

基于模糊聚类的羊群效应检测方法

蔡 炜¹, 徐圣兵¹, 周颖彤², 韦堦瀚¹

¹广东工业大学应用数学学院, 广东 广州

²广东工业大学管理学院, 广东 广州

收稿日期: 2021年9月16日; 录用日期: 2021年11月18日; 发布日期: 2021年11月25日

摘 要

为检测股票市场中有无出现羊群效应, 以及个股的羊群效应程度和股市不同时间段的羊群效应程度, 提出一种基于模糊聚类的羊群效应分析方法。通过计算股票每天的开盘收盘价得到收益率, 计算收益率一阶差分构建具有时间序列属性的收益变化特征。将每支股票作为单个样本, 采用模糊C均值聚类方法得到隶属度矩阵。建立类别判断准则, 计算第一类样本隶属度与第二类样本隶属度的最大/最小差值, 将出现最大/最小差值的样本与收益变化特征均值的方法来判断哪一类是羊群效应。相比计算横截面标准差等传统方法再进行线性回归得到回归系数, 该方法具有能够表现个股羊群效应优势; 相比k-means聚类计算羊群效应, 该方法用隶属度能够得到更多的聚类信息。

关键词

羊群效应, 模糊聚类, 最大 - 最小类别判断准则

Detection Method of Herding Effect Based on Fuzzy Clustering

Wei Cai¹, Shengbing Xu¹, Yingtong Zhou², Yuanhan Wei¹

¹School of Applied Mathematics, Guangdong University of Technology, Guangzhou Guangdong

²School of Management, Guangdong University of Technology, Guangzhou Guangdong

Received: Sep. 16th, 2021; accepted: Nov. 18th, 2021; published: Nov. 25th, 2021

Abstract

In order to detect whether there is herd behavior in the stock market, the degree of herd behavior of individual stocks and the degree of herd behavior in different periods of stock market, this paper proposes an analysis method of herding effect based on fuzzy clustering. By calculating the

daily opening and closing prices of the stock to get the yield, the first-order difference of the yield is calculated to construct the return change characteristics with time series attribute. Taking each stock as a single sample, the membership matrix is obtained by fuzzy c-means clustering. The maximum/minimum difference between the membership degree of the first kind of sample and the membership degree of the second type sample is calculated by establishing the category judgment criteria. The method of the sample with the maximum/minimum difference and the mean value of the income change characteristics are used to determine which kind of herding behavior is. Compared with traditional methods such as calculating cross-section standard deviation and linear regression to get the regression coefficient, this method can show individual stock herding effect; compared with k-means clustering, this method can get more clustering information by membership degree.

Keywords

Herding Effect, Fuzzy Clustering, Maximum Minimum Category Criterion

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

由于投资者缺乏或不了解信息,很难对市场未来的不确定性做出合理的预期,因此,他们往往是通过周围人的行为来获得信息的,这些行为大多与周围人的行为相似,并会相互强化,导致从众心理。经济领域中,投资者的“从众心理”通常被称为“羊群效应”。以往研究发现,金融危机的发生与羊群效应的不稳定性有密切关系,并从羊群效应的角度对系统性问题做出了新的解释,同时也发现了许多有效市场假设无法解释的现象。为了解释这些现象,人们对羊群效应理论产生了兴趣。

自上世纪 90 年代以来,行为金融的研究取得了长足的发展,许多学者提出了羊群效应的理论模型,包括理论模型和实证模型。羊群效应在理论上可分为信息瀑布模型和个人信用模型, Banerjee (1992) [1] 提出的信息瀑布模型是按行为顺序建立的, Bikhchanbani、Hirshleife 和 Welch [2] (1998)提出了投资者形成羊群新效应的条件,而 Avery (1998) [3]则指出,有关固定资产价格的假设与实际不符。Schafstein 和 Stein [4]的个人信贷风险模型(1990)和 Graham [5]的模型(1999)进一步表明,基金管理公司倾向于模仿 Value Line 的投资条件,而 Zwiebel [6] (1995)则通过对基金管理公司相互模仿行为的实证分析,发现除了极少数极好或极差的公司之外,所有基金管理公司在投资活动中都保守地遵循行业标准。除此之外,比较典型的实证模型有:具体投资机构行为研究的 LSV [7]模型、以股票价格分散度为指标的整体股价波动 CH 模型以及 CCK 模型等。Grinblatt, Titman 和 Wermers [8] (1995)在 LSV 模型的基础上做了简单的修正,提出了通过改变组合的数量来克服 LSV 模型的局限性。由 Christie, Huang (1995)提出的 CH 模型将 CSSD [9]值和市场波动的哑变量作为一种回归,该模型将股票收益的横截面标准差(CSSD)作为衡量股票总体羊群行为的指标。但是,以上模型,有些只是用买卖数量来衡量,不能反映不同时期的交易情况,有些只是衡量股票市场上较强的羊群行为,回归结果往往不显著。

