Fuzzy c-means clustering algorithm for directional data (FCM4DD)

Orhan Kesemen*, Özge Tezel, Eda Özkul

Department of Statistics and Computer Sciences, Karadeniz Technical University, 61080 Trabzon, Turkey

Abstract

Cluster analysis is a useful tool used commonly in data analysis. The purpose of cluster analysis is to separate data sets into subsets according to their similarities and dissimilarities. In this paper, the fuzzy c-means algorithm was adapted for directional data. In the literature, several methods have been used for the clustering of directional data. Due to the use of trigonometric functions in these methods, clustering is performed by approximate distances. As opposed to other methods, the FCM4DD uses angular difference as the similarity measure. Therefore, the proposed algorithm is a more consistent clustering algorithm than others. The main benefit of FCM4DD is that the proposed method is effectively a distribution-free approach to clustering for directional data. It can be used for N-dimensional data as well as circular data. In addition to this, the importance of the proposed method is that it would be applicable for decision making process, rule-based expert systems and prediction problems. In this study, some existing clustering algorithms and the FCM4DD algorithm were applied to various artificial and real data, and their results were compared. As a result, these comparisons show the superiority of the FCM4DD algorithm in terms of consistency, accuracy and computational time. Fuzzy clustering algorithms for directional data (FCM4DD and FCD) were compared according to membership values and the FCM4DD algorithm obtained more acceptable results than the FCD algorithm.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

In the statistical analysis of random sampled data, it is assumed that the data came from a random variable. This random variable can exist in various measure spaces such as metric, time, color, angular etc. Univariate data in the angular ($\theta$) space is called circular data. The directions of the winds; the directions of migrating birds or animals (Chang-Chien, Yang, & Hung, 2010); the orientation of objects in the plane can be held up as circular data. On the other hand, data which does not involve orientation but occurs in periodic process can be analyzed in the same class. Periodic data show the same characteristics within a certain period of time. A student’s weekly study schedule and the amount of water consumed daily by living creatures on a yearly basis can be held up as periodic data. Data whose frequency changes periodically can be converted into circular data, although generally it is not circular data.

Generally, angular-based data is called directional data. If directional data has two variables, it is called spherical data. If directional data has more than two variables, it is called hyper-spherical data (Fisher, 1993).

Circular distribution of the data was first examined by von Mises in 1918 (von Mises, 1918). Then, studies on statistical inference from circular data were made by Watson and Williams (1956). After this study, the interest of many researchers in this field increased. Batschelet (1981), Fisher (1993) and Mardia and Jupp (2000) are major books on analysis of circular data which have applications in many fields such as biology, geology, medicine, meteorology, oceanography etc. In addition to these, Money, Helms, and Jolliffe (2003) investigated circular data for a case study involving sudden infant death syndrome (SIDS). Carta, Bueno, and Ramirez (2008) studied statistical modeling of directional wind speeds. Lee (2010) compiled the methods that have been developed the last 50 years. Baayen, Klugkist, and Mechsner (2012) proposed a test of order-constrained hypotheses for circular data with applications to human movement science. Abraham, Molinari, and Servien (2013) studied unsupervised clustering of multivariate circular data which consist of the positions of five separate X-ray beams on a circle. Chen, Singh, Guo, Fang, and Liu (2013) improved a new method to identify flood seasonality and partition the flood season into sub-seasons. A study conducted by Costa, Koivunen, and Poor (2014) estimated directional probability distribution of wavefields observed by sensor arrays. Tasdan and Cetin (2014) carried out a simulation study on the influence of ties on uniform scores test for circular data. Hawkins and Lombard (2015) proposed an optimal method for segmentation of cir-
cular data generated from von Mises distribution. Kitamura et al. (2015) proposed a new hybrid method that combined directional clustering and advanced nonnegative matrix factorization (NMF). They handled the problems in multichannel music signal separation. da Silva (2015) proposed a directional clustering approach based on mixtures of von Mises-Fisher (vMF) distributions to reduce uncertainty in estimating the orientation of neuronal pathways in diffusion magnetic resonance imaging. Yang, Chang-Chien, and Hung (2016) presented an unsupervised clustering algorithm for directional data on the unit hypersphere without initialization, for which it is not necessary to give the number of clusters a priori.

Clustering analysis is one of the most important issues in the data analysis. Clustering is used to separate a data set into a desired number of clusters. In this separation process, the data points in the same cluster are the most similar to each other and the data points in the different clusters are the most dissimilar.

