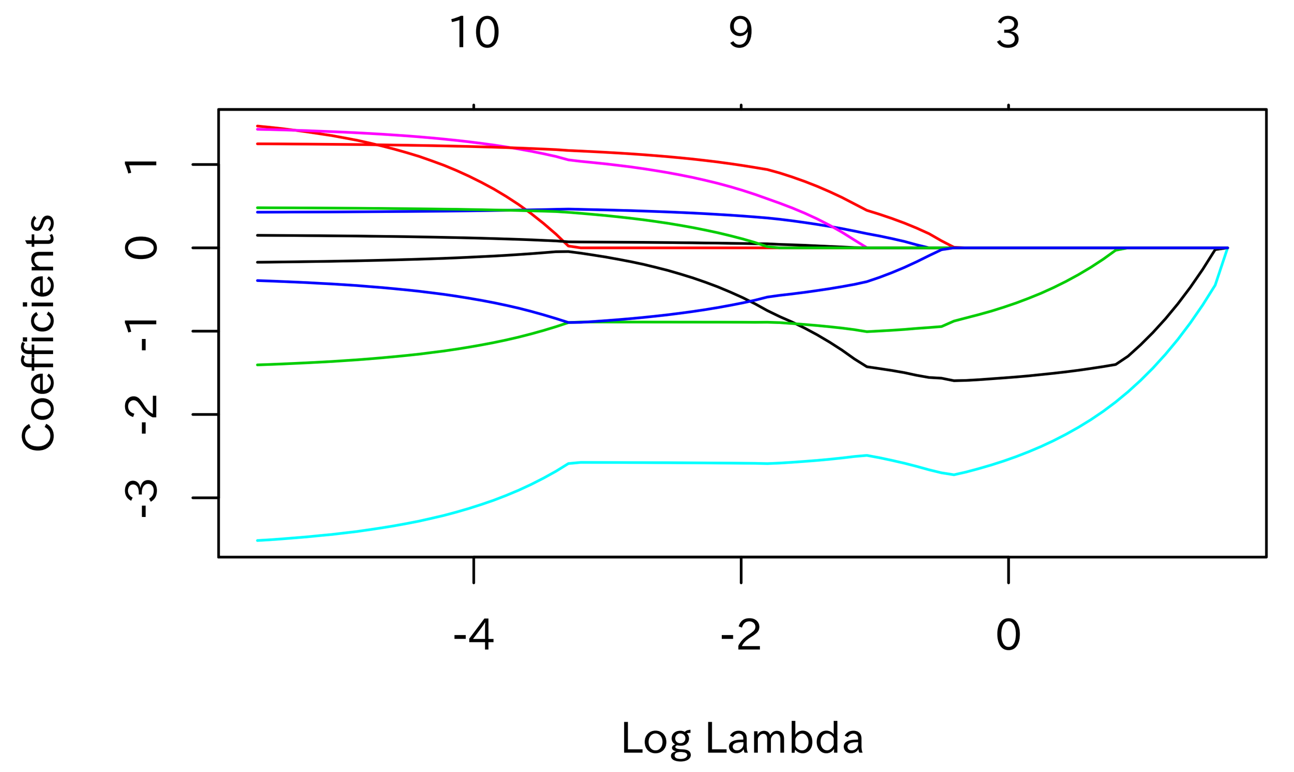
符号制約付き回帰分析によるスパース推定

田上 悠太　**リスク解析戦略研究センター 特任助教**

背景

- データの収集技術の向上によって高次元データ(マイクロアレイデータ、テキストマイニング、画像認識、ファイナンスなど)の分析に注目が集まっている。
- これまでに様々なスパース回帰の方法が提案されている (Lasso, SCAD, MCP など)。
- これらの方法はチューニングパラメータに依存している問題がある。
  - Lasso

$$\hat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}}(\|\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}\|_2^2 + \lambda\|\boldsymbol{\beta}\|_1)$$



- チューニングパラメータによって結果が左右されないようなスパース推定が望まれる。
- 本研究では、チューニングパラメータを含まないスパース推定の方法である符号制約付き回帰モデルの統計的な性質について研究した。

設定

確率空間$(\Omega, \mathcal{F}, P)$を考える。$\mathbf{F}$を実内積空間、$L(f,\omega) : \mathbf{F}\times\Omega \to \mathbb{R}$を損失関数とする。さらに、$L(f) = L(f,\cdot) \in L^1(P)$とする。

- (期待損失) $\bar{L}(f) = E[L(f)], \qquad f \in \mathbf{F}.$
- (最適な関数) $\bar{L}(f^0) = \min_{f\in\mathbf{F}} \bar{L}(f)$
- (超過ロス) $\mathcal{E}(f) = \bar{L}(f) - \bar{L}(f^0), \qquad f \in \mathbf{F}.$

