Les Cahiers du GERAD ISSN: 0711–2440

Convex fuzzy k-medoids clustering

D. N. Pinheiro, D. Aloise,
S. J. Blanchard

G–2019–28

April 2019

La publication de ces rapports de recherche est rendue possible grâce au soutien de HEC Montréal, Polytechnique Montréal, Université McGill, Université du Québec à Montréal, ainsi que du Fonds de recherche du Québec – Nature et technologies.

Dépôt légal – Bibliothèque et Archives nationales du Québec, 2019
– Bibliothèque et Archives Canada, 2019

The publication of these research reports is made possible thanks to the support of HEC Montréal, Polytechnique Montréal, McGill University, Université du Québec à Montréal, as well as the Fonds de recherche du Québec – Nature et technologies.

Legal deposit – Bibliothèque et Archives nationales du Québec, 2019
– Library and Archives Canada, 2019

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Téléc. : 514 340-5665
info@gerad.ca
www.gerad.ca

Convex fuzzy k-medoids clustering

Daniel N. Pinheiro [a]
Daniel Aloise [b]
Simon J. Blanchard [c]

[a] *Center of Technology Federal University of Rio Grande do Norte Natal, RN 1524, Brazil*

[b] *GERAD & Department of Computer and Software Engineering, Polytechnique Montréal (Québec) Canada, H3C 3A7*

[c] *McDonough School of Business, Georgetown University, Washington, DC 20057, USA*

daniel.npinheiro@bct.ect.ufrn.br
daniel.aloise@polymtl.ca
sjb247@georgetown.edu

April 2019
Les Cahiers du GERAD
G–2019–28

Abstract: K-medoids clustering is among the most popular methods for cluster analysis, but it carries several assumptions about the nature of the latent clusters. In this paper, we introduce the Convex Fuzzy k-Medoids (CFKM) model, whose underlying formulation not only relaxes the assumption that objects must be assigned entirely to one and only one medoid, but also that medoids must be assigned entirely to one and only one cluster. Moreover, due to its convexity, CFKM resolution is completely robust to initialization. We compare our model with two fuzzy k-medoids clustering models found in the literature: the Fuzzy k-Medoids (FKM) and the Fuzzy Clustering with Multi-Medoids (FMMdd), both solved approximately by heuristics because of their hard computational complexity. Our experiments in synthesized and real-world data sets reveal that our model can uniquely discover important aspects of clustered data which are inherently fuzzy in nature, besides being more robust regarding the hyperparameters of the fuzzy clustering task.

Keywords: Fuzzy clustering, optimization, clustering models, unsupervised learning, convexity

---

Acknowledgments: This study was financed by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001, and by the Natural Sciences and Engineering Research Council of Canada (NSERC) under grant 2017–05617.

1 Introduction

Clustering involves the sorting of objects into groups (i.e., *clusters*), such that objects in the same group are similar in some way. Clustering is ubiquitous, and there exists countless methods to identify clusters and characterize their composition. Describing the heterogeneity in the available methods, [8] presented a survey of the different decisions that a user of clustering methods needs to make: how objects should be compared (e.g., in pairs), how to assess (dis)similarity (e.g., Euclidean distance, Minkowski metric, Mahalanobis distance), and how to represent the identified clusters and their members (e.g., via hierarchical structures like trees, or partitions). Jain et al. also show that such methods widely apply to contexts ranging context image segmentation, object and character recognition, information retrieval and data mining.

Representative-based algorithms are the simplest form of clustering algorithms. They are distance-based methods for which a set of representative objects is identified, such that the assignment of any object to a cluster is obtained by determining which representative it most resembles. This assessment of similarity, necessary to assign objects to clusters, is based on a dissimilarity (e.g., distance) function that is specified by the user. When each cluster's representative object is required to be exactly one of the objects members of the clusters, one can simultaneously identify clusters and its most representative member via the k-medoids clustering model (KM). This model is formulated as follows:

$$\text{minimize} \quad Z_{KM} = \sum_{i=1}^{n}\sum_{j=1}^{n} d_{ij}e_{ij} \tag{1}$$

$$\text{s.t.} \sum_{j=1}^{n} e_{ij} = 1, \ \forall\, i \in \{1,\dots,n\}, \tag{2}$$

$$e_{ij} \leq e_{jj}, \ \forall\, i,j \in \{1,\dots,n\}, \tag{3}$$

$$\sum_{j=1}^{n} e_{jj} = k, \tag{4}$$

$$e_{ij} \in \{0,1\}, \ \forall\, i,j \in \{1,\dots,n\}. \tag{5}$$

The input necessary to use a k-medoids clustering model is a dissimilarity matrix, D, in which each entry d_{ij} contains a measure of dissimilarity between the data objects o_i and o_j. The decision variables e_{ij}, for all $i,j \in \{1,\dots,n\}$, correspond to the assignment of data object o_i to the cluster for which its representative object (i.e., the medoid) is data object o_j. Constraints (2) state that every object must be assigned to one medoid. Constraints (3) guarantee that a data object o_i is assigned to data object o_j only if the later is a medoid. The constraint (4) requires the selection of exactly k medoids, where k is a hyperparameter of the model. Finally, constraints (5) specify the decision variables to be binary.

One of the strongest assumptions in medoid-based clustering models is that objects must belong to one (and only one) cluster. However, [14] showed that this constraint may be artificial—some objects can be best represented by medoids from multiple clusters. Furthering this idea, [16] showed that objects assigned to a cluster are even sometimes better represented by multiple medoids. Both situations could occur depending on the nature of the data. For instance, imagine an individual grouping a set of food objects (e.g., apple, orange, cucumber, tomato, strawberries, cake) according to their own perceptions. Whereas one could believe that a tomato is either a fruit or vegetable (but not both), strawberries could be reasonably seen as both a fruit and a dessert and be medoids for both. Prior models that incorporate fuzziness via allowing only either hard assignments to multiple medoids or soft memberships may fail to capture the essence of the structure when the data contains examples of both. Moreover, fuzzy clustering open venues for more flexible and sophisticated clustering models and algorithms [12].

In the present paper, we propose a convex fuzzy k-medoids clustering method. Its underlying formulation allows medoids to represent multiple clusters. The proposed model possesses several ad-

vantages. First, unlike prior fuzzy k-medoids formulations, it is convex and thus robust to initialization decisions. Second, as shown in our computational experiments, our method is entirely at the pairwise level, and provides users with full information about the relation between each object and its corresponding clusters through its relations with its medoids. Third, our method is shown to be less sensitive to hyperparameters decisions regarding the number of clusters and the fuzziness factor degree. The remainder of the manuscript is organized as follows. In Section 2, we compare prior fuzzy k-medoids methods and formulations. In Section 3, we introduce our convex method (CFKM). In Section 4, we show how the methods perform in synthesized data sets with controlled cluster structures as well as in a real-world data in which medoid assignments are inherently ambiguous. Finally, Section 5 presents our concluding remarks.

2 Existing fuzzy medoid-based clustering models

Among the numerous representative-based clustering algorithms, k-means is widely considered to be the most popular. Its appeal is sensible. From an initial partition of objects into k clusters, the centroid of each cluster is computed as the average of the features of its members. Then, given the current centroids, the algorithm reassigns each object to its closest centroid based on the Euclidean distance metric. In the sequel, the k centroids are then updated, and this procedure is iterated until stability is obtained.

There exists, however, numerous situations when making the centroid to be the average of the cluster member features is not appropriate. For instance, it may be that object features are on different scales or ordinal values, or the data contain outliers. Moreover, the use of an average as a centroid often leads to its use in describing cluster members (e.g., the most representative object), yet the average representation from the centroid may point to a combination of dimensions that does not exist in the data. As a consequence, scholars have also investigated requiring that the cluster representative is an object that exists within each cluster (i.e., the center-most data object of the cluster), called *medoids*, which are determined to be the objects that minimize the sum of distances to all other points in their respective clusters.

There are many advantages on the use of medoids for clustering. First, research on perception suggests that individuals often use real objects to represent subsets of objects [13, 2]. That is, when individuals assign objects to clusters, the central tendency measure of the cluster tends to be the most representative object and not necessarily an average of its members. Second, the possibility of defining a general dissimilarity matrix makes the KM model flexible to use different distance metrics—even non metric ones. Third, medoid-based models offer classification rates comparable to those of the k-means model for metric data. For a given data set under the same metric (i.e. squared Euclidean distances), the optimal KM solution value is at most twice that of the k-means optimal solution [see e.g., 18]. Finally, medoids-based models tend to be more robust to outliers as approximate clusters with the median object as opposed to the average object [10, 11]. The KM model is NP-hard [9]. Algorithms for finding the optimal solution of the problem for large data sets are presented by [1] and [6], while efficient heuristics are presented by [7] and [19].

2.1 Fuzzy k-medoids (FKM)

The fuzzy k-medoids method (FKM) can be attributed to [14]. FKM optimizes the underlying mathematical optimization model expressed as follows:

$$\text{minimize} \quad Z_{FKM} = \sum_{i=1}^{n}\sum_{j=1}^{n} d_{ij}(e_{ij})^h \tag{6}$$

$$\text{s.t.} \sum_{j=1}^{n} e_{ij} = 1, \ \forall\, i \in \{1,\ldots,n\}, \tag{7}$$

$$e_{ij} \leq e_{jj}, \ \forall\, i,j \in \{1,\ldots,n\}, \tag{8}$$

$$\sum_{j=1}^{n} e_{jj} = k, \tag{9}$$

$$e_{ij} \in [0,1],\ \forall\ i,j \in \{1,\ldots,n\}, i \neq j, \tag{10}$$

$$e_{jj} \in \{0,1\},\ \forall\ j \in \{1,\ldots,n\}, \tag{11}$$

where e_{ij} is the degree of membership of o_i to the cluster whose medoid is o_j ($e_{ij} = 0$ if o_j is not a medoid), and h is the fuzziness factor. The fuzziness factor is a hyperparameter that indicates the desirable level of overlapping between the clusters to be found, or, in other words, how much the degrees of membership will be spread among the clusters. As $h \to 1^+$, the objects tend to be assigned to a single cluster, creating crisp partitions. As $h \to \infty$, the objects tend to be equally spread among clusters. That is, for each non-medoid o_i and each medoid o_j, the degree of membership e_{ij} tends to $1/k$.

[14] present their algorithm, which performs successive swaps of a medoid object with a non-medoid object in order to approach the optimal solution of (6)-(11).