When considered from the statistical point of view, clustering methods can generally be divided into two categories: the not distribution-free approach and the distribution-free approach. The most-used algorithms from the not distribution-free approaches are the expectation and maximization (EM) algorithm (Dempster, Laird, & Rubin, 1977; McLachlan & Basford, 1988) and the fuzzy c-directions (FCD) algorithm (Yang & Pan, 1997). These algorithms can be applied to directional data. Chang-Chien, Hung, and Yang (2012) adapted the mean shift clustering algorithm, used for numeric data, for circular data by determining automatically the number of clusters. Then, Yang, Chang-Chien, and Kuo (2014) applied the mean shift clustering algorithm for circular data to hyperspherical data.

The most-used algorithms from the distribution-free approaches are partitioning clustering methods, k-means (MacQueen, 1967) and fuzzy c-means (FCM) clustering algorithms are the most-common partitioning clustering algorithms. These algorithms are applied to linear data. However, in this study, the FCM clustering algorithm is modified to apply to directional data. Different from the existing studies, the proposed method uses angular difference as the similarity measure. In addition, the proposed algorithm is a distribution-free approach.

In Section 2, classical FCM algorithm is explained. In Section 3, general definitions and similarity measures for directional data are described. In Section 3.1 and Section 3.2, the EM and the FCD algorithms for directional data are introduced. In Section 4, the modified FCM algorithm for directional data is given. In Section 5, the EM, the FCD and the FCM4DD algorithms are applied to the same numerical data, and their performances are compared. In Section 5.1, the membership values of the FCD and the FCM4DD are compared.

2. Fuzzy c-means clustering (FCM) algorithm

The fuzzy c-means clustering (FCM) algorithm was proposed by Dunn in 1973 and improved by Bezdek in 1981 (Höppner, Klawonn, Kruse, & Runkler, 2000). The FCM algorithm, based on objective function, is subject to the principle that each data point belongs to more than one cluster with different membership values, ranging from [0,1]. Additionally, the sum of the membership values for each data point must be one. If a data point is in the cluster center, its membership value is one.

Let \( X = \{x_1, x_2, \ldots, x_N\} \) be a sample of N observations in D-dimensional Euclidean space(\( x_i \in \mathbb{R}^D \)). Clustering is process which separates this data set into \( C \) subsets and their cluster centers which are \( \{v_1, v_2, \ldots, v_C\} \). The desired optimal criterion minimizes the objective function while separating the data set into subsets. The algorithm tries to minimize the following objective function which is the generalized form of the least-squared errors function (Höppner et al., 2000):

\[
j_m = \sum_{i=1}^{N} \sum_{j=1}^{C} \mu_{ij}^m \| x_i - v_j \|^2, \quad 1 < m < \infty
\]

in which \( m \) is the weighting fuzziness parameter and is generally chosen as 2. \( \mu_{ij} \) is the membership value of the \( i^{\text{th}} \) data to the \( j^{\text{th}} \) cluster and \( \mu_{ij} \) must satisfy the following three conditions (Bezdek, 1981):

1. The membership value ranges between zero and one as given in Eq. (2):

\[
\mu_{ij} \in [0, 1], \quad \forall \ i, j
\]

2. The sum of the membership values for each data point must be one as given in Eq. (3):

\[
\sum_{j=1}^{C} \mu_{ij} = 1, \quad \forall i
\]

3. The sum of the all membership values in a cluster must be smaller than the number of data (N) as given in Eq. (4):

\[
0 < \sum_{i=1}^{N} \mu_{ij} < N, \quad \forall N
\]

The FCM algorithm is a simple method and is the most common clustering algorithm in the all fuzzy clustering methods (Bezdek, Ehrlich, & Full, 1984). The FCM algorithm can be summarized as follows:

FCM Algorithm

Step 1. Fix \( C \in \{2, N\} \), \((m > 0)\) and \((\epsilon > 0)\).

Step 2. Give initials \( \mu_{ij}^{(0)} \sim U(0,1) \) and let \( t = 1 \).

