

# 遺伝的ファジィルール選択による動向情報の可視化

能島 裕介 石渕 久生

大阪府立大学大学院工学研究科

## 1. はじめに

一般に知識の精度と複雑性の間にはトレードオフの関係が存在する。たとえば、If-then ルールの形式で知識を獲得した場合、条件部を細かく設定するほど、複雑な表現ができる。逆に、条件部を荒く設定することで抽象的な知識となりうる。

本研究では、統計データから動向情報をファジィ If-then ルール形式で知識として獲得するために、遺伝的ファジィルール選択を適用する[1]。本手法は、二段階の手順で実行される。第一段階として、データ集合から候補となるファジィルールを抽出する。第二段階では、抽出した候補ルールの組み合わせを進化型多目的最適化手法を用いて最適化する。進化型多目的最適化手法は、複数の目的を同時に最適化するために、進化型計算における解の評価に優越関係を導入した手法である。本稿では、多目的知識獲得問題を関数近似問題として定式化し、数値実験により有効性を検討する。

## 2. 遺伝的ファジィルール選択

### 2.1 問題の定式化

本研究では、図1に示すような2次元データを扱う。図1は、1980年から2000年までの地価公示価格である[2]。各都市や主要三大都市の公示価格がいつ頃のくらい変化したかを、次式で示すファジィ If-then ルール形式で知識として獲得する。

$$\text{Rule } R_q: \text{ If } x_1 \text{ is } A_{q1} \text{ and } x_2 \text{ is } A_{q2} \text{ then } y \text{ is } b_q \quad (1)$$
$$q = 1, 2, \dots, N$$

都市に関しては区間集合を用い、時間軸には図2に示す三角型ファジィ集合を用いて条件部を構成する。結論部は、公示価格変動率を実数で表す。具体的には、次式に示す重み付け平均の値とする。

$$b_q = \sum \{ \mu_q(x) \cdot T \} / \sum \mu_q(x) \quad (2)$$

$\mu$ を条件部の適合度、 $T$ を $\mathbf{x}$ のときの真の公示価格変動率とする。本研究では、条件部の全ての組み合わせに対するルールを図1のデータから抽出する。次に、抽出したルールの組み合わせ最適化を行う。

### 2.2 進化型多目的最適化手法の適用

候補ルールの総数を $N$ 個とする。この $N$ 個の候補ルールの部分集合 $S$ を関数とみなし、進化型多目的最適化における個体として扱う。

$$S = s_1 s_2 \dots s_j \dots s_N \quad (3)$$

$s_j = 1$ は $j$ 番目の候補ルールが関数を構成するルール集合 $S$ に含まれることを表し、 $s_j = 0$ は $S$ に含まれないことを表す。各ルール集合 $S$ は、以下の2目的に関して評

価され最適化される。

$$\text{Minimize } f_1(S) \text{ and } f_2(S) \quad (4)$$

ここで、 $f_1(S)$ はルール集合 $S$ の出力 $y$ と $T$ の二乗誤差、 $f_2(S)$ は $S$ に含まれるファジィルールの数である。

本研究では、進化型多目的最適化手法として、探索性能が高く、比較の実装し易いNSGA-II [3]を適用する。NSGA-IIでは、解の選択や世代更新に、解の優越関係によるランキングと解の類似度による混雑度により解を評価する。

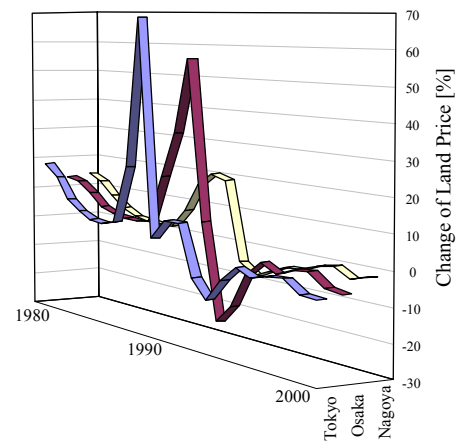


図1 地価公示価格変動率の推移

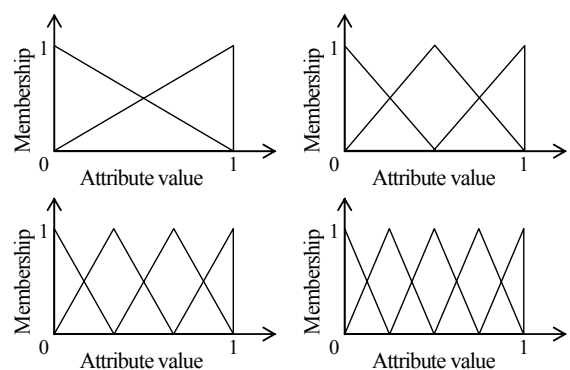


図2 2から5分割のファジィ集合

## 3. 数値実験

本実験では、条件部における都市に対する集合を{東京, 大阪, 名古屋, 三大都市平均}の4つとし、時間軸に対しては2から12分割のファジィ集合(77個)を適用した。また、時間軸を考慮しないdon't careも使用した。つまり、312個(4x78)のルールを抽出した。図3に、抽出したルール全てを用いて、公示価格変動率の推移を近似した結果を示す。図1と比較すると変化の仕方に関

