

データの特性を活かした ファジィ回帰モデル同定法

—モデルに対するデータの帰属度を高める—

田島博之

1. はじめに

回帰分析についての研究は天文学による必要性から考案され、17 世紀後半から 18 世紀初頭にかけて Legendre や Gauss などの数学者によって始められたとされている。回帰分析における誤差は測定誤差として考えられるが、誤差の示す値そのものに実は人間にとって十分意味のある要因が含まれるとしてファジィ回帰モデルが提唱された [1]。ファジィ回帰モデル同定法はモデルが示すファジィ数の内部に、すべてのデータ点を包含するようにモデルの上下辺関数が決定され、モデルが示す中心関数は対称三角型ファジィ数の頂点を表している。

囲碁や将棋の段位、級位、レーティングポイントは棋士にとっての実力であり、棋力ともいう。棋力とは曖昧さを多分に含む数値で低位者が高位者に勝つことが一般的に起こりうる。一方、対局数や勝率、経験年数といった数値は曖昧さを含まない実数値である。棋力の低いものが高いものに勝つことを単に確率で片づけないのがファジィ理論である。上にあげた棋力のようなヒューリスティックで曖昧さを多分に含む値を解析する場合、ファジィ回帰モデル分析は有効な方法である。

田中らの研究では、全てのデータをモデルに包含することを制約条件として目的関数を定め、その問題を解くことによりモデルを同定している。これらの同定法をまとめた研究 [2] [3] には、モデルの幅を最小化する Minimum 問題、最大化する Maximum 問題、conjunction 問題などが紹介されており、今日のファジィ回帰モデル研究の基盤となっている。また、藝内らはファジィ回帰モデルを同定する際のデータセットに存在する特異データを除外する方法として超楕円関数を用いたファジィロバスト回帰分

析[4]を提案している。以上のモデル同定法は、何れも全てのデータを包含することを最重要と考えるためにモデル同定問題の目的関数に従属変数値を使わない。これはファジィ回帰モデルの研究が可能性測度を使った区間解析分析の研究[5]から発展しているところにその原因がある。

著者はこれまでにデータの従属変数値の変化に着目した幾つかのファジィ回帰モデル同定法を提案してきた。本論文では田中らが示した2つのMinimum問題を従来型モデル同定法として取り上げ、著者が提案した2つのモデル同定法[6][7]と比較しながら、著者独自のモデル同定法の有効性を紹介する。

2. 数学的準備

ファジィ数およびモデルに対する定義を以下に示す。従来型モデルではファジィ数を表現する場合、中心 a からの広がり幅を左右それぞれ e として対称三角型ファジィ数 $\tilde{A} = (a, e)$ を定義している。著者はこれを非対称三角型ファジィ数に拡張している。

ファジィ数 \tilde{A} を、その中心 a からの左(減少方向)への広がり幅を l 、右(増加方向)への広がり幅を r として $\tilde{A} = (a, l, r)$ のように表記する。

メンバーシップ関数を定義する。

$$\mu_{\tilde{A}}(x) = \begin{cases} \max\left(0, 1 - \frac{x-a}{r}\right) & \text{for } x \geq a \\ \max\left(0, 1 - \frac{a-x}{l}\right) & \text{for } x < a \end{cases} \quad (2.1)$$

このファジィ数 \tilde{A} の α レベル集合は閉区間となる。

$$[\tilde{A}]_{\alpha} = [(a, l, r)]_{\alpha} = [a - (1-\alpha)l, a + (1-\alpha)r] \quad (0 \leq \alpha \leq 1) \quad (2.2)$$

ファジィ数間の演算を定義する。

$$\tilde{A} + \tilde{B} = (a, s, t) + (b, u, v) = (a + b, s + u, t + v) \quad (2.3)$$

$$k \cdot \tilde{A} = k(a, s, t) = (k \cdot a, k \cdot s, k \cdot t) \quad (2.4)$$

ファジィ回帰モデルを定義する。

$$\begin{aligned} \tilde{Y}(\mathbf{x}) &= \tilde{A}_0 + \tilde{A}_1 x_1 + \cdots + \tilde{A}_n x_n \\ &= (a_0, l_0, r_0) + x_1 \cdot (a_1, l_1, r_1) + \cdots + x_n \cdot (a_n, l_n, r_n) = (a(\mathbf{x}), l(\mathbf{x}), r(\mathbf{x})) \end{aligned} \quad (2.5)$$

