

IV-215 クラスター分析と最尤法の組み合わせによる土地被覆分類

九州大学 工学部 学生員 ○入田健一郎
九州大学 工学部 学生員 杉野 浩茂
宮崎大学 工学部 正 員 出口 近士

1. はじめに

M E S S Rにより収集したデータから土地被覆分類図を作成する方法にクラスター分析を用いる方法がある。これを用いる場合にも、さまざまな誤判別が生じるが、なかでも「山林」など空間的にまとまりの大きい被覆領域の中に、小さな誤判別画素が点在する現象が生じる。そこで、本研究ではクラスター分析に最尤法を組み合わせることによって、この種の誤判別を減少させる手法を提案する。

2. 解析方法

解析手順を図-1のフローチャートに示す。

リモートセンシングデータの解析には、「教師なし分類」の一つであるクラスター分析を用いる場合と、「教師あり分類」の一つである最尤法を用いる場合がある。本研究では、クラスター分析による解析の後に、最尤法による解析を行った。一般に最尤法を用いる場合には、教師データの抽出に労力と時間がかかるという問題が生じる。しかしながら、クラスター分析と最尤法を連結することによって、最尤法を用いる際の教師データの抽出の問題を解消するのである。

（1） クラスター分析

クラスター分析は複数の計測値によって類似性をもとに「似たもの同士」を集める手法である。この場合、複数の計測値とは各バンドのCCTカウントのことである。本研究では、類似性の指標として、個体間の距離をユークリッド距離で定義した。

クラスター分析を行う場合、個体間の距離をすべて計算しなければならず、データ数が増えると計算量は指数関数的に増加する。そこで計算量を軽減するために、まず標本域から機械的にサンプルデータを抽出してクラスター分析を行った。その後、標本域からのデータを最近隣法を用いていずれかのクラスターへ割り付けるという方法を採用した。

以下にユークリッド距離と最近隣の定義を示す。

① ユークリッド距離

$$d_{ab}^2 = \sum_{i=1}^p (x_{ai} - x_{bi})^2 \quad \cdots (1)$$

a, b: 個体番号

i: 特性（バンド番号）

p: 特性の数

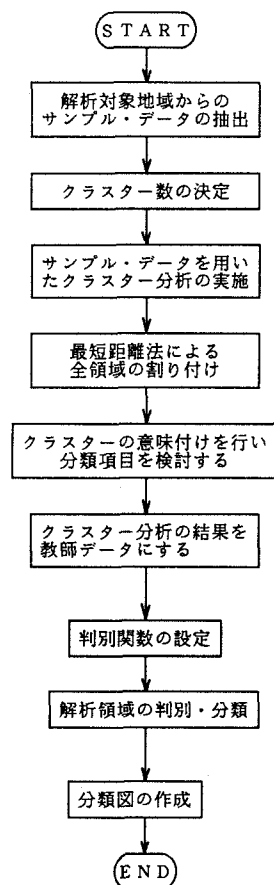


図-1 解析手順

② 最近隣

$$\text{Min} \{ d_{ab} \mid a \in c_f, b \in c_g \} \quad \cdots (2)$$

f, g : クラスター番号

それぞれのクラスターに属する個体間距離の最小値

(2) 最尤法

クラスター分析の結果を教師データにして最尤法を行う。最尤法は、母数の推定値を求める一般的な手法である。いま、解析対象が g 個の母集団から成り立っているとし、その g 個の母集団は p 変量 $\mathbf{x} (x_1, x_2, \dots, x_p)$ の確率密度関数 $f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_g(\mathbf{x})$ で与えられるものと仮定する。測定値 $\mathbf{x}_g$ が得られたとき、この $\mathbf{x}_g$ がどの群に属するかを判定する場合、確率密度関数を用いて $\mathbf{x}_g$ が「群に属する確率が最大となる群に属する」と判定するのである。確率変量 $\mathbf{x}$ が p 変量正規分布に従うとき、判定の基準となる判別関数（これは確率密度関数から導出された。）は第 k 群について

$$S(k) = \text{Ln}(\pi_k) - \frac{1}{2} \text{Ln} |\Sigma_k| - \frac{1}{2} (\mathbf{x} - \mu_k)^t \Sigma_k^{-1} (\mathbf{x} - \mu_k) \quad \cdots (3)$$

π_k : 事前確率 μ_k : 平均値 Σ_k : 母分散・共分散行列 ($k = 1, 2, \dots, p$)

を用いた。この式の $S(k)$ の値が最大となる群に未知データ $\mathbf{x}$ が分類されることになる。

3. 解析結果の一例

解析結果の一例を図-2・図-3に示す。図-2はクラスター分析のみで得た福岡市東部地域の山地部における「山林」の分布を示したものである。図-3はこれに最尤法を追加して再処理したもので、後者は黒いまとまりの中に点在する誤判別画素（白抜き）が減少していることが分かる。

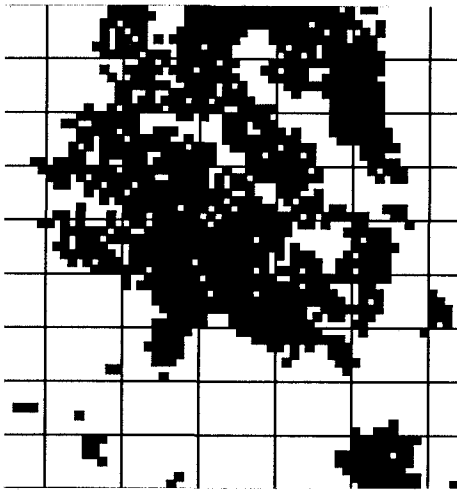


図-2 クラスター分析のみによる
土地被覆分類図

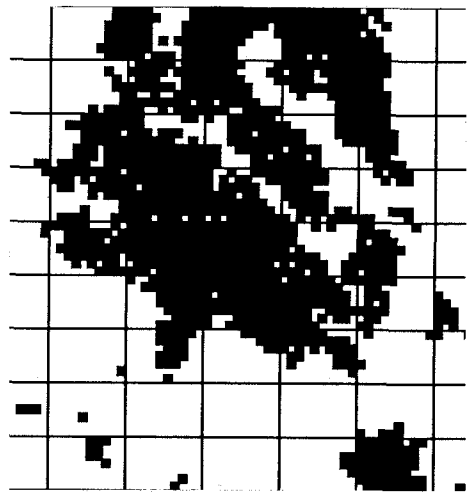


図-3 2法の組み合わせによる
土地被覆分類図

4. まとめ

クラスター分析のみを行った場合と、クラスター分析と最尤法を組み合わせた場合とのそれぞれの土地被覆分類図を航空写真等を基準に比較した。「山林」部分に注目すると、クラスター分析と最尤法を組み合わせた場合の方が、現状に近い土地被覆分類図であることが分かった。

このように、一連の解析手順の中に最尤法を導入し、クラスター分析と最尤法の組み合わせを図って、「誤判別の減少」という目的をある程度達成したと考えられる。