USAS手法による最大尤度法適用時に対する一考察

和歌山工業高等専門学校　正員　星　仰
建設省近畿地方建設局　正員　○福谷晋三

1．はじめに

リモートセンシングのMSSデータ解析に適用されているUSAS手法における教師つき分類法において，最大尤度法を採用した場合に，あるクラスのある変数において分散が零となり，分類図を作成することができなかった事があった。このような場合には本来最大尤度法を使用すべきではない。しかし，ある地域を新しく分析するときには，当初なんらかの分類図をなるべく正確にかつ迅速に求めたい。このような場合には最大尤度法がよく用いられてきている。だが，上記のような分散が零となる場合には他の分類手法を適用せねばならない。このことは少なくとも2段階処理となる。そこで，本研究はこれらの点を考慮して，各クラスの各変数の分散が零でない時は最大尤度法を，その他の時は分散零の代りに全クラスの変数の分散の最小値(COVNUM)を与えるアルゴリズムを作成した。このアルゴリズムによれば，1回の処理によって，必らず分類図を作成できる利点がある。

2．最大尤度法

最大尤度法はベイズの方法によって誤判別の損失を最小にするように判別関数を決定する方法である。ベイズの方法とは，誤判別の損失期待値 $r_g(x)$ が最小となるなら，パターン x はグループ G_g に属すると判別するものである。

いま，$L(g|h)$ が損失関数，$P(x|G_g)$ が確率密度関数，$P(G_g)$ が事前確率であるとき，式(1)によって損失期待値 $r_g(x)$ が与えられる。損失関数は通常，式(2)のように取り扱ってもよいからベイズの方法によりすべての $g \neq h$ に対して式(3)が成立するときは，パターン x は G_g に属すると判定する。

$$r_g(x) = \sum_g L(g|h)\cdot P(x|G_g)P(G_g) \qquad \cdots\cdots (1)$$

$$L(g|h) = 0 \quad (g=h), \quad L(g|h) = 1 \quad (g \neq h) \qquad \cdots\cdots (2)$$

$$P(x|G_g)P(G_g) \; > \; P(x|G_h)P(G_h) \qquad \cdots\cdots (3)$$

ここで式(3)の左辺により判別関数を求めると，確率密度関数 $P(x|G_g)$ が多変量正規分布すると考えるから，その対数をとり g に無関係の項を省略すると判別関数 $y_g(x)$ は式(4)となる。ここで，$x=\{x_1, x_2, \cdots, x_p\}$，$\mu_g=\{\mu_{1g}, \mu_{2g}, \cdots, \mu_{pg}\}$，$\Sigma_g$ = 分散共分散行列とする。

$$y_g(x) = \log_e P(G_g) - \frac{1}{2}\log_e|\Sigma_g| - \frac{1}{2}(x-\mu_g)'\Sigma_g^{-1}(x-\mu_g) \qquad \cdots\cdots (4)$$

パターン x は式(4)によって得られる $y_g(x)$ $(g=1, 2, \cdots, n)$ の値が最大となるグループに割り当てられる。

3．分散を与える改良型最大尤度法プログラムの構成

改良型最大尤度法プログラムは主プログラムと6個のサブルーチンで構成されている。

1)主プログラムMAXLM；トレーニング地区の各クラスの切り出し，各クラスの平均，標準偏差，分散共分散行列等の印刷，および分析地区の分類図の作成などを実行する。

2)サブルーチンSTATS；トレーニング地区データの読み込みとその平均，標準偏差，

分散共分散行列の計算を行なう。また，COVMATにより分散を与えるか否かを判定する。

3)サブルーチンDETERM；分散共分散行列の行列式を計算する。ここで，分散共分散行列が非正則行列の場合，COVMATにより行列式の値を与えるか否かを判定する。

4)サブルーチンMATINV；分散共分散行列の逆行列を計算する。ここで，分散共分散行列が非正則行列の場合，COVMATにより計算を続行するか，中止するかを判定する。

5)サブルーチンFNDMIN；分析地区の各画素単位のデータの判別得点の最小値とそのデータの割り当てられるクラスを見い出すサブルーチンである。

6)サブルーチンLNSKIP；磁気テープ内のデータの中から分析地区の先頭ラインまでを空読みしたり，ライン方向の間引きをするためのサブルーチンである。

7)サブルーチンCCCTBU；磁気テープのデータIA(8ビット)を2次元配列のデータIH(16ビット)にフォーマット変換するサブルーチンである。

4. 数学モデルによる実験結果

このMAXLMプログラムを実際のMSSデータに適用する前に数学モデルによる実験を行ない，その稼動状況を検討することにした。本研究ではクラスの変数xの分散が零であるモデル実験を行なうために，表−1，表−2の平均および標準偏差を有するデータを作成した。この表よりCLASS 2のCH 1の標準偏差が零になっていることがわかる。この標準偏差の値として，この実験モデルの標準偏差の最小値である(0.379)をCOVNUMに代入した。こうして得られた分類図を図−1に示す。ただし，この図で□で囲まれた領域は誤分類された領域を示す。

5. おわりに

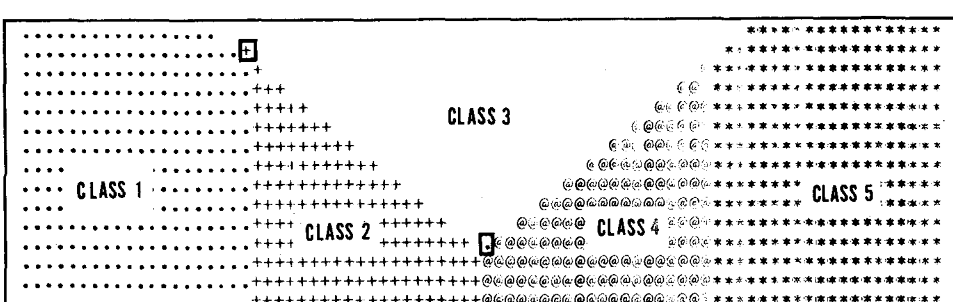
本研究では分散が零の場合でも，最大尤度法を用いて分類図を作成することを目的としてきた。この意味において数学モデルによる実験によってその目的は果されたといえる。しかし，分散の与え方については，その方法について研究の余地があると考える。この点については今後の研究課題としたい。

表−1 各クラスのチャンネル別平均値

	CLASS 1	CLASS 2	CLASS 3	CLASS 4	CLASS 5
CH 1	3.3	0.0	0.313	3.2	6.6
CH 2	6.2	5.9	7.1	0.32	0.5
CH 3	6.2	0.36	3.0	0.2	0.3
CH 4	0.64	5.3	0.137	4.8	4.4
CH 5	0.4	7.2	0.175	7.0	0.22
CH 6	0.3	0.36	5.7	0.2	0.24

表−2 各クラスのチャンネル別標準偏差

	CLASS 1	CLASS 2	CLASS 3	CLASS 4	CLASS 5
CH 1	0.9	0.0	0.644	1.077	0.8
CH 2	0.872	1.136	0.539	0.508	0.755
CH 3	0.872	0.625	1.0	0.529	0.5
CH 4	0.742	0.9	0.379	1.249	0.917
CH 5	0.663	0.98	0.468	1.095	0.576
CH 6	0.608	0.52	1.005	0.566	0.618



図−1 改良型最大尤度法プログラムによる実験モデルの分類結果