3. 従来型ファジィ回帰モデル同定法

実数を用いたデータに対する従来型のファジィ回帰モデル同定法は、与えられた全てのデータ点を、モデルに包括する制約条件下においてモデルの表現する曖昧さの幅を最小にする。以下に田中らの提唱した2つのタイプのモデル同定問題を従来型モデル同定法として紹介する。

分析すべきデータセット：

$$(\mathbf{x}_i; y_i) = ((x_{i1}, x_{i2}, \dots, x_{in}); y_i) \quad i = 1, 2, \dots, m \quad (3.1)$$

以上のデータから同定されるモデルの α カット集合が全てのデータを含むような制約条件を考える。

$$y_i \in [\tilde{Y}(\mathbf{x}_i)]_h, i = 1, 2, \dots, m \quad (3.2)$$

h の値は同定者が決定するモデルの示す曖昧さの幅のパラメータであり、 $0 \leq h < 1$ の範囲をとる。また、これを使い(3.4)～(3.6)の制約条件を設定し、ファジィ回帰モデルの曖昧さの幅を最小にする数理計画問題が定義される[3]。

従来型モデル同定問題 1

$$\text{目的関数: } \min \sum_{i=1}^m e(\mathbf{x}_i) \quad (3.3)$$

制約条件:

$$a(\mathbf{x}_i) - (1-h)e(\mathbf{x}_i) \leq y_i \quad i=1,2,\dots,m \quad (3.4)$$

$$a(\mathbf{x}_i) + (1-h)e(\mathbf{x}_i) \geq y_i \quad i=1,2,\dots,m \quad (3.5)$$

$$e_j > 0, \quad j=0,1,\dots,n \quad (3.6)$$

$$\text{決定関数: } a(\mathbf{x}), e(\mathbf{x}) \quad (3.7)$$

$a(\mathbf{x})$ は中心関数、 $e(\mathbf{x})$ は中心から曖昧さの幅を表す関数である。

本モデルよりも最小二乗法に近いという理由で、曖昧さを表す幅の関数を二乗にした最小化目的関数を用いた同定法も提案されている[8]。

従来型モデル同定問題 2

$$\text{目的関数: } \min \sum_{i=1}^m e(\mathbf{x}_i)^2 \quad (3.8)$$

制約条件:

$$a(\mathbf{x}_i) - (1-h)e(\mathbf{x}_i) \leq y_i \quad i=1,2,\dots,m \quad (3.9)$$

$$a(\mathbf{x}_i) + (1-h)e(\mathbf{x}_i) \geq y_i \quad i=1,2,\dots,m \quad (3.10)$$

$$e_j > 0, \quad j=0,1,\dots,n \quad (3.11)$$

$$\text{決定関数: } a(\mathbf{x}), e(\mathbf{x}) \quad (3.12)$$

3.1 従来型モデル同定法の問題点

独立変数、従属変数が1ずつ与えられた単回帰モデルを例に考える。従来型のモデル同定法では全てのデータを包括する凸多面体内部のデータの従属変数値が凸多面体内部において変化してもモデルの同定に影響を与えない。以下に単回帰モデルを例に説明する。Fig. 3.1は、データセット $\{d_i | i=1,2,\dots,6\}$ の各データ点を \times で表しており、データを挟む直線はデータセットから同定されたファジィ回帰モデルの上下辺関数とする。 $\{d_1, d_2, d_5, d_6\}$ はモデルを決定する上で重要な凸多角形を作るデータである。

この凸多角形内部にある2つのデータ点 $\{d_3, d_4\}$ を、凸多角形内部において \bullet で示される位置 $\{\hat{d}_3, \hat{d}_4\}$ まで従属変数値を変化させても、データセット $\{d_1, d_2, \hat{d}_3, \hat{d}_4, d_5, d_6\}$ から同定されるモデルに変化はない。

この問題は従来型モデル同定法の最小化目的関数に y_i が関与していないからである。ファジィ回帰モデルからファジィ数を推定するにはモデルを示す関数が、データの特徴を活かした形で同定されなければならないと考える。

