



ファジィ回帰分析[†]

石渕 久生*

1. はじめに

システムの入出力関係に関する m 個の入出力対 $(\mathbf{x}_p; y_p)$, $p=1, 2, \dots, m$ が与えられているとする。ただし, $\mathbf{x}_p$ は n 次元入力ベクトル $(x_{p1}, x_{p2}, \dots, x_{pn})$ であり, y_p は出力値である。このようなデータが与えられた場合, 通常回帰分析では, 次のような線形回帰モデルが用いられる。

$$\eta_p = a_0 + a_1 x_{p1} + \dots + a_n x_{pn} \quad (1.1)$$

$$y_p = \eta_p + \varepsilon_p \quad (1.2)$$

ここで, η_p は線形回帰モデルによる推定値であり, ε_p は実際の出力値 y_p とモデルによる推定値 η_p との誤差である。通常回帰分析では, 2 乗誤差 ε_p^2 の総和が最小になるように線形回帰モデルの係数 a_i が求められる。

田中ら[1,2]により提案されたファジィ回帰分析は, 回帰モデルそのものがあいまいさを含むという点で, 従来の回帰分析と大きく異なっている。ファジィ回帰分析では, 次のようなファジィ線形回帰モデルが用いられる。

$$\tilde{Y}(\mathbf{x}_p) = \tilde{A}_0 + \tilde{A}_1 x_{p1} + \dots + \tilde{A}_n x_{pn} \quad (1.3)$$

ここで, 回帰係数 $\tilde{A}_i$ はファジィ数であり, その結果, 回帰モデルにより推定された出力値 $\tilde{Y}(\mathbf{x}_p)$ もファジィ数となる。ファジィ回帰分析の問題は, ファジィ線形回帰モデルにおけるファジィ係数 $\tilde{A}_i$ を求める問題である。

ファジィ係数 $\tilde{A}_i$ としては, 対称三角型ファジィ数が用いられることが多い[1,2]が, 型関数 $L(\cdot)$ を用いた一般的な対称ファジィ数が用いられる場合[3,4] や, 2 次形式のファジィ数が用いられる場合[5,6] もある。また, ファジィ係数 $\tilde{A}_i$ を区間係数に簡略した場合に対する研究[7,8] も行われている。

ファジィ線形回帰モデルの係数が区間である場合は, 特に区間回帰分析と呼ばれ, 次のような線形回帰モデルが用いられる。

$$Y(\mathbf{x}_p) = A_0 + A_1 x_{p1} + \dots + A_n x_{pn} \quad (1.4)$$

ここで, 回帰係数 A_i は区間であり, その結果, 回帰モデルにより推定された出力値 $Y(\mathbf{x}_p)$ も区間となる。

ファジィ回帰分析に関する研究[1-8]では, 線形計画法を用いてファジィ係数 $\tilde{A}_i$ が求められるが, 最小 2 乗法に基づく方法[9-11]も提案されている。

本解説では, まず, 実数値入出力データが与えられている場合に対し, 区間回帰分析の問題を線形計画問題として定式化する。この問題では, 区間線形回帰モデルの係数 A_i が, 次の制約条件を満たすように求められる。

$$y_p \in Y(\mathbf{x}_p), \quad p=1, 2, \dots, m \quad (1.5)$$

すなわち, モデルによる推定値 $Y(\mathbf{x}_p)$ が実際の出力値 y_p を包含するように, 区間線形回帰モデルの係数が求められる。

次に, 出力値が区間であるような入出力データ $(\mathbf{x}_p; Y_p)$, $p=1, 2, \dots, m$ が与えられている場合へ, 区間回帰分析の問題を拡張する。この場合では, 実際の出力区間 Y_p と推定値 $Y(\mathbf{x}_p)$ との間の

[†] Fuzzy Regression Analysis
Hisao ISHIBUCHI

* 大阪府立大学 工学部
College of Engineering, University of Osaka Prefecture

制約として、次の三つの関係が考えられる。

$$Y_p \subseteq Y(\mathbf{x}_p), \quad p=1,2,\dots,m \quad (1.6)$$

$$Y_p \supseteq Y(\mathbf{x}_p), \quad p=1,2,\dots,m \quad (1.7)$$

$$Y_p \cap Y(\mathbf{x}_p) \neq \phi, \quad p=1,2,\dots,m \quad (1.8)$$

この三つの制約条件に基づき、3種類の線形計画問題を定式化する。

最後に、ファジィ回帰分析と区間回帰分析との関係を明らかにし、区間線形回帰モデルから対称三角型ファジィ係数を持つファジィ線形回帰モデルを導出する方法を示す[12]。

