

针茅草原放牧衰退演替阶段 的模糊聚类分析^{**}

李晓军^{**} 赵松岭 张鹏云 陈庆诚

(兰州大学生态学研究室)

王久丽

(北京林业大学森林生态学研究室)

摘 要

本文利用模糊聚类分析的系统聚类法和 ISODATA 分类法,对短花针茅草原因过度放牧而形成的衰退演替系列的各阶段进行了分类,比较成功地在群落分类上联合使用了 PCA 排序与 ISODATA 分类。结果表明,这两种方法都将衰退演替系列分成4个阶段,它们是短花针茅群落、扁穗冰草群落、冷蒿群落和亚氏庭花群落。这与前人用其他方法分类的结果相似。本文为群落分类提供了两种合理、可行的数学分类方法。

关键词: 模糊聚类, 衰退演替, 分析。

广泛分布于我国西北黄上高原与荒漠山地的短花针茅草原,是优良的天然放牧场,同时也是欧亚大陆草原区中的一个重要草原群系。由于长期过度放牧,引起短花针茅草原出现明显衰退现象,严重制约着畜牧业的发展。因此研究短花针茅草原衰退演替动向,对放牧衰退演替阶段进行分类研究,不仅具有理论上的意义,同时对提高畜牧业生产和维持草地生态系统平衡也具有实际意义。

随着放牧强度的增加,针茅草原衰退演替表现为群落中各种群优势度的强烈变化,从而形成不同的群落类型。它们在空间上呈现出随机镶嵌体的形式,各阶段之间又有明显的过度性类型。这给类型划分带来很大的困难。常规的数量分类方法在植物群落分类中的应用常常受到各群落间大量存在的过度类型的影响^[1],以致其应用显得机械和生硬。已报道的模糊数学分类虽然能解决上述问题,但对于处理大量的样方资料又显得过于繁杂。模糊聚类分析却具备上述两种方法的优点,而且其中的 ISODATA 模糊聚类分析又可将排序与分类联合使用,亦即在排序的基础上再进行分类。

作者于1982年7至8月在甘肃省天祝羊场调查了当地天然植被的放牧衰退状况,应用两种模糊聚类分析方法对植物群落演替阶段进行了分类。

* 熊韶峻、李文朝、罗佩华同志协助收集野外资料。

**现在北京林业大学森林生态学教研室。

本文于1986年4月26日收到。

一、研究地点及取样方法

甘肃省天祝藏族自治州天祝羊场位于祁连山东段的山间平原, 海拔 2,000 米左右, $\geq 10^{\circ}\text{C}$ 年活动积温 1,500—2,000 $^{\circ}\text{C}$ 左右, 年降水量 250—260 毫米。土壤为山地栗钙土, 分层明显, 土层平均厚度 90—100 厘米。当地区地带性植被是短花针茅 (*Stipa breviflora*) 草原, 其种类组成单纯。据 308 个样方的统计, 共有种子植物 25 种。由于长期过度放牧的影响, 出现明显的衰退演替迹象。一些研究认为, 本地区放牧衰退演替的顺序是, 短花针茅群落阶段 \rightarrow 扁穗冰草群落阶段 \rightarrow 冷蒿群落阶段 \rightarrow 亚氏旋花群落阶段^[2]。在调查地区, 目前 4 种群落呈随机的镶嵌分布, 边界不明显。其中以冷蒿群落占优势, 说明草场退化已经相当严重。

抽样调查采用随机取样法, 样方面积为 1 平方米, 分别记录各个种的高度、盖度和频度。然后分别计算各样方中每个种的重要值 (重要值 = 相对高度 + 相对盖度 + 相对频度)^[3]。上述 4 种种的群落学指标可以反映出各个种在群落中的作用和地位。

二、聚类分析过程及结果

选取针茅草原中 15 个主要种作为群落分类的指标, 记为 U_i , (其中 $i=1, 2, 3, \dots, 60$ 表示样方, $j=1, 2, 3, \dots, 15$ 表示种) (见表 1)。这样就去掉了一些偶见种在分类过程中的

