

実データによるファジイ最小二乗法の有効性の検証

信州大学工学部 正会員 奥谷巖

信州大学工学部 正会員 高瀬達夫

信州大学工学部 ○小林慶昭

1. はじめに

ファジイ線形システムによる推定方法として、線形計画法を応用した方法、いわゆるファジイ線形回帰分析法が提案されているが、本研究では、観測値と推定値の幅と中心のずれの和を最小にすることとし、その手法として最小二乗法を用いた方法を提案する。実際のデータを用いて予測を行い、他の理論と比較しながら適用性を検討するものである。

本研究では以下の項目についてファジイ最小二乗法とファジイ回帰分析により解析し、相関係数等の値を用いて比較、検討する。

(1) リモートセンシングデータ（神奈川県小田原市周辺の2 km×8 kmの長方形区域における）を用いての土地被覆分類の推定

(2) 長野市統計書による数年分の転入人口、製品出荷額、道路実延長増分、婚姻件数のデータを用いての転入人口の推定

(3) 長野市清掃センターの調べによる平成8年4月分のゴミの種類別排出量、長野市統計書による地区別の総人口、労働力人口、業種別の就業者人口のデータを用いてのゴミの排出量を推定

2. ファジイ最小二乗法

ファジイ数 Y_i の観測値と推定値とのずれの和を最小にすることを、幅と中心のずれの二乗の和を最小にすることと考え、最小二乗法を用いた解法である。

従属変数と説明変数がデータとして与えられた場合 N 組のファジイデータ (Y_i, x_i) 、 $i=1, \dots, N$ 、 n 個の説明変数によるファジイ線形関数を

$$\hat{Y}_i = A_0 + A_1 x_{i1} + A_2 x_{i2} + \dots + A_n x_{in}$$

$$\hat{Y}_i : \hat{Y}_i = (x_i^T \alpha, c^T |x_i|)_L \text{ のファジイ推定値}$$

$A_j : A_j = (\alpha_j, c_j)_L$ 、 $j=0, \dots, n$ のファジイ係数とし、 Y_i と $\hat{Y}_i$ 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るれば、ファジイ最小二乗モデルは次のように定式化さ

れる。

$$\min_{\alpha, c} \sum_{i=1}^N \left\{ (y_i - x_i^T \alpha)^2 + (e_i - c^T |x_i|)^2 \right\} \quad (1)$$

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、この式は

$$\min_{\alpha} \sum_{i=1}^N (y_i - x_i^T \alpha)^2 \quad (2)$$

$$\min_c \sum_{i=1}^N (e_i - c^T |x_i|)^2 \quad (3)$$

なる二つの問題を別々に解くことと等しくなる

(2)、(3)式で表わされる最小化問題はそれぞれ (y_i, x_i) のデータが N 組ある場合及び (e_i, x_i) のデータが N 組ある場合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_i にファジイ性がなくなつて $e_i \rightarrow 0$ となり、したがって $c=0$ とすると、(2)式の最小化問題のみが残りの通常の重回帰分析に一致するということである。

ここで注意しなければならない事は、式(3)を解く場合、つまり、幅 c を求めるとき、線形重回帰モデルの解法や最小二乗法では、すべての解が必ずしも $c \geq 0$ となる保証がなく、ファジイ数の幅は正、もしくは0という定義に反する可能性がある。そこで本研究では、幅 c に負の値が存在する場合も想定し、次に示す二次計画法を用いて幅 c を求めている。

$$\min_c \sum_{i=1}^N \left\{ e_i - (c_0 + c_1 |x_{i1}| + c_2 |x_{i2}| + \dots + c_n |x_{in}|) \right\}^2$$

(4)

$$c \geq 0 \quad c_1 + c_2 \geq 0 \quad (5)$$

式(4)は、最小二乗法による解法の式(3)と同じであるが、ここでは、この目的関数に対して、式(5)のように幅 c が正、もしくは0という制約条件式と、人為的に $c_1 + c_2 \geq 0$ という2つの制約条件式を与え、二次計画法により幅 c を求めている。ここで与えた $c_1 + c_2 \geq 0$ の制約条件式は、 $c \geq 0$ ならばそれが成

