聚类——认识FCM算法

作者：凯鲁嘎吉 - 博客园 http://www.cnblogs.com/kailugaji/

一、FCM概述

FCM算法是基于对目标函数的优化基础上的一种数据聚类方法。聚类结果是每一个数据点对聚类中心的隶属程度，该隶属程度用一个数值来表示。该算法允许同一数据属于多个不同的类。

FCM算法是一种无监督的模糊聚类方法，在算法实现过程中不需要人为的干预。

这种算法的不足之处:首先，算法中需要设定一些参数，若参数的初始化选取的不合适，可能影响聚类结果的正确性;其次，当数据样本集合较大并且特征数目较多时，算法的实时性不太好。

K-means也叫硬C均值聚类（HCM），而FCM是模糊C均值聚类，它是HCM的延伸与拓展，HCM与FCM最大的区别在于隶属函数（划分矩阵）的取值不同，HCM的隶属函数只取两个值：0和1，而FCM的隶属函数可以取[0,1]之间的任何数。K-means和FCM都需要事先给定聚类的类别数，而FCM还需要选取恰当的加权指数α，α的选取对结果有一定的影响，α属于[0,+∞)。

二、FCM算法

C是聚类数，N是样本个数。U是隶属度矩阵，V是聚类中心。

目标函数：

$$\min\ J_{FCM}(U,V)=\sum_{k=1}^{C}\sum_{i=1}^{N}u_{ki}^{m}\left\|x_i-v_k\right\|^2$$

$$s.t.\ \sum_{k=1}^{C}u_{ki}=1,\ u_{ki}\in[0,1]$$

更新公式：

Let cluster centers V be fixed, $F(U)$ is minimized if

$$u_{ki} = \frac{\left(\left\|x_i - v_k\right\|^2\right)^{-\frac{1}{m-1}}}{\sum_{j=1}^{C}\left(\left\|x_i - v_j\right\|^2\right)^{-\frac{1}{m-1}}} \tag{1}$$

$$for\ 1 \le i \le N,\ \ 1 \le k \le C$$

Let membership degrees U be fixed, $F(V)$ is minimized if

$$v_k = \frac{\sum_{i=1}^{N} u_{ki}^m x_i}{\sum_{i=1}^{N} u_{ki}^m} \tag{2}$$

$$for\ 1 \le k \le C$$

三、算法流程

- Fuzzy c-means algorithm

1. Randomly initialize class centers $V^{(0)}$; Give fuzzification parameter m; Set the maximum number of iterations t_{max} and threshold $\varepsilon>0$;
2. Update the membership maxtrix U using Eq.(1)
3. Update the class centers V use Eq.(2)
4. If $||U^{(t+1)}-U^{(t)}||<\varepsilon$ or $t=t_{max}$,then terminate; else $t=t+1$, go to step 2;