2. 実数値データに対する区間回帰分析

2.1 区間演算

まず、区間回帰分析で必要となる区間演算について簡単に述べる。本解説では、実数を英小文字 $a, b, \dots, z$ で、区間を英大文字 $A, B, \dots, Z$ で表わすことにする。区間は、左端(Left limit)と右端(Right limit)を用いて次のように表記される。

$$A = [a_L, a_R] = \{a \mid a_L \leq a \leq a_R\} \quad (2.1)$$

ここで、 a_L と a_R は、区間 A の左端と右端である。また、同じ区間を、中心(Center)と中心から両端までの幅(Width)を用いて次のように表記する。

$$A = \langle a_c, a_w \rangle = \{a \mid a_c - a_w \leq a \leq a_c + a_w\} \quad (2.2)$$

ここで、 a_c と a_w は、区間 A の中心と幅(正確には半径: Radius)である。式(2.1)と式(2.2)の表記法の間には、次の関係がある。

$$a_c = (a_R + a_L) / 2 \quad (2.3)$$

$$a_w = (a_R - a_L) / 2 \quad (2.4)$$

実数に対する演算を区間の場合に一般化することにより、区間演算[13,14]と呼ばれる演算が定義される。区間回帰分析で用いる演算は、次のような区間の和と実数倍である。

$$\begin{aligned} A+B &= \langle a_c, a_w \rangle + \langle b_c, b_w \rangle \\ &= \langle a_c + b_c, a_w + b_w \rangle \end{aligned} \quad (2.5)$$

$$k \cdot A = k \langle a_c, a_w \rangle = \langle k \cdot a_c, |k| a_w \rangle \quad (2.6)$$

ここで、 k は実数である。

この区間演算を用いれば、区間線形回帰モデルは次のように計算される。

$$\begin{aligned} Y(\mathbf{x}_p) &= A_0 + A_1 x_{p1} + \dots + A_n x_{pn} \\ &= \langle a_{0c}, a_{0w} \rangle + \langle a_{1c}, a_{1w} \rangle x_{p1} \\ &\quad + \dots + \langle a_{nc}, a_{nw} \rangle x_{pn} \\ &= \langle y_c(\mathbf{x}_p), y_w(\mathbf{x}_p) \rangle \end{aligned} \quad (2.7)$$

$$y_c(\mathbf{x}_p) = a_{0c} + a_{1c} x_{p1} + \dots + a_{nc} x_{pn} \quad (2.8)$$

$$y_w(\mathbf{x}_p) = a_{0w} + a_{1w} |x_{p1}| + \dots + a_{nw} |x_{pn}| \quad (2.9)$$

2.2 線形計画問題の定式化

入出力関係に関する実数値データ $(\mathbf{x}_p; y_p)$, $p=1,2,\dots,m$ が与えられているとする。このときの区間回帰分析の基本的な問題は、次のような線形計画問題として定式化することができる[7,8]。

[実数値データに対する線形計画問題]

$$\text{目的関数: } \sum_{p=1}^m y_w(\mathbf{x}_p) \rightarrow \text{最小化} \quad (2.10)$$

$$\text{制約条件: } y_p \in Y(\mathbf{x}_p), \quad p=1,2,\dots,m \quad (2.11)$$

$$a_{iw} \geq 0, \quad i=0,1,\dots,n \quad (2.12)$$

$$\text{決定変数: } a_{ic}, a_{iw}, \quad i=0,1,\dots,n \quad (2.13)$$

目的関数(2.10)は、区間線形回帰モデルによる推定区間 $Y(\mathbf{x}_p)$ の幅 $y_w(\mathbf{x}_p)$ の総和の最小化である。また、制約条件(2.11)は、推定区間 $Y(\mathbf{x}_p)$ が実際の出力値 y_p を包含しなければならないという制約であり、制約条件(2.12)は、区間係数 A_i の幅 a_{iw} に関する非負条件である。

2.3 数値計算例

簡単な例題により、前節の定式化の説明を行う。1入力1出力システムの入出力データとして、次のような実数値データが与えられているとする。

$$\begin{aligned} \{(\mathbf{x}_p; y_p)\} &= \{(2; 14), (4; 16), (6; 14), \\ &\quad (8; 18), (10; 18), (12; 22), \\ &\quad (14; 18), (16; 22)\} \end{aligned}$$

入力変数が1個であるので、区間線形回帰モデルは次のようになる。

$$\begin{aligned} Y(x_p) &= A_0 + A_1 x_p \\ &= \langle a_{0c}, a_{0w} \rangle + \langle a_{1c}, a_{1w} \rangle x_p \end{aligned} \quad (2.14)$$

ここでの問題は、与えられた8個のデータ点を包含する幅最小の区間線形回帰モデルを求める問題である。前節で示した方法をこの例題に適用すると、次のような線形計画問題が定式化される。

$$\text{目的関数: } 8a_{0w} + (2+4+\cdots+16)a_{1w} \rightarrow \text{最小化}$$

$$\text{制約条件: } a_{0c} + 2a_{1c} - a_{0w} - 2a_{1w} \leq 14$$

...

$$a_{0c} + 16a_{1c} - a_{0w} - 16a_{1w} \leq 22$$

$$a_{0c} + 2a_{1c} + a_{0w} + 2a_{1w} \geq 14$$

...