Given a known set of medoids, the degrees of membership of each object o_i to a chosen medoid o_j can be found by computing the following expression:

$$e_{ij} = 1 \Bigg/ \sum_{t|e_{tt}=1} \left(\frac{d_{ij}}{d_{it}}\right)^{1/(h-1)}. \tag{12}$$

Other algorithms exist for optimizing (6)—(11) when tackling large data sets. [14] propose the *linearized fuzzy c-medoids* (LFCMdd) and the *robust fuzzy c-medoids* (RFCMdd) algorithms. LFCMdd examines only a subset of objects while updating the medoid for each cluster. This subset is chosen to be the set of objects that possesses the highest degree of membership to that cluster. RFCMdd samples the set of objects in order to ignore outliers in the clustering process. Given a set of medoids, the method computes the cost associated to each object o_i, which is $\sum_{j=1}^{n} d_{ij}(e_{ij})^h$, where e_{ij} is given by Equation (12). At each iteration, RFCMdd ignores the set of objects with highest associated costs when computing the new medoids. [15] and [22] proposed fuzzy medoid-based clustering methods for large data sets, where the data is processed in mini-batch mode. [15] proposes two methods. In the first one, *Online Fuzzy C-Medoids* (OFCMD), each mini-batch is processed separately and an additional step is required to combine the medoids of each mini-batch. In the second one, *History based Online Fuzzy C-Medoids* (HOFCMD), the identified medoids obtained with the previous mini-batch are combined with the current mini-batch as history information and the final set of medoids is given when processing the last data mini-batch. [22] proposed the *Incremental Multiple Medoids based Fuzzy Clustering* (IMMFC) method, which is another variation for large data sets. In this method, multiple weighted medoids are defined for each cluster in each mini-batch. Like in OFCMD, each mini-batch is processed separately and the final set of medoids (one per cluster) can be obtained from the medoids obtained from processing each mini-batch.

The analysis of the scalability of these algorithms for optimizing the FKM model (6)—(11) for large clustering problems is indeed an interesting and relevant topic. However, it is out of the scope of the present paper which is focused on the interpretability of different fuzzy clustering models.

2.2 Fuzzy Clustering with Multi-Medoids (FMMdd)

Recognizing that it may not be sufficient to use only a single medoid to represent a cluster, [16] proposed the Fuzzy Clustering with Multi-Medoids (FMMdd) model. In FMMdd, each cluster may be represented by multiple objects. It is formulated as follows:

$$\text{minimize}\quad Z_{FMMdd} = \sum_{c=1}^{k}\sum_{i=1}^{n}\sum_{j=1}^{n}(e_{ic})^h (v_{jc})^g d_{ij} \tag{13}$$

$$\text{s.t.}\ \sum_{c=1}^{k} e_{ic} = 1,\ \forall\ i \in \{1,\ldots,n\}, \tag{14}$$

$$\sum_{j=1}^{n} v_{jc} = 1, \ \forall\, c \in \{1, \ldots, k\}, \tag{15}$$

$$e_{ic} \in [0, 1], \ \forall\, i \in \{1, \ldots, n\}, \ c \in \{1, \ldots, k\}, \tag{16}$$

$$v_{jc} \in [0, 1], \ \forall\, j \in \{1, \ldots, n\}, \ c \in \{1, \ldots, k\}, \tag{17}$$

where e_{ic} is the degree of membership of object o_i in cluster c, and v_{jc} denotes how representative object o_j is to cluster c. Hyperparameters h and g are the degree of fuzziness of memberships and the level of smoothness of the representativeness of objects in the clusters, respectively. The fuzzy membership allows objects to belong to multiple clusters (up to k), and the representation weights allow each cluster to be represented by multiple objects (up to n). Constraints (14) state that the sum of the membership degrees of each object must be equal to 1, and constraints (15) impose that the sum of representation weights in each cluster must sum exactly one.

[16] also introduce an iterative heuristic for the FMMdd problem based on the Lagrangean method. Assuming that the representativeness of objects in each cluster is known, the first-order optimality conditions assure that optimal degrees of membership are given by

$$e_{ic} = \frac{\left[\sum_{j=1}^{n} (v_{jc})^g d_{ij}\right]^{-1/(h-1)}}{\sum_{f=1}^{k} \left[\sum_{j=1}^{n} (v_{jf})^g d_{ij}\right]^{-1/(h-1)}}, \tag{18}$$

for all $i \in \{1, \ldots, n\}$ and $c \in \{1, \ldots, k\}$. Similarly, assuming that the degrees of membership are know, the optimal representativeness are given by

$$v_{jc} = \frac{\left[\sum_{i=1}^{n} (e_{ic})^h d_{ij}\right]^{-1/(g-1)}}{\sum_{t=1}^{n} \left[\sum_{i=1}^{n} (e_{ic})^h d_{it}\right]^{-1/(g-1)}}, \tag{19}$$

for all $j \in \{1, \ldots, n\}$ and $c \in \{1, \ldots, k\}$. The proposed heuristic starts with a non-negative initialization, and updates alternately e_{ic} and v_{jc} ($i, j \in \{1, \ldots, n\}$, $c \in \{1, \ldots, k\}$) with Equations (18) and (19) until convergence or a maximum number of iterations is reached. During the process, the objective function (13) is successively improved through reassignment of objects to clusters and selection of representative objects for each cluster.

As a local descent method, the quality of the final solution obtained by the heuristic of Mei and Chen for the FMMdd problem depends on its initialization. In Appendix A, we prove that a common initialization with $v_{jc} = 1/n$ and $e_{ic} = 1/k$ for all $i, j \in \{1, \ldots, n\}$ and $c \in \{1, \ldots, k\}$ leads to a poor local minimum, in which all objects are equally spread among the clusters and the representativeness of each object is equally distributed among the clusters.

3 Convex Fuzzy k-Medoids (CFKM)

Insofar, we presented two fuzzy k-medoids clustering models from which several algorithms in the literature were proposed. In FKM a fuzziness factor is used to allow the assignment of an object to multiple clusters. FMMdd was introduced for cases when each cluster could be represented by multiple medoids. These two formulations, and associated algorithms, are suitable formulations yet their usage is limited because they are non-convex optimization problems with multiple local minima of unknown quality. Further, heuristics for FKM and FMMdd are generally sensitive to initialization.

In the present section, we introduce a new model for fuzzy k-medoids clustering, namely Convex Fuzzy k-Medoids (CFKM), formulated as follows:

$$\text{minimize} \quad Z_{CFKM} = \sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij} (e_{ij})^h \tag{20}$$

$$\text{s.t.} \sum_{j=1}^{n} e_{ij} = 1, \ \forall\, i \in \{1, \ldots, n\}, \tag{21}$$

$$e_{ij} \leq e_{jj}, \ \forall\, i, j \in \{1, \ldots, n\}, \tag{22}$$

$$\sum_{j=1}^{n} e_{jj} = k, \tag{23}$$

$$e_{ij} \in [0, 1], \ \forall\, i, j \in \{1, \ldots, n\}. \tag{24}$$

In FKM, only the variables that define the medoids ($e_{jj}, \ j = 1, \ldots, n$) are binary while the other variables (e_{ij},with $i \neq j$) are real-valued. In CFKM, the variables ($e_{jj}, \ j = 1, \ldots, n$) that define the medoids are also fuzzy, such that an object can simultaneously be a medoid for different clusters—with varying degrees of importance (in the interval $[0, 1]$).

Convexity is an important feature in optimization. It allows the solution of optimization problems by means of local minimizers, since local optimality implies optimality for convex problems [4]. We claim that convexity is also desirable for clustering. As a unsupervised machine learning technique, clustering results must be interpreted in order to extract knowledge about the explored data. Such interpretation is indeed compromised if the solution at hand is far from the optimum one, with different assignments and clusters composition. Thus, convex clustering models provide a guarantee that the sole optimal solution obtained is the best and the right one for posterior clustering analysis. In the following theorem, we show that CFKM is a convex optimization problem.

Theorem 1 *CFKM is a convex optimization problem.*

Proof. An optimization problem is said to be convex if the objective function is convex in the set of feasible solutions and if the set of feasible solutions is also convex. If the set of feasible solutions is delimited by convex functions, then it is a convex set. The set of feasible solutions of CFKM is delimited by linear functions, which are convex functions.

The objective function is a sum of functions of $f(z) = az^b$, where $a = d_{ij} \geq 0$, $b = h \geq 1$ and $z = e_{ij} \in [0, 1]$. For such functions, the second derivative of $f(z)$ is

$$\frac{d^2 f(z)}{dz^2} = b(b-1)az^{(b-2)}, \tag{25}$$

which is greater or equal to zero if $a \geq 0$, $z \geq 0$ and $b \geq 1$. Thus, the objective function of CFKM is a sum of convex functions, and hence, a convex function itself. As CFKM has a convex objective function and the set of feasible solutions is also convex, then it is a convex optimization problem. □

4 Computational experiments

In this section, we present a series of experiments with synthesized and real-world data comparing the solutions of the three fuzzy clustering models presented here. Solutions for FKM and FMMdd are the best ones obtained by the associated heuristics executed 100 times. Their codes are implemented in Matlab. CFKM is directly solved by function `fmincon` of Matlab version 2019a. All experiments were performed on a AMD A8-5500B processor running at 3.2 GHz using 16 GB of RAM, running Ubuntu 18.04.2 LTS x86_64.

Synthesized instances were generated with the R package MixSim [17]. MixSim can allow one to generate cluster solutions which vary the average and maximum overlap between clusters, specify if the clusters should have spherical covariance matrices. For our experiments, we generated two-dimensional data sets with $n = 100$ points by setting the number of clusters (k), the average overlap between clusters (ω) and whether the clusters are to be spherical.

4.1 Simulation #1 well-separated clusters

For the simulation, we generated a data set of four well-separated Gaussian clusters with means $(-1,-1)$, $(-1,1)$, $(1,-1)$ and $(1,1)$, each with 25 points within a 0.2 standard deviation from the corresponding means. Figure 1 shows the partitions obtained by the FKM and CFKM for $k = 4$ and $h = 1$.

FMMdd solutions cannot be obtained by the heuristic of [16] because (18) and (19) are not defined for $h = g = 1$. In Figures 1(b) and (c), the lines represent the assignments (variables e) and note that both FKM and CFKM perfectly recover the structure.

