

目录

1 问题重述.....	2
2 基本假设与符号说明.....	2
2.1 基本假设：.....	2
2.2 符号与符号说明：.....	2
3 问题分析和基本思路.....	3
3.1 问题分析和建模思路.....	3
3.2 思路流程图.....	4
4 模型的建立.....	4
4.1 物种多样性评价模型.....	4
4.2 遗传多样性评价.....	7
4.3 生态系统多样性评价.....	8
4.4 熵值法综合评价模型.....	9
5 模型的优缺点.....	10
5.1 模型优点：.....	10
参考文献.....	10
附录.....	12

1 问题重述

2010 年是联合国大会确定的国际生物多样性年。保护地球上的生物多样性已经越来越被人类社会所关注，相关的大规模科研和考察计划也层出不穷。为了更好地建立国际交流与专家间的合作，联合国还建立了生物多样性和生态系统服务政府间科学政策平台（IPBES）。但迄今为止，几乎所有的考察计划都面临着一个基本的困难：如何评价被考察区域的生物多样性。传统的方法是清点物种数量，但现在有许多科学家认为这种方法具有很大的局限性。譬如有人提出应当考虑物种的相似程度。有人则提出有一些物种的基因多样性程度远远超过另一些物种，所以应当考虑基因的多样性等。但现在还缺少一种能全面考虑不同因素的对生物多样性进行测定的方法。

问题： 请你建立合理的数学模型，设计一个全面而有效的评价生物多样性的指标，以利于今后考察和科研工作地开展。

2 基本假设与符号说明

2.1 基本假设：

- ①区域内的生物在一段时间内不会发生太大变动；
- ②种群关系树中兄弟结点间的地位等同，也就是同一层的物种分类对物种多样性影响相同；
- ③所有物种均在物种分类中，不存在特殊物种或是介于两个物种之间的物种；
- ④物种分类可以客观反映出各物种间的亲疏关系；
- ⑤物种基因丰富度以及各物种间基因相似度由目前相关科级手段取得，能够较为客观的反映各物种间的基因情况以及物种基因相互关系；
- ⑥模糊评判矩阵的权系数由相关专家根据各地区的实际情况确定，真实可靠；

2.2 符号与符号说明：

符号	符号说明
x_i	在物种关系树中 i 位置所包含的物种数量
h	树的深度

参赛队号 # 1189

$s(i)$	深度为 i 的生物物种类型之和
P	物种丰富度的综合评价指标
Q	遗传多样性综合评价指标
R	生态系统多样性综合评价指标
a_{ij}	第 i 个物种和第 j 个物种的相似度
z_i	第 i 物种的基因丰富度
$B(i)$	第 i 个物种基因的净丰富度
u_i	生态系统多样性评价的第 i 项指标
v_j	生态系统多样性评价的第 j 个评分级别
r_{ij}	指标 u_i 所得分值对 v_j 级别的隶属度
e_j	第 j 项指标的熵值

3 问题分析和基本思路

3.1 问题分析和建模思路

考虑问题的题设和要求，我们要解决的是评价生物多样性的问题。根据查阅相关资料我们将生物多样性的评价指标分为遗传多样性，物种多样性，生态系统多样性三个方面，并对每个方面建立相应的指标。对于物种的多样性，我们综合考虑物种的种类和物种的类型。根据相关分类标准将物种经行多层次的分类，从而形成物种关系树，物种的多样性不仅指物种个体种类的多样性，而且应当包含各个层次分类的多样性，为此我们对不同层次的物种类型赋予不同权值，通过加权求和得出物种多样性的综合评价指标。

对于遗传的多样性既要考虑到不同物种所包含的基因的不同，又要考虑到许多物种的基因有很大的相似度。我们用相似度矩阵来表示个物种间的基因相似度关系，用一列矩阵表示各物种的基因丰富度，通过相关矩阵运算得出最终的评价指标。对于生态系统的多样性的评价指标的确定，必须通过多方面来衡量，而且