$$a_{0c} + 16a_{1c} + a_{0w} + 16a_{1w} \geq 22$$

$$a_{0w}, a_{1w} \geq 0$$

この線形計画問題を解くことにより、次のような区間線形回帰モデルが得られる。

$$Y(x_p) = \langle 12, 1 \rangle + \langle 0.625, 0.125 \rangle x_p \quad (2.15)$$

この区間線形回帰モデルを、与えられた入出力データと共に図1に示す。図1より、区間線形回帰モデルが全てのデータ点を包含していることがわかる。

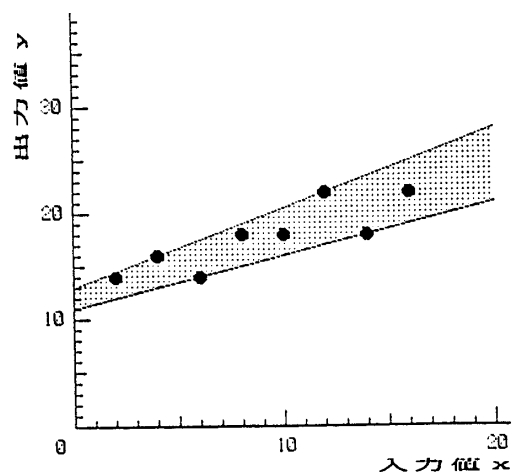


図1 区間線形回帰モデル

2.4 組合せ最適化問題としての定式化

図1の区間線形回帰モデルは、制約条件が等号で成立する最も外側の4個のデータ点で一意的に決定されている。一般に、 n 個の入力変数(説明変数)を持つ区間線形回帰モデルは、最も外側に位置する $2(n+1)$ 個のデータ点で一意的に決定される。したがって、データに異常値が含まれているような場合には、2.2節の線形計画問題から得られる区間線形回帰モデルは、異常値の影響を直接的に受ける。

区間線形回帰モデルが異常値の影響を直接的に受ける原因は、異常値を含めて与えられた全てのデータ点を包含しなければならないという制約条件式(2.11)による。そこで、全てのデータ点を包含しなければならないという制約を、ほとんどのデータ点を包含しなければならないという制約に緩和することにより、次のような組合せ最適化問題が定式化できる[8]。

[組合せ最適化問題としての定式化]

$$\text{目的関数: } \sum_{p \in J} y_w(x_p) \rightarrow \text{最小化} \quad (2.16)$$

$$\text{制約条件: } y_p \in Y(x_p), p \in J \quad (2.17)$$

$$J \subseteq \{1, 2, \dots, m\} \quad (2.18)$$

$$\text{Card}(J) = k \quad (2.19)$$

$$a_{iw} \geq 0, i = 0, 1, \dots, n \quad (2.20)$$

$$\text{決定変数: } a_{ic}, a_{iw}, i = 0, 1, \dots, n \quad (2.21)$$

ここで、集合 J は未知の添字集合である。制約条件式(2.17)は、区間線形回帰モデルが集合 J に含まれるデータ点を包含しなければならないことを示している。また、制約条件式(2.19)は、集合 J に含まれるデータ点の数が k 個($k \leq m$)であることを示している。したがって、この組合せ最適化問題は、 k 個のデータ点を含む幅最小の区間線形回帰モデルを求める問題である。

例題として、次のような10個のデータ点を用いた区間線形回帰分析の問題を考えよう。

$$\{(x_p; y_p)\} = \{(2; 14), (4; 16), (6; 14), (8; 18), (10; 18), (12; 22), (14; 18), (16; 22), (2; 4), (16; 32)\}$$

これは、前節の例題に、人為的に設定した2個の異常値を加えたものである。

まず、2.2節で定式化した線形計画問題により得られた区間線形回帰モデルを図2に示す。図1と図2を比較すれば、区間線形回帰モデルが新たに追加された2個のデータ点の影響を直接的に受けていることがわかる。

次に、 $k=9$ および $k=8$ と設定し、本節で示した組合せ最適化問題を解くことにより得られた区間線形回帰モデルを、それぞれ図3と図4に示す。

図4では、新たに追加された2個のデータが区間線形回帰モデルに含まれていないことがわかる。図4と図2とを比較すれば、図4の区間線形回帰モデルの幅が図2に較べて非常に狭いことがわかる。このように、全てのデータ点を包含しない

ければならないという制約条件を緩和することにより、あいまいさの少ない区間線形回帰モデルが導出できる。

適当な k の値を見つけるためには、 k の値と目的関数の値を2次元上でプロットした図を用いることができる。図5は、本節の例題に対して、 k の値と目的関数の値との関係を図示したものである。この図より、 $k=9$ および $k=8$ と設定することにより、目的関数の値が大きく減少することがわかる。これは、異常値と思われるデータ点が添字集合 J から取り除かれたことを示唆するものである。また、 $k=7$ と設定したときの目的関数の減少がそれほど大きくないこともわかる。したがって、 $k=8$ と設定することにより、異常値以外のデータ点を全て包含する区間線形回帰モデルが得られていることがわかる。