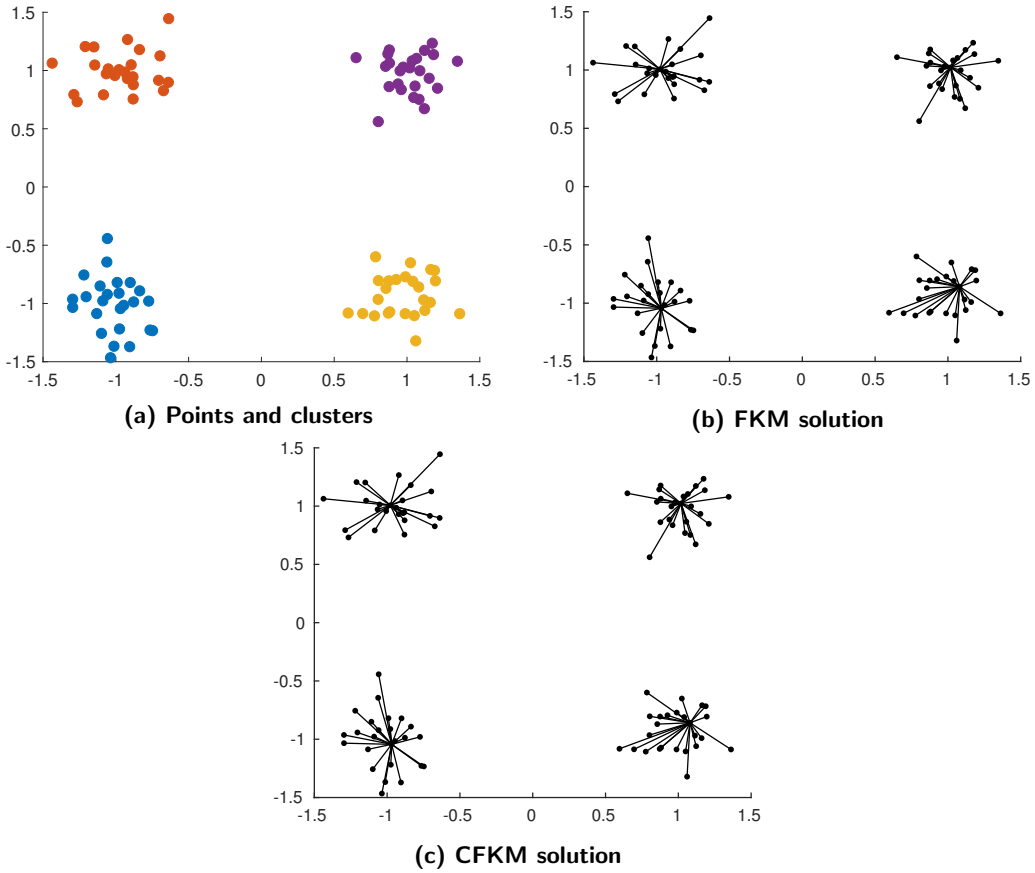


Figure 1: Data points and solutions with $k = 4$ clusters and $h = 1$

In turn, Figure 2 presents the solutions obtained when $h = g = 1.5$. We note that for CFKM, the darkness of the lines indicates the strength of assignment (e_{ij}) for $i, j = 1, \ldots, n$. For FMMdd's solutions, the darkness of the line intensity is proportional to $\sum_{c=1}^{k} e_{ic} v_{jc}$. Moreover, hereafter, we do not plot assignments between two points when they are very weak in value. Specifically, we do not plot assignments between two points if they are 20 times smaller than the maximum observed assignment value.

As when $h = g = 1$, all methods are able to recover the clusters. However, the results illustrated in Figures 2(c) and (d) show that FMMdd and CFKM make use of multiple medoids since we observe that several points are associated to more than a medoid in its own cluster.

In Figure 3, we show how increasing the fuzziness hyperparameters eventually leads to inter-cluster assignments. Specifically, increasing the fuzziness makes all models expand cluster assignments to multiple medoids for the points that are the closest to other clusters.

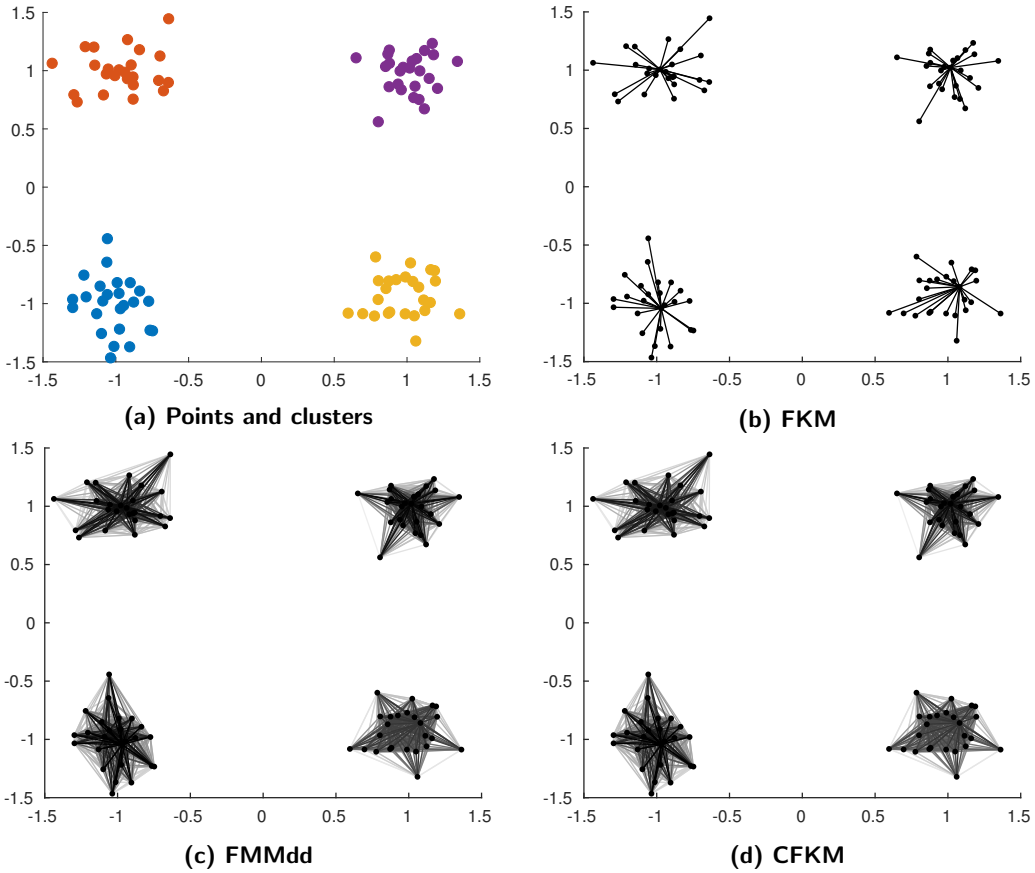


Figure 2: Data points and solutions with $k = 4$ clusters and $h = g = 1.5$

Finally, we note that the prior solutions assumed that the true value of k was done. Figure 4 shows varying k affects the solutions when $h = 1.1$—typically, users of clustering models do not know the true number of clusters present in the data. We can see in the figure that CFKM can help a user determine the most likely value of k. Specifically, when k is less than 4 (the true number of clusters), solutions for both methods appear to have inter-cluster assignments (i.e., assignments of points to clusters seemingly far away). However, for CFKM we no longer see any inter-cluster assignment when $k \geq 4$ and CFKM solutions are quite similar from $k = 4$ to $k = 6$ suggesting that increasing k beyond 4 is unnecessary. In contrast, FKM is obliged to create new medoids for $k > 4$, thus splitting natural clusters.

4.2 Simulation #2 overlapping clusters

The data generating process for the second simulation involved four spherical clusters with a smaller overlap (average overlap $\omega = 0.02$). The solutions obtained for FKM and CFKM for $h = 1$ are shown in Figure 5. Both algorithms obtain the same partitions, where each point is assigned to one and only one cluster. In this case, both CFKM and FKM misclassify a few points due to the error added.

Figure 6 presents the assignment for $h = g = 1.5$, while Figure 7 shows the solutions for $h = g = 2.0$. We note that as h and g increase, FMMdd loses the capacity to separate the yellow and red clusters. When $h = g = 2$, FMMdd seems not to group the yellow cluster as a separated one in the fuzzy partition. We observed something similar in Figure 8 when we looked at non-spherical clusters (average overlap $\omega = 0.02$) for $h = g = 2$. Specifically, FMMdd groups the data into only three clusters. Across our experiments, it does appear as though FMMdd is more sensitive to the fuzziness hyperparameters than CFKM.

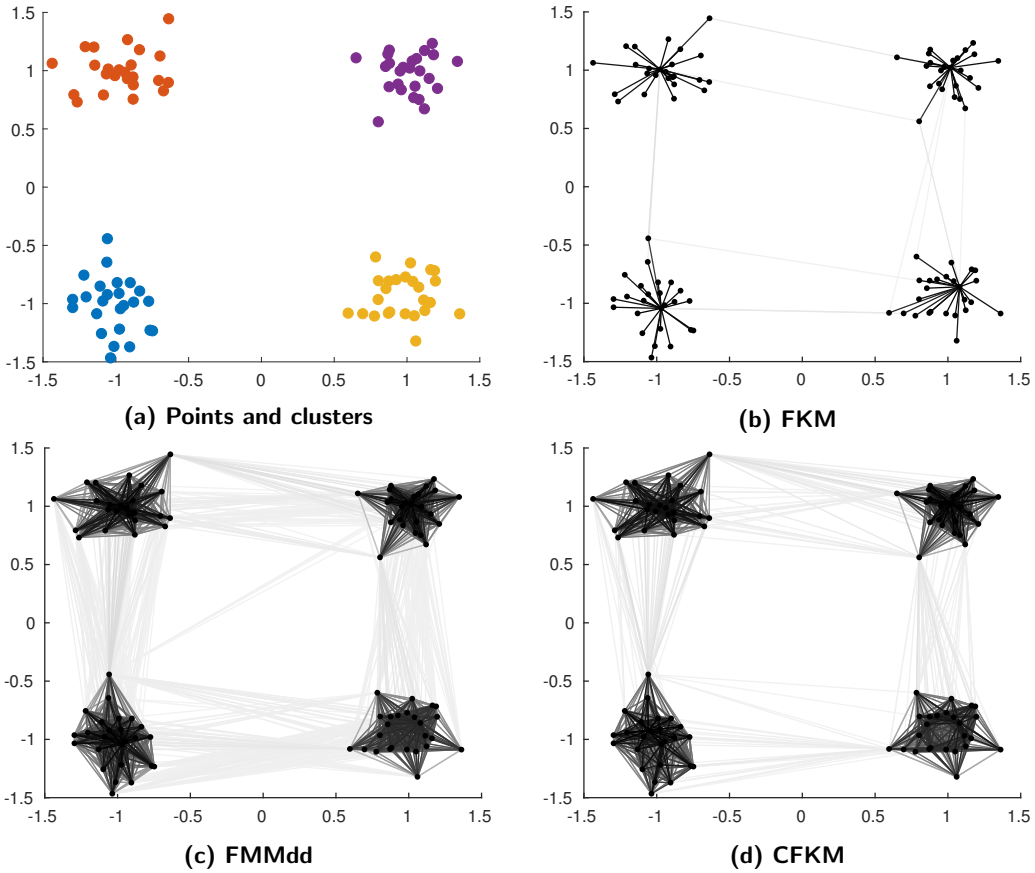


Figure 3: Data points and solutions with $k = 4$ clusters and $h = g = 2$

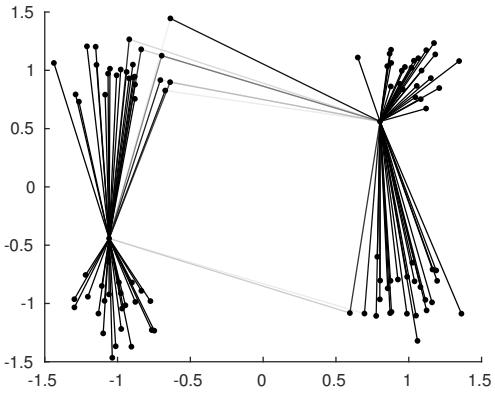
4.3 Empirical illustration 1 - Food object perceptions

To illustrate the implications and differences between the three fuzzy k-medoids models (FKM, FMMdd, and CFKM), and the impact of initialization, we used data by [3] on individuals' perceptions of food objects. The data consists of similarity judgements based on 46 individuals who performed a sorting task.