四、FCM改进算法汇总

Algorithm name	Abbreviation	Objective function	Update formula of degree of membership	Update formula of clustering center
Fuzzy C-Means [11]	FCM	$J_m=\sum_{i=1}^{c}\sum_{j=1}^{n}u_{ij}^m\left\|\mathbf{x}_j-\mathbf{v}_i\right\|^2,\quad s.t.\quad \sum_{i=1}^{c}u_{ij}=1$	$u_{ij}=\left(\sum_{k=1}^{c}\left(\frac{\left\|\mathbf{x}_j-\mathbf{v}_i\right\|^2}{\left\|\mathbf{x}_j-\mathbf{v}_k\right\|^2}\right)^{\frac{1}{m-1}}\right)^{-1}$	$\mathbf{v}_i=\frac{\sum_{j=1}^{n}u_{ij}^m\mathbf{x}_j}{\sum_{j=1}^{n}u_{ij}^m}$
Possibilistic C-Means [12]	PCM	$J_m=\sum_{i=1}^{c}\sum_{j=1}^{n}t_{ij}^m\left\|\mathbf{x}_j-\mathbf{v}_i\right\|^2+\sum_{i=1}^{c}\lambda_i\sum_{j=1}^{n}\left(1-t_{ij}\right)^m$, where $\lambda_i=\frac{\sum_{j=1}^{n}u_{ij}^m\left\|\mathbf{x}_j-\mathbf{v}_i\right\|^2}{\sum_{j=1}^{n}u_{ij}^m}$	$t_{ij}=\left(1+\left(\frac{\left\|\mathbf{x}_j-\mathbf{v}_i\right\|^2}{\gamma_i}\right)^{\frac{1}{m-1}}\right)^{-1}$	$\mathbf{v}_i=\frac{\sum_{j=1}^{n}t_{ij}^m\mathbf{x}_j}{\sum_{j=1}^{n}t_{ij}^m}$
Fuzzy-Possibilistic C-Means [14]	FPCM	$J_m=\sum_{i=1}^{c}\sum_{j=1}^{n}\left(u_{ij}^m+t_{ij}^\theta\right)\left\|\mathbf{x}_j-\mathbf{v}_i\right\|^2,\quad s.t.\quad \sum_{j=1}^{n}t_{ij}=1$	$u_{ij}=\left(\sum_{k=1}^{c}\left(\frac{\left\|\mathbf{x}_j-\mathbf{v}_i\right\|^2}{\left\|\mathbf{x}_j-\mathbf{v}_k\right\|^2}\right)^{\frac{1}{m-1}}\right)^{-1},\quad t_{ij}=\left(\sum_{r=1}^{n}\left(\frac{\left\|\mathbf{x}_j-\mathbf{v}_i\right\|^2}{\left\|\mathbf{x}_r-\mathbf{v}_i\right\|^2}\right)^{\frac{1}{\theta-1}}\right)^{-1}$	$\mathbf{v}_i=\frac{\sum_{j=1}^{n}\left(u_{ij}^m+t_{ij}^\theta\right)\mathbf{x}_j}{\sum_{j=1}^{n}\left(u_{ij}^m+t_{ij}^\theta\right)}$
Possibilistic Fuzzy C-Means [15]	PFCM	$J_m=\sum_{i=1}^{c}\sum_{j=1}^{n}\left(\alpha u_{ij}^m+\beta t_{ij}^\theta\right)\left\|\mathbf{x}_j-\mathbf{v}_i\right\|^2+\sum_{i=1}^{c}\lambda_i\sum_{j=1}^{n}\left(1-t_{ij}\right)^\theta$	$u_{ij}=\left(\sum_{k=1}^{c}\left(\frac{\left\|\mathbf{x}_j-\mathbf{v}_i\right\|^2}{\left\|\mathbf{x}_j-\mathbf{v}_k\right\|^2}\right)^{\frac{1}{m-1}}\right)^{-1},\quad t_{ij}=\left(1+\left(\frac{\beta}{\gamma_i}\left\|\mathbf{x}_j-\mathbf{v}_i\right\|^2\right)^{\frac{1}{\theta-1}}\right)^{-1}$	$\mathbf{v}_i=\frac{\sum_{j=1}^{n}\left(\alpha u_{ij}^m+\beta t_{ij}^\theta\right)\mathbf{x}_j}{\sum_{j=1}^{n}\left(\alpha u_{ij}^m+\beta t_{ij}^\theta\right)}$
FCM with Spatial constraints [16]	FCM_S	$J_m=\sum_{i=1}^{c}\sum_{j=1}^{n}u_{ij}^m\left\|x_j-v_i\right\|^2+\frac{\alpha}{N_R}\sum_{i=1}^{c}\sum_{j=1}^{n}u_{ij}^m\left(\sum_{r\in N_j}\left\|x_r-v_i\right\|^2\right)$	$u_{ij}=\left(\sum_{k=1}^{c}\left(\frac{\left\|x_j-v_i\right\|^2+\frac{\alpha}{N_R}\sum_{r\in N_j}\left\|x_r-v_i\right\|^2}{\left\|x_j-v_k\right\|^2+\frac{\alpha}{N_R}\sum_{r\in N_j}\left\|x_r-v_k\right\|^2}\right)^{\frac{1}{m-1}}\right)^{-1}$	$v_i=\frac{\sum_{j=1}^{n}u_{ij}^m\left(x_j+\frac{\alpha}{N_R}\sum_{r\in N_j}x_r\right)}{\left(1+\alpha\right)\sum_{j=1}^{n}u_{ij}^m}$

