Fuzzy Logic and Behavioral Finance: An Approach Using Fuzzy c-Means Algorithm

Renato Aparecido Aguiar

Electrical Engineering Departament, Centro Universitário da FEI São Bernardo do Campo, Brazil preraquiar@fei.edu.br

Abstract— The aim of this paper is to show the connection between the behavioral finance theory and fuzzy sets theory. Such connection, little explored by researchers, may produce new models for the financial market, leading to a better understanding of anomalies not explained by modern theory of finance. In this paper two techniques based on fuzzy sets, an clustering algorithm and the fuzzy transform shown incorporate, intrinsically, the heuristics of representativeness and anchoring of the behavioral finance theory

Keywords— *Fuzzy Sets; Behavioral Finance; Fuzzy Transform; Fuzzy c-Means Algorithm; Heuristics.*

1. Introduction

The financial market can be viewed as a nonlinear and time varying system, subject to numerous events, such as wars, governmental changes and crises arising from natural phenomena. It is a complex system for which many mathematical models have been developed for a better understand of its dynamics. In particular, some models are the essential foundations for two theories of the financial market: the so called traditional finance theory and the behavioral finance theory. The traditional finance theory considers a rational investor operating in a market that reflects all available information. On the basis of this theory, Harry Markowitz developed a model for portfolio selection in 1952 (Markowitz, 1952) and William Sharpe developed the capital asset pricing model (CAPM) in 1964 (Sharpe, 1964). In 1970 Eugene Fama presented the theory of efficient market hypothesis (Fama, 1970), whereby the market is classified as informationally efficient in the weak form, semi-strong form or strong form. Contrasting the claims of traditional finance theory, the behavioral finance theory proposed by Kahnemann and Tversky (1974), Nobel laureates in economics in 2002, states that the individual is not fully rational and operates in a market that does not reflect all available information, but their decisions are biased

rules of thumb or heuristics, such as bv representativeness, anchoring and availability. Since then, some models were developed based on behavioral finance theory, supporting the idea that the individual is not fully rational and operates in a market not rational. Such models suggest that the stock market, due to biased decisions of investors, produces effects known as overreaction and underreaction. Assuming the influence of the behavioral finance heuristics, DeBondt and Thaler (1985) detected the phenomenon of overreaction in the U.S. stock market. The findings of Kang and others confirm overreaction in short term and underreaction in medium term in the Chinese stock market (Kang, Liu and Ni, 2002) Aguiar and Sales, using an algorithm based on fuzzy sets theory, detected the existence of overreaction in the U.S. stock market as influence of heuristics in the decision making of investors. Beyond the methodology based in the fuzzy sets theory proposed by Aguiar and Sales, the fuzzy sets theory has been widely applied in the analysis of the financial market, as can be observed in Liginlal and Ow (2006), Khcherem and Bouri (2009), Mohamed, Mohamad and Samat (2009). Supporting the application of the fuzzy sets theory in the financial market, Peters (1996) states that there is a strong connection between the fuzzy sets theory and behavioral finance theory, motivating the main focus of the present work: the development of a theoretical context for the connection between fuzzy logic and the theory of behavioral finance. It is noteworthy that a preliminary version of this work was published in (Aguiar and Sales, 2011)

This paper is organized as follows: In section 2 a brief review on fuzzy sets theory and the concepts of fuzzy clustering and Fuzzy Transform are presented; an introduction to behavioral finance theory is presented in section 3. In section 4 the main contribution of this paper is presented, that is, the connection between fuzzy sets theory and behavioral finance theory; the conclusions are presented in section 5.

International Journal of Latest Trends in Finance & Economic Sciences IJLTFES, E-ISSN: 2047-0916

Copyright © ExcelingTech, Pub, UK (http://excelingtech.co.uk/)

2. Fuzzy c- Means Algorithm

The fuzzy set theory proposed by Zadeh (1965) possesses as one of its main characteristics the fact of allowing the treatment of linguistic variables, such as hot, very hot, high, low, advisable, not advisable, highly risky, etc.

The resulting property when considering linguistic variables to characterize objects is that, instead of belonging or not to a certain set, as stated by the classic set theory, these objects will have pertinence indexes associated with different sets. A detailed presentation of the main concepts of the fuzzy theory can be found in Zimmermann (1996).