每一个评价指标不太容易量化。而如何对各个评价指标量化，并将多个因素的评价综合起来。我们先将指标细分为多个再采用模糊评判模型对各个进行打分并通过评判矩阵得出最终的评判结果。由于对生物多样性的评价指标较多，而且各个指标又细分了一些细的指标所以要得出总的综合评价指标必须对各个指标进行统一规划，并分类对待。因此我们采用熵模型综合考虑各个因素得出最终的生物多样性评价指标。

3.2 思路流程图

下面的思路流程图是我们文章结构的一个缩影，它完整而形象的反映了我们文章的建模思路。

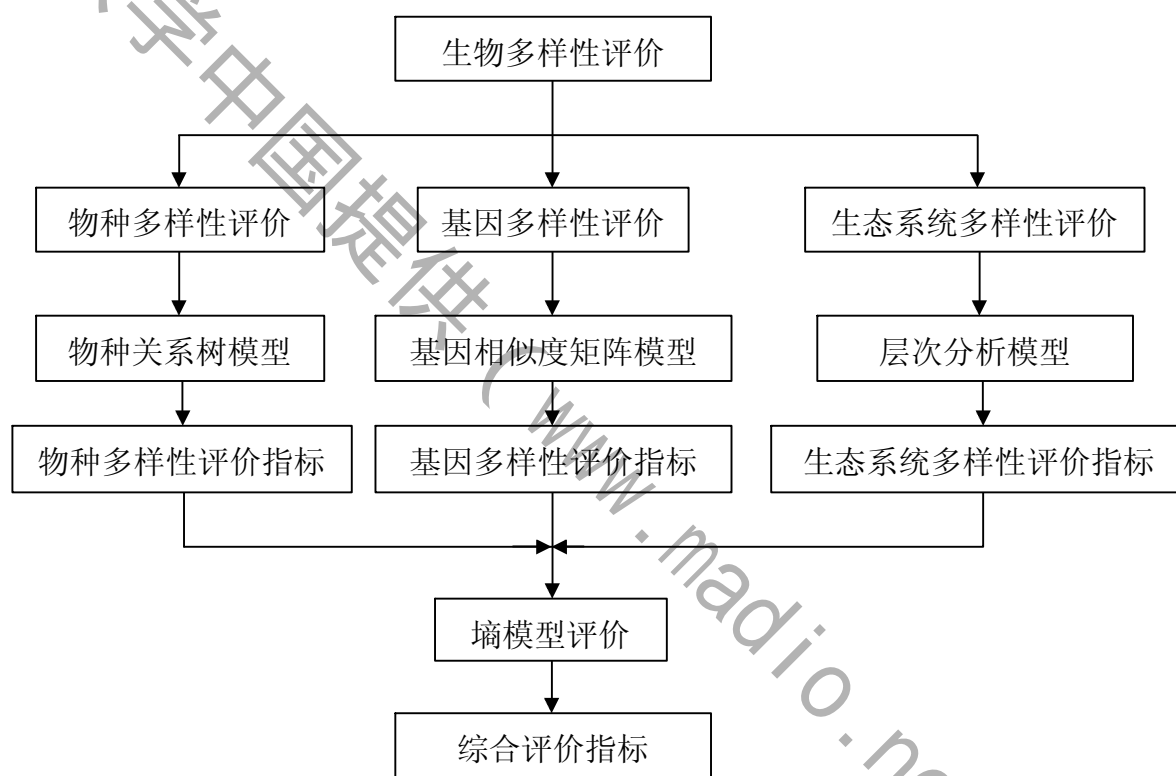


图 3.1 思路流程图

4 模型的建立

4.1 物种多样性评价模型

4.1.1 物种关系树模型

如图 4.1.1 所示：对于某一地区的所采集到的物种数目数据进行分类汇总，得到类似如下的关系树：

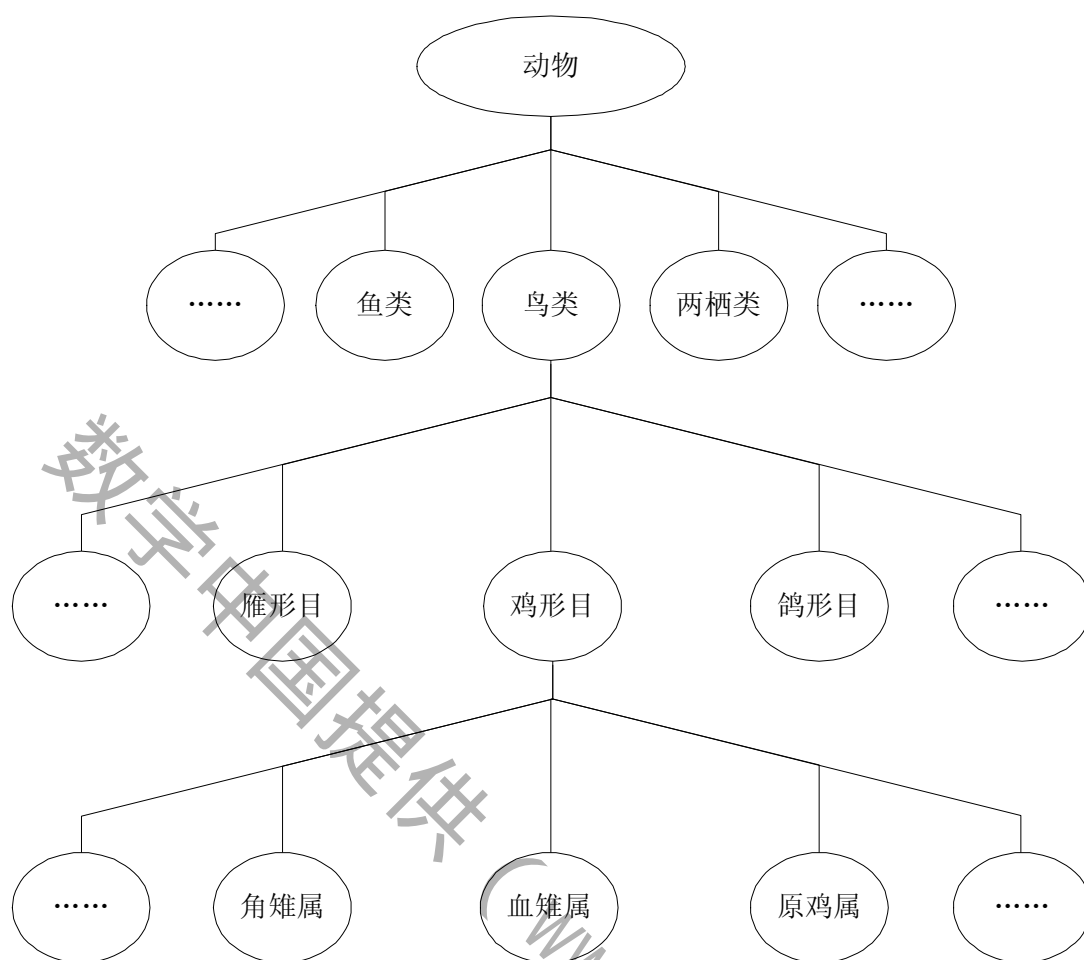


图 4.1.1 物种关系树

树是由 n ($n \geq 0$) 个结点组成的有限集合 T ，其中有且只有一个结点称为根结点，其余结点可以分为 m ($m \geq 0$) 个互不相交的有限集合 T_1, T_2, \dots, T_m ，其中每一个集合 T_i 本身又是一棵树，称为根结点的子树。

