Intuitionistic Fuzzy Clustering Based on Cauchy Kernel Function

Z R Tang [a], F X Zhao [b]

College of Nuclear Technology and Automation Engineering, Chengdu University of Technology, Chengdu 610000, China.

[a]zhirong_tang_1993@126.com, [b]FeixiangZhao_0625@foxmail.com

Abstract

Existing intuitionistic fuzzy C-means clustering algorithm cannot find non-convex clustering structure. To solve this problem, an intuitionistic fuzzy clustering algorithm based on Cauchy kernel function is proposed in this paper. The Cauchy kernel function is used to transform the intuitionistic fuzzy Euclidean distance into an intuitionistic fuzzy number and particle swarm optimization is used to optimize the objective function. In theory, the algorithm can complete intuitionistic fuzzy clustering quickly and accurately.

Keywords

Intuitionistic fuzzy set, Particle Swarm Optimization, C-means.

1. Introduction

Ruspini [1] introduced the Zadeh fuzzy set theory into cluster analysis, first proposed the concept of fuzzy partitioning. Subsequently, the researchers proposed a variety of fuzzy clustering analysis methods, it mainly includes fuzzy equivalence relations-based transitive closure methods, methods based on similarity relations and fuzzy relations, and maximum tree methods based on fuzzy graph theory. However, due to the high computational complexity, these methods are difficult to apply to big data problems and real-time requirements, so they have gradually lost value in practical applications and research [2]. As one of the most important forms of Zadeh's fuzzy theory, the intuitionistic fuzzy set (IFS) increases the hesitation attribute parameters, which further expands and enhances the description and processing functions of fuzzy set theory on complex uncertainty knowledge. At the same time, it provides new ideas and methods for the modeling and processing of fuzzy uncertainty information [3].

In 1995, Cortes and Vapnik [4] proposed the Support Vector Machine (SVM) theory. SVM demonstrates better performance than traditional classifiers in many areas, making kernel methods increasingly valued and applied to all aspects of machine learning [5, 6]. Girolami [7] creatively proposed the fuzzy kernel c-means algorithm (FKCM) algorithm to solve the problem that the FCM algorithm can not find the non-convex clustering structure. Girolami [7] proposed an Intuitionistic Fuzzy Kernel c-means Clustering Algorithm (IFKCM). However, the sum of the memberships of the samples relative to each category is 1, which is inconsistent with the idea of intuitionistic ambiguity. Piciarelli C. et al [8] proposed a kernel clustering method for adaptively determining the number of clusters by extracting the geometric properties of kernel space. In [9, 10], a compromise weight fuzzy factor and kernel distance metric are introduced in the classical FCM algorithm, and a kernel clustering algorithm based on fuzzy factor is proposed and applied to the field of image segmentation. Among them, the closeness can be used as the price index of the optimal attribute value and the worst attribute [11, 12], and can also be used to make attribute simplicity [13], which has attracted wide attention of researchers [14, 15].

Based on the concept of intuitionistic fuzzy class C-means clustering, an intuitionistic fuzzy clustering algorithm based on Cauchy kernel function is proposed in this paper. The Cauchy kernel function is used to transform the intuitionistic fuzzy Euclidean distance into an intuitionistic fuzzy number, and the particle swarm optimization algorithm is used to optimize the objective function. In theory, the algorithm can complete intuitionistic fuzzy clustering quickly and accurately, and solve the problem that the original intuitionistic fuzzy C-means clustering algorithm can not find the non-convex clustering structure.

