G-002

ニューラルネットワーク回帰とグループラッソの組合せによる共非線形変数集合とその代表の発見

Discovery of Sets and Their Representatives of Co-nonlinear Variables by the Combination of Neural Network Regression and Group Lasso

佐々木 捷人 † 大崎 美穂 † 片桐 滋 †
Hayato Sasaki Miho Ohsaki Shigeru Katagiri

1. はじめに

入力変数を用いた出力の予測 (回帰・分類) と予測に寄与する入力変数の同定 (変数選択) は，科学とその応用に広く求められる．しかし，入力変数間に従属関係があると回帰・分類の性能と信頼性が低下する上に，変数選択を誤る恐れが生じる．従属関係を明らかにすればこれらの問題が解消され，従属関係自体も発見された新知識となり得る．ゆえに従属関係は問題視され，研究対象となってきた．

従来は線形な従属関係を仮定し (共線形性)，相関係数や分散拡大要因を測度としていた [1]．多入力変数から共線形関係を持つ集合とその代表を求める場合，これらの測度では組合せ爆発が生じる．また，非線形な従属関係 (共非線形性) を扱えないという根本的な問題がある．共非線形性の測度には従属関係に特定の関数を仮定したもの，区分変数空間の相互情報量の総和，カーネル法と相関の組合せ等がある [2]．これらにも組合せ爆発や，設定の難しさ，感度の低さの問題がある．

我々は従来の問題を解決して共非線形変数集合とその代表を発見することを目指し，以下の考えに基づく手法を提案して原理的な有効性を実験検証する．ニューラルネットワーク回帰 (NNR) [3] により，入力変数間の従属関係の非線形モデルを得る．組合せを探索せず，低寄与の変数の重みを 0 に収束させるグループラッソ (GL) [4] により変数選択する．異なる設定や初期値で得た集合と代表の情報を統合し，平均化効果により結果の再現性を確保する．

2. 提案手法

組合せ爆発を防ぎ，多様な共非線形関係を持つ入力変数集合と代表を導出すべく，提案手法では，まず全変数の 1 つを予測対象とし，残りの変数を予測子とする (図 1 の左参照)．関数族の仮定なく非線形関数を表現可能な NNR [3] を適用し，予測対象と予測子の共非線形関係をモデル化する．正則化により一部の予測子の重みを 0 にする GL [4] を NNR に組み込んでおき，モデル化と同時に，予測対象との共非線形性が強い予測子を選択する．

GL を組み込んだ NNR の目的関数を式 (1) に示す．第 1 項は NNR の残差平方和である．なお，j 番目の変数を予測対象とし，その s 番目の観測値を z_{js}，予測子を用いた推定値を $\hat{z}_{js}(\mathbf{W})$，訓練データ数を N_t とする．$\mathbf{W}$ は予測子の重みである．第 2 項は GL の正則化項である．j 番目の変数を予測対象とする NNR の第 1 層で，k 番目の予測子への重みをグループ化したものが $w_{jk}^{(1)}$，その L2 ノルムが $||w_{jk}^{(1)}||_2$ である．これを 1 から j を除く全予測子数 $N_{z_j}^{(1)}$ で総和して L1 ノルムを得ている．λ は 2 つの項のバランスをとるハイパーパラメータである．

$$J_j(\mathbf{W}) = \frac{1}{N_t}\sum_{s=1}^{N_t}(z_{js} - \hat{z}_{js}(\mathbf{W}))^2 + \lambda \sum_{k=1}^{N_{z_j}^{(1)}} ||w_{jk}^{(1)}||_2 \quad (1)$$

提案手法は予測対象 z_j ごとに重みの初期乱数も変え，z_j に対する共非線形変数集合候補を蓄積する (図 1 の中央参照)．最後に式 (2) を適用し，異なる予測対象と初期乱数で得た全結果を統合する (図 1 の右参照)．P パターンの候補 CS_i の和集合から出現頻度 $\mathrm{Freq}(z)$ が閾値 T_{Freq} 以上の変数 z を抽出し，その集合を最終的な共非線形変数集合と見なすとともに，代表は集合中の最頻出変数とする．この平均化効果が NNR と GL の初期値依存性を抑制し，提案手法の導出結果を安定させる．

$$CS = \left\{ z \mid z \in \bigcup_{i=1}^{P} CS_i,\ T_{\mathrm{Freq}} \le \mathrm{Freq}(z) \right\} \quad (2)$$