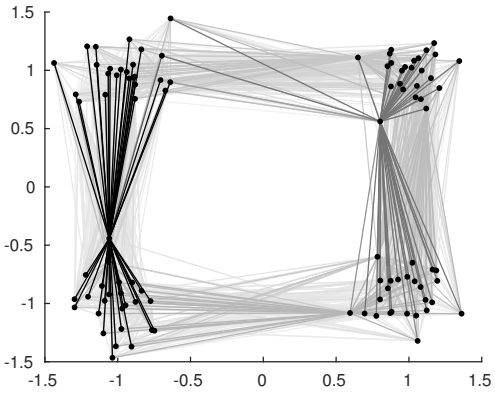
In a sorting task, participants organize objects into piles such that objects in the same pile are similar in some way [5]. Participants could create as many piles (i.e., partitions) as they would like, on the basis of their own perceptions. In this particular instance, the objects were 60 food objects[1] sorted by individuals. The aggregated output from the data is a pairwise count matrix of the number of individuals who put each pair of objects in the same piles—a matrix which in turn can be converted into a pairwise distance matrix.

The data is well suited to study the fuzzy k-medoids models and algorithms for multiple reasons. First, the data is heterogeneous such that the pairwise distances used as input data may reflect different mixed perceptions and thus reflect "fuzzy perceptions" [2, 21]. For instance, we may observe a moderate distance between objects *chocolate milk* and *water* because although some individuals tend to put *chocolate milk* in a *drinks* pile (to which *water* belongs), others tend to put it in a *dairy* pile (to which *water* does not belong). Likewise, whereas some individuals may have grouped all fruits and vegetables into a single *produce* pile, others may have separated them into two. Second, several objects may be

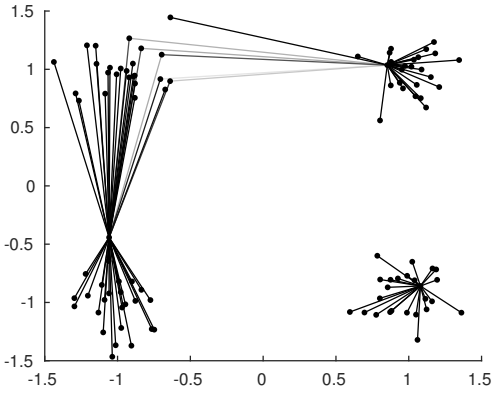
[1]The 60 food objects were: Apple, Bagel, Banana, Bread, Broccoli, Butter, Cake, Carrots, Cereal, Cheese, Chocolate bar, Chicken, Coffee, Cookies, Corn, Crackers, Cucumber, Cupcakes, Dinner roll, Doughnuts, Egg, Granola bar, Chocolate milk, Gummy bears, Hamburger, Ice cream, Kiwi, Lettuce, Lobster, Margarine, Milk, Muffin, Nuts, Oatmeal, Onions, Orange, Pancakes, Pie, Pineapple, Pizza, Popcorn, Popsicles, Pork, Potato chips, Potatoes, Pretzels, Rice, Salmon, Shrimp, Soda, Spaghetti, Steak, Strawberry, Tea, Tofu, Tomato, Waffle, Water, Watermelon, and Yogurt.



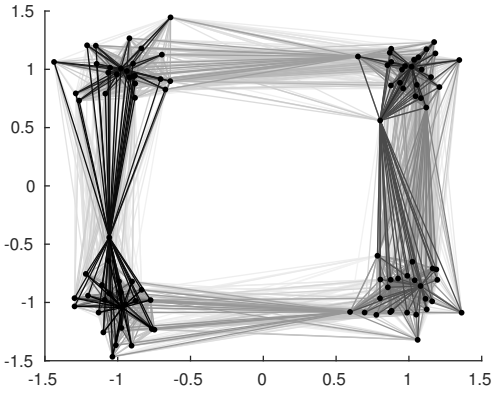
(a) FKM ($k = 2$)



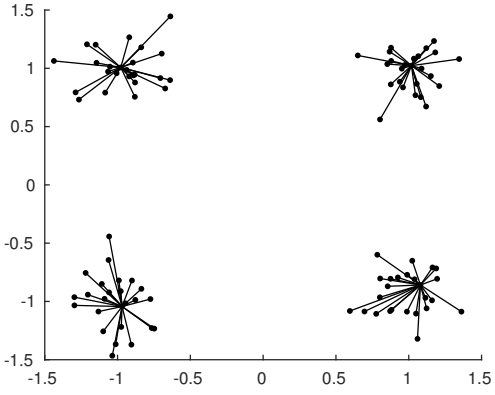
(b) CFKM ($k = 2$)



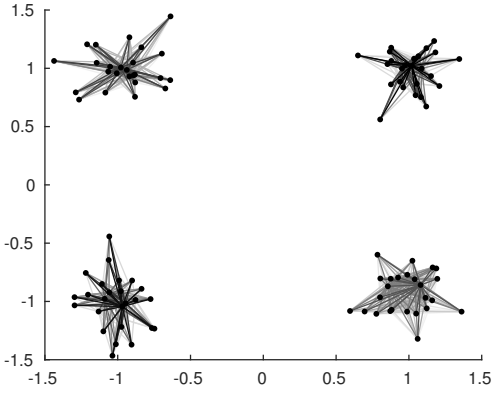
(c) FKM ($k = 3$)



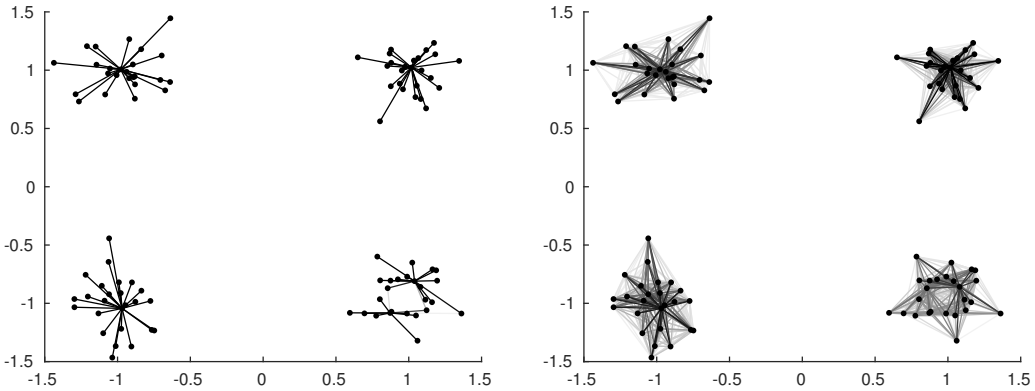
(d) CFKM ($k = 3$)



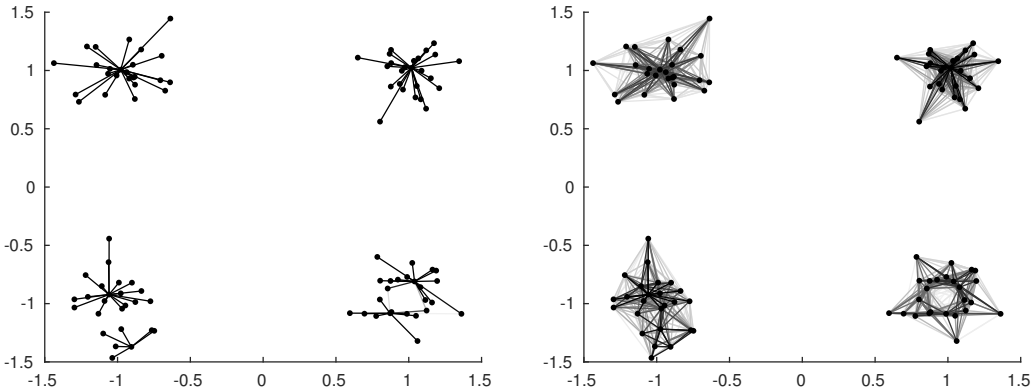
(e) FKM ($k = 4$)



(f) CFKM ($k = 4$)



(g) FKM ($k = 5$) (h) CFKM ($k = 5$)



(i) FKM ($k = 6$) (j) CFKM ($k = 6$)

Figure 4: FKM and CFKM solutions with k varying from 2 to 6 with $h = 1.1$

equally representative of the piles such that some objects may be equally good medoids. For instance, if we observe that most individuals make a pile that combines all vegetables and that all individuals are in agreement into which objects belong to the *vegetable* pile, then all objects would be equally suited to be a medoid. In such cases, the determination of which object is to be assigned the label of a medoid is meaningless. Third, there exists extensive psychological research to aid our interpretation of the proposed solutions. For instance, when [20] asked 1409 individuals to write down categories for a set of food objects similar to the one we used, the participants' list included 312 different partitions. They found that they spanned various types which included taxonomic (i.e., based on properties of objects; e.g., beverages, breads and grains, fruits, meats) scripts (i.e., situations in which the foods are eaten; e.g., breakfast, dinner, appetizers) and macronutrients (e.g., proteins, carbs, and fats) and provide ample evidence for disagreement and differences which should emphasize for the possibility of fuzzy membership of objects to clusters.

Table 1 presents FKM results obtained by 100 executions of the heuristic of [14] for varied values of hyperparameters k and h. Column *best* refers to the best solution value obtained, avg to the average solution value, s to its standard deviation and % to the percentage of executions that converged to the best optimum found. In FKM, we note that it is possible (and even likely) that in the best solution found some objects are equal medoids. That is, multiple assignments of objects to clusters can produce the same optimal solution value. We classified those has having reached the best solution, even though the clustering assignments are different.

