基于模糊 Fisher 准则的自适应降维模糊聚类算法

支晓斌[①]　范九伦[②]
[①](西安电子科技大学电子工程学院　西安　710071)
[②](西安邮电学院信息与控制系　西安　710121)

摘　要：该文指出曹苏群等人提出的基于模糊 Fisher 准则(FFC)的半模糊聚类算法(FFC-SFCA)中的一个推导错误，结合模糊紧性和分离性(FCS)聚类算法提出新的聚类算法：FFC-FCS。FFC-FCS 充分利用 FFC 的特征提取和降维特性，交替运行原始数据空间中 FFC 和投影空间中的 FCS，通过对降维数据的聚类实现对原始数据的聚类。FFC-FCS 不仅对低维数据具有优异的分类性能而且对高维数据也表现出一定的分类优势。实验结果表明，FFC-FCS 的性能明显优于原有的 FCS 算法，FFC-SFCA 算法以及经典的模糊 C-均值(FCM)算法。

关键词：模糊散布矩阵；模糊 Fisher 准则；最优投影矢量；FCS 聚类
中图分类号：TP181　　**文献标识码**：A　　**文章编号**：1009-5896(2009)11-2653-06

Fuzzy Fisher Criterion Based Adaptive Dimension Reduction Fuzzy Clustering Algorithm

Zhi Xiao-bin[①]　Fan Jiu-lun[②]
[①](*School of Electronic Engineering*, *Xidian University*, *Xi'an* 710071, *China*)
[②](*Department of Information and Control, Xi'an Institute of Post and Telecommunications*, *Xi'an* 710121, *China*)

Abstract: The derivation mistake in Cao's Fuzzy Fisher Criterion (FFC) based Semi-Fuzzy Clustering Algorithm (FFC-SFCA) is pointed out. Combining Fuzzy Compactness and Separation (FCS) clustering algorithm, a new clustering algorithm, FFC-FCS, is proposed in this paper. FFC-FCS make full use of the feature extraction and dimension reduction characteristics of FFC, alternately running FFC in the original data space and FCS in the projection space, clustering the original data is accomplished by clustering the dimension reduction data. FFC-FCS not only shows excellent capability of classifying low dimensional data but also has a certain grade classification advantage with respect to high dimensional data. The experimental results show that FFC-FCS has super performance over original FCS, FFC-SFCA and classical Fuzzy C-Means(FCM).

Key words: Fuzzy scatter matrix; Fuzzy Fisher Criterion(FFC); Optimal projection vector; FCS clustering

1　引言

聚类分析是多元统计分析的方法之一，也是无监督模式识别的一个重要分支。聚类方法将相似的数据点分为一类，而将不相似的数据点分为不同的类。在众多的聚类算法中，FCM 聚类算法[1]是最为著名的聚类算法之一，由于它在一定程度上克服了传统硬 C-均值聚类算法对初值敏感，抗噪性差等缺点，得到了学者们的广泛关注[2−4]。然而 FCM 是一种基于紧性致度量的算法，它只考虑了聚类的类内紧致性而没有考虑类间分离性[5]。为了进一步提高 FCM 的性能，Wu 等人提出了模糊散布矩阵的概念[5]，在此基础上通过修改聚类有效性指标 FS(c)[6] 提出了模糊紧性和分离性(FCS)聚类算法。FCS 将 FCM 作为特例，是 FCM 的广义形式。由于 FCS 同时考虑了类内紧致性和类间分离性，聚类性能得到了进一步的提高。然而 FCS 的缺点是需手动选取参数，自适应性差。

Fisher 判别分析(FDA)是监督模式识别的一种重要的特征提取和数据降维方法，FDA 经过寻找最具有判别性的投影方向，使得数据经投影处理后，不仅维数得到降低，而且获得了最大类内紧致性和类间分离性。在此基础上的分类也就变得更加高效和容易。但是 FDA 是一种监督方法，需要知道已有数据的类标号。2002 年，Clausi 提出了 KIF(K-means Iterative Fisher)方法[7]，将 FDA 与硬 C-均值聚类算法结合。该算法通过 K-means 初始化与迭代 FDA 组合实现。但该方法的抗噪性差。鉴于此，曹苏群等人在模糊散布矩阵概念的基础上，提出了模糊 Fisher 准则(FFC)[8]，并以 FFC 为目标函数，提出了一种新的聚类算法：FFC-SFCA。该算法通过

2008-11-25 收到，2009-06-11 改回
国家自然科学基金(60572133)资助课题

对 FFC 迭代优化，不仅能够实现聚类而且同时得到最优投影矢量。

本文指出曹苏群等人提出的 FFC-SFCA 中聚类中心的表达式是错误的，这种错误会导致算法不收敛。在此基础上我们给出新的基于 FFC 的模糊聚类算法： FFC-FCS。FFC-FCS 算法由 FFC 嵌套低维空间中的 FCS 构成。FFC-FCS 充分利用 FFC 的特征提取和降维特性，交替运行原始数据空间中 FFC 和投影空间中的 FCS，通过对降维数据的聚类实现对原始数据的聚类。与 FCS 不同，FFC-FCS 中 FCS 的参数就是最优投影矢量对应的特征值，在算法中是不断迭代寻优的，因此 FFC-FCS 可以看作是参数自适应选取的 FCS 算法。因为 FFC-FCS 中的 FCS 是在原始数据所在空间经 FFC 降维后得到的投影空间中进行的，所以 FFC-FCS 不仅可以获得优异的分类性能并且可以实现对高维数据的聚类。实验表明，FFC-FCS 总体性能明显优于原有的 FCS 算法，FFC-SFCA 算法以及经典的 FCM 算法。