Definition 1: Let the set $X = \{x_1, x_2, ..., x_m\}$, $C_1, C_2, ..., C_n$ subsets of set X and real numbers $0 \le \mu_i(x_j) \le 1, i = 1, 2, ..., n, j = 1, 2, ..., m$, such that, for every j = 1, 2, ..., m, one has $\sum_{i=1}^n \mu_i(x_j) = 1$. Under these conditions, $\mu_i(x_j)$ is denoted membership

degree of the element x_j with respect to fuzzy subset C_i . These subsets can be represented by mathematical functions, called membership functions and defined in a certain interval of real numbers. So, the membership degree obtained through of these membership functions may be understood as a measure of the degree of affinity, similarity or compatibility among elements and a subset C_i .

Among the techniques for obtaining of the membership degree of an element with regard to the subsets are highlighted: techniques of grouping or classification of elements in subsets of a given set (clustering analysis) (Amir and Ganzach, 1998; Bezdek, 1981), rules IF and THEN defined by an expert for classify an element in regard to subsets through of the membership degree

The Fuzzy c-Means (FCM) algorithm (Bezdek, 1981) is a clustering method in which the elements represented by some attributes are classified into more than one group simultaneously. Each group is represented by a center vector and, the distance of each element with respect to the center of the group indicates the similarity between the group and the element.

Considering, for example, m elements represented by set $X = \{x_1, x_2, ..., x_m\}$ and c clusters, the aim is find c-partition of X exhibiting homogeneous subsets. Assuming an element x_i and a group represented by center vector c_i , $||x_j - c_i||$ is defined as the distance between an element and a group. The stronger the proximity of an element to a given subset, that is, the shorter the distance between an element and the center of a given subset, the closer will be the membership degree to the unity of that subset.

The quadratic distance weighted by the membership degree $[\mu_i(x_j)]^m$ produces the objective function represented in (1), which must be minimized.

$$J(U,c) = \sum_{i=1}^{n} \sum_{j=1}^{m} [\mu_i(x_j)^2 \| x_j - c_i \|^2]$$
(1)

where U is a fuzzy c- partition of set X and c is the set of cluster center.

The optimization problem given by (1) can be solved analytically, and the solution is given by Bezdek (Bezdek, 1981):

$$c_{i} = \frac{1}{\sum_{j=1}^{m} (\mu_{i}(x_{j}))^{2}} \sum_{j=1}^{m} (\mu_{i}(x_{j}))^{2} x_{j} \quad i = 1, \dots, n$$
(2)

$$\mu_{i}(x_{j}) = \frac{\left(\frac{1}{\|x_{j} - c_{i}\|^{2}}\right)^{1/(m-1)}}{\sum_{k=1}^{c} \left(\frac{1}{\|x_{j} - c_{k}\|^{2}}\right)^{1/(m-1)}} \quad j = 1, \dots, m, c = 1, \dots, n \quad (3)$$

As one can well observe, the calculation of the vectors of center c_i , given by (2) depends on the membership degrees $\mu_i(x_j)$. These, in turn, depend on c_i , according to (3). The solution can be obtained iteratively by the FCM algorithm, as described below (Bezdek, 1981; Zimmermann, 1996).

Step 1: Initiate the membership matrix in such a way that $\mu_1(x_j) + \mu_2(x_j) = 1, j = 1, 2, ..., m$ and $\mu_1(x_j) \ge 0$ and $\mu_2(x_j) \ge 0, j = 1, 2, ..., m$;

Step 2: Calculate the centers c_i , using (2);

Step 3: Recalculate the new membership matrix via (3) by utilizing the centers obtained in step 2;

Repeat steps 2 and 3 until the value of the objective function represented by (1) does not

decrease any longer according to the adopted precision. For achieve the global minimum of squareerror it takes run the algorithm with different initial partitions until the final partition no more change.

Among the techniques for grouping or classifying elements in subsets of a given set, the Fuzzy c-Means - FCM algorithm has been proved to be an effective tool in those cases in which the features or attributes of the analyzed elements can be represented by a vector of real numbers. In such cases, the FCM algorithm allows identifying clusters of elements from a $n \times p$ matrix, being n the number of elements and p the number of features of these elements, and simultaneously determines membership degrees associated to each element (Zimmermann, 1996). Each cluster is represented by a center, c_i , and the distance of each element in regard to each center defines the membership degree $\mu_i(x_i)$ of the element related to that cluster.