自 2000 年以来,国内学者在借鉴国外一些早期文献的基础上,也开始了关于羊群效应的相关研究。宋军、吴冲锋[10]运用 CH 模型发现我国证券市场存在明显的羊群效应。施东晖(2004) [11] [12]认为单纯使用 CH 模型和 CCK 模型是不能区分羊群在市场中的真实价值。陈浩(2004)应用 LSV 方法检验发现羊群

行为显著存在于投资基金之间, 并且与股票流通股规模、股票历史收益率等属性有一定程度的相关性。针对 CH 模型在衡量市场中小幅度羊群行为时存在的缺陷, Chang, Cheng, Khorana (1998) 提出了 CCK [13] 模型。利用 CSAD 替代了 CSSD 的截面绝对偏差, 并在 CAPM 模型的框架下证明了 CSAD 与市场收益率 R 之间存在线性增长关系。CSAD 和 R 之间的线性相关性在市场上存在羊群效应时消失, 用简单的回归模型就可以检测出这种效应。对 CH 模型、CCK 模型和 CAPM 模型分别进行时序再回归计算, 得到的回归系数只能用于判断该时期股票的羊群效应是否明显 [8] [14], 不可能区分每一时期甚至每一股票是否有羊群效应。此外, 2010 年郭旭超 [15] 还提出了基于聚类的羊群效应分析方法, 即以时间为样本, 以股票的总收益为特征, 对 k-means 进行聚类。相对于传统的 CH 模型, 该方法可以在不同时间区域有效地观察羊群效应的发生情况, 但是与线性回归方法相比, 它丢失了关于羊群效应发生程度的重要信息 [16]。

本研究基于羊群效应的 k-means 聚类分析, 建立了基于模糊聚类的羊群效应分析方法。利用模糊 C-均值聚类法得到隶属度矩阵, 建立了分类判别准则。以 2020 年 11 月股票数据为例, 对基于 k-means 聚类和模糊 C-均值聚类(Fuzzy c-means clustering, FCM)方法的羊群效应进行了创新性的对比。

2. 相关工作

2.1. 检验羊群效应的模型与方法

2.1.1. CH 模型

认为市场中若存在显著的羊群行为, 那么大部分的投资者的投资买卖交易将有一定的趋同性, 因而个股的收益率应该接近股市整体的收益率。故 CH 模型中的个股收益率的横截面标准差(Cross-Sectional Standard Deviation, CSSD)为:

$$CSSD_t = \sqrt{\frac{\sum_{i=1}^n (R_{it} - R_{mt})^2}{n-1}} \quad (1)$$

其中 R_{it} 表示第 i 支股票在 t 时刻的收益率, R_{mt} 表示股市在 t 时刻的整体收益率。

$$R_{mt} = \frac{\sum_{i=1}^n R_{it}}{n} \quad (2)$$

$CSSD_t$ 越小, 说明在 t 时刻股市的羊群效应越明显。根据 D 值线性回归:

$$CSSD_t = \alpha + \beta_1 D_t^H + \beta_2 D_t^L \quad (3)$$

其中 D_t^H 和 D_t^L 分别表示市场在 t 时刻出现股价大幅上涨、下跌时的哑变量。股价大幅下跌, $CSSD_t$ 的值下降, 此时可认为市场存在羊群效应, 故在回归模型中, β_1 和 β_2 应显著为负。