Step 3. Compute cluster centers \( \{v_j\} \) by using Eq. (5):

\[
v_j = \frac{\sum_{i=1}^{N} \mu_{ij}^m x_i}{\sum_{i=1}^{N} \mu_{ij}^m}, \quad (j = 1, 2, \ldots, C)
\]

Step 4. Update \( \mu_{ij} \) with \( v_j \) by using Eq. (6):

\[
\mu_{ij} = \left( \sum_{k=1}^{C} \left( \frac{\| x_i - v_j \|}{\| x_i - v_k \|} \right)^{\frac{2}{m-1}} \right)^{-1}, \quad (i = 1, 2, \ldots, N; \ j = 1, 2, \ldots, C)
\]

Step 5. Compute \( \| \mu_{ij}^{(t)} - \mu_{ij}^{(t-1)} \| \).

IF \( \| \mu_{ij}^{(t)} - \mu_{ij}^{(t-1)} \| < \epsilon, \ STOP \)
ELSE \( t = t + 1 \) and return to Step 3.

3. Clustering methods for directional data

Generally, statistical data analysis is used for data on the linear axis (Kauffman & Rousseseuw, 1990). However, classical statistical methods used for these data cannot be applied inherently to directional data. This is because; directional data have a modular structure. If directional data are defined in the interval \([-\pi, \pi]\), they are continuous between the points \((\pi)\) and \((-\pi)\); if directional data are defined in the interval \([0, 2\pi]\), they are continuous between the points \((2\pi)\) and \((0)\). In terms of numerical values, classical methods cannot be used, when the data is discontinuous within these boundaries. The best example of this is demonstrated by the following: distance between the angles 359° and 1° is 2°, but the numerical subtraction of these angles is 358°. Likewise, it might firstly appear that the mean of these angles is 0°, whereas it is 180° in reality.

Clustering algorithms use the distances between data as the similarity measure. There are angularly two distances between two
angles $\theta_a$ and $\theta_b$ on the circle: clockwise, counter-clockwise. Generally, the shortest distance of these is preferred (Mardia & Jupp, 2000).

One equation is proposed by Ackermann (1997) for the remedy of this confusion in the distance measurement. This circular distance formula is given in Eq. (7). The measure $\delta$ yields the smaller of the two angles between $\theta_a$ and $\theta_b$.

$$\delta = \pi - |\pi - |\theta_a - \theta_b||$$  \hspace{1cm} (7)

The other approach for the distance between two angles is given in Eq. (8) (Lund, 1999):

$$d = 1 - \cos (\theta_a - \theta_b)$$  \hspace{1cm} (8)

The possible range of value of distance $d$ in Eq. (8) is demonstrated in the interval $[0, 2]$. Although the distance measurement in Eq. (7) equals real angular distance, it is insufficient for many circular data applications. Therefore, generally, Eq. (8) is used in the circular data applications.

In this study, we use Eq. (9) to measure angular distance as the similarity measure between two angles.

$$\psi = (((\theta_a - \theta_b) + \pi) \mod 2\pi) - \pi$$  \hspace{1cm} (9)

As opposed to Eqs. (7) and (8), the distance measured in Eq. (9) can be a negative value. Here, according to counter-clockwise, if $\theta_a$ is ahead, $\psi$ has positive values. Otherwise it has negative values. Firstly, this section will be focused on the circular data from the given distribution. Then, we will explain the EM algorithms (Bartels, 1984; Spurr & Koutbeiy, 1991) and the FCD algorithms (Yang & Pan, 1997) algorithm regarding circular data. The angles must be in radians in the interval $[0, 2\pi]$. The statisticians have proposed various distributions for the distribution of the circular data. The most common of these is the von Mises distribution ($VM(v, \kappa)$). The parameters $v$ and $\kappa$ are called the mean direction and the concentration parameter of the distribution, respectively. If the parameter $\kappa$ is too large, the distribution concentrates around the mean. The von Mises distribution is used as a probability model for clustering circular data. The von Mises distribution ($VM(v, \kappa)$) has the probability density function:

$$f(\theta; v, \kappa) = \frac{1}{2\pi I_0(\kappa)} \exp (\kappa \cos (\theta - v)), \quad 0 \leq \theta < 2\pi, \quad 0 \leq \kappa < \infty$$  \hspace{1cm} (10)

where $I_0(\kappa)$ denotes the modified Bessel function of the first kind and order 0, which can be defined by (Mardia & Jupp et al, 2000)

$$I_0(\kappa) = \frac{1}{2\pi} \int_0^{2\pi} \exp (\kappa \cos (\theta - v)) d\theta.$$  \hspace{1cm} (11)