2 FCS 和 FFC-SFCA 算法简介

2.1 FCS

鉴于 FCM 仅考虑了类内紧致性，而没有考虑类间分离性，为了能够同时描述聚类的类内紧致性和类间分离性，Wu 等人提出了模糊总散布矩阵 $\boldsymbol{S}_{\mathrm{FT}}$，模糊类内散布矩阵 $\boldsymbol{S}_{\mathrm{FW}}$ 和模糊类间散布矩阵 $\boldsymbol{S}_{\mathrm{FB}}$ 的概念[5]。它们分别定义为

$$\boldsymbol{S}_{\mathrm{FT}}=\sum_{j=1}^{c}\sum_{i=1}^{n}u_{ij}^{m}(\boldsymbol{x}_i-\overline{\boldsymbol{x}})(\boldsymbol{x}_i-\overline{\boldsymbol{x}})^{\mathrm{T}} \tag{1}$$

$$\boldsymbol{S}_{\mathrm{FW}}=\sum_{j=1}^{c}\sum_{i=1}^{n}u_{ij}^{m}(\boldsymbol{x}_i-\boldsymbol{v}_j)(\boldsymbol{x}_i-\boldsymbol{v}_j)^{\mathrm{T}} \tag{2}$$

$$\boldsymbol{S}_{\mathrm{FB}}=\sum_{j=1}^{c}\sum_{i=1}^{n}u_{ij}^{m}(\boldsymbol{v}_j-\overline{\boldsymbol{x}})(\boldsymbol{v}_j-\overline{\boldsymbol{x}})^{\mathrm{T}} \tag{3}$$

其中 $\boldsymbol{v}_j=\sum_{i=1}^{n}u_{ij}^{m}\boldsymbol{x}_i\Big/\left(\sum_{i=1}^{n}u_{ij}^{m}\right)$ 称为模糊样本均值，$u_{ij}\in[0,1]$，$\sum_{j=1}^{c}u_{ij}=1$，$m>1$；并且 $\boldsymbol{S}_{\mathrm{FT}}$，$\boldsymbol{S}_{\mathrm{FW}}$ 和 $\boldsymbol{S}_{\mathrm{FB}}$ 满足 $\boldsymbol{S}_{\mathrm{FT}}=\boldsymbol{S}_{\mathrm{FW}}+\boldsymbol{S}_{\mathrm{FB}}$。通过修改聚类有效性指标 $\mathrm{FS}(c)=\mathrm{tr}(\boldsymbol{S}_{\mathrm{FW}})-\mathrm{tr}(\boldsymbol{S}_{\mathrm{FB}})=\sum_{j=1}^{c}\sum_{i=1}^{n}u_{ij}^{m}\left\|\boldsymbol{x}_i-\boldsymbol{v}_j\right\|^2-\sum_{j=1}^{c}\sum_{i=1}^{n}u_{ij}^{m}\left\|\boldsymbol{v}_j-\overline{\boldsymbol{x}}\right\|^2$ [6]，Wu 等人提出了 FCS 聚类算法。FCS 的目标函数如下：

$$J_{\mathrm{FCS}}=\sum_{j=1}^{c}\sum_{i=1}^{n}u_{ij}^{m}\left\|\boldsymbol{x}_i-\boldsymbol{v}_j\right\|^2-\sum_{j=1}^{c}\sum_{i=1}^{n}\eta_j u_{ij}^{m}\left\|\boldsymbol{v}_j-\overline{\boldsymbol{x}}\right\|^2 \tag{4}$$

其中 $u_{ij}\in[0,1],\sum_{j=1}^{c}u_{ij}=1$，$m>1$，$\eta_j\geq 0$。当 $\eta_j=0$ 时，$J_{\mathrm{FCS}}=J_{\mathrm{FCM}}$，当 $\eta_j=1$ 时，$J_{\mathrm{FCS}}=\mathrm{FS}(c)$。式(4)的前半部分 $\sum_{j=1}^{c}\sum_{i=1}^{n}u_{ij}^{m}\left\|\boldsymbol{x}_i-\boldsymbol{v}_j\right\|^2=\mathrm{tr}(\boldsymbol{S}_{\mathrm{FW}})$ 衡量了类内紧致性，后半部分 $\sum_{j=1}^{c}\sum_{i=1}^{n}\eta_j u_{ij}^{m}\left\|\boldsymbol{v}_j-\overline{\boldsymbol{x}}\right\|^2$ 衡量了类间分离性。定义 Lagrange 函数。

$$L_{\mathrm{FCS}}=\sum_{j=1}^{c}\sum_{i=1}^{n}u_{ij}^{m}\left\|\boldsymbol{x}_i-\boldsymbol{v}_j\right\|^2-\sum_{j=1}^{c}\sum_{i=1}^{n}\eta_j u_{ij}^{m}\left\|\boldsymbol{v}_j-\overline{\boldsymbol{x}}\right\|^2+\sum_{i=1}^{n}\alpha_i\left(\sum_{j=1}^{c}u_{ij}-1\right) \tag{5}$$