2.1.2. CCK 模型

CCK 模型是与 CH 模型中的度量指标 CSSD 相似的收益率横截面绝对偏离差(Cross-Section Absolute Deviation, CSAD)

$$CSAD_t = \frac{1}{n} \sum_{i=1}^n |R_{it} - R_{mt}| \quad (4)$$

$CSAD_t$ 与市场收益率 R_{it} 线性相关。

2.1.3. CAPM 模型

CAPM 模型提出包含 β_{mt} 系数的指标代替 CSSD、CSAD 作为羊群行为度量指标, 剔除个股信息和市场整体重大信息等因素对羊群行为指标的影响, 有效地检测出市场中的真羊群行为。

$$\beta_m = \frac{1}{n} \sum_{i=1}^n \beta_i \tag{5}$$

$$E(CSAD_t) = \frac{1}{n} \sum_{i=1}^n |\beta_i - \beta_m| \cdot (E(R_{mt}) - R_f) \tag{6}$$

$$CSAD_t = \alpha + \beta_1 |R_{mt}| + \beta_2 R_{mt}^2 \tag{7}$$

2.2. 模糊 C 均值聚类算法

模糊 C 均值聚类算法(FCM)是一种常用的聚类算法[17]。目前, FCM 及其导数聚类方法已广泛应用于计算机视觉、模式识别、数据挖掘等领域。FCM 算法是 Bezdek 于 1973 年提出的一种改进硬 C 均值聚类的方法[17]。与 HCM 和 k-means [18] [19]等硬算法相比, 它最大的优点是用区间[0, 1]的隶属度来表示样本的聚类结果, 以表示样本趋于[19] [20] [21]的某种纯度。

设在 p 维的欧几里得空间 R^p 中的 n 个样本 $X = \{x_1, x_2, x_3, \dots, x_n\} \in R^p$, 其中 $j = 1, 2, \dots, n$; 则 FCM 的聚类准则目标函数为:

$$J(U, V) = \sum_{i=1}^c \sum_{j=1}^n (u_{ij})^m \|x_j - v_i\|^2 \tag{8}$$

其中 c 为预先给定的聚类中心数, 且 $1 < c < n$; m 为预先给定的加权幂指数, 用于影响隶属度矩阵的模糊度; v_i 为第 i 类的聚类中心; u_{ij} 表示样本 j 属于第 i 类的程度, 并且满足 $\sum_{i=1}^c u_{ij} = 1$ 。由目标函数可知需求得当 J 达到最小值时的 U 和 V , 因此由 J 作为目标变量对 u_{ij}, v_i 求偏导, 令它们的导数为零并代入条件 $\sum_{i=1}^c u_{ij} = 1$ 得到隶属度与聚类中心的迭代公式:

$$u_{ij} = \frac{\left(\frac{1}{\|x_j - v_k\|^2} \right)^{\frac{1}{m-1}}}{\sum_{i=1}^c \left(\frac{1}{\|x_j - v_i\|^2} \right)^{\frac{1}{m-1}}} \tag{9}$$

$$v_i = \frac{\sum_{j=1}^n (u_{ij})^m x_j}{\sum_{j=1}^n (u_{ij})^m} \tag{10}$$

通过对上述两个公式进行迭代求得当 J 收敛到极小值 u_{ij} 和 v_i 的值, 即隶属度矩阵 U 和聚类中心 V , 此为最终聚类结果。

3. 基于模糊聚类的羊群效应检测

3.1. 建模动机

CH 模型、CCK 模型等经典模型仅仅可用于判断市场是否存在羊群效应, 并且得到羊群效应的明显程度。而基于 k-means 聚类的方法虽然可以研究某个时间段是否出现羊群效应, 但是这种硬聚类的方法没有得到市场出现羊群效应明显程度这一信息。

综上, 本文提出基于模糊聚类的方法对市场股票或时间段进行建模分析有无出现羊群效应以及程度。

3.2. 构建特征

当股票价格下跌时, 羊群效应表现为刚开始会出现缓慢的下跌, 但是当投资者出现大量抛售, 导致股票价格暴跌, 如图 1 所示:

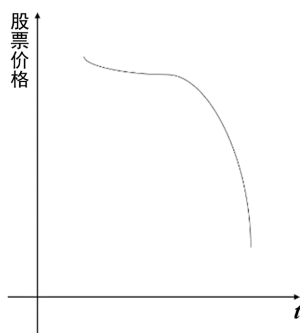


Figure 1. Falling herd effect
图 1. 跟跌羊群效应

因而想要分析股市的羊群效应，可以通过股市在一段时间序列内的信息来判断是否有羊群效应。根据股票构建收益率作为羊群效应检测特征。

$$R_{it} = \frac{P_{it} - P_{i\tau}}{P_{i\tau}} \tag{11}$$

其中 $\tau = t - 1$ ， R_{it} 表示第 i 支股票 t 时刻的收益率， P_{it} 表示第 i 支股票 t 时刻的收盘价。

对样本构建不同的特征，采用相同的机器学习方法会得到不同的结果，如图 2 中的是具有时间序列的三个样本。

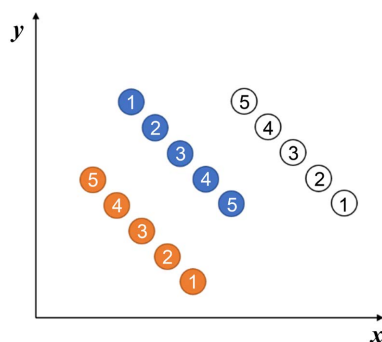


Figure 2. Samples to be clustered
图 2. 待聚类的样本

而以不同规则可以得到不同的聚类结果(相同颜色表示同一类)，如图 3 所示：

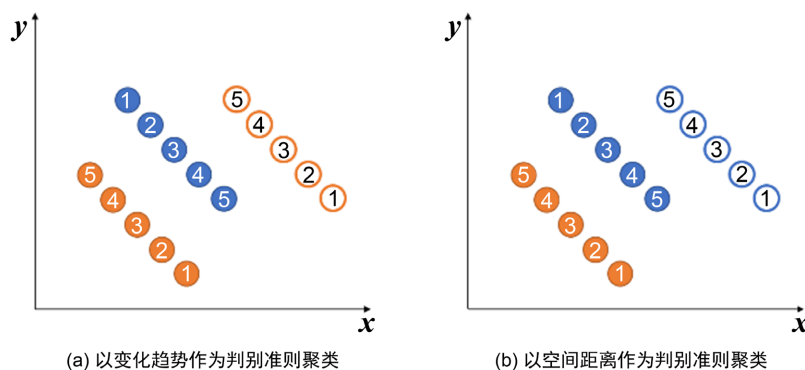


Figure 3. Different clustering results by different criteria
图 3. 不同准则得到不同聚类结果

而描述股市的羊群效应，其实是一个描述时间变化的过程，因而更像是图 3(a)的情况，因而用收益率的一阶差分来表示样本在某段的变化情况。

$$x_{ij} = R_{i,t+1} - R_{i,t} \tag{12}$$

其中 $t = \{1, 2, \dots, T-1\}$ ，故 $j = \{1, 2, \dots, t-1\}$ 。

3.3. 羊群效应检测

3.3.1. 股票羊群效应检测模型

令 $X = [x_{ij}]$ ，其中 $i = \{1, 2, \dots, n\}$ ，进行聚类得到隶属度矩阵 u_{ij} ，但得到隶属度矩阵之后是不知道哪一类是出现羊群效应，哪一类没有出现羊群效应的。因而通过计算股票收益率变化程度来用以类别判断。

建立最大 - 最小类别判断准则：

$$\theta_1 = \arg \max_j (u_{1j} - u_{2j}) \tag{13}$$

$$\theta_2 = \arg \min_j (u_{1j} - u_{2j}) \tag{14}$$

θ_1 和 θ_2 是羊群效应最明显样本或最不明显样本的编号。计算收益率变化程度：

$$\delta_1 = \sum_{j=1}^t (x_{\theta_1 j} - \overline{x_{\theta_1 j}})^2 \tag{15}$$

$$\delta_2 = \sum_{j=1}^t (x_{\theta_2 j} - \overline{x_{\theta_2 j}})^2 \tag{16}$$

