基于直觉模糊熵的直觉模糊聚类

徐小来，　雷英杰，　赵学军
（空军工程大学 导弹学院，陕西 三原　713800）

摘　要：Atanassov直觉模糊集合是对Zadeh模糊集合最有影响的一种扩充和发展，将模糊聚类扩展为直觉模糊聚类，具有重要的应用价值。在研究将模糊聚类扩展为直觉模糊聚类时，提出了一种基于熵最大的直觉模糊聚类，并推导了迭代求解的算法。典型实验表明，直觉模糊聚类的性能优于模糊聚类，提高了聚类的正确率。

关键词：直觉模糊聚类；直觉模糊熵；直觉模糊集合；模糊聚类
中图分类号：TP182　　**文献标识码**：A　　**文章编号**：1009－3516(2008)02－0080－04

传统的聚类分析是一种硬划分，它把每个待辨识的对象严格地划分到某个类中，具有“非此即彼”的“分明概念”，因此这种分类的类别界限是分明的。Dunn[1]按照Ruspini定义的模糊划分的概念，把硬聚类的目标函数推广到模糊聚类的情况，Bezdek[2]又将Dunn的目标函数推广为更普遍的形式，模糊聚类得到了样本属于各个类别的不确定性程度，表达了样本类属的“亦此亦彼”的“模糊概念”，即建立起了样本对于类别的不确定性的描述，能更客观地反映现实世界，从而成为聚类分析研究的主流。

Atanassov直觉模糊集合(Intuitionistic Fuzzy Sets, IFS)[3]是对Zadeh模糊集合最有影响的一种扩充和发展。IFS增加了一个新的属性参数——非隶属度函数，进而还可以描述“非此非彼”的“模糊概念”，更加细腻地刻画客观世界的模糊性本质[4]，因而引起众多学者的关注。本文在直觉模糊熵最大的基础上，将模糊聚类分析扩展为直觉模糊聚类分析，并推导了迭代的算法。

1　直觉模糊集理论

Atanassov对直觉模糊集给出如下定义。

定义1　设X是一个给定论域，则X上的一个直觉模糊集A为

$$A=\{\langle x,\mu_A(x),v_A(x)\rangle \mid x\in X\} \tag{1}$$

式中：$\mu_A(x):X\to[0,1]$和$v_A(x)X\to[0,1]$分别代表A的隶属函数$\mu_A(x)$和非隶属函数$v_A(x)$，且对于A上的所有$x\in X$，$0\leqslant\mu_A(x)+v_A(x)\leqslant 1$成立。直觉模糊集$A$可以简记作$A=\langle x,\mu_A,v_A\rangle$。显然，一般模糊子集对应于下列直觉模糊集$A=\{\langle x,\mu_A(x),1-v_A(x)\rangle \mid x\in X\}$。

对于X中的每一个直觉模糊子集，称$\pi_A(x)=1-\mu_A(x)+v_A(x)$为$A$中$x$的直觉指数(Intuitionistic Index)，它是x对A的犹豫程度的一种测度。显然，对于X中的一般模糊子集A，$\pi_A(x)=0$，$\forall x\in X$。

2　直觉模糊熵

Shannon用概率论作为度量信息的数学工具，将信息定义为消除不确定性的东西，从而把信息与不确定

收稿日期：2007－04－26
基金项目：国家自然科学基金资助项目(60773209)；陕西省自然科学基金资助项目(2006F18)
作者简介：徐小来(1980－)，男，湖南宁乡人，博士生，主要从事智能信息处理与信息融合研究；
E－mail：xxl1024@163.com
雷英杰(1956－)，男，陕西渭南人，教授，博士生导师，主要从事智能信息处理与智能决策等研究.

性关联起来,将熵作为一个度量信息状态不确定性的尺度,提出了信息熵的概念。De Luca 和 Termini 研究了模糊集模糊性的度量,将概率型信息熵扩展为非概率型信息熵,提出了模糊信息熵必须满足的公理[5]。Szmidt 和 Kacprzyk 扩展了 De Luca 和 Termini 公理,将模糊信息熵扩展为直觉模糊信息熵[6]。

定义2　直觉模糊集 $A=\{\langle x,\mu_A(x),v_A(x)\rangle \mid x\in X\}$ 的熵 $H(A)$ 是一个实值函数,$H:H(A)\to[0,1]$,必须满足以下4个公理[5]:

1) $H(A)=0$,当且仅当 A 是一个经典集合,对于所有的 $x_i\in X$ 都有 $\mu_A(x_i)=0$ 或 $\mu_A(x_i)=1$;