用二元组关系来定义树为：

$$Tree = (T, R)$$

物种关系树 $Tree$ 是由数据元素集合 T 和关系集合 R 组成。其中数据集合 T 是具有相同类型的数据元素集合 $T = (x_1, x_2, x_3, \dots, x_i)$ 。其中 x_i 表示在物种关系树中 i 位置所包含的物种数量。关系集合 R 表示各物种之间的分类关系，用三维坐标 (i, j, k) 表示，其含义为物种关系树的第 i 层第 j 个双亲结点的第 k 个分支树结构中常用的术语有：

结点：表示树中的元素；

结点的度：结点拥有的子树数；

叶子：度为 0 的结点，又称为端结点；

孩子：除根节点外，每个结点都是其前趋结点的孩子；

双亲:对应上述孩子结点的上层结点称为这些结点的双亲。

物种关系树中根结点的值表示该地区所拥有的总的物种数目。对于某一地区来说,由于各类物种数目差别较大,单把总的物种数目来作为该区域物种多样性评价指标并不科学。为此我们将关系树中每个双亲结点下的孩子结点赋予相同的权值,而对于树中不同深度赋予不同的权值 h_i ,且权值的由根结点开始由上向下逐渐减小。其中每个结点的数值代表其所下属的子结点的数目。如图 4.1.2 所示:

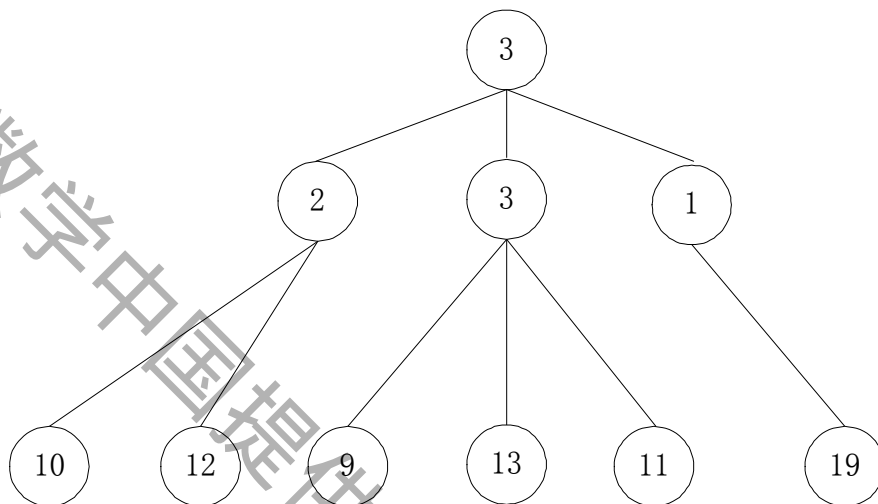


图 4.1.2 (a) 物种关系树一

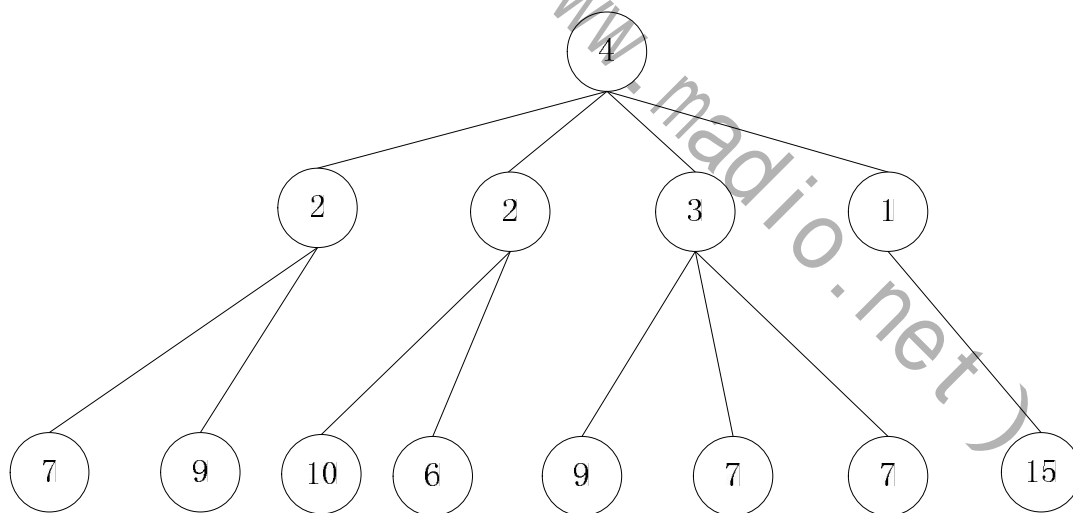


图 4.1.2 (b) 物种关系树二

通过对物种数目和物种种类的耦合关系得出物种丰富度的综合评价指标:

$$P = \sum_{i=1}^h (h-i+1)s(i)$$

其中 h 表示树的深度， $s(i)$ 表示深度为 i 的生物物种类型之和。

对于图 4.1.2 (a) 中，

$$P_a = 3 \times 3 + 2 \times (2 + 3 + 1) + 1 \times (10 + 12 + 9 + 13 + 11 + 19) = 105$$