用 L_{FCS} 对 $\boldsymbol{v}_j$ 和 u_{ij} 分别求偏导数，并令偏导数为零，得到一对方程：

$$\boldsymbol{v}_j=\frac{\sum_{i=1}^{n}u_{ij}^{m}\boldsymbol{x}_i-\sum_{i=1}^{n}\eta_j u_{ij}^{m}\overline{\boldsymbol{x}}}{\sum_{i=1}^{n}u_{ij}^{m}-\sum_{i=1}^{n}\eta_j u_{ij}^{m}} \tag{6}$$

和

$$u_{ij}=\frac{\left(\left\|\boldsymbol{x}_i-\boldsymbol{v}_j\right\|^2-\eta_j\left\|\boldsymbol{v}_j-\overline{\boldsymbol{x}}\right\|^2\right)^{-\frac{1}{m-1}}}{\sum_{k=1}^{c}\left(\left\|\boldsymbol{x}_i-\boldsymbol{v}_k\right\|^2-\eta_j\left\|\boldsymbol{v}_k-\overline{\boldsymbol{x}}\right\|^2\right)^{-\frac{1}{m-1}}} \tag{7}$$

鉴于由式(7)确定的 u_{ij} 可能出现负值，为了使 u_{ij} 的值在[0,1]区间内，文献[5]对 u_{ij} 做了如下的修正：

$$\text{如果}\left\|\boldsymbol{x}_i-\boldsymbol{v}_j\right\|^2\leq\eta_j\left\|\boldsymbol{v}_j-\overline{\boldsymbol{x}}\right\|^2\text{，则 }u_{ij}=1\text{，且对其它的 }j'\neq j\text{，}u_{ij'}=0 \tag{8}$$

另外，Wu 等人给出了用下面的式子确定参数 η_j 的方法：

$$\eta_j=\frac{(\beta/4)\min\limits_{j'\neq j}\left\|\boldsymbol{v}_j-\boldsymbol{v}_{j'}\right\|^2}{\max\limits_{k}\left\|\boldsymbol{v}_k-\overline{\boldsymbol{x}}\right\|},\quad 0\leq\beta\leq 1 \tag{9}$$

FCS 通过交替优化式(6)和式(7) 最小化 L_{FCS}，从而最小化 J_{FCS}，具体算法请参见文献[5]。FCS 算法的缺陷是算法性能依赖于参数 β 的选取，而 β 需要手动选取。

2.2 FFC- SFCA

曹苏群等人在 Wu 等人提出的模糊散布矩阵概念的基础上，提出了模糊 Fisher 准则[8]。记 $\boldsymbol{y}=\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{x}$，在该投影空间，各类样本均值向量记为 $\tilde{\boldsymbol{v}}_j$，有 $\tilde{\boldsymbol{v}}_j=\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{v}_j$。模糊类内散布矩阵 $\widetilde{\boldsymbol{S}}_{\mathrm{FW}}$ 定义为

$$\widetilde{\boldsymbol{S}}_{\mathrm{FW}}=\sum_{j=1}^{c}\sum_{i=1}^{n}u_{ij}^{m}(\boldsymbol{y}_i-\tilde{\boldsymbol{v}}_j)(\boldsymbol{y}_i-\tilde{\boldsymbol{v}}_j)^{\mathrm{T}}=\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{S}_{\mathrm{FW}}\boldsymbol{\omega} \tag{10}$$

模糊类间散布矩阵 $\widetilde{\boldsymbol{S}}_{\mathrm{FB}}$ 定义为

$$\widetilde{\boldsymbol{S}}_{\mathrm{FB}}=\sum_{j=1}^{c}\sum_{i=1}^{n}u_{ij}^{m}(\tilde{\boldsymbol{v}}_j-\overline{\boldsymbol{y}})(\tilde{\boldsymbol{v}}_j-\overline{\boldsymbol{y}})^{\mathrm{T}}=\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{S}_{\mathrm{FB}}\boldsymbol{\omega} \tag{11}$$

定义模糊 Fisher 准则(FCC)函数：

$$J_{\mathrm{FFC}}=\frac{\widetilde{\boldsymbol{S}}_{\mathrm{FB}}}{\widetilde{\boldsymbol{S}}_{\mathrm{FW}}}=\frac{\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{S}_{\mathrm{FB}}\boldsymbol{\omega}}{\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{S}_{\mathrm{FW}}\boldsymbol{\omega}} \tag{12}$$

据此，曹苏群等人提出了基于 FFC 的半模糊聚类算法：FFC-SFCA。FFC-SFCA 以 FFC 为目标函数，即 $J_{\mathrm{FFC}}=\dfrac{\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{S}_{\mathrm{FB}}\boldsymbol{\omega}}{\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{S}_{\mathrm{FW}}\boldsymbol{\omega}}$。定义 Lagrange 函数为

$$L_{\mathrm{FFC}}=\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{S}_{\mathrm{FB}}\boldsymbol{\omega}-\lambda\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{S}_{\mathrm{FW}}\boldsymbol{\omega}+\sum_{i=1}^{n}\lambda_i\left(\sum_{j=1}^{c}u_{ij}-1\right) \tag{13}$$