2) $H(A)=1$,当且仅当对于所有的 $x_i\in X$ 都有 $\mu_A(x_i)=v_A(x_i)$;

3) $H(A)\leqslant H(B)$,如果 $A\subseteq B$,即 $\mu_A(x)\leqslant\mu_B(x)$ 和 $v_A(x)\geqslant\mu_B(x)$;

4) $H(A)=H(A^c)$。

$$H(A)=-\frac{1}{n\ln 2}\sum_{i=1}^{n}[\mu_A(x_i)\ln\mu_A(x_i)+v_A(x_i)\ln v_A(x_i)]-(1-\pi_A(x_i))\ln(1-\pi_A(x_i))-\pi_A(x_i)\ln 2 \tag{2}$$

文献[7]系统研究了直觉模糊信息的测度,提出式(2)作为一种具体的直觉模糊熵,并证明式(2)满足以上4个公理,是一种有效的直觉模糊熵。显然,当 $\pi_A(x_i)=0$,直觉模糊熵就退化为模糊熵,如式(3)所示。

$$H(A)=-\frac{1}{n\ln 2}\sum_{i=1}^{n}[\mu_A(x_i)\ln\mu_A(x_i)+(1-\mu_A(x_i))\ln(1-\mu_A(x_i))] \tag{3}$$

3　直觉模糊聚类模型及算法

设 $\boldsymbol{X}=\{x_1,x_2,\cdots,x_n\}$ 是待聚类分析对象的全体,$\boldsymbol{P}=\{p_1,p_2,\cdots,p_n\}$ 是 c 个聚类原型,c 为聚类类别数,定义直觉模糊集合 $A=\{\langle(\boldsymbol{x}_i,\boldsymbol{p}_j),\mu_{ij},v_{ij}\rangle \mid \boldsymbol{x}_i\in\boldsymbol{X},\boldsymbol{p}_j\in\boldsymbol{P}\}$,$(\boldsymbol{x}_i,\boldsymbol{p}_j)$ 表示第 i 个样本与第 j 类聚类原型的隶属关系。因此聚类分析可以转化为优化问题,使得代价函数 E 最小:

$$E=\sum_{i=1}^{N}\sum_{j=1}^{c}\mu_{ij}d(\boldsymbol{x}_i,\boldsymbol{p}_j) \tag{4}$$

式中 $d(\boldsymbol{x}_i,\boldsymbol{p}_j)=(\boldsymbol{x}_i-\boldsymbol{p}_j)(\boldsymbol{x}_i-\boldsymbol{p}_j)^{\mathrm{T}}$ 表示第 i 个样本和第 j 个聚类原型之间的欧几里德距离的平方,并且必须满足以下的关系式:

$$\sum_{j=1}^{c}u_{ij}\qquad \mu_{ij}\in[0,1] \tag{5}$$

$$\mu_{ij}+v_{ij}+\pi_{ij}=1 \tag{6}$$

根据信息论原理,直觉模糊熵最大能够公正地选取隶属度 μ_{ij} 和非隶属度 v_{ij} 的值,由式(2)可知:

$$H(A)=-\frac{1}{nc\ln 2}\sum_{i=1}^{n}\sum_{j=1}^{c}(\mu_{ij}\ln\mu_{ij}+v_{ij}\ln v_{ij}-(1-\pi_{ij})\ln(1-\pi_{ij})-\pi_{ij}\ln 2) \tag{7}$$

式(7)在式(4)-式(6)的约束下取得极大值,由拉格朗日定理可得目标函数为

$$J(\boldsymbol{U},\boldsymbol{V},\boldsymbol{P})_{\max}=-\sum_{i=1}^{n}\sum_{j=1}^{c}(\mu_{ij}\ln\mu_{ij}+v_{ij}\ln v_{ij}-(1-\pi_{ij})\ln(1-\pi_{ij})-\pi_{ij}\ln 2)-$$
$$\alpha\sum_{i=1}^{n}\sum_{j=1}^{c}\mu_{ij}\cdot d(\boldsymbol{x}_i,\boldsymbol{p}_j)+\sum_{i=1}^{n}\lambda_i\Big(\sum_{t=1}^{c}u_{jt}-1\Big)+\lambda_{ij}(\mu_{ij}+v_{ij}+\pi_{ij}-1) \tag{8}$$

