

IV-172

星座法の判別分類の一方法

筑波大学 正員 ○ 星 仰
 筑波大学大学院 西村 公司
 筑波大学 正員 池辺八州彦

1) はじめに

リモートセンシングの画像データによる解析上で問題となることは多変量で、かつデータ量が膨大であることである。変数を減少させるには変数間の相関をとり類似変数の削除や重回帰分析による説明変数の選択法、あるいは主成分分析による主要主成分の採用などが考えられる。リモートセンシング画像データの場合、これらのいずれの方法を採用しても3変数以上になることが多い。分析者にとってデータの性状を充分把握可能にするには、空間もしくは平面上にデータを表現することが好ましい。そこで本研究では多変量の平面投影法である星座法を取り上げる。星座法による星座グラフは、多変量のデータを平面に表現でき¹⁾、かつデータの散布状態も視覚的に理解しやすい。この星座法は1978年にリモートセンシング画像データに応用されている²⁾。その後、主成分分析による星座法の有効性が具体的に示された³⁾。しかし、星座法に対する判別分類の基準については、曖昧である。そこで本稿では、リモートセンシング画像データをサンプルデータとして、円型・楕円型・扇型の判別の基準について例示して、これらの分類精度について検討する。

2) 星座法とその判別基準

星座法とは変量 X_j ($j=1\sim m$) を夾角変量 η ($0\leq\eta\leq\pi$) に線型変換して、図-1のように各変量を原点移動式の極座標で表現する。最終到達位置とその軌跡がデータの特性を表わす結果となる。図-1のAは判別基準を円型で示したもので、半径 R_k は誤差伝播の法則に従って、各変量の標準偏差 σ_{jk} より求められる。ただし、 i はサンプル数 ($i=1\sim l$)、 j はチャンネル数 ($j=1\sim m$)、 k は分類項目数 ($k=1\sim n$) である。

$$\sigma_{jk}^2 = \frac{1}{n-1} \cdot \sum (\eta_{ijk} - \mu_{jk})^2 \dots\dots\dots (1)$$

$$R_k^2 = \beta \cdot \{ \sigma_{1k}^2 + \sigma_{2k}^2 + \dots + \sigma_{mk}^2 \} \dots\dots\dots (2)$$

図-1のBは楕円にて判別基準を示したもので、その基準の決定方法は最終到達位置Bと原点Oを結ぶ方向OBに上記の円基準の誤差を投影した成分 v_{rjk} を誤差伝播の法則に従って加算し、楕円の短（長）径 b_k を決定するものである。楕円の長（短）径 a_k はOBに垂直な成分である。これらの関係は式(3)となる。

$$\left. \begin{aligned} H a_k^2 &= \beta \cdot \{ H r_{1k}^2 + H r_{2k}^2 + \dots + H r_{mk}^2 \} \\ v b_k^2 &= \beta \cdot \{ v r_{1k}^2 + v r_{2k}^2 + \dots + v r_{mk}^2 \} \end{aligned} \right\} \dots\dots\dots (3)$$

図-1のCは主成分分析をかけ、目標の累積寄与率に到達したところまでの主成分によって統計量をもとめ、星座グラフ上にサンプルを表示するもので、各分類項目の平均値 η_k をソートし、各平均値間で判別領域を決定する。ここでは最も簡易的な標準偏差の比を用いる。いま、分類項目 n と $n+1$ の判別域を求めるために θ_1 と θ_2 を次式(4)のように定める。

$$\theta = \theta_1 + \theta_2 = \frac{\sigma_n}{\sigma_n + \sigma_{n+1}} \theta + \frac{\sigma_{n+1}}{\sigma_n + \sigma_{n+1}} \theta \dots\dots\dots (4)$$

ただし、 θ は各主成分によってもとめられた分類項目 n と $n+1$ の最終到達位置をはさむ夾角である。

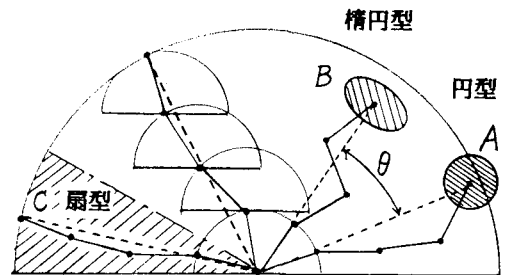


図-1 星座グラフと判別基準

3) 航空機MSS画像データの適用とその結果

教師付き分類ではトレーニング地区データを教師としているので、このデータを判別分類のサンプルデータとして用い、かつ分類の評価にも用いる。MSS画像データは航空機の11chで、分析地区60×60画素の中から、k=8種類の分類項目を選定した。その内容はつぎの通りである。