将 L_{FFC} 分别对 $\boldsymbol{\omega},\boldsymbol{v}_j$ 和 u_{ij} 求偏导数，并令其为零，得到 L_{FFC} 取极大值必须满足的方程为

$$\boldsymbol{S}_{\mathrm{FW}}^{-1}\boldsymbol{S}_{\mathrm{FB}}\boldsymbol{\omega}=\lambda\boldsymbol{\omega} \tag{14}$$

$$\boldsymbol{v}_j=\frac{\sum_{i=1}^{n}u_{ij}^{m}(\boldsymbol{x}_i-(1/\lambda)\overline{\boldsymbol{x}})}{\sum_{i=1}^{n}u_{ij}^{m}(1-1/\lambda)} \tag{15}$$

$$\begin{aligned}u_{ij}=&(\boldsymbol{\omega}^{\mathrm{T}}(\boldsymbol{x}_i-\boldsymbol{v}_j)(\boldsymbol{x}_i-\boldsymbol{v}_j)^{\mathrm{T}}\boldsymbol{\omega}-(1/\lambda)\boldsymbol{\omega}^{\mathrm{T}}(\boldsymbol{v}_j-\overline{\boldsymbol{x}})\\&\cdot(\boldsymbol{v}_j-\overline{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{\omega})^{-\frac{1}{m-1}}\Big/\Bigg[\sum_{k=1}^{c}(\boldsymbol{\omega}^{\mathrm{T}}(\boldsymbol{x}_i-\boldsymbol{v}_k)(\boldsymbol{x}_i-\boldsymbol{v}_k)^{\mathrm{T}}\\&\cdot\boldsymbol{\omega}-(1/\lambda)\boldsymbol{\omega}^{\mathrm{T}}(\boldsymbol{v}_k-\overline{\boldsymbol{x}})(\boldsymbol{v}_k-\overline{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{\omega})^{-\frac{1}{m-1}}\Bigg]\end{aligned} \tag{16}$$

由式(16)确定的 u_{ij} 可能出现负值，为了使 u_{ij} 的值在[0,1]区间内，文献[8]对 u_{ij} 做了如下的修正：如果 $\boldsymbol{\omega}^{\mathrm{T}}(\boldsymbol{x}_i-\boldsymbol{v}_j)(\boldsymbol{x}_i-\boldsymbol{v}_j)^{\mathrm{T}}\boldsymbol{\omega}<(1/\lambda)\boldsymbol{\omega}^{\mathrm{T}}(\boldsymbol{v}_j-\overline{\boldsymbol{x}})(\boldsymbol{v}_j-\overline{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{\omega}$，则

$$u_{ij}=1\text{，且对其它的 }j'\neq j,u_{ij'}=0 \tag{17}$$

FFC-SFCA 以硬 C-均值聚类算法作为初始化，通过交替优化式(14)，式(15)和式(16)最小化 L_{FFC}，从而最小化 J_{FFC}，具体算法参见文献[8]。

需要强调指出的是用式(13) 对 $\boldsymbol{v}_j$ 求偏导数不能导出如式(15)中各类中心 $\boldsymbol{v}_j$ 的表达式。事实上，

$$\begin{aligned}L_{\mathrm{FFC}}=&\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{S}_{\mathrm{FB}}\boldsymbol{\omega}-\lambda\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{S}_{\mathrm{FW}}\boldsymbol{\omega}+\sum_{i=1}^{n}\lambda_i\left(\sum_{j=1}^{c}u_{ij}-1\right)\\=&\boldsymbol{\omega}^{\mathrm{T}}\left(\sum_{j=1}^{c}\sum_{i=1}^{n}u_{ij}^{m}(\boldsymbol{v}_j-\overline{\boldsymbol{x}})(\boldsymbol{v}_j-\overline{\boldsymbol{x}})^{\mathrm{T}}\right)\boldsymbol{\omega}\\&-\lambda\boldsymbol{\omega}^{\mathrm{T}}\left(\sum_{j=1}^{c}\sum_{i=1}^{n}u_{ij}^{m}(\boldsymbol{x}_i-\boldsymbol{v}_j)(\boldsymbol{x}_i-\boldsymbol{v}_j)^{\mathrm{T}}\right)\boldsymbol{\omega}\\&+\sum_{i=1}^{n}\lambda_i\left(\sum_{j=1}^{c}u_{ij}-1\right)\end{aligned}$$

用 L_{FFC} 对 $\boldsymbol{v}_j$ 求偏导数，并令其等于零可得如下方程：

$$2\boldsymbol{\omega}^{\mathrm{T}}\sum_{i=1}^{n}u_{ij}^{m}(\boldsymbol{v}_j-\overline{\boldsymbol{x}})+2\lambda\boldsymbol{\omega}^{\mathrm{T}}\sum_{i=1}^{n}u_{ij}^{m}(\boldsymbol{x}_i-\boldsymbol{v}_j)=0$$

整理得

$$\boldsymbol{\omega}^{\mathrm{T}}\sum_{i=1}^{n}u_{ij}^{m}(\boldsymbol{v}_j-\overline{\boldsymbol{x}}+\lambda(\boldsymbol{x}_i-\boldsymbol{v}_j))=0 \tag{18}$$

我们无法从上式反解出 $\boldsymbol{v}_j$。观察式(18)可知，文献[8]的错误在于直接在上式中将 $\boldsymbol{\omega}^{\mathrm{T}}$ 约去，得方程 $\sum_{i=1}^{n}u_{ij}^{m}(\boldsymbol{v}_j-\overline{\boldsymbol{x}}+\lambda(\boldsymbol{x}_i-\boldsymbol{v}_j))=0$，进而解出关于 $\boldsymbol{v}_j$ 的表达式式(15)。这种错误导致 FFC-SFCA 在迭代求解过程中可能是不收敛的(在实验过程中我们发现常常会出现这样的现象：FFC-SFCA 在迭代过程中总是迭代到设置的最大迭代次数时才停止迭代，见第 4 节的表 1 和表 3)。