表 1 群落聚类分析的特征指标

Table 1 The indexes of the community classification

种	类	记号	样方
短花针茅	<i>Stipa breviflora</i>	U_{i1}	$U_{11}, U_{21}, \dots, U_{601}$
扁穗冰草	<i>Agropyron cristatum</i>	U_{i2}	$U_{12}, U_{22}, \dots, U_{602}$
冷蒿	<i>Artemisia frigida</i>	U_{i3}	$U_{13}, U_{23}, \dots, U_{603}$
亚氏旋花	<i>Convolvulus ammanii</i>	U_{i4}	$U_{14}, U_{24}, \dots, U_{604}$
茵陈蒿	<i>Artemisia capillaris</i>	U_{i5}	$U_{15}, U_{25}, \dots, U_{605}$
阿尔泰紫菀	<i>Aster tataricus</i>	U_{i6}	$U_{16}, U_{26}, \dots, U_{606}$
野葱	<i>Allium sp.</i>	U_{i7}	$U_{17}, U_{27}, \dots, U_{607}$
二色棘豆	<i>Oxytropis bicolor</i>	U_{i8}	$U_{18}, U_{28}, \dots, U_{608}$
二裂萎皮菜	<i>Potentilla bifurca</i>	U_{i9}	$U_{19}, U_{29}, \dots, U_{609}$
心芭	<i>Cybaria mongolica</i>	U_{i10}	$U_{110}, U_{210}, \dots, U_{6010}$
狼毒	<i>Stellera chamaejasmae</i>	U_{i11}	$U_{111}, U_{211}, \dots, U_{6011}$
早熟禾	<i>Poa sp.</i>	U_{i12}	$U_{112}, U_{212}, \dots, U_{6012}$
蚓果芥	<i>Torularia humilis</i>	U_{i13}	$U_{113}, U_{213}, \dots, U_{6013}$
青兰	<i>Dracocephalum heterophyllum</i>	U_{i14}	$U_{114}, U_{214}, \dots, U_{6014}$
瓣叶龙胆	<i>Gentiana squarrosa</i>	U_{i15}	$U_{115}, U_{215}, \dots, U_{6015}$

影响。由于受计算机容量限制, 只随机抽取 60 个样方进行分类, 分别计算种的相对优势度值 ($=$ 种的重要值 / 3), 作为数据的标准化处理, 使得 $0 \leq u_{ij} \leq 1$ 。15 个种和各自的优势度值的不同组合能够基本刻划每个样方在演替系列中的位置。这样每个样方都可以被看成是一个 15 维的向量, 即:

$$U_{ij} = (u_{i1}, u_{i2}, \dots, \dots, u_{i15}) \quad (i=1, 2, 3, \dots, 60)$$

全部运算过程用BASIC语言写成, 在PC-1500机上实现。

1. ISODATA模糊分类

ISODATA分类^[6]是一种软分类方法, 其聚类判据为:

$$J_m(R, V) = \sum_{i=1}^n \sum_{j=1}^c (r_{ij})^m \|u_i - V_j\|^2$$

其中 V_j 表示第 j 类的聚类中心; $\|u_i - V_j\|^2$ 表示样本 u_i 与 V_j 的距离平方; c 为分类之组数; R 为预分类矩阵; m 为参数; V 为聚类中心。

ISODATA分类的结果不仅能将样方合理分类, 而且能给出每个样方对于各类的隶属度, 提供更多的样方信息。同时, 随式中参数 m 的不同取值, 反映出各分类组的模糊度的改变状况, 可以进一步了解处于各模糊集交错区样方的归属情况。ISODATA分类, 首先要求给出一个样方的预分类矩阵。不仅要求预先把所取样方进行人为分类, 而且要求提出分类组数。通常研究者都是根据自己的实际经验给出样方的预分类。对于数量较少, 典型取得的样方这种方法是可行的。但是随着样方数量的增加, 取样方法的随机化, 特别是当研究的植物群落随着某一因子的变化而形成植物群落连续体时, 人为给出预分类的方法会遇到困难, 而且结果的准确性较差。这给下一步的模糊聚类分析带来计算上的复杂性, 使运算时间成倍增加。利用排序方法对样方进行预分类可以解决上述困难。群落排序的实质是在损失最小信息的条件下, 在低维空间合理排出样方的空间相对位置。将排序与分类相结合, 会使分类更为有效。根据排序结果, 按照样方在排序图中的自然结合集团预先定出分类组数。从而避免了人为影响, 使得预分类更加合理。

在各类排序方法中, 主分量分析(Principal components analysis, PCA)^[4]是应用最为广泛的一种, 而且具有严格的数学基础。本文利用PCA排序法, 在坐标图上标出各样方的位点, 根据样方的分布给出ISODATA聚类的预分类矩阵。

对于进行模糊聚类的60个样方按照所选出的15个种的相对优势度进行排序。经计算, 得出特征根(即主分量, 按大小排列)及相应每一轴所占有的信息量的百分比(表2)。从表