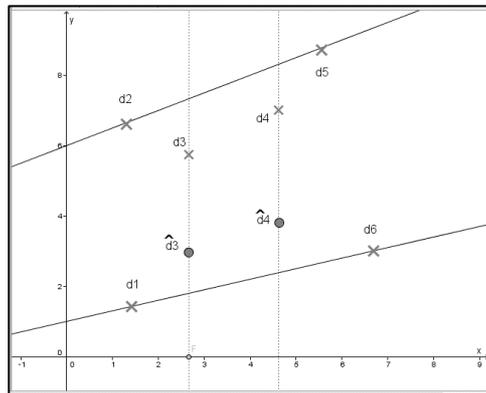


Fig.3.1 凸多角形内部データ概念図

4. 最小化基準に着目した同定法

本章では著者が提案したファジィ回帰モデル同定法を紹介する。このモデル同定法では最小二乗法的考察を導入し、モデルの上下辺関数を決定する。また、中心関数を同定する際にはデータの帰属度を用いる方法と最小二乗法を用いる2つの方法を提案している。

分析すべきデータセット:

$$(\mathbf{x}_i; y_i) = ((x_{i1}, x_{i2}, \dots, x_{in}); y_i) \quad i = 1, 2, \dots, m \quad (4.1)$$

次に最小化基準を定義する。以下に単回帰モデルを例に、その基本概念を説明する。モデルの中心関数を $a(x)$ 、上辺関数を $y(x)^*$ 、下辺関数を $y(x)_*$ と表す。

モデルを作成するためには従来型モデル同定法に使われるパラメータ h の値を固定する必要がある。このパラメータ h が1に近づくほど同定されるモデルの幅が広がることになる。本論文では従来型モデル同定法のパラメータを $h = \mathbf{0}$ として考え提案型モデルと比較する。従来型の上辺関数は (4.2)、下辺関数は (4.3) となる。著者の提案するモデルでは対称三角型ファジィ数から非対称三角型ファジィ数へと拡張するため、上辺関数が (4.4)、下辺関数が (4.5) となる。

$$y(x)^* = a(x) + e(x) \quad (4.2)$$

$$y(x)_* = a(x) - e(x) \quad (4.3)$$

$$y(x)^* = a(x) + r(x) \quad (4.4)$$

$$y(x)_* = a(x) - l(x) \quad (4.5)$$

Fig. 4.1にあるように、いずれのファジィ回帰モデルも与えられた全てのデータ点がモデルの上下辺関数に挟まれる制約条件の下に同定される。提案型モデルの上下辺関数はデータ点からモデルの上辺、また下辺までの距離を考え、それぞれの距離の二乗の総和を最小化目的関数とする問題を解くことにより同定される。

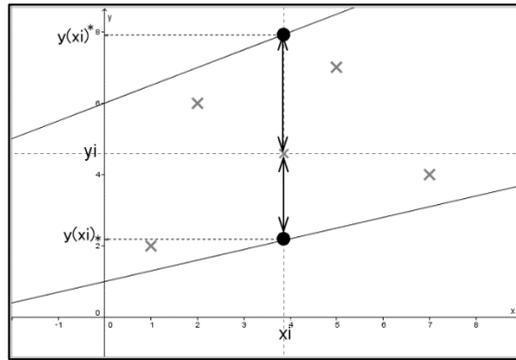


Fig.4.1 提案型モデル概念図

4.1 提案型モデル同定問題1

提案型モデル同定問題はモデルの上下辺関数 $y(\mathbf{x})^*$, $y(\mathbf{x})_*$ を決定した後、モデルの中心関数を求める2段階法となる[6][7]。

First Phase

目的関数1:

$$\min \sum_{i=1}^m \left[\{y(\mathbf{x}_i)^* - y_i\}^2 + \{y(\mathbf{x}_i)_* - y_i\}^2 \right] \quad (4.6)$$

制約条件式:

$$y(\mathbf{x}_i)_* \leq y_i \quad i = 1, 2, \dots, m \quad (4.7)$$

$$y(\mathbf{x}_i)^* \geq y_i \quad i = 1, 2, \dots, m \quad (4.8)$$

決定関数:

$$y(\mathbf{x})^*, y(\mathbf{x})_* \quad (4.9)$$

ここで、決定された、モデルの上下辺の関数を $\hat{y}(\mathbf{x})^*$, $\hat{y}(\mathbf{x})_*$ とする。

次に α_i 値を定義する。

$$\alpha_i = \begin{cases} \frac{y_i - \hat{y}(\mathbf{x}_i)_*}{a(\mathbf{x}_i) - \hat{y}(\mathbf{x}_i)_*} & \text{for } a(\mathbf{x}_i) \geq y_i \\ \frac{\hat{y}(\mathbf{x}_i)^* - y_i}{\hat{y}(\mathbf{x}_i)^* - a(\mathbf{x}_i)} & \text{for } a(\mathbf{x}_i) < y_i \end{cases} \quad (4.10)$$