实验中，FFC-SFCA 表现出一定的分类能力，其原因在于当 $\sum_{i=1}^{n}u_{ij}^{m}(\boldsymbol{v}_j-\overline{\boldsymbol{x}}+\lambda(\boldsymbol{x}_i-\boldsymbol{v}_j))=0$ 时，$\boldsymbol{\omega}^{\mathrm{T}}\sum_{i=1}^{n}u_{ij}^{m}(\boldsymbol{v}_j-\overline{\boldsymbol{x}}+\lambda(\boldsymbol{x}_i-\boldsymbol{v}_j))$ 也必然为零。即 FFC-SFCA 在一个人为缩小的区域内搜寻最优解，这样可能只会得到一个近似最优解，而不是最优解。

3　FCM-FCS 算法

鉴于文献[8]的错误，本文重新考虑基于模糊 Fisher 准则的聚类算法。由 $\widetilde{\boldsymbol{S}}_{\mathrm{FW}}$ 和 $\widetilde{\boldsymbol{S}}_{\mathrm{FB}}$ 的定义和式(13)可知

$$\begin{aligned}L_{\mathrm{FFC}}=&\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{S}_{\mathrm{FB}}\boldsymbol{\omega}-\lambda\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{S}_{\mathrm{FW}}\boldsymbol{\omega}+\sum_{i=1}^{n}\lambda_i\left(\sum_{j=1}^{c}u_{ij}-1\right)\\=&\widetilde{\boldsymbol{S}}_{\mathrm{FB}}-\lambda\widetilde{\boldsymbol{S}}_{\mathrm{FW}}+\sum_{i=1}^{n}\lambda_i\left(\sum_{j=1}^{c}u_{ij}-1\right)\\=&\sum_{j=1}^{c}\sum_{i=1}^{n}u_{ij}^{m}(\tilde{\boldsymbol{v}}_j-\overline{\boldsymbol{y}})(\tilde{\boldsymbol{v}}_j-\overline{\boldsymbol{y}})^{\mathrm{T}}\\&-\lambda\sum_{j=1}^{c}\sum_{i=1}^{n}u_{ij}^{m}(\boldsymbol{y}_i-\tilde{\boldsymbol{v}}_j)(\boldsymbol{y}_i-\tilde{\boldsymbol{v}}_j)^{\mathrm{T}}\\&+\sum_{i=1}^{n}\lambda_i\left(\sum_{j=1}^{c}u_{ij}-1\right)=-\lambda\Bigg(\sum_{j=1}^{c}\sum_{i=1}^{n}u_{ij}^{m}\left\|\boldsymbol{y}_i-\tilde{\boldsymbol{v}}_j\right\|^2\\&-\sum_{j=1}^{c}\sum_{i=1}^{n}\frac{1}{\lambda}u_{ij}^{m}\left\|\tilde{\boldsymbol{v}}_j-\overline{\boldsymbol{y}}\right\|^2+\sum_{i=1}^{n}\left(-\frac{\lambda_i}{\lambda}\right)\left(\sum_{j=1}^{c}u_{ij}-1\right)\Bigg)\end{aligned}$$

其中 $\tilde{\boldsymbol{v}}_j=\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{v}_j$，$\overline{\boldsymbol{y}}=\boldsymbol{\omega}^{\mathrm{T}}\overline{\boldsymbol{x}}$，$\boldsymbol{y}_i=\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{x}_i$。

令 $\lambda'=1/\lambda$，$\lambda_i^{'}=-\lambda_i/\lambda$，则有

$$\begin{aligned}L_{\mathrm{FFC}}=&-\lambda\Bigg(\sum_{j=1}^{c}\sum_{i=1}^{n}u_{ij}^{m}\left\|\boldsymbol{y}_i-\tilde{\boldsymbol{v}}_j\right\|^2-\sum_{j=1}^{c}\sum_{i=1}^{n}\lambda' u_{ij}^{m}\left\|\tilde{\boldsymbol{v}}_j-\overline{\boldsymbol{y}}\right\|^2\\&+\sum_{i=1}^{n}\lambda_i'\left(\sum_{j=1}^{c}u_{ij}-1\right)\Bigg)\end{aligned}$$

令

$$\begin{aligned}L_{\mathrm{FFC}}^{'}=&\sum_{j=1}^{c}\sum_{i=1}^{n}u_{ij}^{m}\left\|\boldsymbol{y}_i-\tilde{\boldsymbol{v}}_j\right\|^2-\sum_{j=1}^{c}\sum_{i=1}^{n}\lambda' u_{ij}^{m}\left\|\tilde{\boldsymbol{v}}_j-\overline{\boldsymbol{y}}\right\|^2\\&+\sum_{i=1}^{n}\lambda_i'\left(\sum_{j=1}^{c}u_{ij}-1\right)\end{aligned} \tag{19}$$