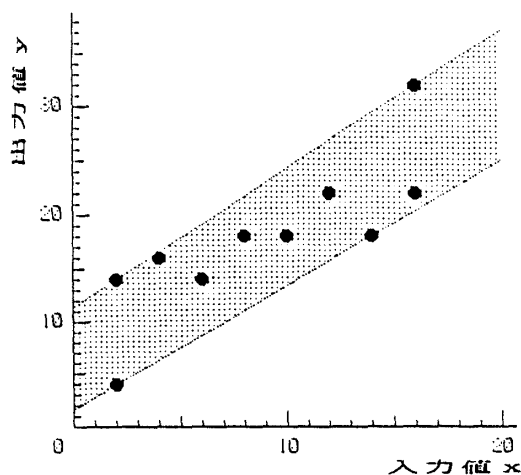


図2 全てのデータ点を包含するモデル

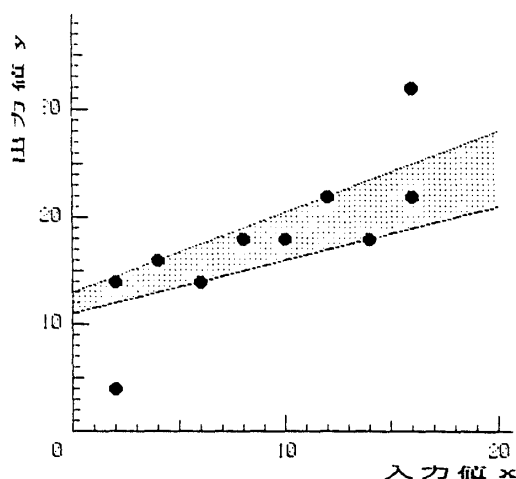


図4 8個のデータ点を包含するモデル

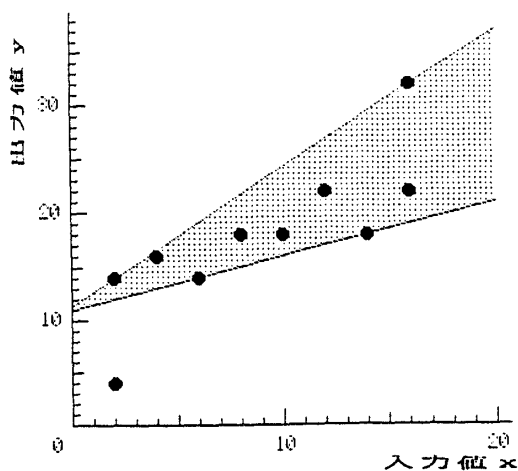


図3 9個のデータ点を包含するモデル

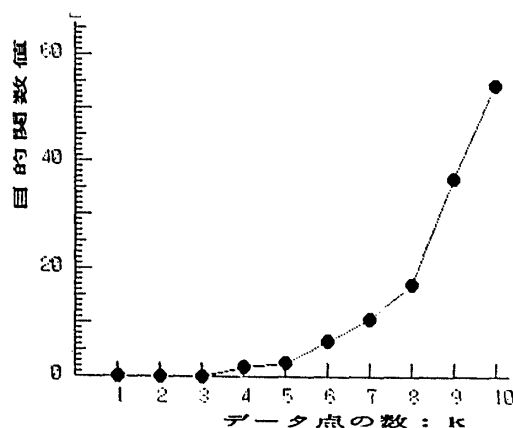


図5 目的関数の値と k の値との関係

3. 区間値データに対する区間回帰分析

ここでは、出力値が区間であるような入出力データ $(x_p; Y_p)$, $p=1, 2, \dots, m$ が与えられている場合での区間回帰分析の問題を解説する。

3.1 区間の包含関係

区間値データに対する区間回帰分析のための準備として、区間に対する包含関係の定義を示す。本解説では、次のような包含関係を用いる。

$$A \subseteq B \Leftrightarrow b_L \leq a_L, \quad a_R \leq b_R \quad (3.1)$$

$$A \cap B \neq \phi \Leftrightarrow a_L \leq b_R, \quad b_L \leq a_R \quad (3.2)$$

区間値データに対する区間回帰分析の問題は、区間線形回帰モデル $Y(x_p)$ と区間出力値 Y_p との間の包含関係により、三つの定式化が可能である。これらの定式化は、田中ら[3] によるファジィ回帰分析の三つの定式化と対応している。

3.2 Min 問題

実数値データに対して 2.2 節で定式化した線形計画問題の最も自然な拡張として、Min 問題と呼ばれる次のような線形計画問題を区間値データに対して定式化することができる。

[Min 問題]

$$\text{目的関数: } \sum_{p=1}^m y_w(x_p) \rightarrow \text{最小化} \quad (3.3)$$

$$\text{制約条件: } Y_p \subseteq Y(x_p), \quad p=1, 2, \dots, m \quad (3.4)$$

$$a_{iw} \geq 0, \quad i=0, 1, \dots, n \quad (3.5)$$

$$\text{決定変数: } a_{ic}, a_{iw}, \quad i=0, 1, \dots, n \quad (3.6)$$

制約条件式(3.4) は、区間線形回帰モデルが与えられた全ての区間出力値を包含しなければならないことを表わしている。したがって、Min 問題は、与えられた区間値データを包含する幅最小の区間線形回帰モデルを求める問題である。