α_i 値の総和から最小化目的関数を設定した中心関数同定問題を次に定義する。

Second Phase

$$\text{目的関数 2 : } \min \left(- \sum_{i=1}^m \alpha_i \right) \quad (4.11)$$

制約条件式 :

$$a(\mathbf{x}_i) \geq \hat{y}(\mathbf{x}_i)_* \quad i = 1, 2, \dots, m \quad (4.12)$$

$$a(\mathbf{x}_i) \leq \hat{y}(\mathbf{x}_i)^* \quad i = 1, 2, \dots, m \quad (4.13)$$

$$\text{決定関数 : } a(\mathbf{x}) \quad (4.14)$$

上の同定問題では上下辺関数を決定する際に最小二乗法的思考を導入しているが目的関数に y_i を用いる点で従来型同定法よりデータの特性を反映していると考えられる。また中心関数を求める際にモデルに対するデータの帰属度に着目した最小化基準を提案しているが、これは従来の同定法にはない独自のものである。

4.2 提案型モデル同定問題 2

ここでは提案型モデル同定問題 1 で求めた上下辺関数を使い、中心関数 $a(\mathbf{x})$ を同定する際に最小二乗法を用いる方法[6]を紹介する。

First Phase

$$\text{目的関数: } \min \sum_{i=1}^m \left[\{y(\mathbf{x}_i)_* - y_i\}^2 + \{y(\mathbf{x}_i)^* - y_i\}^2 \right] \quad (4.15)$$

制約条件:

$$y(\mathbf{x}_i)_* \leq y_i \quad i = 1, 2, \dots, m \quad (4.16)$$

$$y(\mathbf{x}_i)^* \geq y_i \quad i = 1, 2, \dots, m \quad (4.17)$$

$$\text{決定関数: } y(\mathbf{x})^*, y(\mathbf{x})_* \quad (4.18)$$

Second Phase

$$\text{目的関数: } \min \sum_{i=1}^m \{y_i - a(\mathbf{x}_i)\}^2 \quad (4.19)$$

$$\text{決定関数: } a(\mathbf{x}) \quad (4.20)$$

上のタイプの問題から同定されるファジィ回帰モデルは中心を示す関数に最小二乗法を使うため、分析されるデータの変化に対して中心関数が鋭敏に反応する。著者は非対称三角型ファジィ数を使う以前に、対称三角型ファジィ数を用いたモデル同定法の提案を行っている。これは提案型モデル同定問題2と同様に中心関数を最小二乗法によって同定するが、モデルが表現する曖昧さの幅として従属変数値の標準偏差値を用いる場合と、区間推定量を用いる場合の2つのタイプがある[9]。この2つのモデル同定法は本論文で紹介している提案型モデル同定法の基盤となっている。

5. 評価関数

著者は従来型と提案型のファジィ回帰モデルに対するデータの帰属度を比較するために、2つのタイプの評価関数を定義している[6]。

5.1 α レベル平均関数

ファジィ回帰モデルに対してデータセットが、どの程度の帰属度を持っているかを評価するために α レベル平均関数を以下に定義する。其々のモデル同定法によって決定された中心関数、上辺関数、下辺関数を下に示す。

$$\text{中心関数} \quad a(\mathbf{x}) \quad (5.1)$$

$$\text{上辺関数} \quad y(\mathbf{x})^* = a(\mathbf{x}) + r(\mathbf{x}) \quad \text{or} \quad a(\mathbf{x}) + e(\mathbf{x}) \quad (5.2)$$

$$\text{下辺関数} \quad y(\mathbf{x})_* = a(\mathbf{x}) - l(\mathbf{x}) \quad \text{or} \quad a(\mathbf{x}) - e(\mathbf{x}) \quad (5.3)$$