2. Intuitionistic Fuzzy Clustering Based on Cauchy Kernel Function

2.1 Intuitionistic fuzzy set theory

Given a data set $\boldsymbol{X}=\{x_1,x_2,\cdots,x_N\}\subset \boldsymbol{R}^s$, which is a set of finite set of observations for N patterns in the pattern space. $x_i=\left(\langle x\mu_{i1},x\gamma_{i1},x\pi_{i1}\rangle,\langle x\mu_{i2},x\gamma_{i2},x\pi_{i2}\rangle,\langle x\mu_{is},x\gamma_{is},x\pi_{is}\rangle\right)^T$ is the eigenvector of the observed sample, and the assignment $\langle x\mu_{ij},x\gamma_{ij},x\pi_{ij}\rangle$ of the eigenvector on each dimension is an intuitionistic fuzzy number. $P=\{p_1,p_1,p_K\}$ is a set of K clustering prototypes, K is the number of cluster categories and p_k is the clustering prototype vector of the k-th class, where $p_k=\left(\langle p\mu_{k1},p\gamma_{k1},p\pi_{k1}\rangle,\langle p\mu_{k2},p\gamma_{k2},p\pi_{k2}\rangle,\langle p\mu_{ks},p\gamma_{ks},p\pi_{ks}\rangle\right)^T$. The assignment $p_{ik}=\langle p\mu_{ik},p\gamma_{ik},p\pi_{ik}\rangle^T$ on the i-dimensional feature of p_k is also an intuitionistic fuzzy number.

2.2 Similarity criterion

Both the sample x_i and the cluster prototype p_k can be represented by an intuitionistic fuzzy set, and the distance measure function between them can be defined as follows:

$$\begin{aligned}D_\omega(x_i,p_k)&=\frac{1}{\sqrt{2}}\sqrt{(x\mu_i-p\mu_k)W(x\mu_i-p\mu_k)^T+(x\gamma_i-p\gamma_k)W(x\gamma_i-p\gamma_k)^T+(x\pi_i-p\pi_k)W(x\pi_i-p\pi_k)^T}\\&=\frac{1}{\sqrt{2}}\sqrt{\left\|W^{\frac{1}{2}}(x\mu_i-p\mu_k)\right\|^2+\left\|W^{\frac{1}{2}}(x\gamma_i-p\gamma_k)\right\|^2+\left\|W^{\frac{1}{2}}(x\pi_i-p\pi_k)\right\|^2}\end{aligned} \tag{1}$$

Where $x\mu_i$, $p\mu_i$ represents the membership degree vector. $x\gamma_i$, $p\gamma_i$ represents the non-membership degree vector. $x\pi_i$, $p\pi_i$ represents the hesitation degree vector. $x\mu_i+x\gamma_i+x\pi_i=I$, $p\mu_i+p\gamma_i+p\pi_i=I$ (I is the s-dimensional unit vector) is satisfied. The matrix W is the weighted diagonal matrix. $\|\bullet\|^2$ represents the L-2 norm. $\omega_i\,(i=1,2,\cdots,s)$ is the weight added to the i-dimensional feature and ω_i satisfies the normalization condition as shown in equation (2).

$$\sum_{i=1}^{s}\omega(i)\Big/s=1 \tag{2}$$

The matrix W can be expressed as follows:

$$W=diag\{\omega_1,\omega_2,\cdots,\omega_s\} \tag{3}$$

Rewrite equation (1) as a kernel function:

$$\hat{D}_\omega(x_i,p_k)=\frac{1}{\sqrt{2}}\sqrt{\left\|\phi\left(W^{\frac{1}{2}}x\mu_i\right)-\phi\left(W^{\frac{1}{2}}p\mu_k\right)\right\|^2+\left\|\phi\left(W^{\frac{1}{2}}x\gamma_i\right)-\phi\left(W^{\frac{1}{2}}p\gamma_k\right)\right\|^2+\left\|\phi\left(W^{\frac{1}{2}}x\pi_i\right)-\phi\left(W^{\frac{1}{2}}p\pi_k\right)\right\|^2} \tag{4}$$

$\phi(\bullet)$ is the mapping function. It is easy to prove that equation (4) satisfies the four axioms of the intuitionistic fuzzy set dissimilarity measure, which can be used as the distance measure formula between intuitionistic fuzzy sets. The norm in the formula can be further expressed as:

$$\|\phi(x)-\phi(p)\|^2=K(x,x)+K(p,p)-K(x,p)-K(p,x) \tag{5}$$

$K(\bullet)$ is the kernel function. Use the Cauchy kernel function:

$$K(x,p)=\frac{1}{1+\|x-p\|^2/\sigma} \quad (6)$$

σ indicates the standard deviation. Obviously the Cauchy kernel function satisfies the following properties:

$$\begin{cases} K(x,x)=1 \\ K(x,p)=K(p,x) \end{cases} \quad (7)$$

Bring the Cauchy kernel function into equation (5):

$$\|\phi(x)-\phi(p)\|^2=2-2K(x,p) \quad (8)$$

Therefore, equation (4) can be expressed as:

$$\hat{D}_\omega(x_i,p_k)^2=3-K\left(W^{\frac{1}{2}}x\mu_i,W^{\frac{1}{2}}p\mu_k\right)-K\left(W^{\frac{1}{2}}x\gamma_i,W^{\frac{1}{2}}p\gamma_k\right)-K\left(W^{\frac{1}{2}}x\pi_i,W^{\frac{1}{2}}p\pi_k\right) \quad (9)$$

2.3 Parameter solving

2.3.1 Objective function

Firstly, the objective function of intuitionistic fuzzy clustering based on Cauchy kernel function is given:

$$J_m(U_\mu,U_\lambda,P)=\sum_{i=1}^{N}\sum_{k=1}^{K}\left((\mu_{ik})^m/2+(1-\gamma_{ik})^m/2\right)\hat{D}_\omega(x_i,p_k)^2$$
$$m\in[1,\infty),\mu_{ik}\in[0,1],U_\lambda\in[0,1] \quad (10)$$

Where m is the smoothing parameter. U_μ is the kernel-based fuzzy partition membership matrix; U_λ is the kernel-based fuzzy partition non-membership matrix. The constraints of the objective function are:

$$\begin{cases} \sum_{i=1}^{N}\sum_{k=1}^{K}\mu_{ik}=N \\ \mu_{ik}+\gamma_{ik}+\pi_{ik}=1 \end{cases} \quad (11)$$

2.3.2 Cost function

Here the cost function takes the minimum value of the objective function, that is:

$$I(\theta)=\arg\min_\theta J_m(U_\mu,U_\lambda,P) \quad (12)$$

Formula (13) can be obtained from the Lagrange theorem:

$$\mu_{ik}=\left[\frac{\hat{D}_\omega(x_i,p_k)}{\sum_{j=1}^{K}\hat{D}_\omega(x_i,p_j)}\right]^{\frac{m-1}{2}} \quad (13)$$

2.3.3 Parameter Determination Based on Particle Swarm Optimization Algorithm (PSO)

The parameters that need to be solved are $p\mu_k, p\gamma_k, p\pi_k$.

Suppose that in an s-dimensional target search space, there are N particles forming a community, where the i-th particle is represented as a vector of s-dimensional, and the speed at time t is updated as follows:

$$\begin{cases} V_{\mu iD}^{(t+1)} = V_{\mu iD}^{(t)} + c_1 r_1 \left(p_{\mu iD}^{(t)} - p\boldsymbol{\mu}_{kiD}^{(t)}\right) + c_2 r_2 \left(p_{\mu gD}^{(t)} - p\boldsymbol{\mu}_{kiD}^{(t)}\right), i = 1, 2, \cdots, N \\ V_{\gamma iD}^{(t+1)} = V_{\gamma iD}^{(t)} + c_1 r_1 \left(p_{\gamma iD}^{(t)} - p\boldsymbol{\gamma}_{kiD}^{(t)}\right) + c_2 r_2 \left(p_{\gamma gD}^{(t)} - p\boldsymbol{\gamma}_{kiD}^{(t)}\right), i = 1, 2, \cdots, N \end{cases} \tag{14}$$

Where $V_{iD}^{(t)}$ represents the speed change of the i-th particle at time t. c_1, c_2 represents a learning factor, indicating a pseudo-random number uniformly distributed in the region, the value range is [0, 1]. $p_{iD}^{(t)}$ represents the best historical position of the i-th particle experienced at time t. $p_{gD}^{(t)}$ represents the best position experienced by all particles in the group at time t and $W_{iD}^{(t)}$ represents the position of the i-th particle at time t.

