

関連ベクターマシンに基づく不連続な非線形回帰モデリングと多次元への拡張

中央大 松田 和己

中央大 小西 貞則

1 はじめに

非線形回帰モデリングは、データの背後にある複雑な非線形構造を解明する手法として多くの分野で用いられている。通常は、データの平均構造に対して滑らかな非線形構造を仮定するため、突発的な変化点が含まれるデータを扱う場合に構造を適切に捉えることができない。これに対して、不連続な基底関数を用いた関連ベクターマシン (RVM) によって変化点を推定し、その推定結果を利用して非線形構造をモデル化する手法が提案された (Tateishi and Konishi, 2011)。しかしながら、複数の変化点がある場合には、変化点を適切に推定することが困難となり、予測精度が低下することがわかった。

そこで本報告では、変化点の推定精度を向上するための基底関数を提案し、また、提案手法を多次元の問題設定へと拡張する。最後に、数値実験を通して提案手法の有効性を検証する。

2 関連ベクターマシンに基づく変化点推定

目的変数 Y と説明変数 x に関する n 組のデータ $\{(y_i, x_i); i = 1, 2, \dots, n\}$ に対して、不連続な基底関数

$$\phi_j(x) = \begin{cases} \exp\left\{-\frac{(x-x_j)^2}{2h^2}\right\}, & (x < x_j), \\ -\exp\left\{-\frac{(x-x_j)^2}{2h^2}\right\}, & (x \geq x_j), \end{cases}$$

を用いた RVM 回帰によって変化点の位置を推定する (図 1)。ここで、 h^2 は基底関数の拡がりの程度を調節するパラメータである。

次に、この推定結果に基づき、変化点推定に用いた基底関数に加えて、ガウスカーネルを融合した RVM 回帰を適用することで、不連続点を内包する現象を捉える非線形回帰モデリングを構築する。

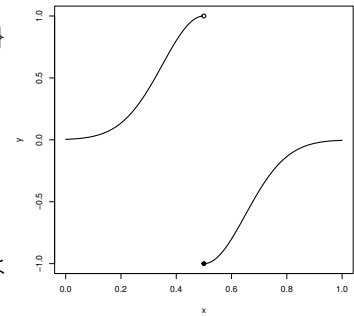


図 1: 不連続な基底関数

3 多次元への拡張

説明変数 $\mathbf{x}$ が p 次元の場合に、曲面の急激な変化を捉える問題を考える。このため、前述の基底関数を次のように拡張する (図 2)。

$$\phi_j(\mathbf{x}) = \begin{cases} \exp\left\{-\frac{\|\mathbf{x}-\mathbf{x}_j\|^2}{2h^2}\right\}, & (\mathbf{x} \in T), \\ -\exp\left\{-\frac{\|\mathbf{x}-\mathbf{x}_j\|^2}{2h^2}\right\}, & (\mathbf{x} \notin T). \end{cases}$$

ただし、 $T = \{\mathbf{x} \in \mathbb{R}^p \mid \mathbf{c}^T(\mathbf{x}-\mathbf{x}_j) < 0, \|\mathbf{c}\| = 1\}$ で、 $\mathbf{c}$ は不連続性が生じている超平面の向きを規定するパラメータである。

最後に、数値実験と実データ解析を通して提案手法の有効性を検証する。詳細については当日報告する。

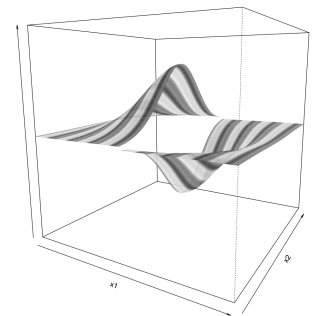


図 2: $p = 2$ の場合

参考文献

- [1] Tateishi, S. and Konishi, S. (2011). Nonlinear regression modeling and detecting change points via the relevance vector machine. *Computational Statistics* **26-3**, 477–490.
- [2] Tipping, M. E. (2001). Sparse Bayesian learning and the relevance vector machine. *Journal of Machine Learning Research*, **1**, 211–244.