この時、それぞれのデータ点 (\mathbf{x}_i, y_i) における $\alpha(\mathbf{x}_i, y_i)$ 値を求める関数を定義する。

$$\alpha(\mathbf{x}_i, y_i) = \begin{cases} \frac{y_i - y(\mathbf{x}_i)_*}{a(\mathbf{x}_i) - y(\mathbf{x}_i)_*} & \text{for } a(\mathbf{x}_i) \geq y_i \\ \frac{y(\mathbf{x}_i)^* - y_i}{y(\mathbf{x}_i)^* - a(\mathbf{x}_i)} & \text{for } a(\mathbf{x}_i) < y_i \end{cases} \quad i = 1, 2, \dots, m \quad (5.4)$$

次に、 $\alpha(\mathbf{x}_i, y_i)$ 値の平均値を α レベル平均関数として定義する。

$$\bar{\alpha} = \frac{1}{m} \sum_{i=1}^m \alpha(\mathbf{x}_i, y_i) \quad (5.5)$$

Fig. 5.1 は、 \mathbf{x}_i に対してモデルが予測する $\tilde{Y}(\mathbf{x}_i)$ と、実数データ y_i を表した図である。

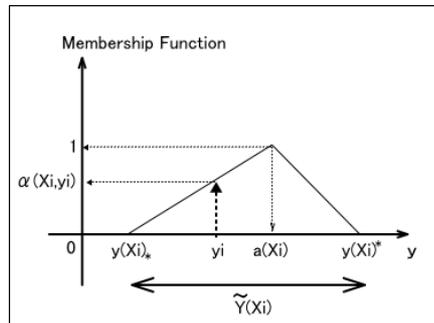


Fig. 5.1 モデル評価関数の概念図

$\alpha(\mathbf{x}_i, y_i)$ の平均値 $\bar{\alpha}$ が 1 に近いほどファジィ回帰モデルに対するデータの帰属度が高いと考えられる。

5.2 α レベル分散値

$\alpha(\mathbf{x}_i, y_i)$ における分散値を α レベル分散値 V_α として以下に定義する。

$$V_\alpha = \frac{1}{m} \sum_{i=1}^m \{\bar{\alpha} - \alpha(\mathbf{x}_i, y_i)\}^2 \quad (5.6)$$

α レベル平均値に近い値の場合、 α レベル分散値が低いモデルがデータに対して適合度が高いと判断する。

6. 数値計算による比較

次に単回帰モデルを使った数値計算例を 2 題紹介する。

ここでは、従来型モデル同定法 1、従来型モデル同定法 2 を prob. 1. 1、prob. 1. 2、また、提案型モデル同定法 1、提案型モデル同定法 2 を prob. 2. 1、prob. 2. 2 とする。

6.1 数値計算 1

数値計算 1 では各モデルの特徴を明確に示すために 2 つのタイプのデータセットを準備した。データセット A, B は平行四辺形の頂点を作る 4 つの点を固定し、辺の周辺にある 8 つの点を平行四辺形の上側、下側にそれぞれ寄せた 11 の点からなる。

Table 6.1 データセット A

data A	No. 1	No. 2	No. 3	No. 4	No. 5	No. 6	No. 7	No. 8	No. 9	No. 10	No. 11
x	0.00	10.0	20.0	30.0	40.0	50.0	60.0	70.0	90.0	110	130
y	1.00	2.00	3.00	4.00	5.00	6.00	7.00	2.00	7.43	7.71	8.00

Table 6.2 データセット B

data B	No. 1	No. 2	No. 3	No. 4	No. 5	No. 6	No. 7	No. 8	No. 9	No. 10	No. 11
x	0.00	10.0	20.0	30.0	40.0	50.0	60.0	70.0	90.0	110	130
y	1.00	1.14	1.29	1.43	1.57	1.71	7.00	2.00	4.00	6.00	8.00

それぞれのデータセットに対して得られた計算結果を以下の表に示す。

Table 6.3 データセット A による計算結果

計算結果 A	上辺関数		中心関数		下辺関数		Ave. α	Ver. α
	傾き	切片	傾き	切片	傾き	切片		
prob. 1. 1	0.100	1.00	0.0571	1.00	0.0143	1.00	0.318	0.183
prob. 1. 2	0.0290	5.26	0.0216	3.13	0.0143	1.00	0.333	0.0858
prob. 2. 1	0.0559	3.65	0.0714	1.00	0.0213	0.510	0.582	0.123
prob. 2. 2	0.0559	3.65	0.0504	2.04	0.0213	0.510	0.535	0.116