3.1. The EM algorithm for circular data

The EM algorithm, one of the most common methods for clustering circular data, was proposed firstly by Dempster et al. (1977). It uses predictive criteria instead of absolute distance criteria in determining which object belongs to which cluster and is used for maximum likelihood estimation in case of incomplete data. On the each iteration of the EM algorithm, there are two steps called the Expectation step (E-step) and Maximization step (M-step). The E-step and the M-step are repeated consecutively until convergence criteria. In the E-step, the best probabilities about incomplete-data are estimated using predictions belonging to parameters of the observed data set. In the M-step, the estimated incomplete-data is put into place and the maximum likelihood is calculated over the completed data set. Thus, the new estimation of the parameters is obtained (Bruzzone & Prieto, 2002).

Suppose that the data set $\Theta = \{\theta_1, \theta_2, \ldots, \theta_N\}$ is a random sample from a mixture of von Mises distribution. The EM algorithm for directional data can be summarized as follows:

**EM Algorithm**

Step 1. Fix $C \in \{2, N\}$ and ($c > 0$).

Step 2. Give initials randomly $z_{ij} \sim U_\theta(0, 1)$. ($\sum_{i=1}^{C} z_{ij} = 1$) and let $\ell = 1$.

Step 3. Compute $\alpha^{(c)}$ with $z^{(c-1)}$ by using Eq. (12):

$$\alpha_j = \frac{1}{N} \sum_{i=1}^{N} z_{ij}, \quad (j = 1, 2, \ldots, C).$$  \hspace{1cm} (12)

Step 4. Compute $\psi^{(c)}$ with $z^{(c-1)}$ by using Eq. (13):

$$v_j = atan2 \left( \sum_{i=1}^{N} z_{ij} \sin (\theta_i), \sum_{i=1}^{N} z_{ij} \cos (\theta_i) \right), \quad (j = 1, 2, \ldots, C).$$  \hspace{1cm} (13)

where $atan2$ is an arc tangent function in the range $[-\pi, \pi]$.

Step 5. Compute $\kappa^{(c)}$ with $z^{(c-1)}$, $\psi^{(c)}$ and $\theta$ by using Eq. (14):

$$k_j = A^{-1} \left( \frac{\sum_{i=1}^{N} z_{ij} \cos (\theta_i - v_j)}{\sum_{i=1}^{N} z_{ij}} \right), \quad (j = 1, 2, \ldots, C).$$  \hspace{1cm} (14)

where $A^{-1}(x)$ can be computed using Batschelet’s table (see. Fisher, 1993).

Step 6. Compute $z^{(c)}$ with $\alpha^{(c)}, \psi^{(c)}, \kappa^{(c)}$ by using Eq. (15):

$$z_{ij} = \frac{\alpha_i f(\theta_i; v_j, k_j)}{\sum_{l=1}^{C} \alpha_l f(\theta_i; v_j, k_j)}, \quad (i = 1, 2, \ldots, N), \quad (j = 1, 2, \ldots, C).$$  \hspace{1cm} (15)

3.2. Fcd algorithm for circular data

Fuzzy c-directions clustering algorithm was proposed by Yang and Pan (1997). In this algorithm, the data point belongs to more than one cluster with different membership values, ranging from $[0, 1]$. The FCD and the EM algorithms are algorithmically very similar to each other. However, membership functions of each data point in all clusters are calculated in the FCD algorithm. This is different from the EM algorithm. Suppose that the data set $\Theta = \{\theta_1, \theta_2, \ldots, \theta_N\}$ is a random sample from a mixture of von Mises distribution. The objective function proposed by Yang and Pan (1997) is given in Eq. (16):

$$B_{m,w}(\mu, \alpha, v, \kappa) = \sum_{i=1}^{N} \sum_{j=1}^{C} \mu_{ij}^{m} \left(-\log_2 \alpha_i f(\theta_i; v_j, k_j) + k_j \cos (\theta_i - v_j) \right) + \sum_{i=1}^{N} \sum_{j=1}^{C} \mu_{ij}^{m} \log_2 \alpha_j$$  \hspace{1cm} (16)

Here, $\alpha_i$ and $\mu_{ij}$ must satisfy the following two conditions:

$$\sum_{j=1}^{C} \alpha_j = 1$$  \hspace{1cm} (17)

$$\sum_{j=1}^{C} \mu_{ij} = 1. \quad (i = 1, 2, \ldots, N)$$  \hspace{1cm} (18)

The FCD algorithm for directional data can be summarized as follows:
4. Fuzzy c-means clustering algorithm for directional data (FCM4DD)

In this study, the fuzzy c-means clustering algorithm was adapted for directional data. The FCM4DD is based on angular difference as given in Eq. (9).