例題として、次のような区間値データを用いた区間回帰分析の問題を考えよう。

$$\begin{aligned} \{(x_p; Y_p)\} = & \{(3; [10, 15]), (6; [10, 15]), \\ & (9; [13, 19]), (12; [16, 22]), \\ & (15; [15, 22]), (18; [19, 28])\} \end{aligned}$$

この例題に対して Min 問題を適用することにより、次のような区間線形回帰モデルが得られる。

$$Y(x_p) = \langle 9.533, 2.867 \rangle + \langle 0.711, 0.156 \rangle x_p \quad (3.7)$$

与えられた区間値データと共に区間線形回帰モデルを図示すれば、図 6 のようになる。図 6 より、与えられた全ての区間値データを包含する区間線形回帰モデルが得られていることがわかる。また、図 6 に示す区間線形回帰モデルは、区間出力値が存在する可能性のある可能的領域を示していると考えることもできる。

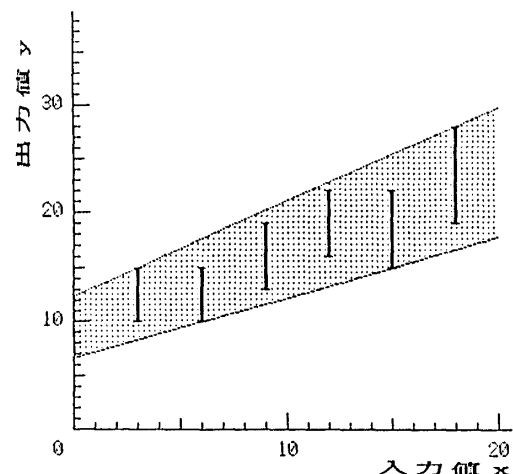


図 6 Min 問題による区間線形回帰モデル

3.3 Max 問題

Min 問題とは逆に、与えられた区間出力値に包含される幅最大の区間線形回帰モデルを求める問題が Max 問題である。Max 問題は、次のような線形計画問題として定式化される。

[Max 問題]

$$\text{目的関数: } \sum_{p=1}^m y_w(x_p) \rightarrow \text{最大化} \quad (3.8)$$

$$\text{制約条件: } Y(x_p) \subseteq Y_p, \quad p=1, 2, \dots, m \quad (3.9)$$

$$a_{iw} \geq 0, \quad i=0, 1, \dots, n \quad (3.10)$$

$$\text{決定変数: } a_{ic}, a_{iw}, \quad i=0, 1, \dots, n \quad (3.11)$$

制約条件式(3.9) は、区間線形回帰モデルが与えられた区間出力値に包含されなければならない

ことを示している。

前節の例題に Max 問題を適用することにより、次のような区間線形回帰モデルが得られる。

$$Y(x_p) = \begin{matrix} <9.167, 1.167> \\ + <0.722, 0.056> x_p \end{matrix} \quad (3.12)$$

この区間線形回帰モデルを、与えられた区間値データと共に図7に示す。図7より、与えられた区間値データに包含される区間線形回帰モデルが得られていることがわかる。また、図7に示す区間線形回帰モデルは、区間値データが必ず存在する必然的領域を表わしているとも考えることもできる。

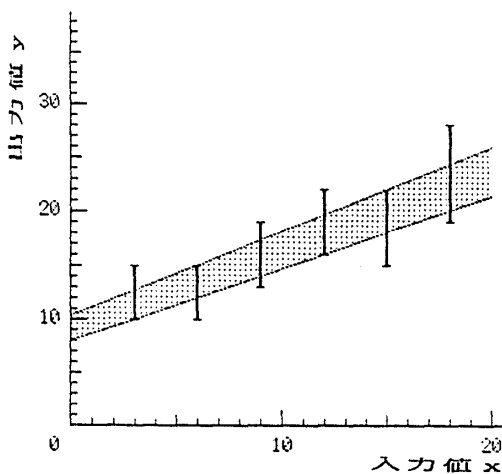


図7 Max 問題による区間線形回帰モデル

3.4 Conjunction 問題

Min 問題には常に実行可能解が存在するが、Max 問題には実行可能解が存在するとは限らない。すなわち、与えられた全ての区間値データと交わりがある直線(入力変数が2個以上の場合では、平面あるいは超平面)が存在しない場合では、Max 問題に実行可能解が存在せず、区間線形回帰モデルを求めることはできない。