表2 PCA排序特征根

Table 2 The eigenvalues of PCA ordination

特征根	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	λ_9	λ_{10}	λ_{11}	λ_{12}	λ_{13}	λ_{14}	λ_{15}	合计
特征根值	201.02	117.44	68.13	18.80	12.86	8.42	6.11	5.22	2.54	2.16	1.69	1.21	0.84	0.47	0.00	434.71
占有信息量(%)	46.24	27.02	15.37	4.28	2.96	1.88	1.41	1.20	0.58	0.50	0.39	0.28	0.19	0.11	0.00	100

2 可以看到, 只取前两个主分量就占全部信息量的73.26%, 可见PCA排序降维的效果很好。为了作图方便, 只取前两个主分量算出各样方的坐标, 并依每个样方的坐标做出排序图(图1)。图中表明, 60个样方大致分成4个集团。另外还有少数是过度样方, 这反映了在取样地区各个群落之间存在过度类型。从原始数据分析, 集团(1)属于短花针茅群落, 集团(2)是冷蒿群落, 集团(3)是亚氏旋花群落, 集团(4)是扁穗冰草群落。从图中还可以看到, 短花针茅群落位于 Y_1 轴的上方, 而冰草群落、亚氏旋花群落位于下方, 说明它

们的划分主要决定于短花针茅（它对 Y_1 轴的负荷量最大，为 0.992），同样亚氏旋花决定了冰草群落与亚氏旋花群落的划分。位于两轴中心的冷蒿群落受这 2 个种的影响是同样的。

根据坐标排序图，把样方分成 4 类，集团间的样方根据它到各集团的距离大小归并到 4 个集团。这样，通过 PCA 排序决定了预分类的组数为 4 ($c=4$)，得出样方的预分类矩阵 $R(r_{ij})$ ($i=1, 2, 3, 4; j=1, 2, \dots, 60$) (表 3)。

令 $m=1.5, \epsilon=0.01$ ，对预分类矩阵 R 及原始数据矩阵 U 进行聚类计算。经过计算机多次迭代计算，得到分类矩阵 (略)。

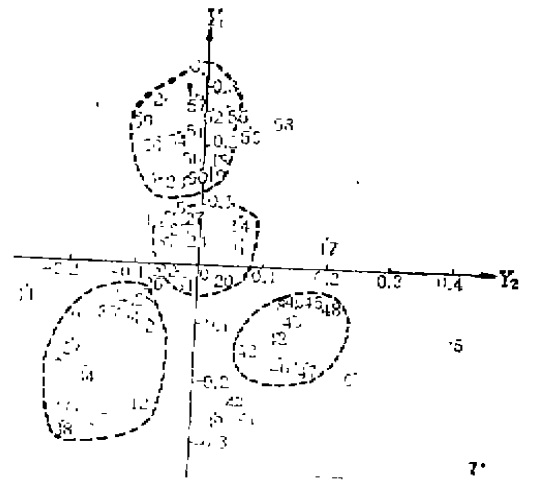


图 1 PCA 排序图
Fig. 1 The map of the PCA ordination

表 3 根据 PCA 排序得到的样方预分类矩阵 (R)

Table 3 The preliminary classification matrix (R) of samples according to the result of the PCA ordination

类	样方																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0
2	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	1
3	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0
类	样方																			
	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	1	1	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
4	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	0	0
类	样方																			
	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
1	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

计算结果也表示出每个样方对于 4 种群落的隶属程度。按照每一列的最大值 (即按照每个样

方的最大隶属度)，可以把60个样方分成4类（表4）。

表 4 ISODATA 模糊分类的结果
Table 4 The result of the ISODATA classification

类	样方																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
3	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0

类	样方																			
	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	1	1	1	1	1	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
4	0	0	0	0	0	0	0	1	1	1	1	0	1	1	1	1	1	1	0	0

类	样方																			
	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
1	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

比较表3和表4，可以看到PCA排序和ISODATA分析的结果是相似的，但是存在一些差异。例如排序把17、18、19、23号样方归为第一类，而ISODATA分析却把它们分到第二类。这是因为二维空间的排序图，歪曲了它们的位置，经过ISODATA分析调正了误差。说明排序和分类的联合使用具有更佳的效果。

利用PCA排序给出预分类，通过ISODATA分析，把研究地区草原群落演替序列分成4个阶段：短花针茅+冷蒿阶段，扁穗冰草+冷蒿阶段，冷蒿+短花针茅阶段，亚氏旋花+冷蒿阶段。