对于图 4.1.2 (b) 中，

$$P_b = 3 \times 4 + 2 \times (2 + 2 + 3 + 1) + 1 \times (7 + 9 + 10 + 6 + 9 + 7 + 7 + 15) = 98$$

由以上结果可得出 A 地区物种比 B 地区物种丰富，即 A 地区物种丰富度较 B 地区物种丰富度高。

4.2 遗传多样性评价

4.2.1 遗传多样性模型

由于各物种的基因多样性不同，而且物种与物种之间基因存在一些相似度，评价物种遗传多样性不能单纯计算物种的数目以及其所包含的基因总数，而是要综合考虑，去除相同的基因。

建立一个列矩阵 λ 来表示各物种所包含的相对基因丰富度，再建立一个方阵 A 表示各物种之间的基因相似度。

假设某地区有 $a_1, a_2, a_3, \dots, a_n$ 等 n 个物种，对于相同物种，其相似度取 1，而对不同物种，其相似度取 $[0, 1]$ 之间不同的值，则：

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{bmatrix} \quad \text{式 (4.2.1)}$$

其中 a_{ij} 表示第 i 个物种和第 j 个物种的相似度。对于任意 $i \in (1, n)$ ，则有

$$a_{ii} = 1。$$

同样的我们可以得到列矩阵 λ 的表达式：

$$\lambda = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ \dots \\ z_n \end{bmatrix} \quad \text{式 (4.2.2)}$$

其中 z_i 表示第 i 物种的基因丰富度。

取

$$B = A \cdot \lambda = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix} \cdot \begin{bmatrix} z_1 \\ z_2 \\ z_i \\ \cdots \\ z_n \end{bmatrix} \quad \text{式 (4.2.3)}$$

其中 B 表示各个物种基因的净丰富度, 即去除一个物种与其他物种重合的基因的丰富度。令

$$B = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \cdots \\ b_n \end{bmatrix}, \text{ 则 } Q = \sum_{i=1}^n b_i \quad \text{式 (4.2.4)}$$

4.3 生态系统多样性评价

模糊综合评判模型的建立

根据相关资料我们将生态系统多样性指标分为生境类型多样性, 生境稀有性, 自然性, 面积适应性, 生态系统稳定性。

4.3.1 确定评判系统指标集和指标权重集

用 $U = \{U_1, U_2, \dots, U_n\}$ 来表示各个评判标准的权重。

4.3.2 确定评判指标评价等级 (备择集)

$$V = \{V_1, V_2, \dots, V_p\} = \{V_1, V_2, V_3, V_4\}$$

$$= \{\text{优, 良, 中, 差}\} = \{90, 80, 70, 60\}$$

4.3.3 评价方法

模糊综合评价是由 $B = A \circ R$ 得出的。其中, B 为评语集 $V = \{I, II, III, IV\}$ 上的一个模糊子集, 称为评价; A 为因子集 U 上的模糊子集, 是权数分配集; “ \circ ” 为算子符号; R 为模糊变换集, 设 $R = (r_{ij})_{n \times 4}$, 则 r_{ij} 为指标 u_i 所得分值对 v_j 级别的隶属度:

$$R = \begin{array}{c|cccc|c} & v_1 & v_2 & v_3 & v_4 & \\ \hline u_1 & r_{11} & r_{12} & r_{13} & r_{14} & \\ \vdots & \vdots & & & \vdots & \vdots \\ \vdots & & & & \vdots & \vdots \\ \vdots & & & & \vdots & \vdots \\ \hline u_n & r_{n1} & \cdots & \cdots & r_{n4} & \end{array} \quad \text{式 (4.3.1)}$$

隶属度由

$$r_{ij} = \frac{n_{ij}}{\sum_{j=1}^4 n_{ij}} \quad \text{式 (4.3.2)}$$

给出。分子表示对 u_i 指标的评分落于 v_j 区间内的个数，分母表示对于指标 u_i 的评分总个数。

对于现状评价、影响评价、项目评价, 以及评价各层次, 因子集 U 都各不相同。如对于物种多样性价值现状评价, $U_{PS} = \{R, S, C, E, V\}$, 其中因子 V 的因子集为 $U_v = \{V_z, V_c, V_e\}$, 详见表 1~4。在影响评价的影响预测中 I_d, I_h, I_c 各指标对各级的隶属度, 均需从远期、近期和可重建性三方面评价得出, 其中 I_d 为对各物种或生态系统影响的等权综合结果。

