

C-01

# ファジィ回帰転移学習における異なる入力変数間マッピングの最適化 Optimization of mappings between different input variables in fuzzy regression transfer learning

清水 歩武† Shimizu Ayumu 謝 孟春† Xie Mengchun 中嶋 崇喜† Nakashima Mitsuki 森 徹† Mori Toru

## 1. はじめに

人工知能には様々な形態があるが、現在広く活用と研究がされているものは機械学習である。その中でも、問題を効果的かつ効率的に解くために、別の関連した問題のデータや学習結果を再利用する転移学習は注目されている<sup>1)</sup>。

転移学習はターゲットドメインの訓練データ数が予測モデルを効果的に学習するために十分ではない場合、ソースドメインの補助データを活用することによって学習を行うことができる<sup>2)</sup>。

近年、転移学習の研究は活発に行われているが、回帰問題における入力空間の異なる転移学習は研究が少なく、その適用は困難であるため、入力変数を削減する従来の転移学習を用いることもある。

本研究ではファジィ回帰転移学習を用いて、ターゲットドメインとソースドメインにおける入力変数が異なる場合のマッピングを行うことで、入力変数を最大限活用し、回帰転移学習の更なる高精度化を目指す。

## 2. 理論

### 2.1 転移学習

転移学習とは新規タスクの効果的な仮説を効率的に見つけ出すために、一つ以上の別のタスクで学習された知識を得て、それを適用する問題である。新規タスクを解くための知識の受け手側をターゲットドメイン、知識の送り手側をソースドメインと呼ぶ<sup>1)</sup>。

### 2.2 高木菅野ファジィ回帰モデル

高木菅野ファジィ回帰モデルは回帰予測モデルの 1 つである。高木菅野ファジィ回帰モデルは式(1)で示すような任意のクラスタ数  $c$  個のファジィルールに従って予測を行う<sup>2)</sup>。

$$\text{if } x \text{ is } A_i(x, v_i) \text{ then } y \text{ is } L_i(x, a_i) \quad i = 1, 2, \dots, c \quad (1)$$

ここで、 $x$  は入力データ、 $v_i$  はクラスタ重心、 $A_i$  はクラスタへの帰属度、 $L_i$  は線形関数、 $y$  は予測モデルの出力、 $a_i$  は線形関数の係数である。最終的な高木菅野ファジィ回帰モデルの出力は  $c$  個のファジィルールの集計であり、式(2)で示される。

$$y = \sum_{i=1}^c A_i(x, v_i) L_i(x, a_i) \quad (2)$$

### 2.3 ファジィ回帰転移学習

ファジィ回帰転移学習とは高木-菅野ファジィ回帰モデルに転移学習を組み合わせたものである<sup>1)</sup>。ファジィ回帰転移学習モデルは図 1 の流れで予測を行う。

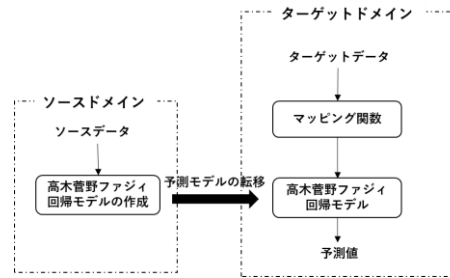


図 1 ファジィ回帰転移学習

まず、ソースデータを用いてソースドメインでのみ使用可能である高木-菅野ファジィ回帰モデルを作成する。次に、ターゲットドメインではターゲットデータをソースドメインのデータにマッピングし、最後にマッピングされたターゲットデータをソースドメインにおいて作成された予測モデルに入力することによって予測を行う。

### 2.4 マッピング関数

マッピング関数はターゲットデータの特徴量をソースデータの特徴量に近づける役割を持ち、1 層の隠れ層を持つニューラルネットによって構成される(図 2)。また、ニューラルネットの各パラメータは最適化アルゴリズムによって最適化される。

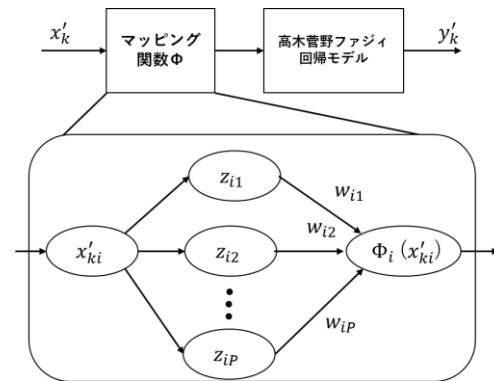


図 2 マッピング関数

マッピング関数の出力  $\Phi_i(x'_{ki})$  は式(3)で表され、活性化関数にシグモイド関数を使用し、その出力は式(4)で表される。

$$\Phi_i(x'_{ki}) = \sum_{i=0}^P w_{ip} z_{ip} \quad (3)$$

$$z_{ip}(x'_{ki}) = \frac{1}{1 + e^{-\alpha_{ip}(x'_{ki} - \beta_{ip})}}$$

$$p = 1, 2, \dots, P, \quad i = 1, 2, \dots, D, \quad 0 < \alpha_{ip} \quad (4)$$

ここで、 $p$  は任意のノード数、 $D$  は入力次元数である。

## 2.5 提案手法

従来のファジィ回帰転移学習ではターゲットドメインとソースドメインにおける同じ入力変数間で一对一のマッピングを行う。本研究では類似性が見られるが、異なる入力変数間で一对一のマッピングを行うことで、本来モデルの学習に使用することができない入力変数を利用しモデルの学習を行う。

