Basis Pursuit に基づく非線形回帰モデリングと変数選択

北大・情報・院　島村 徹平
北大・情報　水田 正弘

はじめに　本報告では，非線形回帰モデルの一つである加法回帰モデルにおける変数選択問題を扱う．一般に加法モデルは，各々の説明変数に関する一次元の滑らかな関数の線形和で表現され，滑らかな関数は既知の基底関数の線形結合で表現される．したがって，加法モデルにおいて，目的変数に関係する説明変数を選択することは，個々の基底関数ではなく，各説明変数に対応する基底関数群全体を選択することに相当する．そこで，LASSO（Tibshirani, 1996）と Basis Pursuit（以下，基底追跡と呼ぶ）（Chen *et. al.*, 1998）の考えを拡張することによって，加法回帰モデルにおける新しい変数選択法を提案する．

基底展開に基づく加法回帰モデル　p 次元の説明変数 $X = (X_1, \ldots, X_p)^{\mathrm{T}}$ と目的変数 Y に関して観測された n 組のデータ $\{(\boldsymbol{x}_i, y_i);\ i = 1, \ldots, n\}$ に対して，誤差項が正規分布に従う回帰モデル

$$y_i = \mu_i + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma^2), \quad i = 1, \ldots, n \tag{1}$$

を仮定する．ただし，μ_i はデータ点 $\boldsymbol{x}_i = (x_{i1}, \ldots, x_{ip})^{\mathrm{T}}$ を与えたときの y_i の条件付き期待値である．以下，y_i は中心化，$\boldsymbol{x}_i$ は標準化されているものとする．本報告では，μ_i の構造に対して，基底展開に基づく加法モデル

$$\mu_i = \eta(\boldsymbol{x}_i) = \sum_{j=1}^{p} \sum_{s=1}^{m_j} \delta_{js} b_{js}(x_{ij}) = \sum_{j=1}^{p} \boldsymbol{\delta}_j^{\mathrm{T}} \boldsymbol{b}_j(x_{ij}) \tag{2}$$

を仮定する．ただし，$\boldsymbol{b}_j(x_{ij}) = (b_{j1}(x_{ij}), \ldots, b_{jm_j}(x_{ij}))^{\mathrm{T}}$ は説明変数 X_j に対する m_j 次元の既知の基底関数ベクトル，$\boldsymbol{\delta}_j = (\delta_{j1}, \ldots, \delta_{jm_j})^{\mathrm{T}}$ は対応する m_j 次元の係数ベクトルである．(1) 式と (2) 式により，y_i の条件付き確率密度関数

$$f(y_i | \boldsymbol{x}_i; \boldsymbol{\delta}, \sigma^2) = \frac{1}{2\pi\sigma^2} \exp\left[-\frac{\{y_i - \sum_{j=1}^{p} \boldsymbol{\delta}_j^{\mathrm{T}} \boldsymbol{b}_j(x_{ij})\}^2}{2\sigma^2} \right], \quad i = 1, \ldots, n \tag{3}$$

を得る．ただし，$\boldsymbol{\delta} = (\boldsymbol{\delta}_1^{\mathrm{T}}, \ldots, \boldsymbol{\delta}_p^{\mathrm{T}})^{\mathrm{T}}$ である．

基底関数群を選択するための基底追跡　基底追跡法は，L_1 ノルムの罰則を用いることによって，基底関数に係る未知の係数ベクトルに関して，スパースな解（推定される要素のいくつかが完全に 0 になる解）を求めるパラメータ推定法である．この手法を用いることにより，(3) 式に含まれる未知パラメータ $\boldsymbol{\delta}$，σ^2 を推定すると同時に，基底関数 $b_{js}(X_j)$ $(j = 1, \ldots, p;\ s = 1, \ldots, m_j)$ を選択することができる．しかし，Y に関係する説明変数 X_j を選択するためには，対応する基底関数ベクトル $\boldsymbol{b}_j(X_j)$ $(j = 1, \ldots, p)$ をまとめて選択する必要がある．そこで，$\boldsymbol{b}_j(X_j)$ を選択するために，$\boldsymbol{\delta}$，σ^2 を以下の罰則付き対数尤度関数を最大化することで推定する手法を提案する．

$$l_\lambda(\boldsymbol{\delta}, \sigma^2) = \sum_{i=1}^{n} \log f(y_i | \boldsymbol{x}_i; \boldsymbol{\delta}, \sigma^2) - n\lambda \sum_{j=1}^{p} \|\boldsymbol{\delta}_j\|_{K_j} \tag{4}$$

ただし，$\lambda \geq 0$ は平滑化パラメータ，$\boldsymbol{K}_j$ は説明変数 X_j に対する既知の $m_j \times m_j$ 非負定値対称行列とする．また，$\|\boldsymbol{\delta}_j\|_{K_j} = (\boldsymbol{\delta}_j^{\mathrm{T}} \boldsymbol{K}_j \boldsymbol{\delta}_j)^{1/2}$ である．

参考文献

Chen, S., Donoho, D., and Saunders, M. (1998). Atomic Decomposition by Basis Pursuit, *SIAM Journal of the Scientific Computing*, 20, 33-61.

Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society*, Ser. B, 58, 267-288.