则当 $\lambda>0$ 时，最大化 L_{FFC} 等价于最小化 L'_{FFC}。用 L'_{FFC} 分别对 $\tilde{\boldsymbol{v}}_j$ 和 u_{ij} 求偏导数，并令偏导数为零，可得如下的一对方程：

$$\tilde{\boldsymbol{v}}_j=\frac{\sum_{i=1}^{n}u_{ij}^{m}\boldsymbol{y}_i-\sum_{i=1}^{n}\lambda' u_{ij}^{m}\overline{\boldsymbol{y}}}{\sum_{i=1}^{n}u_{ij}^{m}-\sum_{i=1}^{n}\lambda' u_{ij}^{m}} \tag{20}$$

$$u_{ij}=\frac{\left(\left\|\boldsymbol{y}_i-\tilde{\boldsymbol{v}}_j\right\|^2-\lambda'\left\|\tilde{\boldsymbol{v}}_j-\overline{\boldsymbol{y}}\right\|^2\right)^{-\frac{1}{m-1}}}{\sum_{k=1}^{c}\left(\left\|\boldsymbol{y}_i-\tilde{\boldsymbol{v}}_k\right\|^2-\lambda'\left\|\tilde{\boldsymbol{v}}_k-\overline{\boldsymbol{y}}\right\|^2\right)^{-\frac{1}{m-1}}} \tag{21}$$

则当 $\lambda>0$ 且值固定时(此时 λ' 自然也为定值)，最小化 L'_{FFC} 可以由交替迭代式(20)和式(21)实现。式(19)，式(20)和式(21)与式(5)，式(6)和式(7)的对比可知：L'_{FFC} 与 FCS 的目标函数 L_{FCS} 是一致的，只是 FCS 算法的参数 η_j 被这里的 λ' 取代。

综上分析有如下结论：当最优投影矢量 $\boldsymbol{\omega}$ 及其对应的特征值 λ (对应于 $\boldsymbol{S}_{\mathrm{FW}}^{-1}\boldsymbol{S}_{\mathrm{FB}}$ 的最大特征值，是大于零的)固定时，交替迭代式(20)和式(21)最小化 L'_{FFC} 就是在投影空间中进行的以 L'_{FFC} 为目标函数的 FCS 算法。又因为最大化 L_{FFC} 等价于最小化 L'_{FFC}，所以当最优投影矢量 $\boldsymbol{\omega}$ 及其对应的特征值 λ 固定时，交替迭代式(20)和式(21)可实现在投影空间中进行的以 L_{FFC} 为目标函数的 FCS 算法。另外，既然 FFC 可以求得最优投影矢量，可以设想在投影空间中进行聚类。这样既容易分类，又由于数据得到降维，使得聚类算法的效率更高。因此，本文考虑利用 FFC 来求得最优投影矢量，在投影空间进行 FCS 的聚类算法。

因为要利用 FFC 来求得最优投影矢量，所以要计算模糊类内散布矩阵 $\boldsymbol{S}_{\mathrm{FW}}$ 和模糊类间散布矩阵 $\boldsymbol{S}_{\mathrm{FB}}$，由模糊类内散布矩阵 $\boldsymbol{S}_{\mathrm{FW}}$ 和模糊类间散布矩阵 $\boldsymbol{S}_{\mathrm{FB}}$ 的定义可知，要计算模糊类内散布矩阵 $\boldsymbol{S}_{\mathrm{FW}}$ 和模糊类间散布矩阵 $\boldsymbol{S}_{\mathrm{FB}}$ 需要知道各类的模糊样本均值 $\boldsymbol{v}_j$ 和隶属度 u_{ij}，事实上只要知道隶属度 u_{ij}，我们可以由 $\boldsymbol{v}_j=\sum_{i=1}^{n}u_{ij}^{m}\boldsymbol{x}_i\Big/\left(\sum_{i=1}^{n}u_{ij}^{m}\right)$ 求得各类的模糊样本均值 $\boldsymbol{v}_j$。而隶属度 u_{ij} 可以由投影空间中的 FCS 聚类算法获得，这样就形成一个迭代优化的过程。这种基于 FCC 的求得最优投影矢量和在投影空间中的 FCS 的交替聚类算法，记为 FFC-FCS。FFC-FCS 可以看作是参数自适应选取的 FCS 算法，参数由最优投影矢量 $\boldsymbol{\omega}$ 对应的特征值 λ 的倒数给出。FFC-FCS 算法在运行过程中的初始隶属度 u_{ij} 由 FCM 给出，在迭代过程中 u_{ij} 由投影空间中 FCS 计算出的结果传递出来。整个算法流程如图 1 所示。

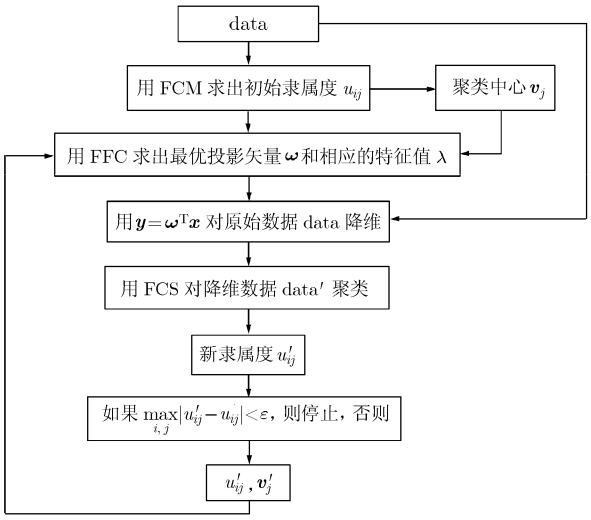

图 1 FFC-FCS 算法流程图

算法步骤：

(1)运行 FCM 算法给出初始隶属度 u_{ij} 和各类模糊样本均值 $\boldsymbol{v}_j$ (即聚类中心)；

(2)利用(1)求得的隶属度 u_{ij} 和各类中心 $\boldsymbol{v}_j$ 和 FFC 求出矩阵 $\boldsymbol{S}_{\mathrm{FW}}^{-1}\boldsymbol{S}_{\mathrm{FB}}$ 的最大特征值 λ，并取 $\boldsymbol{\omega}$ 为矩阵 $\boldsymbol{S}_{\mathrm{FW}}^{-1}\boldsymbol{S}_{\mathrm{FB}}$ 属于 λ 的模为 1 的特征向量(即最优投影矢量)，并用 $\boldsymbol{y}=\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{x}$ 对数据降维；