2. 系统模糊聚类分析

为简便起见，将上述由15个主要种表征的样方记为： U_{ij} （其中 $i, j=1, 2, 3, \dots, 60$ 表示样方）。样方的特征可以用 u_{ij} 刻划，所以它们是15维空间中的点集。每个样方都是15维空间的一点。用 r_{ij} 表示空间中第 i 个样方与第 j 个样方之间的距离，则 r_{ij} 的大小反映了两个样方间差异的大小。这样就可以进行模糊聚类分析。取定论域 $U=\{1, 2, 3, \dots, 60\}$ ，建立模糊相似矩阵 $R=(r_{ij})$ 其中：

$$r_{ij} = \begin{cases} 1 & \text{当 } i=j \text{ 时,} \\ \frac{(u_i, u_j)}{\|u_i\| \cdot \|u_j\|} & \text{当 } i \neq j \text{ 时,} \end{cases}$$

$$(u_i, u_j) = \sum_{k=1}^n u_i^k \cdot u_j^k$$

$$\|u_i\|^2 = (u_i, u_i)$$

由上式⁽⁶⁾可知，当 r_{ij} 愈大时，样方 u_i 与样方 u_j 之间的差异就愈大。因为计算机容量有限，将60个样方分成两组，分别计算 r_{ij} 组成的相似矩阵 R_I 和 R_{II} （略）。

这时，模糊相似矩阵 R_I 、 R_{II} 只具有反身性和对称性，而不具有传递性。将其改造成模糊等价关系 R^* 。经过计算得到：

$$R_I^8 \triangleq R_I^{16} = R_I^*$$

$$R_{II}^{16} \triangleq R_{II}^{32} = R_{II}^*$$

计算结果 R_I^* 、 R_{II}^* 略。根据计算结果，得到动态聚类图（图2、 R_{II}^* 的图略）。从图中

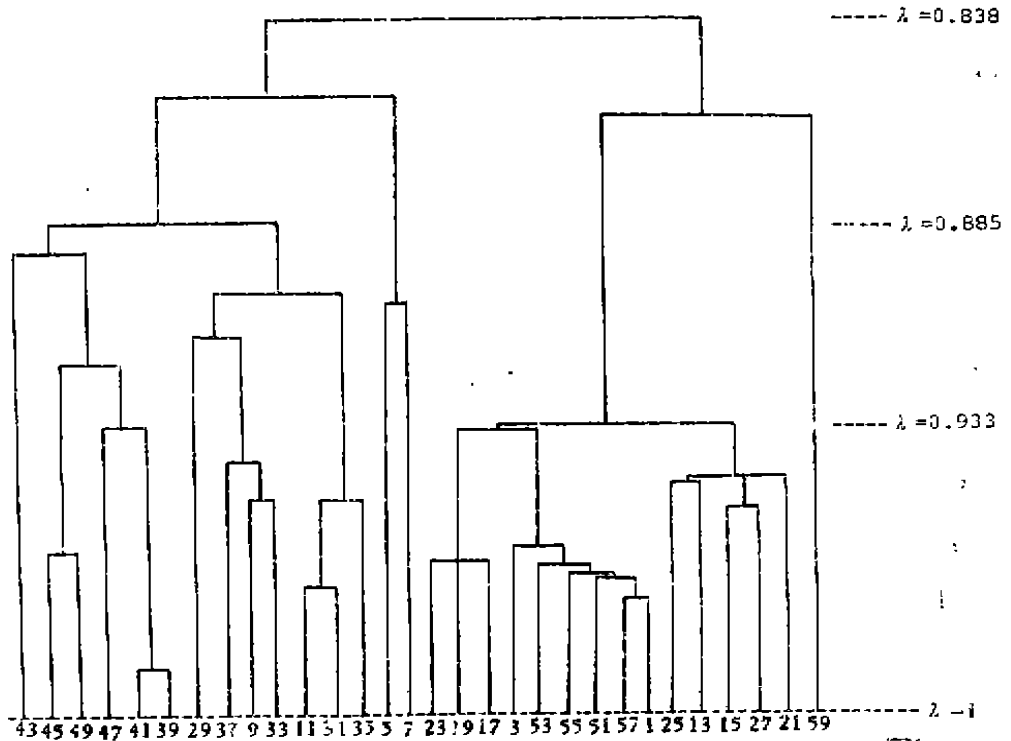


图 2 样方动态聚类图

Fig.2 The dynamic cluster map of samples

可以看出样方的归并情况。取 $\lambda = 0.933$ ，图中的样方有两个明显的集团，即样方21、27、15、13、25和样方1、3、57、51、55、53、17、19、23。从以后的样方归并情况，能够把样方分成4个集团。除上述两组样方外，还有样方35、31、33、11、9、37、29和样方39、41、47、49、45、43。根据 R_{II}^* 作图也可以得到类似的结果。所以系统模糊聚类也把研究