As an example, consider the pattern matrix M formed by six lines (6 objects) and two columns (2 features).

M =	$\begin{bmatrix} x_1 \end{bmatrix}$	_	1	2]
	<i>x</i> ₂		2	1
	<i>x</i> ₃		2	3
	<i>x</i> ₄		4	1
	<i>x</i> ₅		4	3
	_ <i>x</i> ₆ _		_5	2

Applying the FCM algorithm in this data set in order to form two groups, the following data shown in table 1 were obtained:

Observing table 1, note that the smaller the distance of an element with regarding to the center of a particular cluster, the less membership degree in regarding to this group. For example, considering the element xi formed by the ordered pair (1, 2), it is observed that the distance of this element in relation to group 1 is smaller than the distance in relation to group 2 and, consequently, the greater the membership degree of the element xi in relation to group 1.

The clusters are formed from the matrix membership degrees randomly chosen, according Step 1, grouping the elements by similarity measure based on the distance of each element relative to each center. In this case, after several iterations, the vectors of the center are adjusted to produce the final group and, consequently, the final matrix of membership degrees. This means that the element x_i is more similar to the group 1.

Data Matrix		I. MEMBERSHIP Degrees		Distance		
		Cluster 1	Cluster 2	Between the	Between the	
				Object and	Object and	
					Center	Center
					Vector	Vector
					of	of
				Cluster 1	Cluster 2	
X_1	1	2	0.9630	0.0370	0.6552	3.3448
X ₂	2	1	0.8532	0.1468	1.0574	2.5493
X ₃	2	3	0.8530	0.1470	1.0582	2.5450
X_4	4	1	0.1470	0.8530	2.5450	1.0582
X_5	4	3	0.1468	0.8532	2.5493	1.0574
X ₆	5	2	0.0370	0.9630	3.3448	0.6552

Table 1. Application of Fuzzy c-means Algorithm

The center vectors of the cluster 1 and cluster 2 are, respectively, (1.6555; 1.9996) and (4.3448; 2.0004).

One other fuzzy technique, called Fuzzy Transform, developed by Perfilieva (2006), is a transformation between two universes and very useful in many applications, such as image compression, data analysis, and can use several different functions for execute a fuzzy modeling. It's takes a function and produces a set-to-point correspondence between fuzzy sets from the partition and certain average values of that function.

So, let, A_1, \ldots, A_n be membership functions on interval [a,b] and a function f belonging to the set of continuos function on interval [a,b]. The n-tuple of real numbers $[F_1, \ldots, F_n]$ given by (4) is the fuzzy transform of f with respect to A_1, \ldots, A_n .

3

$$F_{K} = \frac{\int_{a}^{b} f(x) A_{K}(x) dx}{\int_{a}^{b} A_{K}(x) dx}, \quad K = 1,...,n$$
(4)

The kth component of the Fuzzy Transform minimizes the function showed in (5).

$$\phi(\mathbf{y}) = \int_{a}^{b} (\mathbf{f}(\mathbf{x}) - \mathbf{y})^{2} A_{K}(\mathbf{x}) d\mathbf{x}$$
 (5)

where $y = F_K$

Once known the Fuzzy Transform components F_K , it is possible (approximately) to reconstruct the original function f using (6):

$$\bar{\mathbf{f}} = \sum_{K=1}^{n} \mathbf{F}_{K} \cdot \mathbf{A}_{K}(\mathbf{x})$$
(6)

Equation (6) is called Fuzzy Transform inversion formula.

More details on the Fuzzy Transform can be obtained in Perfilieva (2006) and Perfilieva, Novák and Dvorák (2008).

3 Behavioral Finance

The modern theory of finance is based on the premise that the investor is rational, risk averse and operates in a market where stock prices reflect all available information. The research developed in (Fama, 1970), Markowitz (1952) and Sharpe (1964) present the fundamental bases for the development of the modern theory of finance.

On the other hand, according to the behavioral theory, individuals make decisions biased by heuristics, with a rationale that deviates from statistical rules. Cognitive psychology, which studies the mechanism of thought, is the basis of this approach and shows that individuals value too recent experience and are overconfident in their own abilities, providing thus the emergence of distortions in their thinkings (Ritter, 2003).