各指标的权数分配集可由层次分析得出, 评价中算子符号“ \odot ”取乘积算子“ \cdot ”和闭合加法算子“ \oplus ”。

4.4 熵值法综合评价模型

(1) n 表示生物多样性评价的三个方面, 每个方面有 m 个指标, 则 x_{ij} 为第 i 个方案的第 j 个指标的数值。 ($i=1, 2, \dots, n; j=1, 2, \dots, m$)

(2) 计算第 j 项指标下第 i 个方面占该指标的比重:

$$p_{ij} = \frac{X_{ij}}{\sum_{i=1}^n X_{ij}}, (i=1, 2, \dots, n; j=1, 2, \dots, m) \quad \text{式 (4.4.1)}$$

(3) 计算第 j 项指标的熵值。

$$e_j = -k \sum_{i=1}^m p_{ij} \ln(p_{ij}), \text{ 其中, } k > 0, k = 1 / \ln(n), e_j \geq 0 \quad \text{式 (4.4.2)}$$

(4) 计算第 j 项指标的差异系数。对于第 j 项指标, 指标值的差异越大, 对方案评价的作用就越大, 熵值就越小。定义差异系数:

$$g_i = \frac{1 - e_j}{m - E_e} \quad \text{式 (4.4.3)}$$

式中 $E_e = \sum_{j=1}^m e_j, 0 \leq g_j \leq 1, \sum_{j=1}^m g_j = 1$ 。

(5) 求权数:

$$w_j = \frac{g_j}{\sum_{j=1}^m g_j} (1 \leq j \leq m) \quad \text{式 (4.4.4)}$$

(6) 计算各个评价方面的综合得分：

$$s_i = \sum_{j=1}^m w_j \cdot p_{ij} (i=1, 2, L, n) \quad \text{式 (4.4.5)}$$

5 模型的优缺点

5.1 模型优点：

- 1、采用物种关系树模型对物种多样性的评价不仅考虑到物种个体类型的多样性而且综合考虑到各个分类类型的多样性，结果更加有代表性；
- 2、运用基因相似度矩阵模型可以综合考虑到各个物种间基因的相似程度以及各个物种所含基因多样性的差异性，得到的结果更能反映物种基因多样性的实际情况
- 3、采用模糊评判模型对生态系统多样性的评价既考虑到各项因素又对各个因素区别对待使结果更加客观。
- 4、采用不同模型求解各项指标，可以减少因一种模型所存在的一些问题而造成的误差。

5.2 模型缺点：

- 1、物种关系树模型需要对所有物种进行细致的分类汇总，资料的考察工作困难；。
- 2、基因的相似度模型当处理较多物种信息时计算复杂，并且对各个物种间的相似度的确定比较困难。

参考文献

- [1]姜启源，数学模型(二版)，北京：高等教育出版社，2002.
- [2]赵静，数学建模与数学实验（二版），北京：高等教育出版社，2003.
- [3]盛骤等，概率论与数理统计，北京：高等教育出版社，2001.
- [4]谢季坚，刘承平，模糊数学方法及其应用（第2版），武汉：华中科技大学出版社，2000. 5.

参赛队号 # 1189

- [5] 沈被娜，计算机软件技术基础，清华大学出版社，2004. 3。
- [6] 王明莉 斯琴高娃 等，一个用于区域物种多样性综合评价的指数，厦门大学学报，2010. 9。
- [7] 曾志新 罗军等，生物多样性的评价指标和评价标准，湖南林业科技，1999. 6。
- [8] 刘晓红，李 校，彭志杰，生物多样性计算方法的探讨，河北林果研究，2008. 6 第 2 期。

数学中国提供 (www.madio.net)