Circular data can be represented by \( \Theta = [\theta_1, \theta_2, \ldots, \theta_N] \) or periodically by \( X = \{x_1, x_2, \ldots, x_1, x_2\} \). If the data is given in the form of a periodic value, it can be converted into circular data by using Eq. (24):

\[
\theta_i = \frac{2\pi x_i}{T} - \pi
\]

where \( T \) denotes the period of the variable \( x_i \).

In this study, circular data can be defined in the range \([-\pi, \pi]\) or \([0, 2\pi]\). The proposed algorithm tries to minimize the following objective function which is the generalized form of the least-squared errors function:

\[
J_m = \sum_{i=1}^{N} \sum_{j=1}^{c} \mu_{ij}^m \| \theta_i - \Phi_j \|^2, \quad 1 < m < \infty
\]

in which \( m \) is the weighting exponent parameter and generally is chosen as 2. \( \Phi_i \) is the center of the \( j \)-th cluster, \( \mu_{ij} \) is the membership value of the \( i \)-th data to the \( j \)-th cluster and \( \mu_{ij} \) must satisfy the conditions given in Eqs. (2)-(4).

The proposed method for circular data can be summarized as follows:

**FCM4DD Algorithm**

Step 1. Fix \( C \in [2, N] \), \( (m > 0) \), \( w > 0 \) and \((\epsilon > 0)\).

Step 2. Give initial cluster centers \( \psi_j^{(0)} \sim U(\pi, \pi) \) and let \( t = 1 \).

Step 3. Compute \( \alpha_j \) with \( \alpha_j^{(t-1)} \) by using Eq. (20):

\[
\alpha_j = \frac{\sum_{i=1}^{N} \mu_{ij}^m}{\sum_{i=1}^{N} \sum_{j=1}^{c} \mu_{ij}^m}, \quad (j = 1, 2, \ldots, C)
\]

Step 4. Compute \( \psi_j \) with \( \psi_j^{(t)} \) by using Eq. (21):

\[
\psi_j = \frac{\sum_{i=1}^{N} \mu_{ij}^m \sin(\theta_i) - \sum_{i=1}^{N} \mu_{ij}^m \cos(\theta_i)}{\sum_{i=1}^{N} \mu_{ij}^m}, \quad (j = 1, 2, \ldots, C)
\]

where \( \text{atan2} \) is an arc tangent function in the range \([-\pi, \pi]\).

Step 5. Compute \( \kappa_j \) with \( \kappa_j^{(t-1)}, \psi_j^{(t)} \) and \( \theta \) by using Eq. (22):

\[
\kappa_j = A^{-1}\left(\frac{\sum_{i=1}^{N} \mu_{ij}^m \cos(\theta_i - \psi_j)}{\sum_{i=1}^{N} \mu_{ij}^m}\right), \quad (j = 1, 2, \ldots, C)
\]

where \( A^{-1}(x) \) can be computed by using Batschelet’s table (see, Fisher, 1993).

Step 6. Compute \( \mu_{ij}^{(t)} \) with \( \alpha_j^{(t)}, \psi_j^{(t)}, \kappa_j \) and \( \theta \) by using Eq. (23):

\[
\mu_{ij}^{(t)} = \left\{ \sum_{k=1}^{C} \left( \frac{\log(2\pi \lambda_0(\kappa_k)) + \kappa_j \cos(\theta_i - \psi_j_i) - \log(\alpha_j)}{\sum_{k=1}^{C} \log(2\pi \lambda_0(\kappa_k)) + \kappa_j \cos(\theta_i - \psi_j_i) - \log(\alpha_j)} \right)^{-1} \right\}^{1/2}, \quad (i = 1, 2, \ldots, N; \; j = 1, 2, \ldots, C)
\]

Step 7. Compute \( \mu^{(t)} - \mu^{(t-1)} \)

IF \( \|\mu^{(t)} - \mu^{(t-1)}\| < \epsilon \), STOP
ELSE \( t = t + 1 \) and return to Step 3.

