最小二乗基準によるファジィ線形回帰

信州大学工学部　正会員　奥谷　巌
信州大学工学部　正会員　高瀬　達夫
信州大学工学部　○松浦　慶明

1．まえがき

変数x_{ji}、y_jの観測データが何組か与えられているとき、y_jなる出現量を最もうまく説明するx_{ji}の線形結合を求める問題は線形回帰分析と呼ばれる周知の問題であり、通常、最小二乗法がその解法として用いられる。しかしながら、y_jやx_{ji}、あるいは係数を表わす量があいまい変量となると、ファジィ線形回帰では考え方が根底的に変わってしまい、たとえファジィ数の幅が0に近づいたとしても従来の最小二乗法を与えないという構造になっている。こうした不合理をなくすために、本研究では新たなファジィ線形回帰モデルを提案する。なお、ここで扱うファジィ数は対称な三角形ファジィ数のみとする。

2．従来の方法[1)]

Y_jをファジィ数$(y_j, e_j)_L$, $x_j = (x_{j1}, x_{j2}, \cdots, x_{jn})^T$とし、$(Y_j, x_j)$なるデータが$N$組与えられたとき、

$$Y_j \cong A_1 x_{j1} + A_2 x_{j2} + \cdots + A_n x_{jn} \tag{1}$$

を満たすファジィ係数$A_i = (\alpha_i, c_i)_L$を求めることを考える。方法論的には二つの方法が一般的に使われるが、その一つは(1)式の右辺の推定値がY_jなる観測値を水準h以上で含むようにする方法で、下記の(2)、(3)、(4)式のような線形計画法を使った最小化問題として定式化される。

もう一つの方法である最大化問題では、(1)式の右辺の推定値がY_jに水準h以上で含まれるようにするという考えから、最大化する目的関数は(2)式、制約条件は (3)、(4)式の不等号の向きが変わる問題である。

<線形計画法による定式化>

$$\min_{\alpha,c} \sum c^T |x_j| \tag{2}$$

$$y_j + |L^{-1}(h)| e_i \le x_j^T \alpha + |L^{-1}(h)| c^T |x_j| \tag{3}$$

$$y_j - |L^{-1}(h)| e_i \ge x_j^T \alpha - |L^{-1}(h)| c^T |x_j| \tag{4}$$

なお、最小化問題の解は常に存在するが、最大化問題の解の存在は保証されていない。

3．最小二乗モデル[2)]

3.1 ファジィ数間の距離

先にも述べたように、ここで扱うファジィ数Xは対称な三角形ファジィ数で、いま、対象とする二つのファジィ数X_1、X_2を

$$X_1 = (m_1, w_1)_L \quad 、 \quad X_2 = (m_2, w_2)_L$$

としたとき、本研究では距離$d = |X_1 - X_2|$を次式で与える。

$$d = \sqrt{(m_1 - m_2)^2 + (w_1 - w_2)^2} \tag{5}$$

上式の定義に従えば、$d = 0$となるのは$m_1 = m_2$かつ$w_1 = w_2$の場合に限定されることになり、たとえ中心m_1とm_2が完全に一致しても幅が異なれば$d > 0$となる。

現在までの二つのファジィ数の近さの考え方では、m_1とm_2が完全に一致すれば、その一致度が最も高くなるというのが従来の考え方である。これに対し本研究での定義では$w_1 \ne w_2$であれば依然として二つのファジィ数には距離があるとする立場をとる。

3.2 モデル

2と同じように(Y_j, x_j)なるデータがN組あるとしよう。このとき、(1)式の右辺を$\hat{Y}_j$とすると、

$$\hat{Y}_j = \left(x_j^T \alpha, c^T |x_j|\right)_L$$

となるから、Y_jと$\hat{Y}_j$の距離の2乗和を最小にするという基準の下にα、cを決定する方法を考えるものとすれば、ファジィ最小二乗モデルは次のように定式化される。

$$\min_{\alpha,c} \sum_{j=1}^{N} \left\{ (y_j - x_j^T \alpha)^2 + (e_j - c^T |x_j|)^2 \right\} \tag{6}$$

この問題は変数αとcによって完全に分離されているから、結局(10)式は

$$\min_{\alpha} \sum_{j=1}^{N} (y_j - x_j^T \alpha)^2 \tag{7}$$

$$\min_{c} \sum_{j=1}^{N} (e_j - c^T |x_j|)^2 \tag{8}$$

なる二つの問題を別々に解くことと等価となる。

(7)、(8)式で表わされる最小化問題はそれぞれ(y_j, x_j)のデータがN組ある場合及び(e_j, x_j)のデータがN組ある場合の通常の最小二乗法の問題に一致することから、解法も通常の方法が適用できること