Table 6.4 データセット B による計算結果

計算結果 B	上辺関数		中心関数		下辺関数		Ave. α	Ver. α
	傾き	切片	傾き	切片	傾き	切片		
prob. 1. 1	0.100	1.00	0.0571	1.00	0.0143	1.00	0.281	0.155
prob. 1. 2	0.0290	5.26	0.0216	3.13	0.0143	1.00	0.162	0.0813
prob. 2. 1	0.0623	3.26	0.0362	-0.531	0.0362	-0.531	0.915	0.0121
prob. 2. 2	0.0623	3.26	0.0521	0.306	0.0362	-0.531	0.681	0.100

(A) 同一モデルにおける data A, B 間の相違

Fig. 6.1 は prob. 1.1 によって data A (●), data B (×) から得られたグラフである。データセットが異なっても同定されるモデルは同一になっていることがわかる。

データの変化にモデルが対応しきれていないのである。これは prob. 1.2 でも同様のことが言える。また、モデルの上下辺の関数に多くのデータ点が重なっていることで、モデルに対するデータの帰属度が低いことが見て取れる。

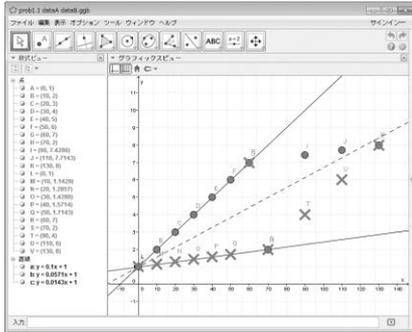


Fig.6.1 prob1.1: data A & data B

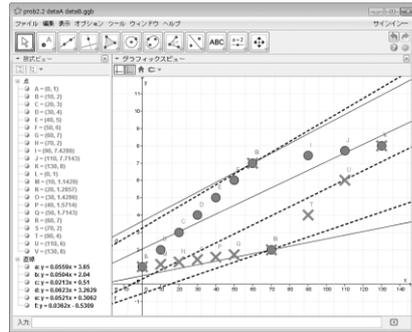


Fig.6.2 prob2.2: data A & data B

Fig. 6. 2 では prob2. 2 によって同定された関数をグラフで表している。data A から同定される関数を実線、data B から同定される関数を破線で表す。従来型と異なりデータの変化によって同定されるモデルが異なり、特に中心関数の変化には顕著に表れている。これは、prob2. 1 も同様であり提案型モデルはデータの変化に敏感に対応してモデルが同定されているといえる。

(B) data A におけるモデル間の相違

ここでは、data A を使った各モデルの違いをグラフと評価関数値によって比較する。以下にある Fig. 6. 3 は従来型 prob1. 1 と提案型 prob2. 1 によって得られた関数をグラフ化している。prob1. 1 が実線、prob2. 1 が破線である。prob1. 1 の α レベル平均値が 0. 318 に対し prob2. 1 では 0. 582 と prob2. 1 のほうが高くなっている。グラフからは、prob1. 1 ではデータ点がモデルの端に寄っているのに対し、prob2. 1 ではモデルの中心関数によっていることがわかる。また、Table. 6. 4 の data B に対する計算結果でも prob1. 1 の α レベル平均値が 0. 281、prob2. 1 では 0. 915 と prob2. 1 が高くなっている。Fig. 6. 4 は提案型の prob2. 1 と prob2. 2 を比較している。prob2. 1 による関数を実線で prob2. 2 による関数を破線で表す。両モデルの上下辺関数は同一関数となる。prob2. 1 の α レベル平均値は 0. 582、prob2. 2 では 0. 535 となっており、prob2. 1 が少し高い値を示している。