若 $\delta_1 > \delta_2$ 则第一类为出现羊群效应，若 $\delta_1 < \delta_2$ 则为第二类出现羊群效应。

3.3.2. 时间段羊群效应检测模型

令 $X = [x_{ij}]^T$ ，进行聚类得到隶属度矩阵 u_{ij} ，此时仍然需要因而通过计算单个样本的 CAPM 值来用以类别判断。

仍然根据表达式(13)和(14)计算 θ_1 和 θ_2 ，进一步计算收益分散度：

$$S_1 = \frac{1}{n} \sum_{i=1}^n (\beta_{i\theta_1} - 1)^2 \tag{17}$$

$$S_2 = \frac{1}{n} \sum_{i=1}^n (\beta_{i\theta_2} - 1)^2 \tag{18}$$

因为当市场中存在羊群效应时，收益分散度显著为 0。所以当 $S_1 < S_2$ 时，第一类为出现羊群效应，当 $S_1 > S_2$ 时，第二类为出现羊群效应。

4. 实验

4.1. 数据的选取与预处理

为了研究中国股票市场中羊群效应的存在性及其存在的时间段，本文实验数据选取国泰安数据¹，计算入上证指数的只股票进行实证分析，样本选取的时间为 2020 年 11 月 2 日到 2020 年 11 月 27 日，共 20 个交易日。因为聚类方法适用于符合正态分布的数据，为了保持本文方法在不同数据中的准确性、有

¹<https://www.gtarsc.com/#/index>。

效性, 故先通过进行 Kolmogorov-Smirnov 正态分布检验[19] [21], 如果不通过则进行中位数绝对偏差 (MAD)剔除出现离群点的异常样本。利用股票的日开盘价格和收盘价格数据, 计算下列信息:

- ① 计算每支股票在不同时间的收益率 R_{it} , 其中 $i = \{1, 2, \dots, n\}, t = \{1, 2, \dots, T\}$;
- ② 整体收益率 R_{mt} ;
- ③ 根据表达式(12)计算变化趋势得到 x_{ij} ;
- ④ 聚类得到隶属度矩阵 u_{ij} ;
- ⑤ 根据判别准则判断哪一类是出现羊群效应的类别;
- ⑥ 结束。

4.2. 检测股票出现的羊群效应实验对比

根据建模求解计算得到部分结果如表 1~4 以及如图 4 所示:

Table 1. Herd effect of different stocks (Week 1)

表 1. 不同股票羊群效应情况(第一周)

股票代码	本文方法		K-means
	出现羊群效应的程度	没出现羊群效应的程度	
000001	0.546236896	0.453763104	1
000002	0.908388375	0.091611625	1
000003	0.379704932	0.620295068	2
000004	0.530905336	0.469094664	2
000005	0.672876655	0.327123345	1
000006	0.695243317	0.304756683	1
000007	0.13934821	0.86065179	2
000008	0.819485048	0.180514952	1
000009	0.876281983	0.123718017	1
000010	0.546236896	0.453763104	2

Table 2. Herd effect of different stocks (Week 2)

表 2. 不同股票羊群效应情况(第二周)

股票代码	本文方法		K-means
	出现羊群效应的程度	没出现羊群效应的程度	
000001	0.870837	0.129163	2
000002	0.326025	0.673975	1
000003	0.69042	0.30958	2
000004	0.457802	0.542198	1
000005	0.153475	0.846525	1
000006	0.546717	0.453283	2

Continued

000007	0.282614	0.717386	1
000008	0.493802	0.506198	2
000009	0.453328	0.546672	1
000010	0.261689	0.738311	1

Table 3. Herd effect of different stocks (Week 3)

表 3. 不同股票羊群效应情况(第三周)

股票代码	本文方法		K-means
	出现羊群效应的程度	没出现羊群效应的程度	
000001	0.460344	0.539656	1
000002	0.295121	0.704879	1
000003	0.515498	0.484502	2
000004	0.560762	0.439238	2
000005	0.190533	0.809467	1
000006	0.268053	0.731947	1
000007	0.125041	0.874959	1
000008	0.757405	0.242595	2
000009	0.549151	0.450849	2
000010	0.612012	0.387988	2

Table 4. Herd effect of different stocks (Week 4)

表 4. 不同股票羊群效应情况(第四周)

股票代码	本文方法		K-means
	出现羊群效应的程度	没出现羊群效应的程度	
000001	0.120832	0.879168	1
000002	0.450139	0.549861	1
000003	0.333998	0.666002	1
000004	0.06171	0.93829	1
000005	0.157347	0.842653	1
000006	0.240242	0.759758	1
000007	0.045671	0.954329	1
000008	0.155757	0.844243	1
000009	0.222187	0.777813	1
000010	0.289069	0.710931	1