Among the heuristics of the behavioral finance, the representativeness and anchoring are heuristics of particular interest for this paper and are briefly described below:

i) Heuristic of representativeness: refers to a kind of mental shortcut in which there is a tendency to assume that something belongs to a particular group, based on the similarity with a member of that category. Many probabilistic questions with which people are concerned are: what is the probability that Vol-2 No. 1 March, 2012

4

the object A belongs to Class B? What is the probability that event A originates from process B? To answer these questions people use the representativeness heuristic, in which probabilities are assessed by the degree to which A is representative of B. In a classic example of literature, some individuals must answer what is the occupation of a person from a group of ten people, knowing that eight people in the group are truck drivers and two are brokers. In the first case the ten people are equally dressed and, after choosing randomly one person from the ten participants, based on the known probability, judged that this person would be a truck driver. In the second case, an element of ambiguity was added: the ten people were dressed differently and, when a person was chosen a person, wearing suits, sunglasses and carrying a folder. In this case, most participants identified this person as a broker, although the likelihood of this person be a truck driver to overcome the likelihood, known a priori, to be a broker. In this example, the man wearing suit, goggles and carrying a folder has more similarity to the set of brokers and less similarity to the set of truck drivers. The individuals make an association based on similarity, without conducting an analysis of the structure of probabilities, responding that the person selected was a broker (Peters, 1996). In the context of decisions on the economy, the individual under the influence of the representativeness heuristic has a strong tendency to overvalue recent information. As in the previous example, there is new ambiguous information that reduces the accuracy of the analysis, thereby producing a biased decision (Amir and Ganzach, 1998). The existence of such heuristic in decision making tends to produce overreaction, meaning that past losers tend to be winners in the future and vice versa (Fama, 1998).

ii) Heuristic of Anchoring: refers to a kind of mental shortcut which verifies the use of a standard as a starting point, adjusting decisions on the basis of this initial anchor. In many situations individuals make estimates supported by an initial value, which is adjusted to produce the final answer. The initial value may be suggested by the formulation of the problem or may be the result of a calculation. An experiment conducted by Tversky and Kahneman (1974) shows the influence of anchoring in the decision of an individual. In this experiment, two student groups must estimate the value of an expression in five seconds:

 $8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1$ for the group 1; $1 \times 2 \times 3 \times 4 \times 5 \times 6 \times 7 \times 8$ for the group 2.

Although the correct answer to the two sequences is 40.320, the average estimate obtained for the sequence 2 was 512, while the average estimate obtained by the group 1 (descending sequence) was 2.250. This occurs because in the descending sequence the first steps of multiplication (from left to right) produce a number greater than the ascending sequence. Thus, when individuals are faced with complex situations, they use to make decisions supported by the information available. In the stock market, where the amount of information is very extensive and dispersed, individuals tend to use mental shortcuts, or heuristic judgments in decisionmaking, transforming a complex trial in a simple task. The heuristic of anchoring is usually present when decisions are based on facts or terms of reference It is associated to conservative decisions, causing people to resist sudden changes in their decisions when faced with new information (Peters, 1996).

In terms of stock investments, market prices are usually a reference in the decision of an investor, since the information is extensive and scattered. The existence of such heuristic in decision making tends to produce sub-reaction, in which past winners tend to be future winners and losers in the past tend to be losers in the future (Fama, 1998).

4 Fuzzy c- Means Algorithm, Fuzzy Transform and Behavioral Finance

In this section the connection between the behavioral finance theory and the theory of fuzzy sets is explored. More specifically, it is shown that the fuzzy c- means (FCM) algorithm incorporates, intrinsically, some heuristics of the behavioral finance theory.

Concerning the heuristic of representativeness, it is present in the fuzzy clustering algorithm in the separation and grouping of objects. The groupings are made based on the similarity of each object with respect to each group. In the same way, the representativeness heuristic is mainly based on the similarity between objects as described in section 3.

As an example, the figure 1 shows the distribution of 100 random data, each one with two characteristics.



Figure 1. Data Matrix

Applying the FCM algorithm to this data set a clustering consisting of two groups is obtained as shown in figure 2.