附录

附录一：模糊评价代码

```

1. function
[U,P,Dist,Cluster_Res,Obj_Fcn,iter]=fuzzycm(Data,C,plotflag,M,epsm)
2. % 模糊 C 均值聚类 FCM: 从随机初始化划分矩阵开始迭代
3. % [U,P,Dist,Cluster_Res,Obj_Fcn,iter] =
fuzzycm(Data,C,plotflag,M,epsm)
4. % 输入:
5. % Data: N×S 型矩阵, 聚类的原始数据, 即一组有限的观测样本集,
6. % Data 的每一行为一个观测样本的特征矢量, S 为特征矢量
7. % 的维数, N 为样本点的个数
8. % C: 聚类数, 1<C<N
9. % plotflag: 聚类结果 2D/3D 绘图标记, 0 表示不绘图, 为缺省值
10. % M: 加权指数, 缺省值为 2
11. % epsm: FCM 算法的迭代停止阈值, 缺省值为 1.0e-6
12. % 输出:
13. % U: C×N 型矩阵, FCM 的划分矩阵
14. % P: C×S 型矩阵, FCM 的聚类中心, 每一行对应一个聚类原型
15. % Dist: C×N 型矩阵, FCM 各聚类中心到各样本点的距离, 聚类中
16. % 心 i 到样本点 j 的距离为 Dist(i, j)
17. % Cluster_Res: 聚类结果, 共 C 行, 每一行对应一类
18. % Obj_Fcn: 目标函数值
19. % iter: FCM 算法迭代次数
20. % See also: fuzzydist maxrowf fcmlplot
21. if nargin<5
22.     epsm=1.0e-6;
23. end
24. if nargin<4
25.     M=2;
26. end
27. if nargin<3
28.     plotflag=0;
29. end
30. [N,S]=size(Data);m=2/(M-1);iter=0;

```

参赛队号 # 1189

```
31. Dist(C,N)=0; U(C,N)=0; P(C,S)=0;
32. % 随机初始化划分矩阵
33. U0 = rand(C,N);
34. U0=U0./(ones(C,1)*sum(U0));
35. % FCM 的迭代算法
36. while true
37.     % 迭代计数器
38.     iter=iter+1;
39.     % 计算或更新聚类中心 P
40.     Um=U0.^M;
41.     P=Um*Data./(ones(S,1)*sum(Um'))';
42.     % 更新划分矩阵 U
43.     for i=1:C
44.         for j=1:N
45.             Dist(i,j)=fuzzydist(P(i,:),Data(j,:));
46.         end
47.     end
48.     U=1./(Dist.^m.*(ones(C,1)*sum(Dist.^(-m))));
49.     % 目标函数值：类内加权平方误差和
50.     if nargout>4 | plotflag
51.         Obj_Fcn(iter)=sum(sum(Um.*Dist.^2));
52.     end
53.     % FCM 算法迭代停止条件
54.     if norm(U-U0, Inf)<epsm
55.         break
56.     end
57.     U0=U;
58. end
59. % 聚类结果
60. if nargout > 3
61.     res = maxrowf(U);
62.     for c = 1:C
63.         v = find(res==c);
64.         Cluster_Res(c,1:length(v))=v;
65.     end
66. end
```

参赛队号 # 1189

```
67. % 绘图
68. if plotflag
69.     fcmplot(Data,U,P,Obj_Fcn);
70. end
71. m 文件 2/7:
72. function
[U,P,Dist,Cluster_Res,Obj_Fcn,iter]=fuzzycm2(Data,P0,plotflag,M,epsm)
73. % 模糊 C 均值聚类 FCM: 从指定初始聚类中心开始迭代
74. % [U,P,Dist,Cluster_Res,Obj_Fcn,iter] =
fuzzycm2(Data,P0,plotflag,M,epsm)
75. % 输入: Data,plotflag,M,epsm: 见 fuzzycm.m
76. % P0: 初始聚类中心
77. % 输出: U,P,Dist,Cluster_Res,Obj_Fcn,iter: 见 fuzzycm.m
78. % See also: fuzzycm
79. if nargin<5
80.     epsm=1.0e-6;
81. end
82. if nargin<4
83.     M=2;
84. end
85. if nargin<3
86.     plotflag=0;
87. end
88. [N,S] = size(Data); m = 2/(M-1); iter = 0;
89. C=size(P0,1);Dist(C,N)=0;U(C,N)=0;P(C,S)=0;
90. % FCM 的迭代算法
91. while true
92.     % 迭代计数器
93.     iter=iter+1;
94.     % 计算或更新划分矩阵 U
95.     for i=1:C
96.         for j=1:N
97.             Dist(i,j)=fuzzydist(P0(i,:),Data(j,:));
98.         end
99.     end
100.     U=1./(Dist.^m.*(ones(C,1)*sum(Dist.^(-m))));
```

