

## GAによるファジイ推論ルール最適化とその応用

九州大学 譚 康融 TAN Kangrong 九州大学 \*時永 祥三 TOKINAGA Shozo

## 1 はじめに

本報告では、債券格付けを自動化システムにおいて [1], ファジイ推論を用いるメンバーシップ関数の形状を遺伝的アルゴリズム (Genetic Algorithm) を用いて最適化する方法を示し、日本で行われて格付けへ応用する。

## 2 ファジイルールによる推論

## 2.1 ファジイルールの出力

ファジイ推論は一般的に次のような *if...then...* ルールで記述される。

If  $x_1$  is  $A_{i1}$  and ... and  $x_m$  is  $A_{im}$  then  $y$  is  $w_i$  (1)

ここで  $x_j$  ( $j = 1, 2, \dots, m$ ) 入力変数であり,  $A_{ij}$  ( $i = 1, 2, \dots, n, j = 1, 2, \dots, m$ ) はファジイ集合である。更に,  $w_i$  はルールの付与された重みであり,  $n$  はファジイルールの総数である。

与えられた入力に対して, それぞれのルールのもつ適合度を次の式により計算する。

$$\mu_i = \prod_{j=1}^m \mu_{A_{ij}}(x_j) \quad (2)$$

ここで,  $\mu_{A_{ij}}$  はファジイ集合  $A_{ij}$  のメンバーシップ関数である。この値は, いわばそれぞれのルールが出力に寄与する割合と解釈される。

計算された適合度を, それぞれのルールの重みでウェイトづけして, ルール全体 (システムの全体) の出力を計算する。

$$y = \frac{\sum_{i=1}^n \mu_i w_i}{\sum_{i=1}^n \mu_i} \quad (3)$$

ここで,  $\mu_i$  はルールの適合度であり, 適合度をウェイト  $w_i$  で荷重平均したものが  $y$  が, ルールの出力となる。

## 2.2 ファジイルール最適化

最適化とは, ある企業  $p$  について格付け機関の与えたランク  $y^p$  (判別分析における外的基準) と, その企業の財務指標  $x_1^p, x_2^p, \dots, x_m^p$  の組を作成し, 財務指標を式 (1) に示すファジイ推論ルールに入れた場合の結論  $y_k^p$  が, できるだけ  $y^p$  に一致するようにする方法である。  $H_p = (y_k^p - y^p)^2 / 2$  の最小化である。式 (1) のファジイ推論ルールには, 次のような未知パラメータが含まれている。

- 1) メンバーシップ関数の数
- 2) メンバーシップ関数の形状
- 3) ルールのウェイト

1), 2) が与えている場合には,  $H_p$  が最小となる方向に数値  $w_i$  を調整していく。

上述した最適化は通常の最急降下法により行なうことができる [3]。

$$w_i(t+1) = w_i(t) - \alpha \frac{\mu_i^p}{\sum \mu_i^p} (y_k^p - y^p) \quad (4)$$

ただし,  $\alpha$  は収束を加速するための定数であり,  $t$  は繰り返し計算回数である。また,  $\mu_i$  は  $i$  番目のルールの適応度を表す。

## 3 メンバーシップ関数の最適化

## 3.1 メンバーシップ関数の形状

ルールの総数は入力変数の個数に対しファジイ集合の冪乗を取ったものとなるので, 数値を適当に選ん

でしまうと、ルール数は極めて大きなものとなる場合もある。以上のことを考慮して本報告では、ファジ集合の数を4に限定している。以下のように遺伝的アルゴリズムを用いてメンバーシップ関数の形状を最適化した。

1) 1入力変数に対して、メンバーシップ関数が4個であるために、10個の端点を決める必要があるが、後の推論では、3つの入力変数があるので、遺伝子のビット数は30である。

2) 1世代において、50個の個体がある。

3) 保存戦略としては、適応度、即ち、判別効率(正解率)が最も高いの20個を次世代に残す。

4) 残した20個間、及び新たな生成した個体から交叉、変異によって新個体30個を生成する。

以上のようなGAを用いてメンバーシップ関数の形状を最適化した。全部で11000世代を計算し、その結果として判別効率は79%に達した。

## 4 ファジ推論の応用

### 4.1 格付値と財務指標の計算

企業の財務状況を最も反映する19個の財務指標を計算した。分布の形状が著しく不均一である財務指標は、入力変数から除外した。その結果、9個の財務指標のみが入力変数として用いられている。なお、これら企業の選択に当たっては、最初、約140社の企業の財務諸表を計算し、入力変数となる9個の財務指標の分布の $5\sigma$ の外に分布する企業は除外した。

このように計算された財務指標を更に主成分分析して、3つの主成分を抽出して企業の指標として用いる。このようにもともと計算された財務指標ではなく主成分を用いる理由としては入力変数の数を少なくすることにより、生成されるルールの総数を制限することがある。3つの主成分を入力変数 $x_1, x_2, x_3$ として用いた。累積寄与率は約70%である。なお、主成分スコアは0から1まで正規化されている。

### 4.2 推論結果

最初に、学習に用いる企業のサンプルと、推論に用いる企業のサンプルを別にした場合の結果を示す。

表1は、No.1,2,3,4、それぞれ、電気機器39社、機械30社、複数にまたがる31社、およびこれらのすべてを対象とした場合のファジ推論による正しい判別

効率(正解率)を示している。表1では、学習に用いた企業のサンプル数と推論に用いたサンプル数による判別率の違いを示している。さらに正解率1、正解率2はそれぞれGAによる最適化される前、後の判別効率を示す。

これらの結果より分かるように、学習に用いる企業のサンプル数が増加するに従って、判別効率は増加する傾向が見られ、更に、業種を限定した認識システムのほうが、業種を混合したケースよりも良好な結果を与えている。

表1 各業種の正解率(%)

| NO | 学習数 | 推論数 | 正解率1 | 正解率2 |
|----|-----|-----|------|------|
| 1  | 20  | 19  | 68   | 74   |
| 2  | 20  | 10  | 60   | 70   |
| 3  | 20  | 11  | 55   | 64   |
| 4  | 70  | 30  | 60   | 67   |

表2 各業種の正解率(%)

| NO. | 学習数 | 推論数 | 正解率1 | 正解率2 |
|-----|-----|-----|------|------|
| 1   | 20  | 39  | 72   | 79   |
| 2   | 20  | 30  | 70   | 77   |
| 3   | 20  | 31  | 61   | 71   |
| 4   | 70  | 100 | 69   | 74   |

次に、学習に用いる企業のサンプルと、推論に用いる企業のサンプルが一部オーバーラップする場合の結果を表2に示す。

これらの結果より分かるように、学習に用いる企業のサンプル数をより多く推論サンプルに取り入れた方が、判別効率は増加する傾向が見られ、学習サンプルの数が多くなるほうが、判別効率も増加する。さらにGAによるメンバーシップ関数の最適化により、正解率が向上されていることが確認した。

### 参考文献

- [1] 譚 康融, 時永 祥三(1997): "ファジルールを用いた債券格付けの一手法", 経営情報学会秋期全国大会予稿集(1997)pp104-107
- [2] JBRI Ratings, No.2, April, 1997, Japan Bond Research Institute
- [3] 村上周太: "ファジ制御", 日本ファジ学会