Max 問題に解が存在するためには、次のような線形計画問題として定式化される Conjunction 問題の最適値が0となることが必要十分条件である。

[Conjunction 問題]

$$\text{目的関数: } \sum_{p=1}^m y_w(x_p) \rightarrow \text{最小化} \quad (3.13)$$

$$\text{制約条件: } Y_p \cap Y(x_p) \neq \phi, \quad p=1, 2, \dots, m \quad (3.14)$$

$$a_{iw} \geq 0, \quad i=0, 1, \dots, n \quad (3.15)$$

$$\text{決定変数: } a_{ic}, a_{iw}, \quad i=0, 1, \dots, n \quad (3.16)$$

制約条件(3.14)は、与えられた区間値データと区間線形回帰モデルとの間に交わりがあることを表わしている。この問題の最適値が0であれば、区間線形回帰モデル $Y(x_p)$ の幅が0となり、式(3.14)から次式を得ることができる。

$$Y(x_p) \subseteq Y_p, \quad p=1, 2, \dots, m \quad (3.17)$$

すなわち、Conjunction 問題の最適値が0であれば、この問題の最適解は Max 問題の実行可能解となる。また、逆に、Max 問題に実行可能解が存在すれば、Conjunction 問題の最適値は0となる。例えば、図7に示す例題に Conjunction 問題を適用すれば、最適値は0となる。

Min 問題は区間出力値の存在する可能性のある領域を求める問題、Max 問題は必然性のある領域を求める問題と考えることができるのに対し、Conjunction 問題から得られる区間線形回帰モデルの表わす領域の解釈は明確ではない。したがって、Conjunction 問題は区間線形回帰モデルを求めるための定式化ではなく、Max 問題における実行可能解の存在を検討するための問題と考えることができる。

3.5 Min 問題と Max 問題との関係

Min 問題から次のような区間線形回帰モデルが得られているとする。

$$Y^*(x_p) = A_0^* + A_1^* x_{p1} + \dots + A_n^* x_{pn} \quad (3.18)$$

また、Max 問題から次のような区間線形回帰モデルが得られているとする。

$$Y^*(x_p) = A_0^* + A_1^* x_{p1} + \dots + A_n^* x_{pn} \quad (3.19)$$

このとき、Min 問題と Max 問題における制約条件から明らかなように、次のような関係が成立する。

$$Y^*(x_p) \subseteq Y_p \subseteq Y^*(x_p), \quad p=1, 2, \dots, m \quad (3.20)$$

この関係は、ラフ集合[15,16]における二つの近似の概念と対応している。ラフ集合の言葉を用いれば、区間 $Y_*(x_p)$ は区間 Y_p の下からの近似であり、区間 $Y^*(x_p)$ は上からの近似となる。

入出力データとして与えられている入力ベクトル x_p に対しては式(3.20)のような包含関係が成り立つが、Min 問題と Max 問題が独立に解かれるため、一般の入力ベクトル x に対して次式が成立する保証はない。

$$Y_*(x) \subseteq Y^*(x) \quad (3.21)$$

すなわち、必然的領域 $Y_*(x)$ が可能的領域 $Y^*(x)$ に包含される保証はない。

この問題点を解決するため、Min 問題と Max 問題を独立に解くのではなく、次の制約条件のもとで同時に解くという方法が提案されている[17]。

$$A_{i*} \subseteq A_i^*, i=0,1,\dots,n \quad (3.22)$$

区間演算の包含関係に関する単調性[13]より、式(3.22)の制約のもとでは、式(3.21)が任意の入力ベクトル x に対して成立する。

4. ファジィ線形回帰モデルの導出

4.1 ファジィ線形回帰モデル

ファジィ線形回帰モデルのファジィ係数 $\tilde{A}_i$ を対称三角型ファジィ数とし、その中心 a_i と広がり e_i により $\tilde{A}_i = (a_i, e_i)_L$ と表記する。このファジィ数のメンバーシップ関数は、次のように定義される。

$$\mu_{\tilde{A}_i}(a) = \max \{1 - |a - a_i|/e_i, 0\} \quad (4.1)$$

すなわち、対称三角型ファジィ数 $\tilde{A}_i$ のメンバーシップ値は、中心 a_i で1となり、开区間 $(a_i - e_i, a_i + e_i)$ で正となる。

ファジィ係数 $\tilde{A}_i$ が対称三角型ファジィ数の場合、ファジィ線形回帰モデル $\tilde{Y}(x_p)$ のメンバーシップ関数も対称三角型となり、拡張原理[18]により次のように計算される。

$$\begin{aligned} \tilde{Y}(x_p) &= \tilde{A}_0 + \tilde{A}_1 x_{p1} + \dots + \tilde{A}_n x_{pn} \dots \\ &= (a(x_p), e(x_p))_L \end{aligned} \quad (4.2)$$

$$a(x_p) = a_0 + a_1 x_{p1} + \dots + a_n x_{pn} \quad (4.3)$$

$$e(x_p) = e_0 + e_1 |x_{p1}| + \dots + e_n |x_{pn}| \quad (4.4)$$

4.2 実数値データに対するファジィ回帰分析

実数値入出力データ $(x_p; y_p)$, $p=1,2,\dots,m$ が与えられた場合での基本的なファジィ回帰分析の問題は、次の制約を満たすファジィ線形回帰モデル $\tilde{Y}(x_p)$ を求める問題である[1-4]。