参赛队号 # 1189

```
101.      % 更新聚类中心 P
102.      Um=U.^M;
103.      P=Um*Data./(ones(S,1)*sum(Um'))';
104.      % 目标函数值：类内加权平方误差和
105.      if nargout>4 | plotflag
106.          Obj_Fcn(iter)=sum(sum(Um.*Dist.^2));
107.      end
108.      % FCM 算法迭代停止条件
109.      if norm(P-P0, Inf)<epsm
110.          break
111.      end
112.      P0=P;
113.  end
114.  % 聚类结果
115.  if nargout > 3
116.      res = maxrowf(U);
117.      for c = 1:C
118.          v = find(res==c);
119.          Cluster_Res(c,1:length(v))=v;
120.      end
121.  end
122.  % 绘图
123.  if plotflag
124.      fcmlplot(Data,U,P,Obj_Fcn);
125.  end
126.  m 文件 3/7:
127.  function fcmlplot(Data,U,P,Obj_Fcn)
128.  % FCM 结果绘图函数
129.  % See also: fuzzycm maxrowf ellipse
130.  [C,S] = size(P); res = maxrowf(U);
131.  str = 'po*x+d^v><.h';
132.  % 目标函数绘图
133.  figure(1), plot(Obj_Fcn)
134.  title('目标函数值变化曲线','fontsize',8)
135.  % 2D 绘图
136.  if S==2
```

参赛队号 # 1189

```
137.         figure(2),plot(P(:,1),P(:,2),'rs'),hold on
138.         for i=1:C
139.             v=Data(find(res==i),:);
140.             plot(v(:,1),v(:,2),str(rem(i,12)+1))
141.             ellipse(max(v(:,1))-min(v(:,1)), ...
142.                     max(v(:,2))-min(v(:,2)), ...
143.                     [max(v(:,1))+min(v(:,1)), ...
144.                     max(v(:,2))+min(v(:,2))]/2,'r:')
145.         end
146.         grid on,title('2D 聚类结果图','fontsize',8),hold off
147.     end
148.     % 3D 绘图
149.     if S>2
150.         figure(2),plot3(P(:,1),P(:,2),P(:,3),'rs'),hold on
151.         for i=1:C
152.             v=Data(find(res==i),:);
153.             plot3(v(:,1),v(:,2),v(:,3),str(rem(i,12)+1))
154.             ellipse(max(v(:,1))-min(v(:,1)), ...
155.                     max(v(:,2))-min(v(:,2)), ...
156.                     [max(v(:,1))+min(v(:,1)), ...
157.                     max(v(:,2))+min(v(:,2))]/2, ...
158.                     'r:',(max(v(:,3))+min(v(:,3)))/2)
159.         end
160.         grid on,title('3D 聚类结果图','fontsize',8),hold off
161.     end
162.     m 文件 4/7:
163.     function D=fuzzydist(A,B)
164.     % 模糊聚类分析：样本间的距离
165.     % D = fuzzydist(A,B)
166.     D=norm(A-B);
167.     m 文件 5/7:
168.     function mr=maxrowf(U,c)
169.     % 求矩阵 U 每列第 c 大元素所在行,c 的缺省值为 1
170.     % 调用格式：mr = maxrowf(U,c)
171.     % See also: addr
172.     if nargin<2
```


参赛队号 # 1189

```
173.      c=1;
174.  end
175.  N=size(U,2);mr(1,N)=0;
176.  for j=1:N
177.      aj=addr(U(:,j),'descend');
178.      mr(j)=aj(c);
179.  end
180.  m 文件 6/7:
181.  function ellipse(a,b,center,style,c_3d)
```