5. Experimental results

The EM, the FCD and the FCM4DD algorithms are applied to some numerical examples, and their performances are compared. Also, it is shown that the FCM4DD algorithm can be applied to the spherical data.

**Example 1.** To compare the performances of the three kinds of clustering algorithms, EM, FCD and FCM4DD, we consider the simulation data from the mixture of von Mises distributions by generating 100 data points from \( 0.7\text{VM}(\pi/3, 2.3) + 0.3\text{VM}(4\pi/3, 3.8) \) using Best and Fisher’s simulation method (1979). This process is repeated 1000 times and the means of cluster centers (\( \bar{v}_1; \bar{v}_2 \)) for each method are shown in Table 1.

<table>
<thead>
<tr>
<th>Algorithms</th>
<th>EM</th>
<th>FCD</th>
<th>FCM4DD</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \bar{v}_1 ) (deg)</td>
<td>59.7129</td>
<td>60.4142</td>
<td>59.9552</td>
</tr>
<tr>
<td>( \bar{v}_2 ) (deg)</td>
<td>239.6176</td>
<td>239.2764</td>
<td>240.2741</td>
</tr>
<tr>
<td>MSE</td>
<td>0.00006966</td>
<td>0.00021177</td>
<td>0.00023250</td>
</tr>
<tr>
<td>ACT (sec)</td>
<td>0.20307981</td>
<td>0.10858994</td>
<td>0.00094319</td>
</tr>
</tbody>
</table>

According to MSE and the average computation time, the FCM4DD algorithm gives the best results.

**Example 2.** To compare performances of the three kinds of clustering algorithms, EM, FCD and FCM4DD, we consider the simulation data from the mixture of triangular distributions by generating 100 data points from \( 0.5\Lambda(0, \pi/2, \pi/2) + 0.5\Lambda(\pi, \pi/2, 3\pi/2) \) and the means of which are \( (\pi/3, 4\pi/3) \). This process is repeated 1000 times and the means of cluster centers (\( \bar{v}_1; \bar{v}_2 \)) for each method are shown in Table 2.

The FCM4DD algorithm shows better performance than the other algorithms, and we observe that the MSE values of the
FCM4DD and the EM are almost the same in many trials. The FCM4DD algorithm maintains its superiority according to the average computation time as in the previous example.

**Example 3.** In this example, we consider a real data set. Table 3 shows the directions of 76 turtles after laying eggs which are given by Stephens (1969).

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Directions of 76 turtles after laying eggs.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Numbers (deg)</td>
</tr>
<tr>
<td>8-9</td>
<td>12 13 14 15 18 22 27 30 34</td>
</tr>
<tr>
<td>38-50</td>
<td>65 66 67 70 73 78 78 83</td>
</tr>
<tr>
<td>83-88</td>
<td>90 92 92 93 95 96</td>
</tr>
<tr>
<td>98-100</td>
<td>103 106 113 118 138 153 153 155</td>
</tr>
<tr>
<td>204-213</td>
<td>223 226 237 243 244 250 251</td>
</tr>
<tr>
<td>257-268</td>
<td>285 319 343 350</td>
</tr>
</tbody>
</table>

**Fig. 1.** Rose diagram of the directions of 76 turtles.

FCM4DD and the EM are almost the same in many trials. The FCM4DD algorithm maintains its superiority according to the average computation time as in the previous example.

**Example 3.** In this example, we consider a real data set. Table 3 shows the directions of 76 turtles after laying eggs which are given by Stephens (1969).

Fig. 1 shows a rose diagram of the data in Example 3. The rose diagram shows that the data has two cluster centers (one is about 60°, another is about 240°) (Chang-Chien et al., 2012).

The directions of 76 turtles are clustered using the three clustering algorithms. This process is repeated 10 times and the cluster centers are then estimated. Then, we compare the estimated cluster centers with real cluster centers in order to compare the performances of the three clustering algorithms. Their performances are shown in Table 4.

According to MSE and the average computation time, the FCM4DD algorithm gives the best result.

While the directions of the turtles are evaluated according to maximum crisp membership values, we find that, in all trials using the FCD algorithm, 61 turtles move the direction of 60° and that 15 turtles move the direction of 240°. We find that, in all trials using the FCM4DD algorithm, 59 turtles move the direction of 60° and that 17 turtles move the direction of 240°. According to the results of the EM algorithm, we find that, in four trials, 60 turtles move the direction of 60° and that 16 turtles move the direction of 240°, that, in six trials, 63 turtles move the direction of 60° and that 63 turtles move the direction of 240°. According to these results the EM algorithm is not consistent.