$$y_p \in [\tilde{Y}(x_p)]_h, p=1,2,\dots,m \quad (4.5)$$

ここで、 $0 < h < 1$ であり、 $[\tilde{Y}(x_p)]_h$ は推定ファジィ数 $\tilde{Y}(x_p)$ の h レベル集合である。なお、推定ファジィ数 $\tilde{Y}(x_p)$ の h レベル集合は、次のように定義される。

$$[\tilde{Y}(x_p)]_h = \{y \mid \mu_{\tilde{Y}(x_p)}(y) \geq h\} \quad (4.6)$$

いま、2.で述べた区間回帰分析の手法により、次のような区間線形回帰モデル $Y(x_p)$ が得られているとする。

$$\begin{aligned} y_p \in Y(x_p) &= A_0 + A_1 x_{p1} + \dots + A_n x_{pn}, \\ p &= 1, 2, \dots, m \end{aligned} \quad (4.7)$$

このとき、式(4.5)の制約条件を満たすファジィ線形回帰モデルを、次のように導出することができる。

$$a_i = a_i c, e_i = a_i w / (1-h), i=0,1,\dots,n \quad (4.8)$$

式(4.8)より、次のような関係が導出され、ファジィ線形回帰モデルが式(4.5)の制約条件を満たすことがわかる。

$$A_i = [\tilde{A}_i]_h, i=0,1,\dots,n \quad (4.9)$$

$$Y(x_p) = [\tilde{Y}(x_p)]_h, p=1,2,\dots,m \quad (4.10)$$

すなわち、式(4.8)のようにファジィ係数 $\tilde{A}_i$ を定めることにより、区間線形回帰モデル $Y(x_p)$ に対して成り立つ条件が、ファジィ線形回帰モデルの h レベル集合 $[\tilde{Y}(x_p)]_h$ に対しても同様に成

立することになる。したがって、2.4 で述べた区間回帰分析の方法と式(4.8)を用いることにより、 h レベル集合 $[\tilde{Y}(x_p)]_h$ 内に k 個のデータ点を含むようなファジィ線形回帰モデル $\tilde{Y}(x_p)$ を求めることもできる。

4.3 ファジィデータに対するファジィ回帰分析

出力値がファジィ数 $\tilde{Y}_p$ であるようなファジィデータ $(x_p; \tilde{Y}_p)$ に対するファジィ回帰分析の問題は、3. で述べた区間回帰分析の場合と同様に、次の三つの制約条件に基づき定式化される[3,4].

$$[\tilde{Y}_p]_h \subseteq [\tilde{Y}(x_p)]_h, \quad p=1,2,\dots,m \quad (4.11)$$

$$[\tilde{Y}(x_p)]_h \subseteq [\tilde{Y}_p]_h, \quad p=1,2,\dots,m \quad (4.12)$$

$$[\tilde{Y}(x_p)]_h \cap [\tilde{Y}_p]_h \neq \phi, \quad p=1,2,\dots,m \quad (4.13)$$

ファジィ数の h レベル集合が区間であることから、式(4.11)－(4.13)の制約は、区間回帰分析の問題における三つの制約条件と本質的には同じであることがわかる。

式(4.11)－(4.13)の制約を満たすようなファジィ線形回帰モデルは、次のようにして区間線形回帰モデルから導出することができる。まず、次の三つの制約条件を満たすような区間線形回帰モデルを、それぞれ3. で述べた方法により求める。

$$[\tilde{Y}_p]_h \subseteq Y(x_p), \quad p=1,2,\dots,m \quad (4.14)$$

$$Y(x_p) \subseteq [\tilde{Y}_p]_h, \quad p=1,2,\dots,m \quad (4.15)$$

$$Y(x_p) \cap [\tilde{Y}_p]_h \neq \phi, \quad p=1,2,\dots,m \quad (4.16)$$

式(4.11)－(4.13)の制約条件を満たすような三つのファジィ線形回帰モデルは、式(4.14)－(4.16)の制約条件から得られる三つの区間線形回帰モデルを用いて、式(4.8)により導出することができる。

5. おわりに

本解説では、ファジィ回帰分析の基本的な考え方を説明するために、回帰モデルとして区間線形関数が用いられる場合での定式化を中心に解説を行った。また、区間回帰分析により得られた区間

線形回帰モデルからファジィ線形回帰モデルを導出する方法を示した。

本解説では、紙面の都合もあり、最も基本的な定式化のみを示した。いくつかの代替的な定式化に関する研究が文献[19]で行われている。また、入出力値が共にファジィ数であるような入出力データへの拡張が文献[20]で、各々のデータ点に重要度が付けられている場合への拡張が文献[21]で行われている。さらに、非線形ファジィ回帰モデルや非線形区間回帰モデルへの拡張として、GMDH を用いた方法[22,23]やニューラルネットを用いた方法[24,25]も提案されている。