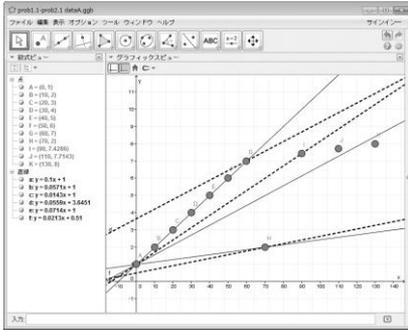


Fig.6.3 data Aprob1.1&prob2.1

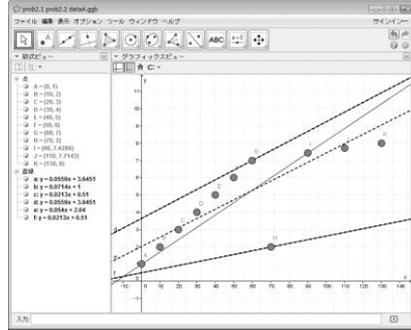


Fig.6.4 data Aprob2.1&prob2.2

6.2 数値計算 2

ここでは将棋、囲碁、チェスのような対戦型オンライン・ゲームに着目し、ゲーム参加者の勝率と、レーティングポイントを例とした計算例を紹介する。これらのゲームでは参加者一人一人の試合数が大きく異なり、それぞれの能力も異なる。本項では、あるオンライン・ゲームの参加者 100 人をランダムに抽出し、その勝率とレーティングポイントの関係をファジィ回帰モデルに表現する。

参加者の勝率 = 勝数 / 試合数 $\times 100$ (%) であり、この値は曖昧さを含まない実数値となる。また、レーティングポイントは実力という曖昧さを多分に含んだ実数値と考える。以下に、参加者の勝率を独立変数 x 、レーティングポイントを従属変数 y としたデータセット C を Table. 6. 5 に示す。データセット C を使い prob1. 1 と prob2. 1 からモデルを同定した結果を Table. 6. 6 に示す。

Table. 6. 5 データセット C

No	x	y	No	x	y	No	x	y	No	x	y
1	10.0	255	26	28.3	714	51	54.6	1115	76	71.5	1785
2	12.5	306	27	29.6	900	52	54.2	1377	77	71.5	2010
3	12.6	352	28	29.2	1050	53	55.4	1405	78	72.3	2011
4	13.5	258	29	30.0	765	54	57.8	1377	79	73.0	1700
5	13.9	320	30	31.5	816	55	57.5	1479	80	74.5	1887

6	13.3	458	31	32.0	1150	56	58.9	1256	81	74.4	1887
7	14.1	450	32	33.5	1000	57	58.2	1825	82	75.5	1700
8	14.5	500	33	35.5	758	58	59.5	1225	83	76.8	1938
9	15.7	300	34	36.0	918	59	59.6	1432	84	77.4	1800
10	15.6	600	35	37.0	950	60	60.4	1530	85	77.9	1968
11	16.5	480	36	39.8	700	61	60.7	1800	86	81.5	2125
12	17.7	400	37	39.5	800	62	61.2	1415	87	82.7	2091
13	17.5	800	38	40.7	800	63	62.5	1450	88	80.0	2400
14	18.6	400	39	41.0	1200	64	62.9	1734	89	84.5	1850
15	18.5	459	40	43.9	750	65	63.2	1052	90	84.0	2400
16	20.0	300	41	43.3	1260	66	63.5	1358	91	87.4	2150
17	20.8	510	42	44.5	1500	67	64.6	1245	92	90.5	1958
18	21.9	700	43	45.0	1156	68	65.4	1530	93	91.8	2158
19	20.0	1000	44	47.9	1590	69	67.5	1680	94	92.9	2000
20	22.5	561	45	49.2	1358	70	68.6	1052	95	92.1	2005
21	24.4	550	46	50.5	1275	71	68.5	1805	96	93.5	2451
22	24.7	612	47	50.0	2000	72	68.4	1215	97	94.2	2397
23	25.0	750	48	51.1	1302	73	69.2	1765	98	100	2000
24	27.1	750	49	52.5	1450	74	70.0	1000	99	100	2485
25	28.4	600	50	53.2	1800	75	70.0	1785	100	100	2250

Table 6.6 データセットCによる計算結果

計算結果 C	上辺関数		中心関数		下辺関数		Ave. α	Ver. α
	傾き	切片	傾き	切片	傾き	切片		
prob. 1. 1	13.33	1333	13.67	676.6	14.00	19.77	0.5661	0.08560
prob. 2. 1	20.91	954.3	21.19	160.0	14.00	20.00	0.6959	0.06180