Figure 2. Clustering Fuzzy

Clearly, elements more similar in the sense of their membership degrees are grouped. This grouping occurs due to the heuristic of representativeness contained, intrinsically, in the fuzzy c- means algorithm that, based on the similarity between the objects and each group, assigns a membership degree to each element with regard to each cluster.

The anchoring heuristic is also present, intrinsically, in the fuzzy c- means algorithm. As already defined in section 3, a decision based on this heuristic is adjusted for an anchor, in other words, an initial value used to produce the final answer. Similarly, as described in the section 2, the first step of the algorithm starts with a matrix of membership degrees that will be adjusted at each iteration of the algorithm, to produce the final answer or the final grouping.

Since the fuzzy c- means algorithm is based on fuzzy sets theory, there is a strong connection, or a great similarity between the fuzzy sets theory and the theory of behavioral finance. In the model developed by Aguiar and Sales, called Behavioral Fuzzy Model (Aguiar et al, 2008), stocks of the financial market are classified by fuzzy c-means algorithm and two groups (winner and loser) are defined, being each group represented by a center. The grouping of the stocks is based on the similarity of each stock with regard to the center of each group, forming a winning stock portfolio and a loser portfolio. Aguiar and Sales find that the Behavioral Fuzzy Model is biased by representativeness and anchoring heuristics in decision making (Aguiar et al, 2008), exploring, in this way, anomalies (overreaction or underreaction) present in the stock market. In Aguiar and Sales (2010), applying the methodology developed em (Aguiar et al, 2008) in the American stock market, the link between fuzzy logic and behavioral finance was found in the practice.

Similarly, there is an apparent connection between the behavioral finance theory and the theory of fuzzy sets. More specifically, the Fuzzy Transform possesses, intrinsically, some heuristics of the behavioral finance theory.

Concerning the representativeness heuristic, it is present in the Fuzzy Transform in the following sense: as can be seen in (4), the components of the Fuzzy Transform are the weighted mean values of the given function, where the weights are gives by the membership degree, $A_K(x)$, obtained through of the membership functions. The membership degree represents the similarity of each object with regard to each group. Surprisingly, the representativeness heuristic is mainly based on the similarity between objects as described in the section 3.

The anchoring heuristic is also present, intrinsically, in the Fuzzy Transform in the following sense: a already defined in the section 3, a decision based on this heuristic is adjusted for an anchor, an initial value used for produce the final answer. Similarly, as described in the section 2, the components of the Fuzzy Transform must minimize (5); otherwise the components are not valid. Thus, (5) represents an anchor for the Fuzzy Transform.

Since that the Fuzzy Transform is based on fuzzy sets theory, there is a strong connection, or a great similarity between the fuzzy sets theory and the theory of behavioral finance. In the methodology developed by Perfilieva, Novák and Dvorák (2008) the Fuzzy Transform is utilized for detection and characterization of dependencies among attributes. In Perfilieva, Novák, Dvorák, (2008) was created an optimal mathematical model of the gross domestic product (GDP), in other words, was founded a minimal set of attributes that determine the dynamic of the GDP. The results showed that the GDP can be represented by just three variables: gross product, gross capital and final consumption.

Of course, the Fuzzy Transform may be utilized in the financial market for forecast, dependence analysis and portfolio formation, allowing to explore the anomalies (overreaction or underreaction) present in the stock market (Ritter, 2003; Sharpe, 1964), once that the Fuzzy Transform has biases contained in the behavioral finance theory. The connection between fuzzy logic and behavioral finance evidenced in this paper, not only serve as a demonstration of approach of fuzzy techniques in stock market, but also for provide new methodologies for to model and analyze the stock market.

5 Conclusion

Many models have been developed to assist and improve the performance of investors in decision making. Some models, based on the theory of modern finance, assume that the investor is rational and risk averse. On the other hand, behavioral finance theory claims that the investor is biased by heuristics, such as representativeness, anchoring and availability in the decision making in the stock market.

This paper shows the strong connection between fuzzy sets theory and behavioral finance. The comparison between these two theories shows that two fuzzy techniques, called fuzzy c- means (FCM) algorithm and Fuzzy Transform (F- Transform), which has constructive base in the theory of fuzzy sets, incorporates, intrinsically, the heuristics of representativeness and anchoring derived from the theory of behavioral finance.