Table 2 reports the same metrics for the solutions obtained in 100 executions of the heuristic of [16] using different values of k, h and g. For FMMdd, its heuristic is an iterative method that finishes when

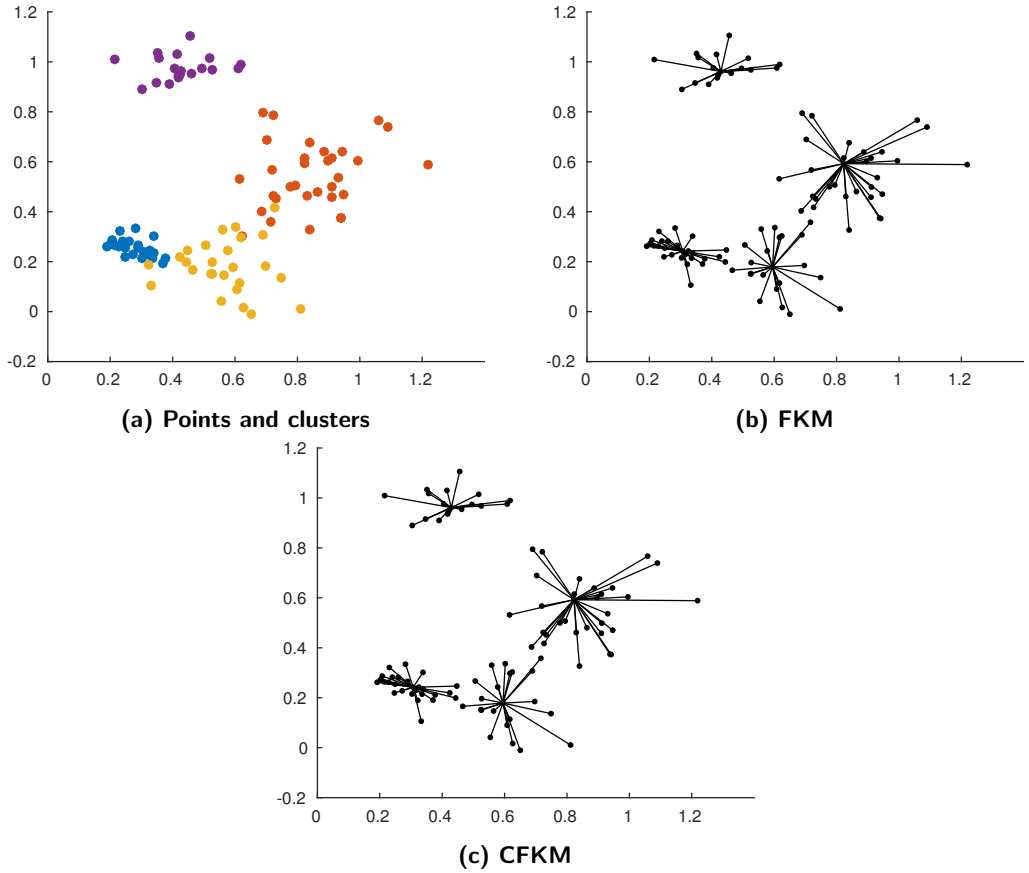


Figure 5: Data points and solutions with $k = 4$ spherical overlapping clusters ($\omega = 0.02$) and $h = 1$

the current solution is too similar to the previous one. As there are no discrete variables, column % refers to the number of solutions whose cost differs by up to 10^{-6} from the best solution.

Finally, with respect to CFKM, the problem is convex and as such it can be solved to the global optimum in one execution. Table 3 reports the optimal solution value and the objective function improvement (in absolute and percentage terms) to aid in determining the value of k.

4.3.1 Performance and stability

Given that two problems being compared have different objective functions, and given that there is no ground-truth to the assignment of food objects to clusters, we cannot directly compare the solutions in terms of recovery. We can, however, compare the stability of the solutions.

First, we observe that all formulations are equally consistent when $k = 1$ and $h = 1$. However, we can observe that when either k or h increased, the percentage of executions that reached the best solution found diminishes. For instance for FKM, even when the solution is not fuzzy (i.e., $h = 1$), only approximately 20%-30% of the executions lead to the best solution found when k was greater than four. As h and k both increase, for example to when $h = 1.5$ and $k = 10$, 99 executions was not sufficient to replicate the best solution found. Such convergence results suggest that a large number of executions is necessary, particularly when $k > 4$ and $h > 1$. FMMdd suffers from similar problem, with few executions converging to the same objective function value as k increases, however the solutions found appear to be having similar deviation from the best optimum. In contrast, CFKM only needs one execution to achieve the global optima. It is thus stable and converges regardless of the value of k or h.

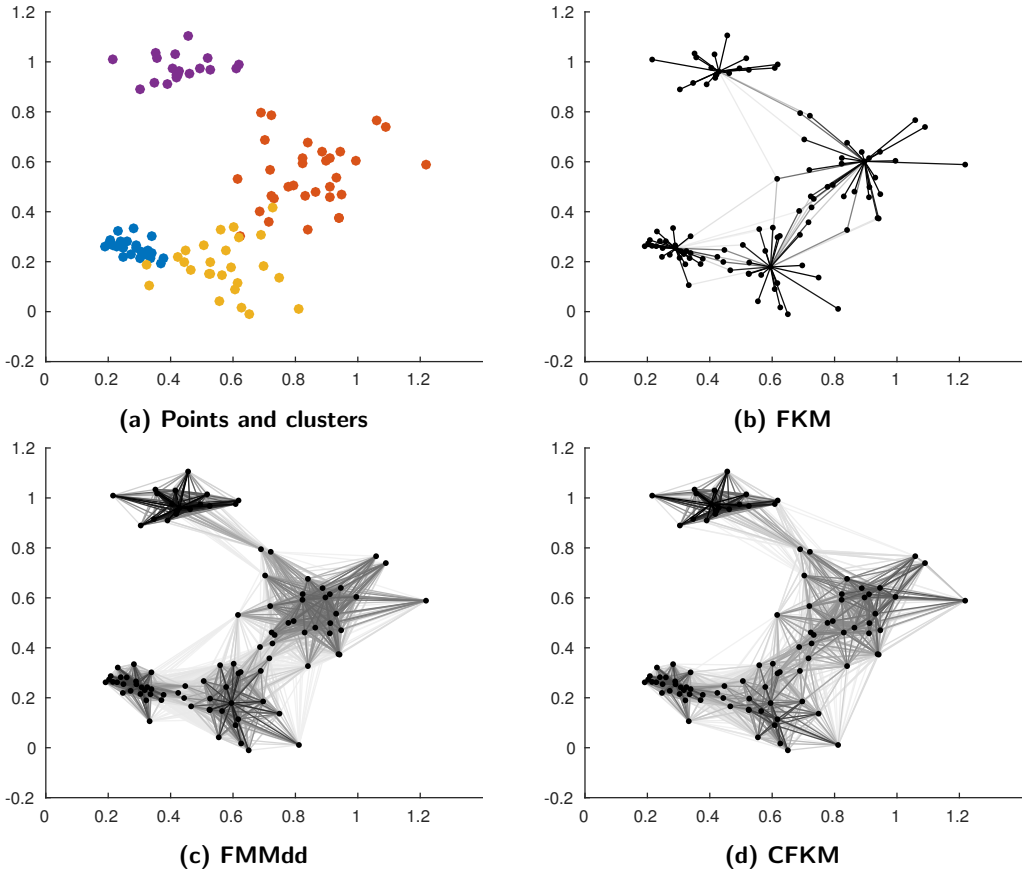


Figure 6: Data points and solutions with $k = 4$ spherical overlapping clusters ($\omega = 0.02$) and $h = g = 1.5$

Table 1: FKM model: Results for 100 executions of the heuristic of [14]

	$h = 1$				$h = 1.1$			
k	*best*	$\bar{x}$	s	%	*best*	$\bar{x}$	s	%
1	12.556	12.556	0.000	100	12.556	12.556	0.000	100
2	8.754	8.754	0.000	100	8.730	8.735	0.007	68
3	5.909	5.912	0.017	96	5.902	5.903	0.000	97
4	4.294	4.489	0.289	42	4.281	4.468	0.287	48
5	3.652	3.719	0.104	24	3.639	3.707	0.099	22
6	3.337	3.375	0.029	20	3.315	3.353	0.028	18
7	3.064	3.099	0.041	26	3.034	3.064	0.026	23
8	2.897	2.940	0.040	27	2.882	2.917	0.028	26
9	2.778	2.818	0.037	14	2.757	2.799	0.033	15
10	2.670	2.715	0.036	6	2.657	2.695	0.027	2
	$h = 1.5$				$h = 2$			
k	*best*	$\bar{x}$	s	%	*best*	$\bar{x}$	s	%
1	12.556	12.556	0.000	100	12.556	12.556	0.000	100
2	7.956	8.008	0.168	91	6.084	6.159	0.178	69
3	5.231	5.232	0.008	99	3.792	3.831	0.065	57
4	3.984	4.046	0.094	34	2.698	2.717	0.017	37
5	3.251	3.300	0.062	12	2.081	2.104	0.016	18
6	2.899	2.925	0.026	9	1.695	1.712	0.016	8
7	2.569	2.610	0.029	10	1.406	1.416	0.009	16
8	2.319	2.334	0.018	15	1.206	1.214	0.006	9
9	2.143	2.166	0.017	4	1.049	1.057	0.007	19
10	2.007	2.024	0.013	1	0.925	0.934	0.006	3

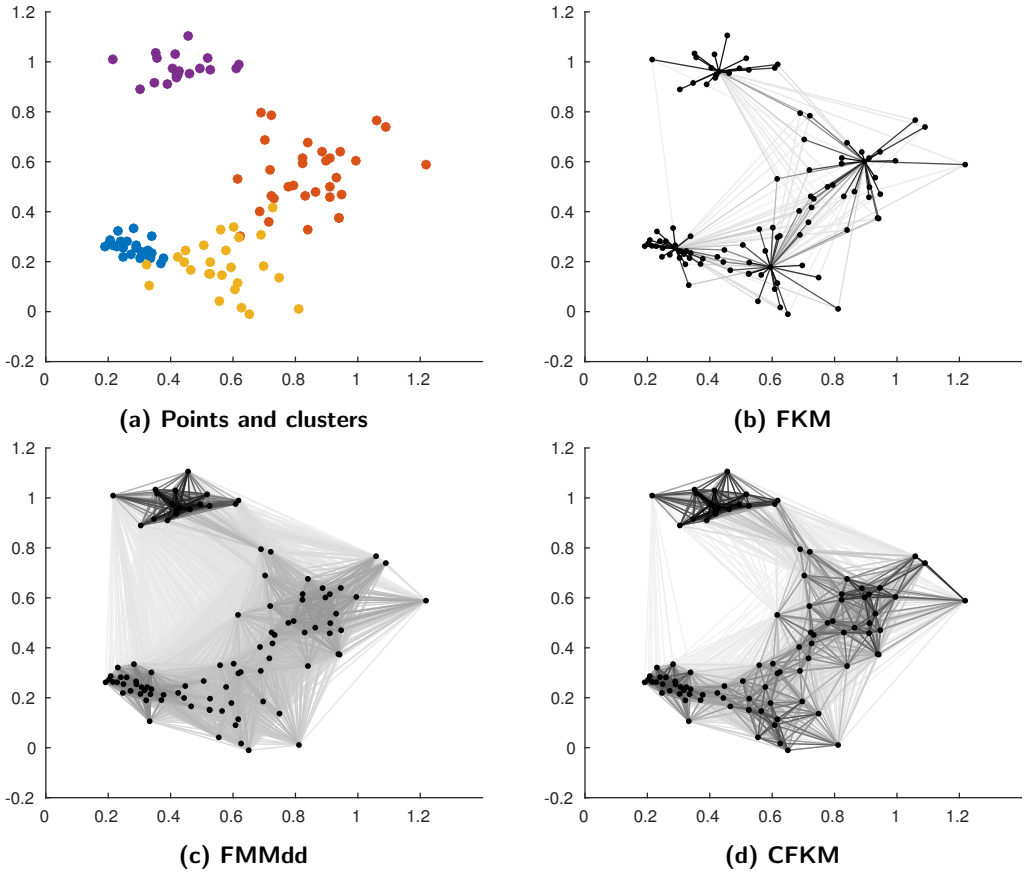


Figure 7: Data points and solutions with $k = 4$ spherical overlapping clusters ($\omega = 0.02$) and $h = g = 2$