†同志社大学

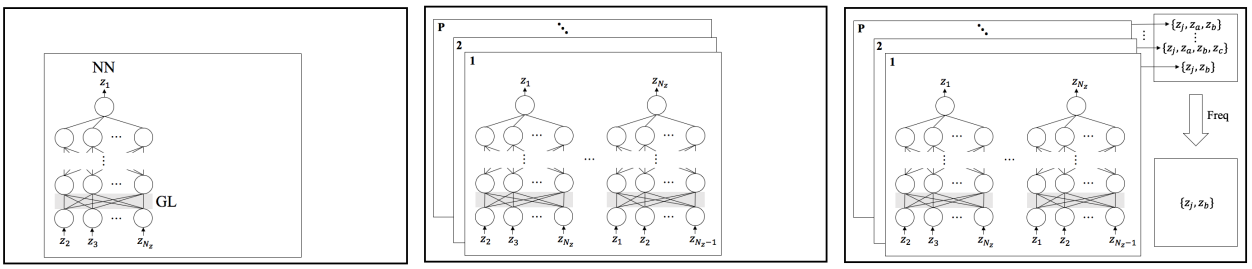

図 1: 提案手法の概念図.

3. 評価実験

提案手法を従属関係が既知の人工データに適用して出力を既知の従属関係と比較することで，提案手法の原理的な有効性を検証する．人工データに関しては文献 [2] を参考に，従属な 2 変数 x_1, x_2 と独立な 1 変数 x_3 を仮定した (表 1 の左と図 2 の上参照)．5 種類の従属関係うち，Line は線形，他は非線形 (Exp: 指数，Parab: 2 次，Cubic: 3 次，Sine: 正弦) の関数である．x_1 を等間隔に取り，関数に代入して x_2 の値を求めるとともに一様乱数で x_3 を生成し，3000 点から成る 5 種類の人工データを求めた．

ハイパーパラメータは次のように設定した．予備実験を行った結果，提案手法は NNR の層の深さと中間層のノードの数には敏感であり，GL の正則化項の重み，および，共非線形変数集合と代表の導出の閾値・要素数には頑健であった．ゆえに，重みと閾値・要素数は予備実験で得た定数に固定し，層の深さと中間層のノードの数のみ，本実験にて探索して設定した．

訓練から結果出力への一連の流れを以下に述べる．1 つの人工データを訓練用，検証用，試験用に三分割し，NNR と GL の組合せのパラメータ設定，NN の層の深さと中間層のノードの数の設定，共非線形変数集合と代表の導出に用いた．ただし，安定した結果が得られることを検証すべく，NNR と GL の組合せに異なる 3 種類の初期設定を与え，1 つの人工データあたり 3 種類の結果を得た．

提案手法が導出した集合と代表を表 1 の右に，学習した従属関係のグラフを図 2 の下に示す．初期設定 1 の結果を見ると，どの従属関係でも x_1 と x_2 を 1 つの集合に，x_3 を別の集合に正しく分割している．Line では同値である x_1 と x_2 の両方を，他の従属関係では x_1 から x_2 への生成過程を反映して x_1 を代表としており，これらも正しい．グラフからも従属関係のモデル化が成功したと分かる．初期設定 2，3 でも完全に同じ正解の集合と代表が得られた．よって，提案手法が高い再現性のもとに共非線形変数集合と代表を発見できることが確認された．

ここで，提案手法がデータの生成過程を反映できた理由を検討する．Sine を例に挙げると，$x_1 \to x_2$ の写像は 1 対 1 であるが，逆写像，すなわち $x_1 \leftarrow x_2$ の写像は 1 対多となる (図 2 参照)．NN と GL の組合せによる x_2 の予測性能は高いため，提案手法は集合 $\{x_1, x_2\}$ を作り出した上で，$x_1 \to x_2$ の写像元である x_1 を代表に推す．一方，1 対多写像の影響で x_1 の予測性能は低く，x_2 を代表に推すことはない．この仕組みが正しい集合と代表の導出を実現したと考えられる．