在优化过程中,我们假定 π_{ij} 是已知的,并令 $m_{ij}=1-\pi_{ij}$,最优化的一阶必要条件为

$$\frac{\partial J}{\partial\mu_{ij}}=-(\ln\mu_{ij}+1)-\alpha d(\boldsymbol{x}_i,\boldsymbol{p}_j)+\lambda_{ij}+\lambda_i=0 \tag{9}$$

$$\frac{\partial J}{\partial v_{ij}}=-(\ln v_{ij}+1)+\lambda_{ij}=0 \tag{10}$$

$$\frac{\partial J}{\partial\lambda_{ij}}=\mu_{ij}+v_{ij}+\pi_{ij}-1=0 \tag{11}$$

$$\frac{\partial J}{\partial\lambda_i}=\sum_{t=1}^{c}\mu_{jt}-1=0 \tag{12}$$

通过求解上面的方程组可得

$$\sum_{t=1}^{c} \frac{m_{it}}{e^{\alpha d(x_i,p_j)-\lambda_i}+1} = 1 \tag{13}$$

$$\mu_{ij} = \frac{m_{it}}{e^{\alpha d(x_i,p_j)-\lambda_i}+1} \tag{14}$$

用牛顿迭代法求解式(13)中的e^{λ_i},然后再代入式(14)求μ_{ij}。用类似的方法可以求得$J(\boldsymbol{U},\boldsymbol{V},\boldsymbol{P})$最大时$\boldsymbol{p}_j$的值。令

$$\frac{\partial J}{\partial \boldsymbol{p}_j} = \alpha\mu_{ij}\frac{\partial}{\partial \boldsymbol{p}_j}[(\boldsymbol{x}_i-\boldsymbol{p}_j)(\boldsymbol{x}_i-\boldsymbol{p}_j)^{\mathrm{T}}] = 0 \tag{15}$$

$$\boldsymbol{p}_j = \frac{1}{\sum_{i=1}^{n}\mu_{ij}}\sum_{i=1}^{n}\mu_{ij}\boldsymbol{x}_i \tag{16}$$

根据上面的推导,我们采用交替优化的策略,求解此最优化问题,提出了如下算法:

步骤1 给定聚类类别数c,$2 \leqslant c \leqslant n$,$n$是样本个数,设定迭代停止阈值$\varepsilon$,初始化聚类原型模式$\boldsymbol{P}^{(0)}$,设置迭代计数器$b = 0$,设定$\alpha$的值和$\pi$。

步骤2 $i = 1:n$,分别利用式(17)迭代求e^{λ_i},然后代入式(18)求μ_{ij},即

$$e^{\lambda_i} = \frac{1}{\sum_{t=1}^{c}\frac{m_{it}}{e^{\alpha d(x_i,p_j)}e^{\lambda_i}}} \tag{17}$$

$$\mu_{ij}^{(b)} = \frac{m_{ij}}{e^{\alpha d(x_i,p_j)-\lambda_i}+1} \tag{18}$$

步骤3 利用式(19)更新聚类原型模式矩阵$\boldsymbol{P}^{(b+1)}$:

$$\boldsymbol{p}_j^{(b+1)} = = \frac{1}{\sum_{i=1}^{n}\mu_{ij}^{(b)}m_{ij}}\sum_{i=1}^{n}\mu_{ij}^{(b)}\boldsymbol{x}_i \tag{19}$$

步骤4 如果$\|\boldsymbol{P}^{(b+1)}-\boldsymbol{P}^{(b)}\| < \varepsilon$,则算法停止并输出矩阵$\boldsymbol{U}$、$\boldsymbol{V}$和聚类原型$\boldsymbol{P}$,否则令$b = b+1$,转向步骤2。其中$\|\cdot\|$为某种合适的矩阵范数。

4 算例和算法性能分析

实验采用IRIS数据作为验证数据,IRIS数据是国际公认的比较无监督聚类方法效果好坏的典型数据,它包含3类,每一类各有50个样本点,每个样本点有4个属性,数据特点是第一类和其它类离的较远,第二类和第三类数据离的较近,且有部分重叠。

实验中,$\alpha = 0.8$,$\pi_{ij} = 0.4(1-\exp(\|x_i-p_j\|_2^2)/4))$,在决定样本类的归属时,当$\mu_{ij} > 0.55$时,认为样本$x_i$属于第$j$类,当$\mu_{ij} < 0.55$时,分别计算直觉模糊熵$H(A = \{\langle (x_i,p_j),\mu_{ij},v_{ij}\rangle | p_j \in P\})$,其中$j = 1,2,\cdots,c$,认为样本$x_i$属于熵最大的那一类,直觉模糊聚类与模糊聚类的算法性能比较如表1所示。在步骤2中增加了一个用牛顿迭代求解方程,所以运算复杂度较传统的模糊聚类增大,但总的迭代次数减小;因为求得样本类归属的非隶属度和犹豫因子,获取了更多的关于样本分类的信息,因此可以通过有效的决策方法,提高分类的精度。