しては類似しているが、312 個のルールで表現されているため大変複雑である。

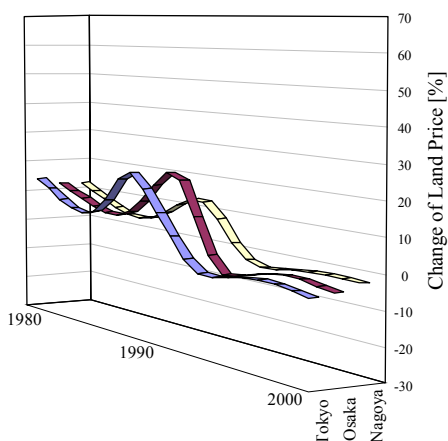


図3 地価公示価格変動率の近似 (全てのルール)

このルールの組み合わせ最適化に用いた NSGA-II の各種パラメータを以下に示す。

- 個体群サイズ: 100
- 交叉確率(一様交叉): 0.8
- 突然変異確率(ビット反転): 1/312
- 終了条件: 1000 世代

図4に進化型多目的最適化により得られた非劣解集合を示す。手法の1回の実行で精度と複雑性の異なるルール集合が獲得できていることがわかる。非劣解の中で、最も単純なルール集合(ルール数:5)と、ある程度複雑かつ精度の高いルール集合(ルール数:13)を図5と図6に示す。最も単純なルール集合は、三大都市平均とファジィ集合を組み合わせた5つのルールで構成されている。公示価格変動率を最も抽象的に表している。

図6に示すルール集合は、三大都市平均とファジィ集合の組み合わせで構成される6つのルールで大まかな動向を表現し、都市別の7つのルールで個々の都市の大きな変化を表している。図1と比較して、変動率の値域が異なるものの、都市ごとに異なる特徴も可視化できていることがわかる。

実際に動向情報を知りたい使用者から、知識の精度や複雑性を事前を知るできない、或いは、得られたものを評価することで、選好が変わる場合がある。このような問題に対し、進化型多目的最適化は複数のルール集合を提示することができるという利点がある。

#### 4. おわりに

本研究では、統計データから言語的な知識を If-then ルールの形式で獲得するために、数値データをファジィルールで近似する関数近似問題とみなし、遺伝的ファジィルール選択手法を適用した。数値実験により、精度と複雑性の異なる知識が獲得できることを示した。

今後の課題としては、得られたファジィルール集合に対応する新聞記事から、どのように言語化されているかを解析する予定である。

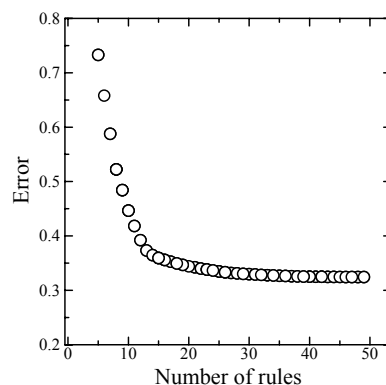


図4 知識の精度と複雑さのトレードオフ

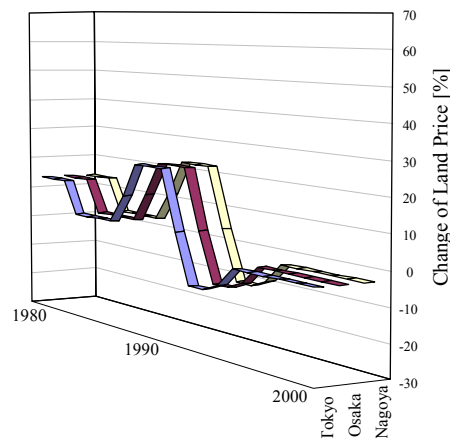


図5 地価公示価格変動率の近似 (ルール数: 5)

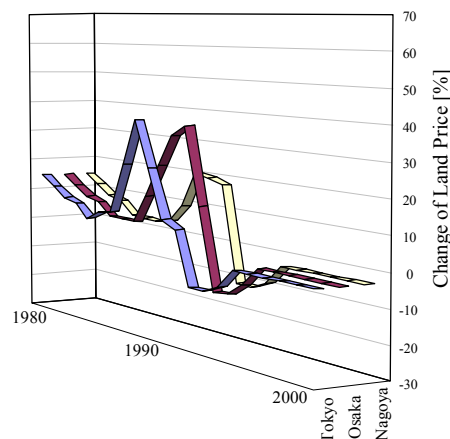


図6 地価公示価格変動率の近似 (ルール数: 13)

#### 参考文献

- [1] H. Ishibuchi, K. Nozaki, N. Yamamoto, and H. Tanaka, "Selecting fuzzy if-then rules for classification problems using genetic algorithms," *IEEE Trans. on Fuzzy Systems*, vol. 3, no. 3, August 1995, pp. 260-270.
- [2] 土地総合情報ライブラリー, <http://tochi.mlit.go.jp/>
- [3] K. Deb, A. Pratap, S. Agarwal, T. Meyarivan: A fast and elitist multiobjective genetic algorithm: NSGA-II, *IEEE Trans. on Evolutionary Computation*, Vol. 6, No. 2, pp. 182-197, 2002

能島 裕介  
 大阪府立大学 大学院工学研究科  
 〒599-8531 大阪府堺市中央区学園町1番1号  
 E-mail: nojima@cs.osakafu-u.ac.jp