Two variants of FCM_S [17]	FCM_S1	$J_m=\sum_{i=1}^{c}\sum_{j=1}^{n}u_{ij}^{m}\left\|x_j-v_i\right\|^2+\alpha\sum_{i=1}^{c}\sum_{j=1}^{n}u_{ij}^{m}\left\|\overline{x}_j-v_i\right\|^2$		$u_{ij}=\left(\sum_{k=1}^{c}\left(\frac{\left\|x_j-v_i\right\|^2+\alpha\left\|\overline{x}_j-v_i\right\|^2}{\left\|x_j-v_k\right\|^2+\alpha\left\|\overline{x}_j-v_k\right\|^2}\right)^{\frac{1}{m-1}}\right)^{-1}$	$v_i=\frac{\sum_{j=1}^{n}u_{ij}^{m}\left(x_j+\alpha\overline{x}_j\right)}{(1+\alpha)\sum_{j=1}^{n}u_{ij}^{m}}$
	FCM_S2	$J_m=\sum_{i=1}^{c}\sum_{j=1}^{n}u_{ij}^{m}\left\|x_j-v_i\right\|^2+\alpha\sum_{i=1}^{c}\sum_{j=1}^{n}u_{ij}^{m}\left\|\widehat{x}_j-v_i\right\|^2$		$u_{ij}=\left(\sum_{k=1}^{c}\left(\frac{\left\|x_j-v_i\right\|^2+\alpha\left\|\widehat{x}_j-v_i\right\|^2}{\left\|x_j-v_k\right\|^2+\alpha\left\|\widehat{x}_j-v_k\right\|^2}\right)^{\frac{1}{m-1}}\right)^{-1}$	$v_i=\frac{\sum_{j=1}^{n}u_{ij}^{m}\left(x_j+\alpha\widehat{x}_j\right)}{(1+\alpha)\sum_{j=1}^{n}u_{ij}^{m}}$
Enhanced FCM [18]	EnFCM	$J_m=\sum_{i=1}^{c}\sum_{j=1}^{l}\gamma_j u_{ij}^{m}\left(\xi_j-v_i\right)^2$	$\xi_j=\frac{1}{1+\alpha}\left(x_j+\frac{\alpha}{N_R}\sum_{r\in N_j}x_r\right)$	$u_{ij}=\left(\sum_{k=1}^{c}\left(\frac{\xi_j-v_i}{\xi_j-v_k}\right)^{\frac{2}{m-1}}\right)^{-1}$	$v_i=\frac{\sum_{j=1}^{q}\gamma_j u_{ij}^{m}\xi_j}{\sum_{j=1}^{q}\gamma_j u_{ij}^{m}}$
Fast Generalized FCM [19]	FGFCM		$\xi_j=\frac{\sum_{r\in N_j}S_{jr}x_r}{\sum_{r\in N_j}S_{jr}}$, where $S_{jr}=\begin{cases}e^{-\frac{\max\left(\left\|p_j-p_r\right\|,\left\|q_j-q_r\right\|\right)}{\lambda_s}-\frac{\left\|x_j-x_r\right\|^2}{\lambda_g\sigma_j^2}}, & j\neq r,\\ 0, & j=r\end{cases}$ and $\sigma_j=\sqrt{\frac{\sum_{r\in N_j}\left\|x_j-x_r\right\|^2}{N_R}}$		

Fuzzy Weighted C-Means [20]	FWCM	$J_m = \sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^m \left\Vert \mathbf{x}_j - M_{ij} \right\Vert^2$ $M_{ij} = \frac{\sum_{r=1,r\neq j}^{n} \left\Vert \mathbf{x}_j - \mathbf{x}_r \right\Vert^{-1} u_{ir}\mathbf{x}_r}{\sum_{r=1,r\neq j}^{n} \left\Vert \mathbf{x}_j - \mathbf{x}_r \right\Vert^{-1} u_{ir}}$	$u_{ij} = \zeta_j^{1/(m-1)} \left(\left\Vert \mathbf{x}_j - M_{ij} \right\Vert^2 m \sum_{i=1}^{c} u_{ij}^{m-1} \right)^{1/(1-m)}$, where $\zeta_j = \left(\sum_{i=1}^{c} \left(\left\Vert \mathbf{x}_j - M_{ij} \right\Vert^2 m \sum_{i=1}^{c} u_{ij}^{m-1} \right)^{1/(1-m)} \right)^{1-m}$	
New Weighted Fuzzy C-Means [21]	NW_FCM	$M_{ij} = \frac{\sum_{r=1,r\neq j}^{n} \left\Vert \mathbf{x}_r - \mathbf{v}_i \right\Vert^{-1} u_{ir}\mathbf{x}_r}{\sum_{r=1,r\neq j}^{n} \left\Vert \mathbf{x}_r - \mathbf{v}_i \right\Vert^{-1} u_{ir}}$		$\mathbf{v}_i = \frac{\sum_{j=1}^{n} u_{ij}^m \mathbf{x}_j}{\sum_{j=1}^{n} u_{ij}^m}$
Fuzzy Local Information C-Means [22]	FLICM	$J_m = \sum_{i=1}^{c}\sum_{j=1}^{n} \left(u_{ij}^m \left\Vert x_j - v_i \right\Vert^2 + G_{ij} \right)$, where $G_{ij} = \sum_{\substack{r\in N_j \\ j\neq r}} \frac{1}{1+d_{jr}} (1-u_{ir})^m \left\Vert x_r - v_i \right\Vert^2$	$u_{ij} = \left(\sum_{k=1}^{c} \left(\frac{\left\Vert x_j - v_i \right\Vert^2 + G_{ij}}{\left\Vert x_j - v_k \right\Vert^2 + G_{kj}} \right)^{\frac{1}{m-1}} \right)^{-1}$	$v_i = \frac{\sum_{j=1}^{n} u_{ij}^m x_j}{\sum_{j=1}^{n} u_{ij}^m}$
Generalized FCM [23]	GFCM	$J_m = \sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^m \left\Vert \mathbf{x}_j - \mathbf{v}_i \right\Vert_P^P, \quad s.t. \quad \sum_{i=1}^{c} u_{ij} = 1$	$u_{ij} = \left(\sum_{k=1}^{c} \left(\frac{\left\Vert \mathbf{x}_j - \mathbf{v}_i \right\Vert_P^P}{\left\Vert \mathbf{x}_j - \mathbf{v}_k \right\Vert_P^P} \right)^{\frac{1}{m-1}} \right)^{-1}$	$\mathbf{v}_i = \frac{\sum_{j=1}^{n} u_{ij}^m \mathbf{x}_j}{\sum_{j=1}^{n} u_{ij}^m}$
Generalized FCM with Improved Fuzzy Partition [24]	GIFP_FCM	$J_m = \sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^m \left\Vert \mathbf{x}_j - \mathbf{v}_i \right\Vert^2 + \sum_{j=1}^{n} a_j \sum_{i=1}^{c} u_{ij} \left(1 - u_{ij}^{m-1}\right)$, where $a_j = \alpha \cdot \min\left\{ \left\Vert \mathbf{x}_j - \mathbf{v}_k \right\Vert^2 \middle\vert k \in \{1,\cdots,c\} \right\}$	$u_{ij} = \left(\sum_{k=1}^{c} \left(\frac{\left\Vert \mathbf{x}_j - \mathbf{v}_i \right\Vert^2 - a_j}{\left\Vert \mathbf{x}_j - \mathbf{v}_k \right\Vert^2 - a_j} \right)^{\frac{1}{m-1}} \right)^{-1}$	$\mathbf{v}_i = \frac{\sum_{j=1}^{n} u_{ij}^m \mathbf{x}_j}{\sum_{j=1}^{n} u_{ij}^m}$