Table 2: FMMdd Problem: Results for 100 executions of the heuristic of [16]

	$h = g = 1.1$				$h = g = 1.5$				$h = g = 2$			
k	*best*	$\bar{x}$	s	%	*best*	$\bar{x}$	s	%	*best*	$\bar{x}$	s	%
1	12.286	12.286	0.000	100	3.222	3.222	0.000	100	0.430	0.430	0.000	100
2	7.831	8.018	0.637	92	1.902	1.902	0.000	100	0.213	0.213	0.000	100
3	5.173	5.339	0.554	90	1.423	1.427	0.022	97	0.142	0.142	0.000	100
4	3.764	4.384	0.640	39	1.199	1.213	0.015	45	0.106	0.106	0.000	99
5	3.132	3.660	0.616	43	1.014	1.042	0.044	71	0.085	0.085	0.000	89
6	2.838	3.111	0.361	13	0.920	0.933	0.019	20	0.071	0.071	0.000	100
7	2.625	2.951	0.328	10	0.848	0.858	0.014	16	0.061	0.061	0.000	33
8	2.500	2.792	0.200	2	0.779	0.800	0.014	24	0.053	0.053	0.000	30
9	2.445	2.673	0.191	1	0.733	0.753	0.017	19	0.047	0.047	0.000	50
10	2.348	2.575	0.204	1	0.696	0.713	0.018	9	0.043	0.043	0.000	31

4.3.2 Illustrative comparisons for k = 5

We also wish to determine if the fuzzy k-medoids models can be used to represent a set of partitions in a manner that is sensible and consistent with the theories surrounding the objects used. As is commonly assumed in clustering, we use a scree plot (i.e., "elbow" in the curve strategy) to find the value of k that leads to the greatest improvement over the solution with $k - 1$ medoids. For CFKM, by looking at Table 3, one can easily see that the improvement by incorporating additional clusters diminishes when $k > 5$. For instance, when $h = 1$ and when $h = 1.1$, moving from 5 to 6 clusters dropped the improvement to the prior solution from double digits (e.g., 14.9%) to single digits (e.g., 8.6%). It also appears that for FKM and FMMdd, solutions for which $k > 5$ were considerably less stable. As such, in this section, we compare the solutions for $k = 5$. We begin with $h = 1$, contrasting our interpretation

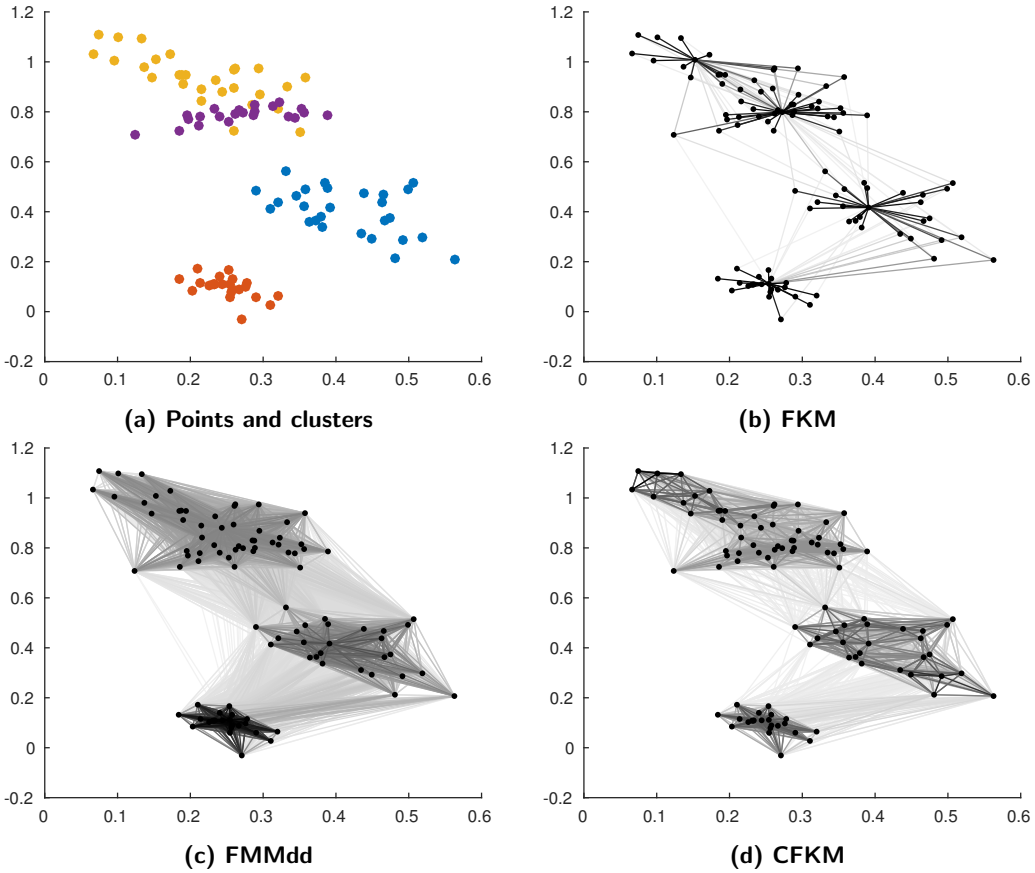


Figure 8: Data points and solutions with $k = 4$ non-spherical overlapping clusters ($\omega = 0.02$) and $h = g = 2$

Table 3: CFKM results (only one execution)

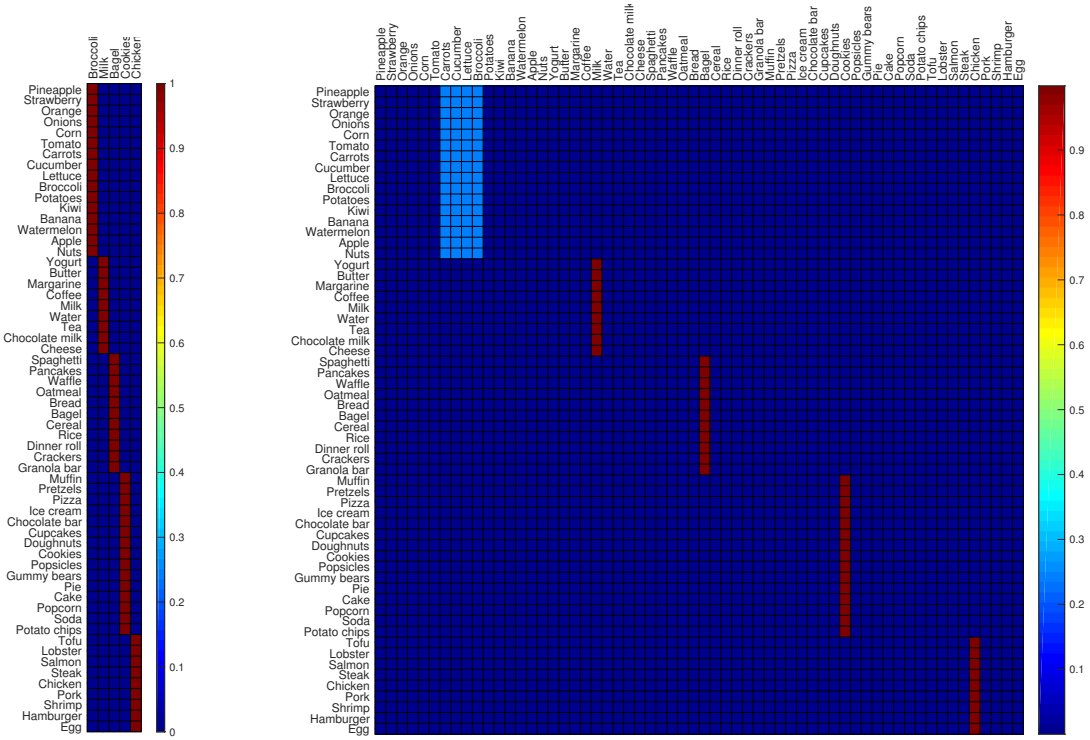
	$h = 1$			$h = 1.1$		
k	*opt*	*improv.*	*%improv.*	*opt*	*improv.*	*%improv.*
1	12.556	-	-	12.344	-	-
2	8.754	3.803	30.3	7.056	5.288	42.8
3	5.909	2.845	32.5	5.340	1.717	24.3
4	4.294	1.615	27.3	3.634	1.706	32.0
5	3.652	0.642	14.9	3.061	0.572	15.7
6	3.337	0.315	8.6	2.773	0.289	9.4
7	3.064	0.274	8.2	2.583	0.190	6.9
8	2.897	0.167	5.4	2.465	0.118	4.6
9	2.778	0.119	4.1	2.461	0.004	0.2
10	2.670	0.108	3.9	2.368	0.092	3.8
	$h = 1.5$			$h = 2$		
k	*opt*	*improv.*	*%improv.*	*opt*	*improv.*	*%improv.*
1	3.222	-	-	0.430	-	-
2	1.615	1.607	49.9	0.221	0.209	48.6
3	1.215	0.400	24.8	0.186	0.035	15.9
4	1.039	0.175	14.4	0.176	0.010	5.2
5	0.954	0.085	8.2	0.167	0.009	5.1
6	0.910	0.044	4.6	0.160	0.007	4.3
7	0.882	0.028	3.1	0.153	0.006	4.0
8	0.844	0.037	4.2	0.147	0.006	4.0
9	0.814	0.030	3.6	0.142	0.006	3.9
10	0.786	0.029	3.5	0.136	0.006	4.0

of how the solutions (in)correctly reflect our understanding of how individuals think about the food objects, and eventually show how the solutions differ when h increases.

4.3.3 h = 1: crisp (non-fuzzy) solutions

Figure 9 presents a graphical summary of the solutions obtained by both FKM (left) and CFKM (right) for $h = 1$ where a value of e_{ij} close to 1 indicates a high degree of assignment of the object o_i to medoid o_j. The degree of assignment of objects to clusters is represented by means of a RGB scale indicated in the right of each figure, where colors closer to dark red indicates higher degrees of membership.