(3)在降维的投影空间中执行以 $1/\lambda$ 为参数的 FCS 并更新隶属度 u'_{ij}；

(4)如果 $\max_{i,j}\left|u'_{ij}-u_{ij}\right|<\varepsilon$，停止；否则，更新各类中心 $\boldsymbol{v}'_j$ 并返回到(2)。

4 实验结果及分析

为了验证本文提出的 FFC-FCS 算法的有效性，本节用 FCM，FCS，FFC-SFCA 和 FFC-FCS 4 个算法分别对 1 个人造数据和两个标准数据 Iris 和 Glass[9]进行仿真实验。在本文的实验中，为避免算法陷入局部最优，算法在随机初始化后运行 100 次取最优分类结果，并取相应的迭代次数和 CPU 运行时间。最大迭代次数设为 200，停止阈值设为 10^{-5}。4 个算法中的参数 m 都取为 2，FCS 算法中的参数 η_j 取为 0.5。

实验 1 人造数据

图 2 所示的长条型数据来自文献[10]中的数据“fourlines”，选其中长度接近的两类数据，为了增大聚类的难度，将两类数据间的距离变小，新数据记为“twolines”。4 个算法对“twolines”的分类结果如图 2 和图 3 所示。其中，“•”和“+”分别表示两类数据，“∘”表示聚类中心。图 2 中的 3 幅图 2(a)，2(b)和 2(c)分别是 FCM，FCS 和 FFC-SFCA 对“twolines”的分类结果。图 3 中的 3 幅图 3(a)，

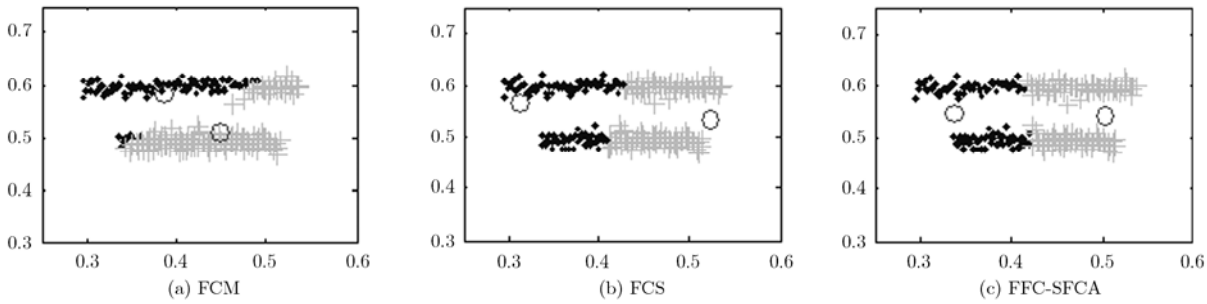

图 2 FCM, FCS 和 FFC-SFCA 对“twolines”的分类结果

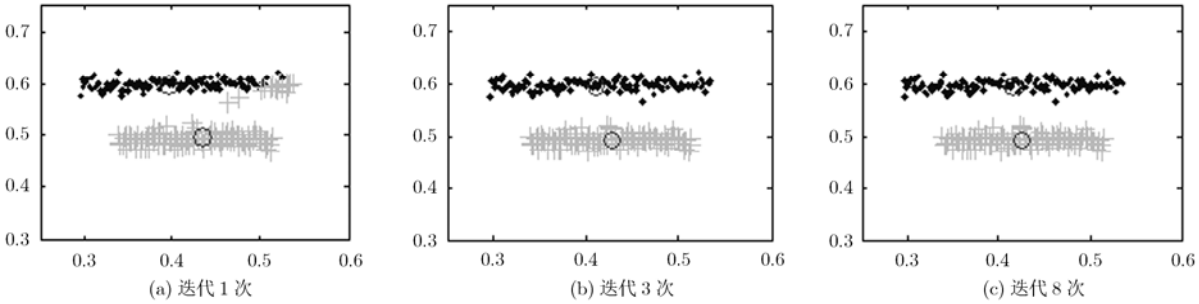

图 3 FFC-FCS 不同迭代次数时对“twolines”的分类结果

3(b)和 3(c)分别是 FFC-FCS 对“twolines”聚类迭代过程中迭代次数为 1 次，3 次和 8 次时分类结果。由图 2 可知，FCM，FCS 和 FFC-SFCA 都不能对“twolines”正确分类。由图 3 可知，本文提出的 FFC-FCS 算法可以在迭代过程中自适应地优化投影方向，最终找到最优投影方向并对“twolines”正确分类。由表 1 可知，FFC-FCS 迭代次数和运行时间比 FCM 和 FCS 略多一些。但少于 FFC-SFCA。FFC-SFCA 的迭代次数达到最大迭代次数，导致用时也较多。

