安徽大学

本科毕业论文

题　　目：　模糊最大熵模型及其应用

学生姓名：　覃浩蓝　　学号：　A01714003

院（系）：　数学科学学院　　专业：　信息与计算科学

入学时间：　二〇一七年　九月

导师姓名：　吴涛　　职称/学位：　教授

导师单位：　数学科学学院

完成时间：　二〇二一年　四月

模糊最大熵模型及其应用

摘　要

在模糊 C 均值（Fuzzy C-means）算法中，通过与模糊数学的融合，给出了相比于 K-means 硬聚类更灵活的聚类结果。

关键字:　模糊熵，最大熵，聚类

Fuzzy Maximum Entropy Model and Its Application

Abstract

In the.

Keywords: **Fuzzy Sets;**

目 录

1　绪论

1.1　研究背景

人类从原始社会一步步走到现在，经历了漫长的进化和发展，站到了食物链的顶端，步入了信息时代，这一切都得益于我们对信息的获取和加工能力在不断进步。进入信息时代后，我们周围的信息越来越多，五花八门各式各样的信息充斥在我们的生活中，这些信息有的是确定的，但更多的是不确定的、带有模糊性的信息。所谓模糊性是指不确定的，介于是和不是两者之间的性质。例如，对于优等生的判定，有的人觉得 90 就可以了，有的人却觉得需要达到 95 分以上才算优秀，所以我们很难这种非此即彼的性质去衡量一个人是不是优等生。

我们所处的是一个复杂多变、时刻在运动的世界，大到星系运动，小到粒子碰撞，里面蕴含的规律都是复杂多样的。信息本身就包含着确定性和不确定性，无所谓的好坏之分，它取决于我们如何认识信息，了解信息和使用信息。比如，我们在评价某一个菜品时，会用“好吃”、“还行“、”难吃“来形容；描述天气时说“多云”、“晴朗”；说一个人的衣服搭配好看等等。这些问题很难用统一的标准去衡量，但是我们却可以得到清晰的结论，我们已经习惯在生活中运用模糊性所谓语言描述事物，用模糊的方法认识生活中的事物。虽然信息带着不确定性，但是我们所处的客观世界是确定的，所以我们需要一种方法研究模糊的信息，得到清晰的结论。于是数学诞生了一个新的分支：模糊数学。1965 年，L.A.Zadeh 在期刊 Information and Control 上发表了论文《Fuzzy Sets》，标志着模糊理论的诞生。

1.2　研究内容

最大熵模型是一种分类学习模型，模糊熵是模糊数学里面的概念，本文在模糊理论框架下，将最大熵推广到模型信息情形，与传统的模糊 C 均值聚类（FCM）进行比较，建立模糊最大熵模型并应用于实际的分类问题中，通过进一步的研究探索模糊最大熵模型在实际问题中的应用。

1.3　研究意义

随着人工智能的大热，机器学习开始迅速应用于我们的生活中，比如商品推荐、语音识别和智能导航等。其中，分类问题是机器学习领域的一个重要问题。生活中许多的分类问题是模糊的，计算机无法直接处理这些模糊信息，而我们的人脑却可以很好地从这些模糊信息中得到精确的结论。随着熵理论和模糊数学的发展，模糊数学和最大熵模型的应用范围也越来越广泛，为了处理分类问题中的不确定性，国内外的许多这方面的学者也进行了许多研究，寻找分类问题模糊性的度量方式，探寻新的实际应用。本文将模糊熵与最大熵原理结合，

2 模糊数学理论

在经典集合理论里面，一个集合就是某一个概念的内涵。对于论域上的一个对象，它要么属于这个集合，要么不属于这个集合，两者只能选一个，不能两者兼之，也不能有模棱两可的情况。而对模糊数学研究的对象来说，我们不能简单地用是或否来描述一个对象是否属于一个集合。由此，我们把集合的特征函数的取值从 $\{0,1\}$ 这个集合扩充到 $[0,1]$ 这个区间上的连续取值。越靠近 1，说明该对象属于集合的程度越大，反之，越靠近 0 就越小。这样我们就把经典集合扩充到带有模糊边界的模糊集了，从而我们可以用这样的集合表示模糊概念。

