

# ロバスト関連ベクター回帰モデルにおける予測情報量規準

中央大 松田 和己

電気通信大 川野 秀一

中央大 小西 貞則

## 1 はじめに

関連ベクター回帰 (RVR; Tipping, 2001) は、ベイズアプローチに基づく非線形モデリング手法として広く用いられている。しかしながら、RVR モデルは、誤差項に対して等分散性を仮定しているため、学習データに外れ値が含まれる場合に汎化能力が低下することが知られている。これに対して、等分散性を仮定しないロバスト関連ベクター回帰 (RRVR) が提案された (Han and Zhao, 2010)。RRVR モデルは、ベイズアプローチによって推定するが、カーネル関数に含まれる調整パラメータの値が汎化能力に強く影響する。

本報告では、RRVR モデルを評価するための予測情報量規準 (Kitagawa, 1997) を組み込んだロバストモデリングを提案する。また、提案手法の有効性を数値実験を通して検証する。

## 2 ロバスト関連ベクター回帰モデル

目的変数  $Y$  と  $p$  次元説明変数  $\mathbf{x} = (x_1, \dots, x_p)^T$  に関する  $n$  組のデータ  $\{(y_i, \mathbf{x}_i); i = 1, \dots, n\}$  に対して、ガウス型カーネル関数を用いた RRVR モデルは次式で与えられる。

$$y_i = w_0 + \sum_{j=1}^n w_j \exp\left\{-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2h^2}\right\} + \varepsilon_i, \quad i = 1, 2, \dots, n.$$

ここで、 $w_j$  ( $j = 0, \dots, n$ ) は回帰係数パラメータ、 $h^2$  はカーネル関数の拡がり調節するパラメータである。このとき、外れ値に対してロバストなモデルを実現するため、誤差項  $\varepsilon_i$  はそれぞれ平均 0、分散  $\sigma^2/\beta_i$  の正規分布に従うと仮定する。

## 3 モデル評価基準

図 1 に示すように、RRVR モデルはカーネル関数に含まれる調整パラメータの値に強く依存し、その値を適切に決定する必要がある。この問題に対してベイズアプローチによるモデル評価基準である予測情報量規準を導出して、モデルの評価を通して調整パラメータの値を決定する方法を提案する。RRVR における調整パラメータを含む予測分布は、

$$h(\mathbf{z}|\mathbf{y}) = N(\mathbf{z}; \boldsymbol{\mu}_p, \boldsymbol{\Sigma}_p),$$

で与えられる。ここで、 $\mathbf{z}$  は  $\mathbf{y}$  とは独立に観測される将来のデータとし、 $\boldsymbol{\mu}_p = \Phi \hat{\boldsymbol{\mu}}$ 、 $\boldsymbol{\Sigma}_p = \hat{\sigma}^2 B^{-1} + \Phi \hat{\boldsymbol{\Sigma}} \Phi^T$ 、 $\hat{\boldsymbol{\mu}} = \hat{\sigma}^{-2} \hat{\boldsymbol{\Sigma}} \Phi^T \hat{B} \mathbf{y}$ 、 $\hat{\boldsymbol{\Sigma}} = (\hat{A} + \hat{\sigma}^{-2} \Phi^T \hat{B} \Phi)^{-1}$ 、 $\hat{A} = \text{diag}(\hat{\alpha}_0, \dots, \hat{\alpha}_n)$  and  $\hat{B} = \text{diag}(\hat{\beta}_1, \dots, \hat{\beta}_n)$  である。

この予測分布を評価するための予測情報量規準の導出と数値実験による検証の詳細については当日報告する。

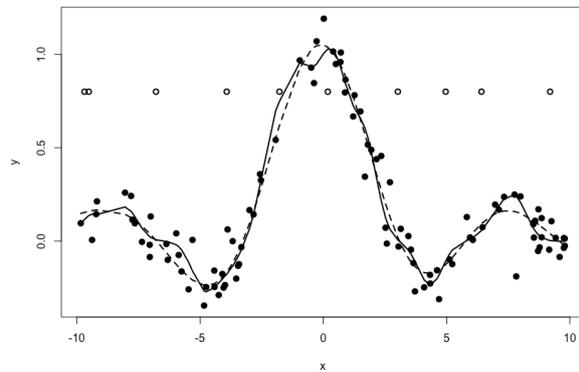


図 1: RRVR モデルによる推定曲線 (実線:  $h^2 = 0.5^2$ , 破線:  $h^2 = 2.5^2$ ).

## 参考文献

- [1] Han, M., and Zhao, Y. (2010). Robust relevance vector machine with noise variance coefficient. In: *Neural Networks (IJCNN), The 2010 International Joint Conference on*, (pp. 1-6). IEEE.
- [2] Kitagawa, G. (1997). Information criteria for the predictive evaluation of Bayesian models. *Communications in Statistics-Theory and Methods*, **26**, 2223-2246.
- [3] Tipping, M. E. (2001). Sparse Bayesian learning and the relevance vector machine. *Journal of Machine Learning Research*, **1**, 211-244.