表 1 4 个算法对“twolines”分类时的迭代次数和运行时间

	迭代次数	运行时间(s)
FCM	67	0.234000
FCS	5	0.187000
FFC-SFCA	200	8.875000
FCC-FCS	8	0.647000

实验 2　标准数据

本节选取的两个标准数据 Iris 和 Glass 均来源于 UCI repository of machine learning databases[9]，数据的特性如表 2 所示。本节用这两个数据对 4 个聚类算法的分类性能进行比较，用错分数，迭代次数和运行时间 3 个指标综合衡量聚类算法的优劣。结果如表 3 和表 4 所示。其中“×”表示该算法不能运行。

表 2 数据描述

	样本数	维数	类数
Iris	150	4	3
Glass	214	9	7

表 3 4 个算法对 IRIS 数据分类结果

	错分数	迭代次数	运行时间(s)
FCM	16	24	0.047000
FCS	16	28	0.188000
FFC-SFCA	5	200	10.531000
FCC-FCS	2	14	0.547000

表 4 4 个算法对 Glass 数据分类结果

	错分数	迭代次数	运行时间(s)
FCM	109	200	0.594000
FCS	109	200	7.500000
FFC-SFCA	×	×	×
FCC-FCS	16	17	9.438000

由表 3 可知，FFC-FCS 对 IRIS 的分类精度最高，要明显优于 FCM 和 FCS。FFC-SFCA 也达到很高的分类精度，但是 FFC-SFCA 的迭代次数达到最大迭代次数，导致用时也较多。

由表 4 可知，FFC-FCS 对 Glass 的分类精度最

高，要明显优于 FCM 和 FCS，FFC-FCS 的用时相对多一些。FFC-SFCA 对 Glass 分类时算法不能运行。

综上分析，本文提出的 FFC-FCS 虽然用时稍微多一些，但是在分类精度上表现出优异的性能，FFC-FCS 算法总体上要优于 FCM，FCS 和 FFC-SFCA。

5 结论

本文首先指出曹苏群等人提出的 FFC-SFCA 聚类算法中的一个推导错误，在分析该算法错误根源的基础上提出一个新的基于 FFC 的聚类算法：FFC-FCS。FFC-FCS 充分利用 FFC 的特征提取和降维优势，由原始数据空间的 FFC 和投影空间中的 FCS 聚类交替进行实现，可以在迭代过程中不断优化投影方向，同时完成最优投影矢量的寻找和对原始数据的聚类，FFC-FCS 可以看作是参数自适应选取的 FCS。实验表明，本文提出的 FFC-FCS 的算法性能在总体上要优于 FCM，FCS 和 FFC-SFCA。本文提出的 FFC-FCS 本质上是一种线性聚类算法，如何通过“核”技巧构造核版本的 FFC-FCS 聚类算法是本文下一步的工作。

参 考 文 献

[1] Bezdek J C. Pattern Recognition with Fuzzy Objective Function Algorithms[M]. New York, Plenum Press, 1981: 95–107.

[2] Yang Miin-Shen and Wu Kuo-Lung, *et al.*. Alpha-cut implemented fuzzy clustering algorithms and switching regressions[J]. *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, 2008, 38(3): 588–603.

[3] Cai Wei-ling, Chen Song-can, and Zhang Dao-qiang. Fast and robust fuzzy c-means clustering algorithms incorporating local information for image segmentation[J]. *Pattern Recognition*, 2007, 40(3): 825–838.

[4] Xing Hong-jie and Hu Bao-gang. An adaptive fuzzy c-means clustering-based mixtures of experts model for unlabeled data classification [J]. *Neurocomputing*, 2008, 71(4): 1008–1021.

[5] Wu Kuo-lung, Yu Jian, and Yang Miin-Shen. A novel fuzzy clustering algorithm based on a fuzzy scatter matrix with optimality test [J]. *Pattern Recognition Letters*, 2005, 26(5): 639–652.

[6] Yoshiki Fukuyama and Michio Sugeno. A new method of choosing the number of clusters for the fuzzy c-means method [C]. Proceedings of the 5th Fuzzy System Symposium, in Japanese, 1989: 247–250.

[7] Clausi David. K-means Iterative Fisher (KIF) unsupervised clustering algorithm applied image texture segmentation [J]. *Pattern Recognition*, 2002, 35(9): 1959–1972.

[8] 曹苏群, 王士同, 陈晓峰等. 基于模糊 Fisher 准则的半模糊聚类算法[J]. 电子与信息学报, 2008, 30(9): 2162–2165.
Cao Su-qun, Wang Shi-tong, and Chen Xiao-feng, *et al.*. Fuzzy fisher criterion based semi-fuzzy clustering algorithm[J]. *Journal of Electronics & Information Technology*, 2008, 30(9): 2162–2165.

[9] Blake C L and Merz C J. UCI repository of machine learning databases, Irvine. CA: University of California, Department of Information and Computer Science, http://www.ics.uci.edu/~mlearn /MLRepository.html, 1998, 7.

[10] 王玲, 薄列峰, 焦李成. 密度敏感的谱聚类[J]. 电子学报, 2007, 35(8): 1577–1581.
Wang Ling, Bo Lie-feng, and Jiao Li-cheng. Density-sensitive spectral clustering [J]. *Acta Electronica Sinica*, 2007, 35(8): 1577–1581.

支晓斌： 男，1976 年生，讲师，研究方向为模式识别、模糊集理论及其应用.

范九伦： 男，1964 年生，教授，博士生导师，研究方向为模糊集理论、模式识别与图像处理、智能信息处理.