は自明である。

本方法の最大の特徴は、データY_jにファジィ性がなくなって$e_j \to 0$となり、したがって$c=0$とするとき、(7)式の最小化問題のみが残り通常の重回帰分析に一致するということである。これに対して従来の方法は$c=0$とすれば問題自体が成立しない。

4. 乱数を用いたモデルによる有効性の検討

4.1 データの作成

まずN個の正規乱数を、説明変数x_{ji}、誤差εそれぞれについて発生させ、係数a_{ji}については、三角形に従う乱数として発生させる。次にj番目の乱数の組で$y_j = a_{j0} + a_{j1}x_{j1} + \cdots + a_{jn}x_{jn} + \varepsilon$を満たすようなに$y_j$を求める。これをN組すべて行い、$y_j$の平均$\bar{y}$・標準誤差$s$を求める。同様のことを多数回実行する事により$\bar{y}$、$s$の関係（$s = F(\bar{y})$）を見出す。

再び乱数を発生させ、説明変数x_{ij}、それらより得られたy_jと$s = F(\bar{y})$を満たすsを求めファジィ数Y_j（$Y_j = (y_j, bs)_L$、bは定数）を決定する。

以上の手順を繰り返す事によりファジィ数である独立変数Y_j、非ファジィ数である説明変数x_{ij}を得る。

4.2 モデルによる推定及び結果の比較

4.1 で作成したデータを用いてファジィ線形回帰モデルとファジィ最小二乗法モデルによる推定を行ない、推定パラメータ$\hat{A}_i$を求める。そして、それぞれのモデルで得られた推定パラメータ$\hat{A}_i$が係数a_{ji}の三角形にしたがう分布にどれほど近いかを検証する。

次に再度 4.1 の手順を行ない、N個のy_jを得る。また推定パラメータ$\hat{A}_i$の中心$\hat{\alpha}_i$を用いて（9）式を満たすような$\hat{y}_j$をN個求め、y_jと$\hat{y}_j$から誤差指標をもとめ、整合性の検討を行なう。

$$\hat{y}_j = \hat{\alpha}_0 + \hat{\alpha}_1 x_{j1} + \cdots + \hat{\alpha}_n x_{jn} \qquad (9)$$

なおファジィ線形回帰モデルでは上限と下限の二つの値が存在するので、二つのファジィ数の中心の平均値を$\hat{y}_j$のかわりに用いた。

また、変数の数や、乱数発生時の平均及び分散の値を変化させて同様の手順で分析を行なった。その結果をまとめたものが表 1, 2であるが、ここでは、変数が 3 個、乱数発生時の分散の値が大と小の場合について示している。本表において、RME は平均相対誤差、WRE は重み付き平均相対誤差、MAE は平均絶対値誤差、RMSE は平方根平均二乗誤差、ηは一致係数、ρは相関係数を表わしている。この表より、すべての指標に関して提案方法の方が優れていることがわかる。

様々な条件による結果は講演時に示す。

表 1-1（分散－大）　推定パラメータ$\hat{A}_i$の比較

	従来法	新方法
RME	0.918	0.757
WRE	1.356	0.541
MAE	3.364	1.465
RMSE	6.2291889	2.2866642
η	0.614	0.818
ρ	0.9326807	0.9934584

表 1-2（分散－大）　y_jと$\hat{y}_j$の比較

	従来法	新方法
RME	0.180	0.170
WRE	0.213	0.204
MAE	8.661	8.155
RMSE	10.714643	10.163218
η	0.891	0.898

表 2-1（分散－小）　推定パラメータ$\hat{A}_i$の比較

	従来法	新方法
RME	2.789	1.994
WRE	2.935	2.086
MAE	7.616	5.763
RMSE	16.322	9.880
η	0.370	0.509
ρ	0.877	0.984

表 2-2（分散－小）　y_jと$\hat{y}_j$の比較

	従来法	新方法
RME	0.169	0.166
WRE	0.204	0.200
MAE	8.144	7.905
RMSE	10.104	9.793
η	0.896	0.900

5. むすび

本研究では、線形計画法を用いたいわゆるファジィ線形回帰モデルに対し、最小二乗法を用いた新たな方法を提案した。そして乱数を用いて従来モデルと新たなモデルによる比較を行いその有効性を検証した。

参考文献

1) 寺野他：ファジィシステム入門、オーム社
2) 奥谷・高瀬：ファジィ最小二乗法の適用性について、土木学会中部支部研究発表会講演概要集Ⅳ-72.1997.
3) 奥谷・松浦：ファジィ線形回帰における最小二乗モデル、第１２回信州地区計測制御研究講演会講演論文集.pp.25~28,1998.