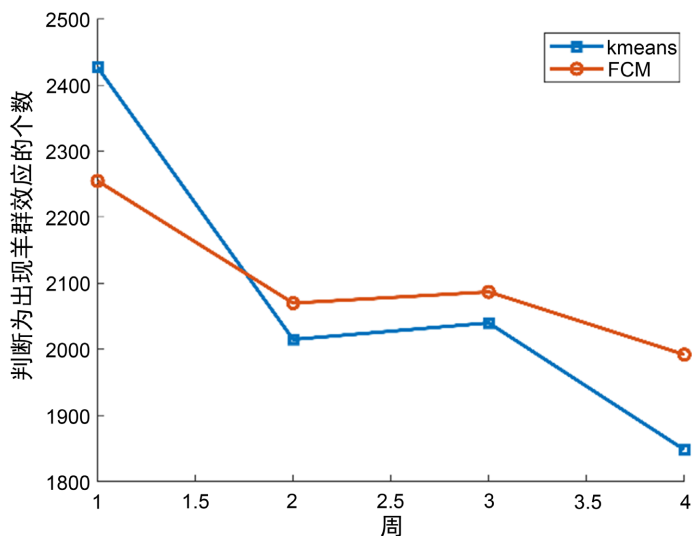


Figure 4. Experimental comparison of stock herd effect
图 4. 股票出现羊群效应实验对比

4.3. 检测时间段出现的羊群效应实验对比

根据建模求解计算得到检测时间段出现的羊群效应实验对比表 5 所示：

Table 5. The herd effect in different time periods
表 5. 不同时间段羊群效应情况

时间段	本文方法		K-means
	出现羊群效应的程度	没出现羊群效应的程度	
2020-11-02	0.500327	0.499673	1
2020-11-03	0.499502	0.500498	2
2020-11-04	0.500473	0.499527	1
2020-11-05	0.499488	0.500512	2
2020-11-6	0.500485	0.499515	1
2020-11-09	0.499533	0.500467	2
2020-11-10	0.500095	0.499905	1
2020-11-11	0.500202	0.499798	1
2020-11-12	0.5	0.5	2
2020-11-13	0.500132	0.499868	1
2020-11-16	0.499686	0.500314	2
2020-11-17	0.500347	0.499653	1
2020-11-18	0.49992	0.50008	1
2020-11-19	0.500026	0.499974	1
2020-11-20	0.499894	0.500106	2

Continued

2020-11-23	0.50003	0.49997	1
2020-11-24	0.499561	0.500439	2
2020-11-25	0.500383	0.499617	1
2020-11-26	0.499967	0.500033	2

5. 结论

本文通过提出一种基于模糊聚类的羊群效应分析方法，以检测股票市场中有无出现羊群效应、个股的羊群效应程度以及股市不同时间段的羊群效应程度。

本文通过计算收益率一阶差分来构建具有时间序列属性的收益变化特征，分别对在同时间段不同股票收益变化特征和整体股市不同时间段变化收益特征作为样本进行模糊聚类，得到隶属度来表示股票或时间段趋向于出现羊群效应的程度。

本文将每支股票作为单个样本，采用模糊 C 均值聚类，以得到隶属度矩阵。接着，建立类别判断准则。判断哪一类是羊群效应，可以通过计算第一类样本隶属度与第二类样本隶属度的最大/最小差值，将出现最大/最小差值的样本与收益变化特征均值的方法来判断。

相比计算横截面标准差等传统方法再进行线性回归得到回归系数，该方法具有能够表现个股羊群效应优势；相比 k-means 聚类计算羊群效应，基于模糊聚类的羊群效应分析方法用隶属度能够得到更多的聚类信息。

相比用硬聚类的方法，模糊聚类的方法得到的结果具有更多的信息，更符合实际应用。但本文方法少数样本会出现羊群效应和不出现羊群效应的程度相同，即无法判断有无羊群效应的情况。对于该问题，可以通过加入已有的先验信息，用半监督模糊聚类的方法来检测羊群效应提高羊群效应检测的准确度和稳定性。