- 【i】送電線の直下地 ○ 【iv】アメリカ・スギ × 【vii】天然林（樹高31～40m）X
 【ii】アメリカ・ツガ △ 【v】針 葉 樹 ◇ 【viii】天然林（樹高40m以上）Z
 【iii】森林開拓地 + 【vi】アメリカ・マツ △

航空機MSS画像データ（1画素8bit）の各分類項目別に統計量を求め、平均値を星座法でプロットしたものが図-2～図-5である。図-2は原データを表現したものである。図-3はデータ域を十分に活用する目的で、正規化データを表示したもので、その効果が現れている。このときの判別基準は図-1のAに相当する円型である。図-3は図-1のBに相当するもので、楕円を判別基準にしている。これらの正規化データでも類似パターンが多いと隣接する分類項目間ではデータの重複が見受けられる。そこで、正規化データに主成分分析をかけ、累積寄与率が99.5%を越える第4主成分（表-1を参照）まで採用することにし、これを星座グラフで描写したのが図-5である。図-5の左すみでは分類平面が狭くなり誤分類を生じやすくなっているが、中央右の3項目は誤分類を生じにくい性状を呈している。図-5における判別基準は式（4）により求められる。

4) 結論と今後の課題

トレーニング地区のデータを分類項目ごとに上記の三種類の判別基準で分類したところ、円型基準では68.7%、楕円基準では74.9%、主成分分析の扇型基準では78.6%となった。以上の分類では、8種類の分類項目以外に“その他の分類項目”を考慮せざるを得なかった。しかし、主成分分析の扇型の基準では、いずれの分類項目にも属さない“その他の項目”を取り扱わないことも可能である。この方法により分類したところ、81.5%の分類精度を得た。この結果から、主成分分析の扇型の基準はデータの密集域では分離効果があるといえよう。しかし、星座グラフの左すみのように精度が低下する部分もあるので、この部分の処置については今後の課題といえよう。

参 考 文 献

- 1) 脇本和昌：“Graphical Method による多次元データ解析”，文部省特定研究，pp. 3～4，1977.
- 2) 星 仰：“リモートセンシングによる都市域の地物判定に関する研究”，pp. 86～93，1978.2.
- 3) 星 仰、西村公：“リモートセンシング画像データを用いた星座法による判別・分類法、情報処理学会第34回全国大会講演論文集、III-1D-8，pp.1693～1694，1987.3.

表-1

主成分	累積寄与率
1	84.29
2	94.02
3	97.20
4	99.59
5	99.81
6	99.86
7	99.90
8	99.93
9	99.96
10	99.98
11	100.00

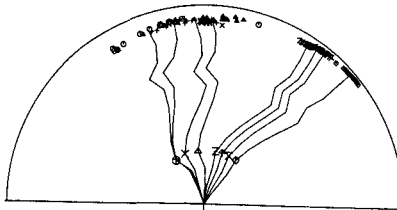


図-2 原データの性状

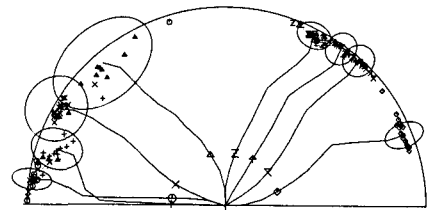


図-4 正規化データ（楕円型基準）

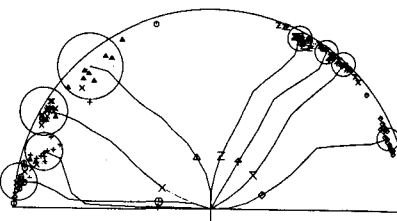


図-3 正規化データ（円型基準）

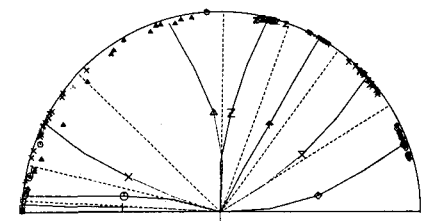


図-5 主成分データ（扇型基準）