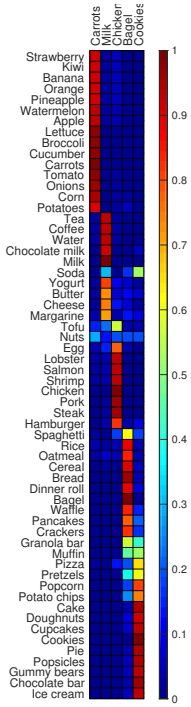
The solutions are fairly consistent. For FKM, the five clusters can be interpreted as follows: Produce (medoid: broccoli), Dairies and Liquids (medoid: milk), Grains (medoid: bagel), Carbs (medoid: cookies), and main dishes (medoid: chicken). There is little information as to which objects are better medoids. Concerning CFKM, we do not have much additional information other than it seems like objects of the produce cluster are not fit well by the data. Most objects have low representativeness (i.e., value of e_{ij} near 0). However, it still provides important information. Although in FKM broccoli was choice as medoid, the solution from CFKM reveals that any of carrots, cucumber or broccoli could have been medoids. As a consequence, an expert relying on the FKM solution is misled when considering broccoli to be uniquely suited to be a medoid.



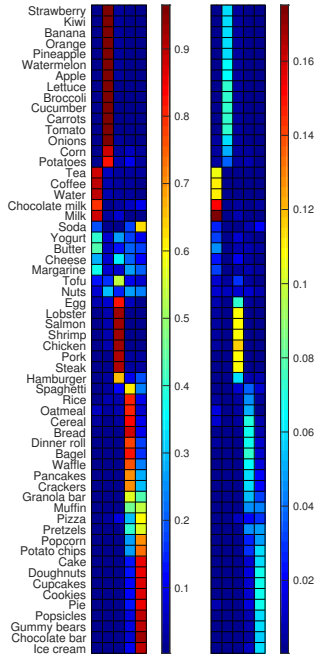
(a) FKM: Best of 100 executions **(b) CFKM: One execution**

Figure 9: Solutions for $k = 5$ medoids and $h = 1$

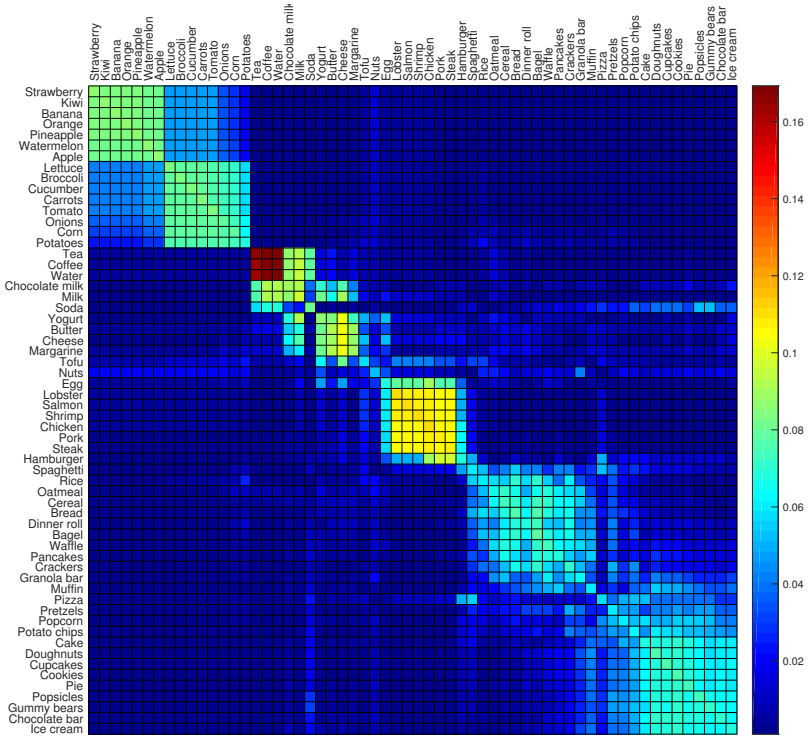
One could look at the solution in Figure 9 and conclude that the crisp solution presented by FKM ($k = 5$) is a good way to represent the data. However, using CFKM and increasing the fuzziness factor (e.g., $h = 1.5$), we show that some of the prior results may be misleading. Once we increase the fuzziness parameter ($h = 1.5$), as shown in Figure 10, we quickly see differences emerging and understand how $h = 1$ solutions obscure important aspects of the data.



(a) FKM: Best of 100 executions

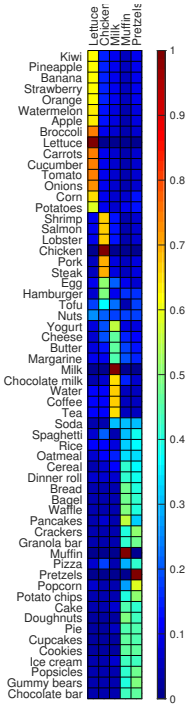


(b) FMMdd: Best of 100 executions. Memberships (left) and weighted medoids (right)

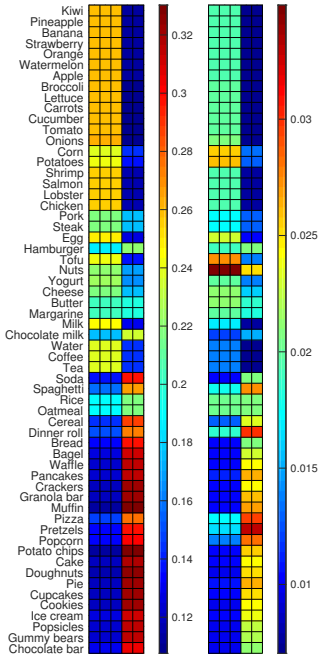


(c) CFKM: one execution

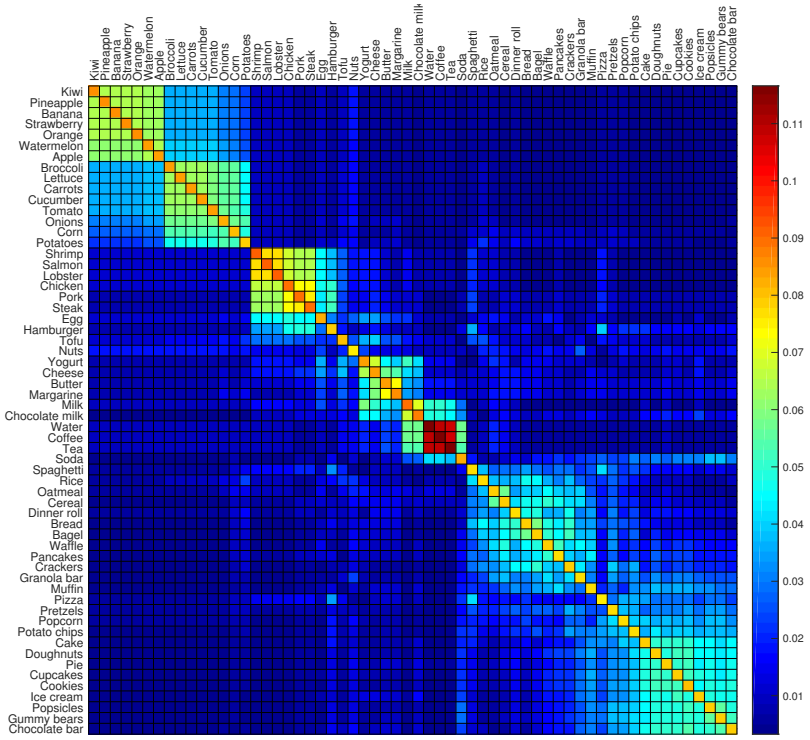
Figure 10: Solutions for $k = 5$ medoids and $h = g = 1.5$



(a) FKM: Best of 100 executions



(b) FMMdd: Best of 100 executions. Memberships (left) and weighted medoids (right)



(c) CFKM: One execution

Figure 11: Solutions for $k = 5$ medoids and $h = g = 2$

4.3.4 h = 1.5: moderately fuzzy solutions

With respect to FKM, one would conclude that the produce category is a good cluster, with carrot being the most representative object. From that, we can see that fruits and vegetables fit well with carrots, and that vegetables fit better (i.e., squares are darker). There is now some evidence of soda being assigned to two clusters, and nuts fitting poorly. We also find that spaghetti does not fit well with the grains cluster, and that some snack objects (i.e., granola bars, muffins, pretzel, popcorns, chips) seem to be belonging to both the grains and the dessert clusters. In sum, we find that at $h = 1.5$, there is some differences emerging as some objects fit to multiple clusters. Still, we find that the determination of which object is a better medoid is rather arbitrary.

With respect to FMMdd, two panels are presented. In the left panel, we provide information regarding the strength of membership. We notice that FMMdd provides nearly the same cluster assignments as FKM, with the exception that yogurt, butter, cheese, margarine, and eggs are poorly classified and appear to be members of several clusters. On the right panel, the weighted medoids data provides valuable information about which items are likely medoids. We find clear evidence that milk is a strong medoid for a liquid cluster, and some evidence that chicken is a good medoid for main dish. In summary, we find that at $h = 1.5$, the solution of FMMdd shows nearly the same clusters as was found when $h = 1$. However, we find some ambiguity in determining which objects are medoids, and numerous objects seems to be poorly classified.

Finally with respect to CFKM, we find clear evidence that the produce cluster is composed of two subsets, fruits and vegetables, and that there is no object that appears to be a better medoid. We also find evidence of a drink cluster, however we note that any of the non-dairy (i.e., tea, coffee, water) would be better medoids than the two milks (i.e., regular and chocolate) partially because there seems to be some association with the non-liquid dairy products (i.e., butter, margarine, yogurt, cheese). Just as in the prior solution, we find evidence of a main dish cluster for which we find that eggs, and hamburger (all lighter blue) are only of low representativeness. We also find evidence for grains and dessert clusters. However, the low values in the figure suggest that not only are the respective objects poor medoid choices, there is significant overlap among the clusters. We also find an interesting distinction in that whereas soda is associated strongly only with the tea, coffee and water from the drinks, it is also different in that it is associated with some of the high sugar items like gummy bears and chocolate bar. We also find that egg is associated with both the main dish and some of the breakfast products, like yogurt and cheese.

4.3.5 h = 2: highly fuzzy solutions

With respect to FKM, increasing h to 2 does not help with the interpretation of the solution. Several objects (e.g., all main dishes) are assigned to more than one cluster as nuts becomes a cluster of one. Furthermore, the distinction between types of drinks is lost and some of the grain and dessert cluster objects appear to equally fit both. In general, the FKM solution is difficult to interpret even though the medoids remained the same.