Thus, a model for the financial market that has as constructive base the theory of fuzzy sets, leads an investor to make a biased decision by heuristics of representativeness and anchoring. The presence of these heuristics in decision making of an investor can produce some anomalies in the stock market, known as overreaction and underreaction.

References

 Aguiar, R. A., Sales and R. M., Sousa, L. A.. (2008), Um Modelo Fuzzy Comportamental para Análise de Sobrereação e Subreação no Mercado de Ações Brasileiro. *Revista de Administração de Empresas* - RAE, v. 48, n.3.

- [2] Amir, E. and Ganzach, Y. (1998), Overreaction and Underreaction in Analysts' Forecasts. *Journal of Economic Behavior & Organization*, v. 37, p. 333-347.
- [3] Bezdek, J. C. (1981), *Pattern Recognition with Fuzzy Objective Function Algorithm*. Plenum Press, New York and London.
- [4] Fama, E. (1970), Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal* of *Finance*, v. 25, p. 383-417.
- [5] Fama, E. (1998), Markets Efficiency, Long-Term Returns, and Behavioral Finance. *Journal of Finance Economics*, v. 49, p. 283-306.
- [6] Tversky, A. and Kahneman, D. (1974), Judgement under Uncertainty: Heuristics and biases. *Science*, v. 185, n. 4157, pp. 1124-1131.
- [7] Markowitz, H. M. (1952), Portfolio Selection. *The Journal of Finance*, Chicago: American Finance Association, v. VII, n. 1, p. 77-91.
- [8] Peters, E. (1996), *Chaos and Order in the Capital Markets*, 2nd edition. New York: Wiley.
- [9] Peters, E. (2003), Simple and Complex Market Inefficiencies: Integrating Efficient Markets, Behavioral Finance, and Complexity. *The Journal of Behavioral Finance*, v. April, n. 4.
- [10] Ritter, J. R. (2003), Behavioral Finance. Pacific-Basin. *Finance Journal*, v. 11, p. 429-437.
- [11] Sharpe, W. F. (1964), Capital Asset Prices: a theory of market equilibrium under conditions of risk. *Journal of Finance*, Chicago: American Finance Association, p. 425-442, set.
- [12] Zadeh, L. A. (1965), Fuzzy sets. *Information and Control*, vol. 8, p. 338-353.
- [13] Zimmermann, H. J. (1996), *Fuzzy Set Theory* and Its Application. Kluwer Academic, Boston.
- [14] Aguiar, R. A. and Sales, R. M. (2010), Overreact Analysis in the American Stock Market: A Fuzzy c-Means Algorithm Approach. International Journal of Trade, Economics and Finance, v.1, n. 4, p. 325-330.
- [15] Perfilieva, I. (2006), Fuzzy Transform: Theory and Applications. *Fuzzy Sets and Systems*, v.157, p. 993-1023, 2006.
- [16] Perfilieva, I. Novák, V. and Dvorák, A. (2008), Fuzzy Transform in the Analysis of Data.

International Journal of Approximate Reasoning, v.48, p. 36-46.

- [17] DeBondt, W. F. M., Thaler, R. H. (1985), Does the Stock Market Overreact? *Journal of Finance*, v.40, p. 793-807.
- [18] Kang, J., Liu, M. and Ni, S. X. (2002), Contrarian and Momentum Strategies in the China Stock Market: 1993-2000. *Pacific-Basin Finance Journal*, v.10, p. 243-265.
- [19] Liginlal, D. and Ow, T.T. (2006), Modeling Attitude to Risk in Human Decision Processes: An Application of Fuzzy Measures. *Fuzzy Sets* and Systems, v.157, p. 3040-3054.
- [20] Khcherem, F. and Bouri, A. (2009), Fuzzy Logic and Investment Strategy. *Global Economy & Finance Journal*, v.2, p. 22-37.
- [21] Mohamed, Z., Mohamad, D. and Samat, O. (2009), A Fuzzy Approach to Portfolio Selection. *Sains Malaysian*, v.38, p. 895-899.
- [22] Aguiar, R. A. and Sales, R. M. (2011), Fuzzy Logic and Behavioral Finance: A Connection. International Conference on Economics and Finance Research, *Proceedings*.

7