立することは明白である。これを最小にする c を求めるのがここでの問題である

3. リモートセンシングを用いた土地被覆分類の推定への応用

<4.1>カテゴリ分解

従来の推定方法では、土地被覆カテゴリにおけるバンドデータがそれぞれカテゴリ固有の多次元正規分布に従うという前提のもとに、与えられたデータに対する確立が最大になるカテゴリを選び出すという方法を採用しており、一画素は必ず一つのカテゴリに分類されることになっていた。しかし、開発が進展している市街地等では、一画素に対応するような小区域であっても複数のカテゴリが存在する可能性が高く、従来法では不具合が発生する。そこで本研究では、各画素は基本的に複数カテゴリから構成されているという前提のもとに、その混在比を推定するというカテゴリ分解原理に基礎を置いている。

<4.2>土地被覆状態の推定法

本研究ではカテゴリ分光特性にファジィ集合の概念を導入した点が特色になっている。

土地被覆状態の推定は、テストエリアにおいて $200\text{m} \times 200\text{m}$ の基本メッシュ単位で $k=1,2,\dots,K$ の合計 K 個の小区域に分割して行うものとする。ファジィ最小二乗法及びファジィ回帰分析を用いる方法ではデータとして推定メッシュのバンド値の平均値 $y_j(k)$ と標準偏差 $q_j(k)$ を与え当該メッシュのカテゴリ別面積割合を一気に求めるが、比較に用いる従来法では、分光特性値として各カテゴリの代表区域のデータから求めたものを用い、基本メッシュの土地被覆状態はメッシュ内のピクセル単位で推定した結果を集計して求める。

ここでのファジィ最小二乗法は m_{ji} = カテゴリ i のサンプルデータのバンド j の平均値

$w_{ji} = \alpha \cdot (\text{同上データの標準偏差})$

$$F = \sum_{k=1}^K \sum_{j=1}^N \left\{ \left[y_j(k) - z(k)^T m_j \right]^2 + \left[q_j(k) - z(k)^T w_j \right]^2 \right\}$$

$$\rightarrow \min_z \quad (6)$$

s.t. $z_1(k) + z_2(k) + \dots + z_m(k) = 1$
 $z_i(k) \geq 0 \quad (k=1,2,\dots,K; i=1,2,\dots,M)$ (7)
 より、カテゴリ別面積割合 z_i を求めるものである。

<4.3>結果の比較

表1は土地被覆カテゴリを5項目とした場合の推定制度を6つの誤差指標について比較整理したものである。本表より、あらゆる指標に関してファジィ集合の概念を用いた方法が判別分析法及び最尤法を凌駕しているが、ファジィ最小二乗法とファジィ回帰分析を比較すると、若干最小二乗法が勝っているが、それほど大きな差はないようである。表2は分類項目別に推定制度を計算し整理したものである。ただし、推定誤差指標としては RMSE を代表的指標として採用した。5項目の中、農地、果樹園、森林、密集地の4項目でファジィ概念を用いた方法が従来法より良い結果を出している。しかし、水域については従来法と比べて遜色がある。また、ファジィ最小二乗法とファジィ回帰分析とを比較すると、ここでも大きな差は見受けられない。

表1 種々の誤差指標に基づく提案法と既存法の比較

誤差指標	ファジィ最小二乗	ファジィ回帰分析	判別分析法	最尤法
RME	0.5539	0.5836	0.7260	0.7373
WRE	0.6708	0.5947	0.9629	1.1628
MAE	0.1227	0.1276	0.2120	0.2663
RMSE	0.1264	0.1276	0.1831	0.2119
η	0.8030	0.8122	0.7554	0.7208
ρ	0.8732	0.9059	0.8387	0.7941

表2 RMSE による分類項目別制度比較

分類項目	ファジィ最小二乗	ファジィ回帰分析	判別分析法	最尤法
水域	0.0561	0.1363	0.0487	0.0499
農地	0.1307	0.1037	0.1652	0.1341
果樹園	0.1253	0.1100	0.3163	0.3647
森林	0.1841	0.1518	0.2588	0.3257
密集地	0.1007	0.1303	0.1754	0.2005
合計	0.1264	0.1276	0.2133	0.2449

4. おわりに

紙面上の都合によりゴミ問題、転入人口の推定について扱えなかったため、本稿についての詳しい説明とともに講演時に発表したい。ランドサットデータに誤差を含む場合についても講演時に発表する予定である。