2.1 模糊集及其表示方法

2.1.1 模糊集的定义

定义 2.1.1 (**模糊子集**[1]). 设 U 为我们所研究的论域，

$$\mu_{\tilde{A}} : \mathrm{U} \longrightarrow [0,1]$$

称 μ 确定了 U 上的一个模糊子集，记为 $\tilde{A}$。μ 称为 $\tilde{A}$ 的隶属函数，把 $\mu_{\tilde{A}}(u)(u \in \mathrm{U})$ 的值称为 u 对于模糊子集 $\tilde{A}$ 的隶属度。$\mu_{\tilde{A}}(u)$ 越大，代表 u 隶属于 $\tilde{A}$ 的程度越高。通常，我们也把模糊子集简称为模糊集。

2.1.2 模糊集的表示方法

设有限集 $\mathrm{U} = \{u_1, u_2, \dots, u_n\}$，则有限集可以用如下几种方法表示[2]。

- Zadeh 表示法

$$\tilde{A} = \frac{\tilde{A}(u_1)}{u_1} + \frac{\tilde{A}(u_2)}{u_2} + \cdots + \frac{\tilde{A}(u_n)}{u_n}.$$

  虽然我们以分式和的方式表示，但是其中的 $\tilde{A}(u_i)/u_i$ 并不表示分数，“+”也不表示和。$\tilde{A}(u_i)/u_i$ 表示的是元素 u_i 与对 $\tilde{A}$ 的隶属度的一一对应关系；“+”表示的是 $\tilde{A}$ 在论域 U 上的整体。

- 序偶表示法

$$\tilde{A} = \{(\tilde{A}(u_1), u_1), (\tilde{A}(u_2), u_2), \dots, (\tilde{A}(u_n), u_n)\}.$$

  序偶表示法是从例举法演变而来，由元素的隶属度和对应的元素组成的有序对列出。

- 向量表示法

$$\tilde{A} = (\tilde{A}(u_1), \tilde{A}(u_2), \dots, \tilde{A}(u_n)).$$

  向量表示法是用 n 维数组来实现的，在论域中的元素按一定的顺序排列时，按此顺序记录元素的隶属度。此时也称 $\tilde{A}$ 为模糊向量。

2.2 模糊集的运算及其性质

我们先给出模糊幂集的定义：

定义 2.2.1. 论域 U 上的模糊子集的全体称为模糊幂集，记为 $\mathcal{F}(U)$，即

$$\mathcal{F}(U) = \{\tilde{A} \mid \tilde{A}(u) : \mathrm{U} \to [0,1]\}$$

模糊集的包含与相等：

定义 2.2.2. 设 $\tilde{A}, \tilde{B} \in \mathcal{F}(U)$, 如果对 $\forall u \in \mathrm{U}$ 都成立 $\tilde{B}(u) \geqslant \tilde{A}(u)$, 则称 $\tilde{B}$ 包含 $\tilde{A}(u)$，记作 $\tilde{B}(u) \supseteq \tilde{A}(u)$。

定义 2.2.3. 设 $\tilde{A}, \tilde{B} \in \mathcal{F}(U)$, 如果对 $\forall u \in \mathrm{U}$ 都成立 $\tilde{B}(u) = \tilde{A}(u)$, 则称 $\tilde{B}$ 等于 $\tilde{A}(u)$，记作 $\tilde{B}(u) = \tilde{A}(u)$。