また、本研究の提案手法との比較を行うため、以下のルールに基づいて4つのモデルの作成を行う。

- $Q_1$ : 転移学習を利用せずにモデルを作成
- $Q_2$ : 一致しない入力変数を削減してモデルを作成
- $Q_3$ : 異なる入力変数をマッピングしてモデルを作成
- $Q_4$ : 全ての入力変数をマッピングしてモデルを作成

## 2.6 評価方法

モデルの学習には最適化アルゴリズムが用いられるため、作成されたモデル毎に性能のばらつきがあると考えられる。手法の汎化性能を評価するため10回の交差検証を行った(図3)。ターゲットデータを無作為に10のセグメントに分割し、9つのセグメントを学習データ、1つをテストデータとして実験を行った。各モデルの評価には平均二乗誤差を使用した。

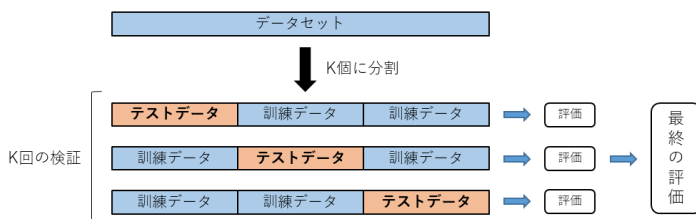


図3 交差検証

## 3. 実験

### 3.1 データセット

実験を行うためのデータセットとして、USAの不動産情報をスクレイピングして作られた住宅価格の実データセットを使用した<sup>3)</sup>。出力変数を住宅価格とし、以下の説明変数を考慮した。説明変数 $x_1$ は寝室数、 $x_2$ はバスルーム数、 $x_3$ はリビング面積、 $x_4$ は地下室面積、 $x_5$ は屋根裏部屋面積、敷地面積、 $x_6$ はフロア数、 $x_7$ は建築年。シアトルの1560件のデータをソースデータ、ベルビューの280件のデータをターゲットデータとした。

### 3.2 実験内容

この実験では、ソースドメインとターゲットドメインにおける入力変数が異なることを表現するために、ソースドメインでは $x_2$ 、ターゲットドメインでは $x_1$ を持たないと仮定した。そして、ソースドメインの入力変数 $x_2$ と、ターゲットドメインの入力変数 $x_1$ をマッピングして予測モデルを作成し、作成された予測モデルの性能を評価を行った。各モデルの学習に使用された入力変数を表1に示す。

表1 各モデル学習データ

	ソースドメイン	ターゲットドメイン
$Q_1$	-	$x_2, x_3, x_4, x_5, x_6, x_7$
$Q_2$	$x_3, x_4, x_5, x_6, x_7$	$x_3, x_4, x_5, x_6, x_7$
$Q_3$	$x_1, x_3, x_4, x_5, x_6, x_7$	$x_2, x_3, x_4, x_5, x_6, x_7$
$Q_4$	$x_1, x_2, x_3, x_4, x_5, x_6, x_7$	$x_1, x_2, x_3, x_4, x_5, x_6, x_7$

また、マッピング関数の最適化には粒子群最適化法を使用した。実験のハイパーパラメータを表2に示す。

表2 ハイパーパラメータ

ファジィ回帰転移学習		粒子群最適化法	
ファジィ度	2.1	反復回数	300
クラスタ数	7	粒子数	1000
ノード数	10	慣性係数	0.9
		加速係数	0.6

### 3.3 実験結果

各モデルの予測性能評価を表3に示す。

表3 実験結果

	MSE (平均±標準偏差) [10 <sup>11</sup> ]
$Q_1$	1.48 ± 0.91
$Q_2$	11.52 ± 24.71
$Q_3$	1.31 ± 0.99
$Q_4$	1.01 ± 0.67

$Q_2$ の一致しない入力変数を削減して作成したモデルの平均二乗誤差の平均が大きく、標準偏差がさらに大きいことから、いくつかのモデルでは外れ値に対応できなかったことが分かる。 $Q_3$ の提案手法のMSEの平均値が $Q_1, Q_2$ より小さいことから、提案手法は、転移学習を利用しない手法と一致しない入力変数を削減する手法と比べて、予測精度が向上したことが分かった。また、 $Q_4$ の全ての入力変数を利用して作成したモデルのMSEの平均値が1.01、 $Q_3$ によるモデルのMSEの平均値が1.31と大きく差が開いているため、異なる変数間と同じ変数間でのマッピング性能に大きな差があると考えられる。

## 4. まとめ

本研究では回帰転移学習においてターゲットドメインとソースドメインの入力変数が異なる場合に、異なる入力変数間のマッピングを行う、新たなモデルの作成手法を提案した。

提案手法は一致しない入力変数を削減する手法や、転移学習を使用しない手法と比べ、作成された予測モデルの精度が向上した。しかし、全ての入力変数を利用した学習手法と比べて予測精度に大きな差があることから、異なる入力変数間のマッピングは完全には行われていないことが分かった。

今回の実験では異なる入力変数に類似性が大きい入力変数を使用した。類似性が小さい入力変数間でのマッピングでも正の転移を行うことができるか確認する必要がある。また、この実験ではターゲットドメインとソースドメインの入力変数の数が同じ場合の転移学習を試みたが、ターゲットドメインとソースドメインで入力変数の数が異なる場合の回帰転移学習を扱うことが今後の課題である。

## 参考文献

- 1) 神島 敏弘 転移学習 人工知能学会誌 2010, vol25, 4
- 2) Hua Zuo, et al., Fuzzy Regression Transfer Learning in Takagi-Sugeno Fuzzy Models, IEEE Transactions on Fuzzy System vol.25 no.6(2016)
- 3) Kaggle : <https://www.kaggle.com/datasets/gpandi007/usa-housing-dataset>