地区草原群落演替系列分成4个阶段,结果与ISODATA方法的基本相同。

三 讨 论

1. 组成植物群落的种类及其数量指标是植物群落最本质的特征,它们决定了群落的性质。因此,本文以构成群落的主要种类作为群落模糊数学分类的指标是完全符合植被分类原则的。在分析过程中去掉少数偶见种,可以大大减少计算量而不影响分类结果的可靠性。两种模糊数学方法都将研究地区的短花针茅草原放牧衰退演替系列分成4个阶段,其结果与前人的研究相吻合。

2. 与其他常见的数学分类方法相比较,模糊聚类分析能够深刻地揭示样方的本质,为研究者提供更多的信息。ISODATA分类法不仅能够合理地把样方分类,同时又详细地刻划了每个样方对各阶段的隶属程度,可以了解它们在演替序列中的位置。以模糊等价关系为分类依据的系统模糊聚类法,在分类的同时,通过动态聚类图,清晰地反应出样方的合并过程。

3. 排序和分类是数量分类的两大类方法,它们各有所长。“在排序的基础上再进行分类可能效率更高”〔4〕。在此以前,因为数学方法未能提供这种可能性以及受计算技术的限制,这方面的工作尚少见。ISODATA分类要求先对样方预分类,还要求给出合理的分类组数。这就给排序与分类的联合使用提供了良好的前提。特别是当对分类对象了解不多时,先利用排序方法分析样方,在此基础上得到预分类矩阵和分类组数,可以提高以后的聚类分析效率,又可以避免人为预分类时的盲目和主观影响。同时,排序图低维空间反映高维空间点的位置时产生的误差,又可以通过聚类分析得到修正。两种方法的结合取得了相互补充、校正的良好效果。

4. 与其他模糊数学分类方法相比较,模糊聚类分析的两种方法具有运算迅速,分类合理、明确的优点,特别适用于处理大量的样方资料。

模糊聚类分析不仅可以定量地描述各种群落类型间的差异,而且由于它们是建立在样本间模糊性的基础之上,所以比常规的数量分类方法更加适合于处理来自植被连续体的样方。它是群落分类的理想方法之一。随着对植物群落的深入研究,对于作为基础性研究的群落分类的要求越来越高,模糊聚类分析会得到广泛的应用。

参 考 文 献

- 李守虔、陈塞琳、张忠奎 1984 亚高山草甸蒿草植被放牧衰退演替阶段的数值分类。植物学报 26, 202—208。
 赵松岭、杨凤翔、陈庆诚 1982 针茅草原放牧衰退演替阶段的模糊数学分类。植物学报 24, 366—370。
 考克斯, E.W. 1972 普通生态学实验手册(蒋有绪译), 35页, 科学出版社。
 阳含熙、卢泽恩 1981 植物生态学的数量分类方法, 233—252, 349页, 科学出版社。
 贺仲雄 1983 模糊数学及其应用, 152—187页 天津科学技术出版社。
 冯德益、楼世博等 1983 模糊数学方法及其应用, 地震出版社。

USING FUZZY CLUSTERING ANALYSIS IN THE STUDY OF CLASSIFICATION OF SUCCESSION STAGES IN THE *SIPA* STEPPE

Li Xiaojun* Zhao Songling Zhang Pengyun Chen Qingcheng
(Ecological Research Institute, Lan Zhou University)

Wang Jiuli
(Ecological Research Institute, Beijing Forestry University)

With the aid of two methods of Fuzzy Clustering Analysis—Systematic Fuzzy Clustering Analysis and Iteration Self-Organizing Data Classification (ISODATA), this paper divides the plant communities of the retrogressive succession of the *Stipa* Steppe. The results show that it is more successful and ideal to combine the Principal Component Analysis (PCA) with the ISODATA to classify the plant communities. Both methods of Fuzzy Clustering Analysis divide the retrogressive successional stages into four types: *Stipa breviflora* Assoc., *Agropyron cristatum* Assoc., *Artemisia frigida* Assoc., and *Convolvulus ammannii* Assoc.. These results are similar to those results obtained by traditional methods. This paper provides two reasonable and feasible mathematic methods for studying the classification of plant communities.

Key words, fuzzy clustering, succession, classification.

*Note, at present in the Forestry Department, Beijing Forestry University.