表1 性能比较

Tab. 1 Performance comparision

聚类算法	误分个数	误分率	迭代次数	聚类中心
传统模糊聚类	16	10.67%	24	$\begin{pmatrix}5.889\,0 & 2.761\,2 & 4.364\,0 & 1.397\,3\\ 5.003\,6 & 3.403\,0 & 1.485\,0 & 0.251\,9\\ 6.774\,9 & 3.052\,4 & 5.646\,6 & 2.053\,5\end{pmatrix}$
直觉模糊聚类	9	6%	21	$\begin{pmatrix}6.474\,2 & 2.945\,9 & 5.201\,6 & 1.813\,3\\ 6.060\,1 & 2.832\,9 & 4.573\,8 & 1.523\,4\\ 5.006\,6 & 3.385\,5 & 1.518\,4 & 0.267\,4\end{pmatrix}$

5 结束语

本文提出了一种基于直觉模糊熵的直觉模糊聚类,推导了迭代算法,并用典型的实验数据验证了模型的可靠性和高效性。步骤2中使用了牛顿迭代法求解系数,所以当数据中存在野点时,对算法的收敛和性能有较大的影响,在后续的研究中将着重解决这个问题,并研究算法的快速性和样本类的归属决策的方法,进一步提高分类精度。

参考文献:

[1] Dunn J C. A Graph Theoretic Analysis of Pattern Classification Via Tamura's Fuzzy Relation[J]. IEEE Trans. SMC, 1974, 4(3):310-313.

[2] Bezdek J C. Pattern Recognition With Fuzzy Objective Function Algorithms[M]. New York: Plenum Press, 1981.

[3] Atanassov K. Intuitionistic Fuzzy Sets [J]. Fuzzy Sets and Systems, 1986, 20(1): 87-96.

[4] 雷英杰,王　涛,赵　晔,等. 直觉模糊匹配的语义距离与贴近度[J]. 空军工程大学学报:自然科学版,2005, 6(1): 69-72.
LEI Yingjie, WANG Tao, ZHAO Ye, et al. Semantic Distance and Close Degree of Intuitionistic Fuzzy[J]. Journal of Air Force Engineering University: Natural Science Edition, 2005, 6(1): 69-72. (in Chinese)

[5] De Luca A, Termini S. A Definition of a Nonprobabilistic Entropy in the Setting of Fuzzy Sets Theory[J]. Inform. Control, 1972, 3(2): 301-312.

[6] Eulalia Szmidt, Janusz Kacprxyk. Entropy for Intuitionstic Fuzzy Sets[J]. Fuzzy Sets and Systems, 2001, 35(4): 467-477.

[7] Vlachos I K, Sergiadis G D. Intuitionistic Fuzzy Information - Applications to Pattern Recognition[J]. Pattern Recognition Letters, 2007, 28: 197-206.

[8] Gao Xinbo, Ji Hong bing, Xie wei xin. A Novel FCM Clustering Algorithms: Proceedings of International Conference on Signal Processing(ICSP2000)[C]. Beiging: [s. i]. 2000: 1457-1461.

[9] Pal N R, Bezdek J C. On Clustering for the Fuzzy Gmeans Model[J]. IEEE Trans FS, 1995, 13(3): 370-379.

[10] Pezdrcy W, Waltezky J. Fuzzy Clustering with Partial Supervision[J]. IEEE Trans SMC, 1997, 27(5): 787-795.

(编辑:田新华)

Intuitionistic Fuzzy Clustering Based on Intuitionistic Fuzzy Entropy

XU Xiao-lai, LEI Ying-jie, ZHAO Xue-jun

(The Missile Institute, Air Force Engineering University, Sanyuan 713800, Shaanxi, China)

Abstract: Intuitionistic fuzzy sets defined by Atanassov are of the most influential extension and evolution of fuzzy sets defined by Zadeh, so making the fuzzy clustering expand into an intuitionistic fuzzy clustering is of significant value in application research. This paper presents an intuitionistic fuzzy clustering based on maximum entropy, and deduces an iterative algorithm. At last, the typical experiment indicates that the intuitionistic fuzzy clustering is superior to the fuzzy clustering in performance and the clustering right rate is improved.

Key words: intuitionistic fuzzy clustering; intuitionistic fuzzy entropy; intuitionistic fuzzy sets; fuzzy clustering