Table 6.6 の計算結果を Fig. 6.5 にグラフ化する。prob1.1 が破線で prob2.1 が実線である。prob1.1 と prob2.1 の下辺関数は近いが、上辺関数、中心関数は prob1.1 より prob2.1 の方がデータ群に近い。これは従来の対称三角型ファジィ数から非対称三

角型ファジィ数へ拡張した効果である。prob1.1では勝率の低い人がモデルの中心関数から下に勝率の高い人が上に集中している。

prob2.1では中心関数の上下ほぼ均等にデータ点が密集している。データ点が中心関数周辺に密集したファジィ回帰モデルから推定されるファジィ数には説得力がある。このことはTable. 6.6の α レベル平均値からもわ

かる。prob1.1では0.5661、prob2.1では0.6959とprob2.1のほうが高い。以上の理由からprob2.1モデルが表現するファジィ数には説得力があり参加者のレーティングをよりよく推定できると考える。

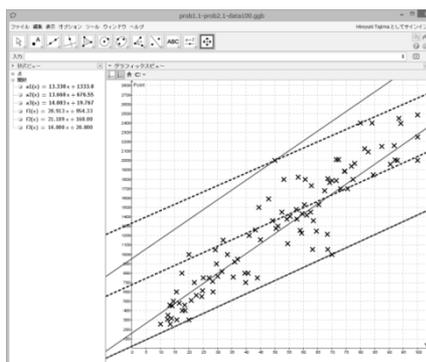


Fig.6.5 data C:prob1.1&prob2.1

7. まとめ

著者の提案するファジィ回帰モデル同定法はモデルの示すファジィ数を対称型から非対称型に拡張し、その中心関数に強い意味を持たせている。また、最小二乗法の概念を導入することによってデータの変化に敏感に対応する上下辺関数を同定することを可能にしている。Prob2.1ではモデルに対するデータの帰属度を高くする中心関数を同定し、prob2.2ではデータ点と中心関数の距離の二乗を最小にする同定法を提案している。2つの同定法は利用する実験者によって使い分けられるべきであるが、いずれのモデルからも優れた中心関数を持つファジィ数が推定されると考えている。

ファジィ回帰モデルに関する研究は少ない。これはデータ分析者が問題を解くごとに難解なコーディングを行う必要があるためである。実用に当たっては更に多くの研究が必要と思われるが、本研究をさらに発展させていくためにはEXCELや算術計算ソフト等の組み込み関数を構築する必要がある。誰もがファジィ回帰モデルを容易に利用できる環境を整えることが今後の研究課題となる。

参考文献

- [1] 田中, 和多田, 林, ファジィ線形回帰分析の三つの定式化, 計測自動制御学会論文集, Vol. 22, No. 10, pp. 1051-1057 (1986)
- [2] 石淵, ファジィ回帰分析, 日本ファジィ学会誌 Vol. 4, No. 1, pp. 52-60 (1992)
- [3] 石淵, ファジィ回帰分析, 講座ファジィ 6 ファジィ OR, 日本工業新聞社, pp. 141-185 (1993)
- [4] 藪内, 和多田, 超楕円関数に基づくファジィロバスト回帰分析, Journal of the Operations Research Society of Japan, Vol. 39, No4, pp. 512-524 (1996)
- [5] 田中, 林, 長坂, 可能性測度による区間回帰分析, 行動計量学, Vol. 16, No. 1, pp. 1-7 (1988)
- [6] 田島, ファジィ回帰モデル同定アルゴリズムに関する基礎的研究, 京都大学数理解析研究所・研究集会「不確実性の下での数理解析モデルの構築と最適化」京都大学数理解析研究所. 数理解析研究所講究録 1194, pp. 65-66 (2000)
- [7] 田島, 古川, 非対称型ファジィ回帰モデルの研究, 情報処理学会第 34 回数理解析モデル化と問題解決研究会, 情報処理学会研究報 2001-MPS-34, pp. 35-38 (2001)
- [8] 田中, 小山, 李, 二次計画法による可能性回帰分析, 第 12 回ファジィシステムシンポジウム講演論文, pp. 845-846 (1996)
- [9] 田島, ファジィ回帰分析における新たな最小化基準の研究, 数理解析研究所講究録 891 「連続と離散の最適化数理解析」 京都大学数理解析研究所, pp. 89-94 (1997)

(たじま ひろゆき・講師)