Algorithm name	Abbreviation	Objective function	Update formula of degree of membership	Update formula of clustering center
Two Kernel variants of GIFP_FCM [27]	KGFCM_S1	$J_m = 2\sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^m \left(1 - K(x_j, v_i)\right) + \sum_{j=1}^{n} a_j \sum_{i=1}^{c} u_{ij}\left(1 - u_{ij}^{m-1}\right)$ $+2\beta \sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^m \left(1 - K(\overline{x}_j, v_i)\right)$	$u_{ij} = \left(\sum_{k=1}^{c} \left(\frac{2\left(1-K(x_j,v_i)\right) - a_j + 2\beta\left(1-K(\overline{x}_j,v_i)\right)}{2\left(1-K(x_j,v_k)\right) - a_j + 2\beta\left(1-K(\overline{x}_j,v_k)\right)} \right)^{\frac{1}{m-1}} \right)^{-1}$	$v_i = \frac{\sum_{j=1}^{n} u_{ij}^m \left(K(x_j,v_i)\cdot x_j + \beta K(\overline{x}_j,v_i)\cdot \overline{x}_j \right)}{\sum_{j=1}^{n} u_{ij}^m \left(K(x_j,v_i) + \beta K(\overline{x}_j,v_i) \right)}$
	KGFCM_S2	$J_m = 2\sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^m \left(1 - K(x_j, v_i)\right) + \sum_{j=1}^{n} a_j \sum_{i=1}^{c} u_{ij}\left(1 - u_{ij}^{m-1}\right)$ $+2\beta \sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^m \left(1 - K(\widehat{x}_j, v_i)\right)$	$u_{ij} = \left(\sum_{k=1}^{c} \left(\frac{2\left(1-K(x_j,v_i)\right) - a_j + 2\beta\left(1-K(\widehat{x}_j,v_i)\right)}{2\left(1-K(x_j,v_k)\right) - a_j + 2\beta\left(1-K(\widehat{x}_j,v_k)\right)} \right)^{\frac{1}{m-1}} \right)^{-1}$	$v_i = \frac{\sum_{j=1}^{n} u_{ij}^m \left(K(x_j,v_i)\cdot x_j + \beta K(\widehat{x}_j,v_i)\cdot \widehat{x}_j \right)}{\sum_{j=1}^{n} u_{ij}^m \left(K(x_j,v_i) + \beta K(\widehat{x}_j,v_i) \right)}$

推荐：Clustering - Fuzzy C-means

FCM改进版本汇总参考：J. Gu, L. Jiao, S. Yang and F. Liu, "Fuzzy Double C-Means Clustering Based on Sparse Self-Representation," in IEEE Transactions on Fuzzy Systems, vol. 26, no. 2, pp. 612-626, April 2018.