我们规定 $a \vee b = MAX(a,b), a \wedge b = MIN(a,b)$, 所以我们可以这样描述模糊集的并、交、余：

定义 2.2.4. 如果对于任意一个 $u \in \mathrm{U}$，有 $\tilde{C}(u) = \tilde{A}(u) \vee \tilde{B}$，则称 $\tilde{C}$ 为 $\tilde{A}$ 与 $\tilde{B}(u)$ 的并，记为 $\tilde{C} = \tilde{A} \cup \tilde{B}$。如果对于任意一个 $u \in \mathrm{U}$，有 $\tilde{C}(u) = \tilde{A}(u) \wedge \tilde{B}$，则称 $\tilde{C}$ 为 $\tilde{A}$ 与 $\tilde{B}(u)$ 的交，记为 $\tilde{C} = \tilde{A} \cap \tilde{B}$。
它们的隶属度函数定义为：

$$(\tilde{A} \cup \tilde{B})(u) \overset{\text{def}}{=} \tilde{A}(u) \vee \tilde{B}(u)\ \forall u \in \mathrm{U}$$

$$(\tilde{A} \cap \tilde{B})(u) \overset{\text{def}}{=} \tilde{A}(u) \wedge \tilde{B}(u)\ \forall u \in \mathrm{U}$$

定义 2.2.5. 如果对于 $\forall u \in \mathrm{U}$，有 $\tilde{B}(u) = 1 - \tilde{A}(u)$，则称 $\tilde{B}$ 为 $\tilde{A}$ 的余，记为 $\tilde{B} = \tilde{A}^c$。

2.3 模糊集的截集

模糊集可以很好地描述模糊概念，但是在客观现实里，我们需要在最后把模糊集变成我们需要的各种集合，从而得出确定的结论，于是我们引入了 λ-截集的概念。

定义 2.3.1. 设 $\tilde{A} \in \mathcal{F}(U)$，对于 $\forall \lambda \in [0,1]$，记

$$(\tilde{A})_\lambda = A_\lambda \overset{\text{def}}{=} \{u \mid \tilde{A}(u) \geqslant \lambda\}$$

λ-截集本质上是一个普通集合，是通过对 $\tilde{A}$ 进行截取得出的集合。转化为特征函数即为

$$A_\lambda(u) = \begin{cases} 1, & A(u) \geqslant \lambda \\ 0, & A(u) < \lambda \end{cases}$$

3　模糊最大熵模型

3.1　最大熵原理

熵原本是物理学中的概念，是由热力学第二定律引出的一个物质系统的状态参量。

3.2　模糊熵

自从模糊理论提出

定义 3.2.1 (**模糊熵**[3]). 对于离散的模糊变量

$$H(\tilde{A}) = -\sum_{i=1}^{n}\left(\tilde{A}(u_i)\ln\tilde{A}(u_i) + \left(1-\tilde{A}(u_i)\right)\ln\left(1-\tilde{A}(u_i)\right)\right)$$

对于连续的模糊变量

$$H(\tilde{A}) = -\int_{-\infty}^{\infty}(\tilde{A}(u)\ln\tilde{A}(u) + (1-\tilde{A}(u))\ln(1-\tilde{A}(u)))du$$

3.3　模糊最大熵模型

4 模糊最大熵模型在 iris 数据集上的表现

4.1 模型建立

巴拉巴拉

4.2 模型求解

巴拉巴拉打算拿彩票

4.3 与实际值和 FCM 算法分类结果比较

5 总结与展望

5.1 总结

5.2 展望

参考文献

[1] ZADEH L A. Fuzzy sets[J]. Information and Control, 1965, 8(3): 338-353.

[2] 李安贵, 张志宏. 模糊数学及其应用 [M]. （第 2 版）. 北京: 冶金工业出版社, 2005.

[3] LI P, LIU B. Entropy of credibility distributions for fuzzy variables[J]. IEEE Transactions on Fuzzy Systems, 2008, 16(1): 123-129.

[4] ZADEH L A. Probability measures of fuzzy events[J]. Journal of mathematical analysis and applications, 1968, 23(2): 421-427.

附录 A　致谢

谢谢！