**Example 4.** To compare the performances, of the three kinds of clustering algorithms, EM, FCD and FCM4DD, we consider the simulation data from the mixture of von Mises distributions by generating 100 data points from 0.3V/\( \pi /2 \), 25 + 0.3V/\( \pi /2 \), 25 + 0.4V/\( 5\pi /3 \), 25 using Best and Fisher’s simulation method (Best et al., 1979). This process is repeated 1000 times and the means of cluster centers \( (\bar{v}_1; \bar{v}_2; \bar{v}_3) \) for each method are shown in Table 5.

We evaluate the MSE to see the performance for the three algorithms. As a conclusion, the FCM4DD algorithm results in considerably better performance than the other algorithms. According to the average computation time, the FCM4DD algorithm maintains its superiority and the EM algorithm does not. The main reason for this situation is that the EM algorithm is an unstable method.

**Example 5.** To compare the performances of the three kinds of clustering algorithms, EM, FCD and FCM4DD, we consider the simulation data from the mixture of three triangular distributions by generating 150 data points from 0.33A(\( \pi /2 \), \( \pi /2 \), \( 3\pi /2 \)) + 0.33A(\( 5\pi /3 \), \( 2\pi /3 \), \( 5\pi /3 \)) and their means which are (\( \pi /2 \), \( \pi /2 \), \( 5\pi /3 \)). This process is repeated 1000 times and the means of the cluster centers, \( (\bar{v}_1; \bar{v}_2; \bar{v}_3) \), for each method are shown in Table 6.

We observe that the MSE for the FCM4DD and the EM have almost the same results in all trials. According to the average computation time, the EM algorithm results in negative outcomes, while the FCM4DD algorithm maintains its superiority. The main reason for this situation is that the EM algorithm is an unstable method.

**Example 6.** In this example, to show the performance of the FCM4DD algorithm with spherical data, we consider the simula-
Algorithm 13.1

The algorithm is designed to...
Although the EM algorithm delivered an adequate amount of results on the directional data, it was not usually stable. Therefore, the EM algorithm was run several times because of the deadlock. But in this case, each runtime was time-intensive. The FCD algorithm was more consistent than the EM algorithm. Membership values, which are obtained from the FCD algorithm, sometimes did not satisfy necessary conditions. While membership values must be within the interval \([0, 1]\), the membership values obtained by the FCD algorithm were outside of this interval. In the present case, the FCD algorithm had the ability to result in failure within fuzzy systems.

The existing and the proposed clustering algorithms are sensitive to initial values, but the proposed algorithm is less sensitive. However, the main weakness of the proposed method like the EM and the FCD algorithm is that it is necessary to assign a cluster number.

The FCM4DD algorithm was applied to spherical data and the obtained results were very effective. Moreover, the FCM4DD algorithm can be applied to N-dimensional directional data as well as circular and spherical data. However, the results of the FCM4DD algorithm upon for spherical data were not compared with other algorithms in the existing literature. This comparison, along with the comparison of the EM and the FCD algorithms results to the available literature, can be explored in future studies.

During the recent years, many researchers have attempted to predict the future trend of financial time series. Therefore, they have applied the fuzzy rule-based expert system to predict stock price movement. (Shakiri, Fazel Zaranidi, Tarimordi, & Turkseven, 2015). Additionally, usage of fuzzy directional clustering algorithms contributes to the rule-based expert system for measuring the effects of seasonality, as in all-time series. One of the most important points of these algorithms is that the calculation of the membership values is completed properly. Therefore, this paper shows how sensitively accurate this point is.

Some diseases have seasonal features (Money et al., 2003). The effects of seasonal variation should also be considered in the diagnosis of diseases related with seasons in the rule-based expert systems. Since the diseases have periodic variation, the usage of the directional clustering algorithms has become popular for the determination of the seasonal effects. However, fuzzy clustering methods have been developed as a result of stability problems within the crisp methods. Nevertheless, some problems have occurred for calculation of the membership values in this method. During the automatic bombing of unmanned air vehicles, some factors should be considered in order to shoot at the target. The most important factors are average wind direction and gale force. Directional clustering algorithms can be used to determine the average wind direction accurately by eliminating its variation.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.eswa.2016.03.034.

References