4. おわりに

本研究では，共非線形変数集合とその代表を発見する手法を，ニューラルネットワーク回帰と，グループラッソ，平均化効果に基づき実現した．そして，共非線形関係を模擬した人工データを用いた実験の結果，提案手法が共非線形変数集合とその代表を正確に導出できることが示された．今後は，従来の共非線形性の測度との比較や，出力予測への寄与の観点から提案手法の拡張を試みる．

参考文献

[1] T. W. Anderson, An Introduction to Multivariate Statistical Analysis, 3rd Edition, Wiley, New York (2003).

[2] D. N. Reshef et. al, Detecting Novel Associatoins in Large Data Sets, Science, vol.344, no.6062, pp.1518–1524 (2011).

[3] C. M. Bishop, Pattern Recognition and Machine Learning, Springer, Berlin (2006).

[4] M. Yuan and Y. Lin, Model Selection and Estimation in Regression with Grouped Variables, Journal of Royal Statistical Society B, vol.68, Part 1, pp.49–67 (2006).

表 1: 評価実験に用いた人工データ (左) と提案手法の適用結果 (右)．$\star$ 付きの変数は集合の代表．

従属関係	x_1：等間隔	x_2：x_1 代入	x_3：一様乱数
Line	$x_1 \in [0, 1]$	$x_2 = x_1$	$x_3 \in [0, 1]$
Exp	$x_1 \in [0, 10]$	$x_2 = 10^{x_1}$	$x_3 \in [0, 10]$
Parab	$x_1 \in [-0.5, 0.5]$	$x_2 = 4x_1^2$	$x_3 \in [-0.5, 0.5]$
Cubic	$x_1 \in [-1.3, 1.1]$	$x_2 = 4x_1^3 + x_1^2 - 4x_1$	$x_3 \in [-1.3, 1.1]$
Sine	$x_1 \in [0, 1]$	$x_2 = sin(8\pi x_1)$	$x_3 \in [0, 1]$

初期設定 1	初期設定 2	初期設定 3
$\{\star x_1, \star x_2\}, \{\star x_3\}$	$\{\star x_1, \star x_2\}, \{\star x_3\}$	$\{\star x_1, \star x_2\}, \{\star x_3\}$
$\{\star x_1, x_2\}, \{\star x_3\}$	$\{\star x_1, x_2\}, \{\star x_3\}$	$\{\star x_1, x_2\}, \{\star x_3\}$
$\{\star x_1, x_2\}, \{\star x_3\}$	$\{\star x_1, x_2\}, \{\star x_3\}$	$\{\star x_1, x_2\}, \{\star x_3\}$
$\{\star x_1, x_2\}, \{\star x_3\}$	$\{\star x_1, x_2\}, \{\star x_3\}$	$\{\star x_1, x_2\}, \{\star x_3\}$
$\{\star x_1, x_2\}, \{\star x_3\}$	$\{\star x_1, x_2\}, \{\star x_3\}$	$\{\star x_1, x_2\}, \{\star x_3\}$

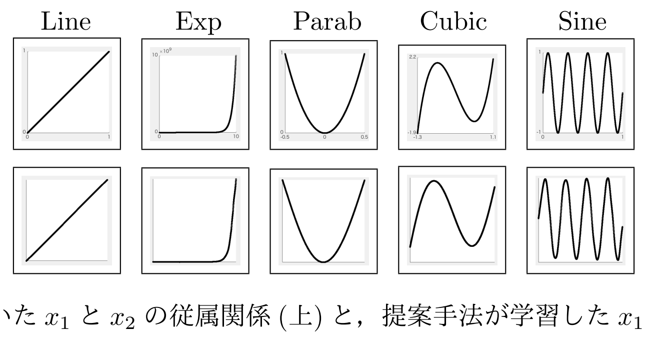

図 2: データ生成に用いた x_1 と x_2 の従属関係 (上) と，提案手法が学習した x_1 と x_2 の従属関係 (下)．