After the speed update, the update of position of current time is performed. The parameter location is updated as follows:

$$\begin{cases} p\boldsymbol{\mu}_k^{(t+1)} = p\boldsymbol{\mu}_k^{(t)} + V_{\mu iD}^{(t+1)} \\ p\boldsymbol{\gamma}_k^{(t+1)} = p\boldsymbol{\gamma}_k^{(t)} + V_{\gamma iD}^{(t+1)} \end{cases} \tag{15}$$

According to $p\mu_k + p\gamma_k + p\pi_k = I$, the following formula can be obtained:

$$p\boldsymbol{\pi}_k = I - p\boldsymbol{\mu}_k - p\boldsymbol{\gamma}_k \tag{16}$$

Repeat the above steps until convergence.

3. Conclusion

In this paper, the kernel method and the intuitionistic fuzzy clustering algorithm are effectively combined and an intuitionistic fuzzy Euclidean distance based on Cauchy kernel is proposed. Then an intuitionistic fuzzy c-means clustering algorithm based on Cauchy nucleation distance is proposed. The algorithm solves the problem that the classic intuitionistic fuzzy c-means clustering algorithm can not find the non-convex clustering structure.

References

[1] E. H. Ruspini: A new approach to clustering, Information & Control, Vol. 15(1969)No. 1, p. 22-32.

[2] M. Ceccarelli , A. Maratea: Improving fuzzy clustering of biological data by metric learning with side information, International Journal of Approximate Reasoning, Vol. 47(2008)No. 1, p.45-57.

[3] Y. Song, X. Wang, L. Lei: Combination of interval-valued belief structures based on intuitionistic fuzzy set, Knowledge-Based Systems, Voi. 67(2014), p.61-70.

[4] C. Cortes , V. Vapnik: Support-Vector Networks, Machine Learning, Vol. 20(1995)No. 3, p.273-297.

[5] A. Saket, M. Sushil, T. Oncel, M. Peter: Semi-supervised ker- nel mean shift clustering, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 36(2014)No. 6, p.1201-1215.

[6] R. P. Nikhil, S. Kaushik: What and when can we gain from the kernel versions of c-means algorithm?, IEEE Transactions on Fuzzy Systems, Vol. 22(2014)No. 2, p.363-379.

[7] M. Girolami: Mercer kernel-based clustering in feature space, IEEE Transactions on Neural Networks, Vol. 13(2002)No. 3, p.780-784.

[8] C. Piciarelli, C. Micheloni, G. L. Foresti: Kernel-based clustering, Electronics Letters, Vol.

49(2011)No. 2, p.113-114.

[9] D. Xiang, T. Tang, C. Hu: A Kernel Clustering Algorithm With Fuzzy Factor: Application to SAR Image Segmentation, IEEE Geoscience and Remote Sensing Letters, Vol. 11(2014)No. 7, p.1290-1294.

[10] M. G. Gong: Fuzzy c-means clustering with local information and kernel metric for image segmentation, IEEE Transactions on Image Processing, Vol. 22(2013)No. 2, p.573-584.

[11] C. Sunil: Hybrid FAHP (fuzzy analytical hierarchy process)-FTOPSIS (fuzzy technique for order preference by similarity of an ideal solution) approach for performance evaluation of the V down perforated baffle roughened rectangular channel, Energy, Vol. 84(2015), p.432-442.

[12] S. H. Zyoud, D. Fuchshanusch: A bibliometric-based survey on AHP and TOPSIS techniques, Expert Systems with Applications, Vol. 78(2017), p.158-181.

[13] S. Chebrolu, S. G. Sanjeevi: Attribute Reduction on Continuous Data in Rough Set Theory using Ant Colony Optimization Metaheuristic, *The third International Symposium on Women in Computing & Informatics*. (ACM, India, August 10-13, 2015), p.17-24.

[14] X. Zhang, J. Dai, Y. Yu: On the union and intersection operations of rough sets based on various approximation spaces, Information Sciences, Vol. 292(2015), p.214-229.

[15] J. Li, H. Zhao, W. Zhu: Fast randomized algorithm with restart strategy for minimal test cost feature selection, International Journal of Machine Learning and Cybernetics, Vol. 6(2015)No. 3, p.435-442.