参考文献

- [1] Banerjee, A.V. (1992) A Simple Model of Herd Behavior. *The Quarterly Journal of Economics*, **107**, 797-817. <https://doi.org/10.2307/2118364>
- [2] Bikhchandani, S., Hirshleifer, D. and Welch, I. (1992) A Theory of Fads, Fashion, Custom and Cultural Changes Informational Cascades. *Journal of Political Economy*, **100**, 992-1026. <https://doi.org/10.1086/261849>
- [3] Avery, C. and Zemsky, P. (1998) Multidimensional Uncertainty and Herd Behavior in Financial Markets. *American Economic Review*, **88**, 724-748.
- [4] Graham, J.R. (1999) Herding among Investment Newsletters: Theory and Evidence. *Journal of Finance*, **54**, 237-268. <https://doi.org/10.1111/0022-1082.00103>
- [5] Zwiebel, J. (1995) Corporate Conservatism and Relative Compensation. *Journal of Political Economy*, **103**, 1-25. <https://doi.org/10.1086/261973>
- [6] Lakonishok, J., Shleifer, A. and Vishny, R.W. (1992) The Impact of Institutional Trading on Stock Prices. *Journal of Financial Economics*, **32**, 23-43. [https://doi.org/10.1016/0304-405X\(92\)90023-Q](https://doi.org/10.1016/0304-405X(92)90023-Q)
- [7] Gdnblatt, M., Titman, S. and Wermers, R. (1995) Momentum Investment Strategies, Portfolio Performance and Herding: A Study of Mutual Fund Behavior. *American Economic Review*, **85**, 1088-1105.
- [8] Christie, W.G. and Huang, R.D. (1995) Following the Pied Piper: Do Individual Returns Herd around the Market? *Financial Analysts Journal*, **51**, 31-37. <https://doi.org/10.2469/faj.v51.n4.1918>
- [9] 宋军, 吴冲锋. 基于分散度的金融市场的羊群行为研究[J]. 经济研究, 2001(11): 21-27.
- [10] 张锐昕. 基于 CAPM 模型对中国股市羊群行为的实证研究[J]. 科教导刊, 2010(35): 64+68.
- [11] 孙培源, 施东晖. 中国证券市场羊群行为实证研究[J]. 证券市场导报, 2004(8): 39-43.

-
- [12] Chang, E.C., Cheng, J.W. and Khorana, A. (2000) An Examination of Herd Behavior in Equity Markets: An International Perspective. *Journal of Banking and Finance*, **24**, 1651-1679.
- [13] 宋军, 吴冲锋. 中国股评家的羊群行为研究[J]. 管理科学学报, 2003(1): 68-74.
- [14] Scharfstein, David, S. and Stein, J.C. (1990) Herd Behavior and Investment. *American Economic Review*, **80**, 465-479.
- [15] 郭旭超. 基于聚类的羊群效应分析模型[D]: [硕士学位论文]. 广州: 中山大学, 2010.
- [16] 蒋学雷, 陈敏, 吴国富. 中国股市的羊群效应的 ARCH 检验模型与实证分析[J]. 数学的实践与认识, 2003, 33(3): 56-63.
- [17] Bezdek, J.C., Ehrlich, R. and Full, W. (1984) FCM: The Fuzzy C-Means Clustering Algorithm. *Computers & Geosciences*, **10**, 191-203. [https://doi.org/10.1016/0098-3004\(84\)90020-7](https://doi.org/10.1016/0098-3004(84)90020-7)
- [18] Hathaway, R.J. and Davenport, J.W. (1989) Relational Duals of the C-Means Clustering Algorithms. *Pattern Recognition*, **22**, 205-212. [https://doi.org/10.1016/0031-3203\(89\)90066-6](https://doi.org/10.1016/0031-3203(89)90066-6)
- [19] Steinley, D. (2011) K-Means Clustering: A Half-Century Synthesis. *British Journal of Mathematical and Statistical Psychology*, **59**, 1-34. <https://doi.org/10.1348/000711005X48266>
- [20] Bora, D.J. and Gupta, A.K. (2014) A Comparative Study Between Fuzzy Clustering Algorithm and Hard Clustering Algorithm. *International Journal of Computer Trends and Technology*, **10**, 108-113. <https://doi.org/10.14445/22312803/IJCTT-V10P119>
- [21] Pawan, K. and Deepika, S. (2010) Comparative Analysis of FCM and HCM Algorithm on Iris Data Set. *International Journal of Computer Applications*, **5**, 33-37.