参 考 文 献

- [1] H.Tanaka, S.Uejima and K.Asai : Linear Regression Analysis with Fuzzy Model, IEEE Trans. Syst. Man Cybernet. Vol.SMC-12, No.6, pp.903-907 (1982)
- [2] 田中, 上嶋, 浅居 : ファジィ関数による線形回帰モデル, J. Operations Research Society of Japan, Vol.25, No.2, pp.162-173 (1982)
- [3] 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)
- [4] H.Tanaka : Fuzzy Data Analysis by Possibilistic Linear Models, Fuzzy Sets and Systems, Vol. 24, pp.363-375 (1987)
- [5] 田中, 石淵 : 2次形式メンバシップ関数による可能性線形システムの同定, 計測自動制御学会論文集, Vol.26, No.1, pp.93-100 (1990)
- [6] H.Tanaka and H.Ishibuchi : Identification of Possibilistic Linear Systems by Quadratic Membership Functions of Fuzzy Parameters, Fuzzy Sets and Systems, Vol.41, pp.145-160 (1991)
- [7] 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)
- [8] 石淵, 田中 : 混合 0-1 整数計画問題による区間回帰分析, 日本経営工学会誌, Vol.40, No.5, pp.312-319 (1989)
- [9] A.Celmins : Least Squares Model Fitting to Fuzzy Vector Data, Fuzzy Sets and Systems, Vol.22, pp.245-269 (1987)
- [10] A.Celmins : Multidimensional Least-Squares

- Fitting of Fuzzy Models, Mathematical Modeling, Vol.9, No.9, pp.669-690 (1987)
- [11] P.Diamond: Fuzzy Least Squares, Information Sciences, Vol.46, pp.141-157 (1988)
- [12] 石渕, 田中: 区間回帰モデルによるファジー係数の同定, 電子情報通信学会論文誌 A, Vol.J 72-A, No.12, pp.1958-1964 (1989)
- [13] G.Alefeld and J.Herzberger: Introduction to Interval Computations, Academic Press, New York (1983)
- [14] R.E.Moore: Methods and Applications of Interval Analysis, SIAM (Society for Industrial and Applied Mathematics), Philadelphia (1979)
- [15] Z.Pawlak: Rough Sets, Int. J. Computer and Information Sciences, Vol.11, No.5, pp.341-356 (1982)
- [16] Z.Pawlak: Rough Classification, Int. J. Machine Studies, Vol.20, pp.469-483 (1984)
- [17] 石渕, 田中, 長坂: 改良型区間回帰モデルによる区間データの解析, 計測自動制御学会論文集, Vol.25, No.11, pp.1218-1224 (1989)
- [18] L.A.Zadeh: The Concept of a Linguistic Variable and Its Application to Approximate Reasoning: Part 1, 2 and 3, Information Sciences, Vol.8, pp.199-249, pp.301-357 and Vol.9, pp.43-80 (1975)
- [19] 石渕, 田中, 黄: 作業時間解析におけるファジィ回帰手法の比較, システム制御情報学会論文誌, Vol.3, No.3, pp.90-92 (1990)
- [20] 坂和, 矢野: ファジィ入出力データに対する多目的ファジィ線形回帰分析, 日本ファジィ学会誌, Vol.1, No.1, pp.107-116 (1989)
- [21] 田中, 石渕, 黄: 類似度を導入したファジィ回帰分析による職員数モデル, 日本経営工学会誌, Vol.41, No.2, pp.99-104 (1990)
- [22] 田中, 林, 浅居: ファジィ GMDH の定式化, システムと制御, Vol.30, No.9, pp.581-587 (1986)
- [23] 石渕, 田中, 長坂: 区間 GMDH とその応用, システム制御情報学会論文誌, Vol.2, No.6, pp.191-199 (1989)
- [24] H.Ishibuchi and H.Tanaka: Regression Analysis with Interval Model by Neural Networks, Proc.IJCNN'91-Singapore, pp.1594-1599 (1991)
- [25] 石渕, 田中: ニューラルネットによるファジィ回帰分析とその応用, 計測自動制御学会論文集, Vol.28, No.1, pp.116-124 (1992)

(1991年10月30日 受付)

[問い合わせ先]

〒591 堺市百舌鳥梅町4-804

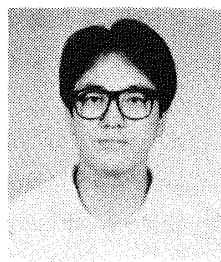
大阪府立大学 工学部 経営工学科第3講座

石渕 久生

☎: 0722-52-1161 (内線) 2377

FAX: 0722-59-3340

—— 著 者 紹 介 ——



石渕 久生 (いしぶち ひさお)

大阪府立大学 工学部経営工学科

1963年2月7日生。'85年, 京都大学工学部機械系精密工学科卒業。'87年, 同大学大学院修士課程精密工学専攻修了。同年, 大阪府立大学工学部経営工学科助手となり現在に至る。ファジィ理論とニューラルネットの応用に関する研究に従事。システム制御情報学会, 電子情報通信学会, 日本機械学会, JNNS, INNSなどの会員。