モデルとして以下のf_βでf^0を推定する問題を考え、$(\psi_j \in \mathbf{F}, (j = 1, \cdots, p))$。

$$f_\beta := \sum_{j=1}^{p} \beta_j \psi_j$$

以下のような経験ロス$L(f_\beta)$を最小にする$\hat{\beta}$を考える。

$$L(f_{\hat{\beta}}) = \min_{\beta\succeq 0} L(f_\beta).$$

更に、$S \subset \{1, \cdots, p\}$でアクティブセットを表し、それを用いて以下の$\beta_{j,S}$を定義する。

$$\beta_{j,S} = \beta_j 1_{\{j\in S\}}, j = 1, \ldots, p \qquad \beta_S = (\beta_{1,S}, \cdots, \beta_{p,S}).$$

また$\Sigma = (\Sigma^{ij})_{1\le i,j\le p} = (\langle \psi_i, \psi_j \rangle)_{1\le i,j\le p}$で、以下のcompatibility constantを定義し、

$$\phi^2(L,S) = \min\{|S|\beta^\top \Sigma \beta : \|\beta_S\|_1 = 1, \|\beta_{S^c}\| \le L\}$$

Σのnon-negatively constrainde minimal ℓ_1-eigenvalue

$$\phi_{pos}^2(\Sigma) = \min\{\beta^\top \Sigma \beta : \|\beta\|_1 = 1, \beta \succeq 0\}.$$

を定義する。さらに、

- $\beta^* = \arg\min_{\beta\succeq 0} \bar{L}(f_\beta)$
- $v(\beta) = L(f_\beta) - \bar{L}(f_\beta)$
- $\mathbf{Z}_M = \sup_{\beta:\|\beta-\beta^*\|_1\le M} |v(\beta) - v(\beta^*)|$
- $f^* = f_{\beta^*}$, $\hat{f} = f_{\hat{\beta}}$, $S_* = \{j : \beta_j^* \neq 0\}$, $s_* = |S_*|$.

とする。

主な結果

仮定1(Convexity) $\beta \mapsto L(f_\beta, \omega)$は任意の$\omega \in \Omega$に対して、凸である。

仮定2(Quadratic margin) 以下を満たす$c > 0$、$\mathbf{F}_{\text{local}} \subset \mathbf{F}$が存在する。

$$\forall f \in \mathbf{F}_{\text{local}},\ \mathcal{E}(f) \ge c\|f - f^0\|^2$$

定理

仮定1、2、さらに$\nu := \phi_{pos}^2(\Sigma) > 0$、$\phi_*^2 := \phi^2(3C/\nu, S_*) > 0$、ただし、$C := \max_{1\le j\le p} \|\psi_j\|^2$が満たされている時、$\lambda > 0$で

$$\varepsilon_\lambda = 4\mathcal{E}(f^*) + \frac{12}{c\phi_*^2}\left(1 + \frac{3C}{\nu}\right)^2 s_*\lambda^2, \qquad M_\lambda = \varepsilon_\lambda/\lambda.$$

とする。また、$f_\beta \in \mathbf{F}_{\text{local}}$で$\beta \in \mathbb{R}^p$が$\|\beta - \beta^*\|_1 \le M$を満たすとする。この時、$\{\mathbf{Z}_{M_\lambda} \le \varepsilon_\lambda\}$において、

$$\|\hat{\beta} - \beta^*\|_1 \le M_\lambda = \frac{4}{\lambda}\mathcal{E}(f^*) + \frac{12}{c\phi_*^2}\left(1 + \frac{3C}{\nu}\right)^2 s_*\lambda \tag{1}$$

さらに、

$$\mathcal{E}(\hat{f}) \le \frac{5}{4}\varepsilon_\lambda = 5\mathcal{E}(f^*) + \frac{15}{c\phi_*^2}\left(1 + \frac{3C}{\nu}\right)^2 s_*\lambda^2 \tag{2}$$

数値実験

- Toeplitz design $(\Sigma_{kk'} = \rho^{|k-k'|/p}, \rho \in (0,1))$
- ロジスティック回帰
  - $P(y = 1|\boldsymbol{x}) = \frac{1}{1+\exp(-\boldsymbol{x}^\top\beta)}$
  - $y \in \{0, 1\}$
  - $\boldsymbol{x} \sim N_p(0, \Sigma)$
  - βはs個$(s \in \{3, 10, 20\})$が1で他は0とする。
  - $p \in \{20, 50, 100, 200, 500, 1000\}$.
  - $n = 100$