With respect to FMMdd, we see a large change from the solution obtained when $h = g = 1.5$. Specifically, we see a two cluster solution where some items are equally representative of two or three clusters. For instance, all items from kiwi to tea (top 60%) become assigned to the same three clusters (largest weight medoid: nuts), each with equal membership to the three clusters. We also find that the second cluster (mostly comprising unhealthy items) groups everything else together (largest weight medoid: pretzel).

With respect to CFKM, its solution solidifies what was learned when $h = 1.5$. Specifically, we find that vegetables and fruits have strong differentiation, such that they were perhaps misassigned to a single cluster due to the constraint on k. We also find evidence that the main dish category previously identified really comprised three subsets: seafood/fish (shrimp, salmon, lobster), meat (pork, steak and chicken) and to an extent the other main portions of a meal (hamburger, tofu, egg). We continue to find evidence that nuts and tofu are singled out, that milk and chocolate milk belong both to a cluster

of dairies and of drinks. We still find evidence that soda and spaghetti are poorly associated with clusters, and a similar breakdown of grains and desserts. No such insight can be found by inspecting the solutions from FKM and FMMdd.

4.3.6 Summary of k = 5

Several interesting differences emerged comparing the solutions of the three different fuzzy k-medoids models. First, when $h = 1$, all three formulations provided sensible clustering solutions that were easy to interpret. However, increasing the fuzziness of the models suggested that doing so may have masked substantial heterogeneity from the observed distances. Specifically, FKM provided little guidance as to why some objects were selected as medoids. In fact, in some cases, it appeared that any object part of the cluster would have been equally likely. Second, it failed to capture the true nature of the data whereby some distance scores were reflecting disagreement between participants (e.g., some people saw spaghetti as a main dish, others as a grain). For instance in FKM, we found no evidence that any of the participants acknowledged that some objects were dairies. Third, we found that FMMdd's solution became less clear as h increase to 2. The $k = 5$ solution reflected something more akin to a $k = 2$ solution, which moreover obscured a lot of important information that we could learn looking at CFKM solution. Indeed, an important advantage of CFKM over FMMdd is that the first allows to recover, for any object, which are its possible cluster medoids. Similar information is simply aggregated in FMMdd and as a consequence, unavailable to aid interpretation.

5 Concluding remarks

There are numerous ways in which one can incorporate fuzziness into a clustering model. Often, fuzziness is added by allowing non-medoid objects to be determined as varying in degree of assignment to the different clusters but still requiring that all medoids are themselves assigned to one and only one cluster. In this manuscript, we propose an alternative specification that relaxes the assumption that medoids (i.e., most representative objects) must themselves be assigned to one and only one cluster. We were able to show several advantages of our proposed fuzzy clustering model (i.e., CFKM) over the current ones found in the literature:

- The CFKM formulation is convex, and as such a method that reach a local minimum necessarily achieves the global optimum. It is thus robust to starting conditions and initializations.
- CFKM helps maintain the true clustering solution if k is increased beyond the true number of clusters in the data. In contrast, FKM forces each medoid to represent one different cluster which may not be an accurate representation of the data. Further, it seems that CFKM can help facilitate the identification of k.
- It is known that as fuzzy hyperparameters tend to infinity, medoids-based fuzzy clustering models spread data objects membership equally among the clusters. Our experiments showed that CFKM is less sensitive to increases in fuzzy hyperparameters.
- CFKM preserves more information for cluster analysis and practical interpretation (size: $n \times n$), compared to FKM (size: $n \times k$) and FMMdd (size: $2 \times n \times k$). As we show using an illustrative example of clustering of perceptions of food objects, the formulation is particularly helpful at understanding clustering structures when some several objects are equally plausible as medoids. In such cases, other medoid-based formulations tend to assign one randomly and in doing so bias our interpretation of the structure.

We expect that CFKM could be used as a tool for data practitioners who wish to perform in-depth cluster analysis. The tradeoff of keeping more information from the clustering itself is that CFKM requires more time for the analysis of the clusters found. However, we argue in this paper that such compromise might be profitable in terms of insight gain. All codes are available at https://github.com/danielnopinheiro/CFKM.

Appendix A.

[16] propose to approach FMMdd using a heuristic. In this appendix, we show that the heuristic can converge to a trivial local optimum if it begins with all objects as equally spread among the clusters, or if the representativeness of each object is equally distributed among the clusters.

Formally, assume that we initialize with a solution of $v_{jc} = 1/n$ for all $j \in \{1, \ldots, n\}$ and $c \in \{1, \ldots, k\}$. Following the heuristic steps through Equation (18), degrees of membership are updated to

$$\begin{aligned} e_{ic} &= \frac{\left[\sum_{j=1}^{n} (1/n)^g d_{ij}\right]^{-1/(h-1)}}{\sum_{f=1}^{k} \left[\sum_{j=1}^{n} (1/n)^g d_{ij}\right]^{-1/(h-1)}} \\ &= 1/k, \end{aligned}$$

for all $i \in \{1, \ldots, n\}$ and $c \in \{1, \ldots, k\}$. Now, through Equation (19), the representativeness of each object in each cluster are updated to

$$\begin{aligned} v_{jc} &= \frac{\left[\sum_{i=1}^{n} (1/k)^h d_{ij}\right]^{-1/(g-1)}}{\sum_{t=1}^{n} \left[\sum_{i=1}^{n} (1/k)^h d_{it}\right]^{-1/(g-1)}} \\ &= \frac{(1/k)^{-h/(g-1)} \left(\sum_{i=1}^{n} d_{ij}\right)^{-1/(g-1)}}{(1/k)^{-h/(g-1)} \sum_{t=1}^{n} \left(\sum_{i=1}^{n} d_{it}\right)^{-1/(g-1)}} \\ &= \frac{\left(\sum_{i=1}^{n} d_{ij}\right)^{-1/(g-1)}}{\sum_{t=1}^{n} \left(\sum_{i=1}^{n} d_{it}\right)^{-1/(g-1)}} \equiv \beta_j, \end{aligned}$$

for all $j \in \{1, \ldots, n\}$ and $c \in \{1, \ldots, k\}$.

One can notice that the new values of v_{jc} are being expressed as β_j, which makes it explicit that the new representativeness does not depend on c. Consequently, the next updating of the cluster membership leads to

$$\begin{aligned} e_{ic} &= \frac{\left(\sum_{j=1}^{n} \beta_j^g d_{ij}\right)^{-1/(h-1)}}{\sum_{f=1}^{k} \left(\sum_{j=1}^{n} \beta_j^g d_{ij}\right)^{-1/(h-1)}} \\ &= 1/k. \end{aligned}$$

Thus, if the heuristic for the FMMdd problem begins with a solution where all objects are equally spread among the clusters or if the representativeness of each object is equally distributed among the clusters, then the algorithm converges to a solution where $e_{ic} = 1/k$ and $v_{jc} = \beta_j$ for all $i, j \in \{1, \ldots, n\}$ and $c \in \{1, \ldots, k\}$, where $\beta_j = \frac{\left(\sum_{i=1}^{n} d_{ij}\right)^{-1/(g-1)}}{\sum_{t=1}^{n} \left(\sum_{i=1}^{n} d_{it}\right)^{-1/(g-1)}}$.

References

[1] Pasquale Avella, Antonio Sassano, and Igor Vasil'ev. Computational study of large-scale p-median problems. Mathematical Programming, 109(1):89–114, 2007.

[2] Simon J. Blanchard, Daniel Aloise, and Wayne S. DeSarbo. The heterogeneous p-median problem for categorization based clustering. Psychometrika, 77(4):741–762, 2012.

[3] Simon J. Blanchard and Ishani Banerji. Evidence-based recommendations for designing free-sorting experiments. Behavior research methods, 48(4):1318–1336, 2016.

[4] Stephen Boyd and Lieven Vandenberghe. Convex optimization. Cambridge university press, 2004.

[5] Anthony Peter Macmillan Coxon. Sorting data: Collection and analysis, volume 127. Sage, 1999.

[6] Sergio García, Martine Labbé, and Alfredo Marín. Solving large p-median problems with a radius formulation. INFORMS Journal on Computing, 23(4):546–556, 2011.

[7] Pierre Hansen and Nenad Mladenović. Variable neighborhood search for the p-median. Location Science, 5(4):207–226, 1997.

[8] Anil K. Jain, M. Narasimha Murty, and Patrick J. Flynn. Data clustering: A review. ACM Comput. Surv., 31(3):264–323, September 1999.

[9] Oded Kariv and Seifollah L. Hakimi. An algorithmic approach to network location problems. II: the p-medians. SIAM J Appl Math, 37:539–560, 1979.

[10] Leonard Kaufman and Peter Rousseeuw. Clustering by means of medoids. In Data Analysis based on the L1-Norm and Related Methods, pages 405–416, 01 1987.

[11] Leonard Kaufman and Peter Rousseeuw. Finding groups in data: An introduction to cluster analysis. In Wiley, New York. ISBN 0-471-87876-6, 01 1990.

[12] Frank Klawonn, Rudolf Kruse, and Roland Winkler. Fuzzy clustering: More than just fuzzification. Fuzzy sets and systems, 281:272–279, 2015.

[13] Hans-Friedrich Köhn, Douglas Steinley, and Michael J Brusco. The p-median model as a tool for clustering psychological data. Psychological Methods, 15(1):87, 2010.

[14] Raghu Krishnapuram, Anupam Joshi, Olfa Nasraoui, and Liyu Yi. Low-complexity fuzzy relational clustering algorithms for web mining. IEEE transactions on Fuzzy Systems, 9(4):595–607, 2001.

[15] Nicolas Labroche. New incremental fuzzy c medoids clustering algorithms. In Fuzzy Information Processing Society (NAFIPS), 2010 Annual Meeting of the North American, pages 1–6. IEEE, 2010.

[16] Jian-Ping Mei and Lihui Chen. Fuzzy relational clustering around medoids: A unified view. Fuzzy Sets and Systems, 183(1):44–56, 2011. Theme : Information processing.

[17] Volodymyr Melnykov, Wei-Chen Chen, and Ranjan Maitra. MixSim: An R package for simulating data to study performance of clustering algorithms. Journal of Statistical Software, 51(12):1–25, 2012.

[18] Frank Nielsen. Introduction to HPC with MPI for Data Science. Springer, 2016.

[19] Mauricio G. C. Resende and Renato F. Werneck. A fast swap-based local search procedure for location problems. Annals of Operations Research, 150(1):205–230, 2007.

[20] Brian H. Ross and Gregory L. Murphy. Food for thought: Cross-classification and category organization in a complex real-world domain. Cognitive psychology, 38(4):495–553, 1999.

[21] Éverton Santi, Daniel Aloise, and Simon J. Blanchard. A model for clustering data from heterogeneous dissimilarities. European Journal of Operational Research, 253(3):659–672, 2016.

[22] Yangtao Wang, Lihui Chen, and Jian-Ping Mei. Incremental fuzzy clustering with multiple medoids for large data. IEEE transactions on fuzzy systems, 22(6):1557–1568, 2014.