

# グラフ上の近単調回帰

東京大・情報理工 南 賢太郎 東京大・情報理工 / 理研・脳センター 駒木 文保

与えられた順序制約を満たすようにパラメータを推定する問題は単調回帰 (isotonic regression) と呼ばれ、長い歴史をもつ [1].

順序制約は、添字集合  $V = \{1, \dots, n\}$  を頂点とする有向非巡回グラフ (DAG) を用いて記述できる. すなわち、 $V$  上の任意の半順序  $\preceq$  に対して、有向グラフ  $G_{\preceq} = (V, E)$  が存在して、「 $i \preceq j \Rightarrow \theta_i \leq \theta_j$ 」という条件と、「すべての枝  $(i, j) \in E$  に対して  $\theta_i \leq \theta_j$ 」という条件が同値になる. 最も単純な例として、単調順序制約  $\theta_1 \leq \dots \leq \theta_n$  に対応する DAG は鎖状グラフ  $E = \{(i, i+1) \mid i = 1, \dots, n-1\}$  である.

$y \sim N(\theta, \sigma^2 I)$  を観測し、順序制約についての事前知識のもとに  $\theta$  を推定する問題を考える. 単調回帰における古典的な推定量は順序制約を満たす凸錐の上への射影推定量であり、最小化問題

$$\text{minimize } \|y - \vartheta\|_2 \quad \text{s.t. } \vartheta_i \leq \vartheta_j \text{ for all } (i, j) \in E$$

の解として与えられる. 一方、グラフ  $G$  が複雑かつ巨大であるような場合などには、一部の枝に対して局所的に制約を満たさないことを許容した推定量を考えたい. 本発表では、次の正則化型の推定量について考察する:

$$\hat{\theta}_\lambda \in \operatorname{argmin}_{\vartheta \in \mathbb{R}^n} \frac{1}{2} \|y - \vartheta\|_2^2 + \lambda \sum_{(i, j) \in E} (\vartheta_i - \vartheta_j)_+ \quad (1)$$

ここで、 $\lambda \geq 0$  は正則化パラメータであり、 $(z)_+ = \max\{0, z\}$  である. これは、Tibshirani, et al. [2] によって提案された近単調回帰 (Nearly-isotonic regression) を一般の DAG 上に拡張したものである. (1) は順序制約を乱す枝  $(i, j)$  に対して罰則を課す形になっており、特に、十分大きな  $\lambda$  に対しては射影推定量と一致した解が得られる. また、無向グラフに対する fused lasso [3] と比較すると、順序制約が満たされている枝では過度に縮小しないことを意図したものと見える. 実際、順序制約が局所的に正しい場合に、(1) は射影推定量と fused lasso に対して中間的な挙動を示すことが数値的にも確認される.

本発表では、真の  $\theta$  に対する局所誤特定の状態を考える. すなわち、 $G$  の連結成分からなる分割  $A_1, \dots, A_k$  が存在して、各  $A_l$  上に制限した順序制約が満たされているとする. この場合には、正則化パラメータ  $\lambda$  をデータから適切に決定することが重要である. しかし、グラフ構造の自然な分割が一般には明らかではないため、cross-validation などの手法を適用することはできない. 一方、例えば  $C_p$  型の規準

$$\lambda^* \in \operatorname{argmin}_{\lambda \geq 0} \frac{1}{n} \|y - \hat{\theta}_\lambda\|_2^2 + \frac{2\sigma^2}{n} \hat{\text{df}}(\hat{\theta}_\lambda)$$

を考えることができる. ここで  $\hat{\text{df}}(\hat{\theta}_\lambda)$  は  $\hat{\theta}_\lambda$  の自由度 (degrees of freedom) の不偏推定量であるが、具体的な表現の導出は自明ではない. 本発表ではまず、 $\hat{\text{df}}(\hat{\theta}_\lambda)$  がグラフ  $G$  に依存しない統一的な表現をもつことを、劣モジュラ関数の一般論によって示す. また、数値例を紹介しつつ、 $C_p$  型の規準によるパラメータ選択の性能について議論する.

## 参考文献

- [1] Robertson, T., Wright, F. T., and Dykstra, R. L. (1988) *Order Restricted Statistical Inference*. Wiley.
- [2] Tibshirani, R. J., Hoefling, H. and Tibshirani, R. (2011) Nearly-isotonic regression. *Technometrics*, 53, 54–61.
- [3] Tibshirani, R., Saunders, M., Rosset, S., Zhu, J., and Knight, K. (2005) Sparsity and smoothness via the fused lasso. *J. R. Statist. Soc. B*, 67, 91–108.