Balanced Fair K-Means Clustering

Pan R, Zhong C and Qian J

IEEE Transactions on Industrial Informatics, 2023

Jinwon Park, Jihu Lee

February 22, 2024

Seoul National University

Abstract

Abstract

- Many existing fair clustering algorithms generate numberous small clusters
- This paper presents a balanced fair K-means clustering algorithm that prevents the generation of small clusters

Main contribution

- A fairness constraint that can handle a multi-value sensitive attribute
- Proposes a control of the trade off between fairness and K-means objective
- Avoids to generate some empty or small clusters
- The optimization problem is discontinuous, hence proposes an iterative update solution

Notations and preliminaries

- $\mathcal{X} = \mathcal{G}_1 \cup \mathcal{G}_2 \cup \cdots \cup \mathcal{G}_m = \{x_1, \ldots, x_n\}, x_i \in \mathbb{R}^d$: given set of data points
- $\mathcal{C} = \{\mathcal{C}_1, \ldots, \mathcal{C}_k\}$: set of k clusters where $\mathcal{C}_j = \{x \in \mathcal{X} : f(x) = j, 1 \leq j \leq k\}$
- Demographic Parity

$$P(f(x) = i | x \in \mathcal{G}_1) = p(f(x) = i | x \in \mathcal{G}_2)$$

$$\forall j \in \{1, \ldots, k\} \; \forall l \in \{1, \ldots, m\}$$

- $Y = [y_1, \ldots, y_n]^T \in \mathbb{R}^{n \times k}$: label matrix with

$$y_{ij} = \begin{cases} 1, & \text{if } x_i \in \mathcal{C}_j \\ 0, & \text{otherwise} \end{cases}$$

Fair K-Means Clustering

$$\min_{Y,M} \|X - MY^T\|_F^2$$

$$\text{s.t. } y_{ij} \in \{0, 1\}$$
$$\sum_{j=1}^{k} y_{ij} = 1$$

where $X = [x_1, x_2, \ldots, x_n] \in \mathbb{R}^{d\times n}$ is the matrix form of $\mathcal{X}$

and $M = [\mu_1, \mu_2, \ldots, \mu_k] \in \mathbb{R}^{d\times k}$ is the cluster centroid matrix

Fair K-Means Clustering

- Let $F = [f_1, f_2, \dots, f_n]^t \in \mathbb{R}^{n \times m}$ be group matrix with

$$f_{ij} = \begin{cases} 1, & \text{if } x_i \in \mathcal{G}_j \\ 0, & \text{otherwise} \end{cases}$$

- Then, DP can be written as

$$\frac{Y_{:j}^T F_{:l}}{F_{:l}^T F_{:l}} = \frac{Y_{:j}^T \mathbf{1}}{n} \quad \forall j \in \{1, \dots, k\} \; \forall l \in \{1, \dots, m\}$$

- Hence, the fairness constraint can be written as

$$\frac{Y^T F}{F^T F} - A = 0$$

where $A \in \mathbb{R}^{k \times m}$ is a matrix with all columns are $\frac{Y^T \mathbf{1}}{n}$

As a result, by employing Lagrangian multiplier and relaxing the fairness constraints,

$$\min_{Y,M} \|X - MY^T\|_F^2 + \rho \left\| \frac{Y^T F}{F^T F} - A \right\|_F^2$$

$$\text{s.t. } f_{ij} \in \{0, 1\}$$

$$\sum_{j=1}^{k} f_{ij} = 1$$

FrKM

Fair K-Means clustering algorithm is to solve the optimization problem above

To solve the problem, Coordinate descent is needed since gradient descent is not an option for discrete values

By taking derivatives w.r.t each variables,

Fixing Y, update M

$$
\begin{aligned}
&\frac{\partial}{\partial M}\|X - MY^T\|_F^2 \\
=&\frac{\partial}{\partial M} tr((X - MY^T)(X - MY^T)^T) \\
=&2(MY^TY - XY) \\
\rightarrow& M = XY(Y^TY)^{-1}
\end{aligned}
$$

Fixing M, update Y

$$
y_{ij} = \begin{cases} 1, & \text{if } j = \text{argmin}\,\|X - MY^T\|_F^2 + \rho \left\| \frac{Y^F}{F^TF} - A \right\|_F^2 \\ 0, & \text{otherwise} \end{cases}
$$

Algorithm 1: FrKM.

Input: Data matrix X, cluster number k, parameters ρ.
Output: Binary indicator matrix Y.

1: Initialize Y randomly.
2: **repeat**
3: Update M by (13).
4: Calculate Y by (14).
5: **until** convergence

BFKM

The balanced clusters can be generated by minimizing $\sum_{i=1}^{k} \frac{1}{s_i}$ where $s_i = |\mathcal{C}_j|$

For Balanced Fair K-Means clustering algorithm,

$$\min_{Y,M} \|X - MY^T\|_F^2 + \rho \left\| \frac{Y^T F}{F^T F} - A \right\|_F^2 + \lambda tr((Y^T Y)^{-1})$$

$$\text{s.t. } f_{ij} \in \{0, 1\}$$

$$\sum_{j=1}^{k} f_{ij} = 1$$

with λ as a balance parameter

Algorithm 2: BFKM.

Input: Data matrix X, cluster number k, parameters ρ and λ.

Output: Binary indicator matrix Y.

1: Initialize Y randomly.

2: **repeat**

3: Update M by (13).

4: Calculate Y by (17).

5: **until** convergence

By taking derivatives w.r.t each variables, y_i is optimized by

$$y_{ij} = \begin{cases} 1, & \text{if } j = \operatorname{argmin} \|X - MY^T\|_F^2 + \rho \left\| \frac{Y^F}{F^T F} - A \right\|_F^2 + \lambda tr((Y^T Y)^{-1}) \\ 0, & \text{otherwise} \end{cases}$$

Experiments

TABLE III
CLUSTERING FAIRNESS

Dataset	FR ↑							AWD ↓						
	Lloyd	FSCUN	FSCN	FAC	VFC	FrKM	BFKM	Lloyd	FSCUN	FSCN	FAC	VFC	FrKM	BFKM
Elliptical (k=2)	0	0.8878	0.8835	0.8000	0.8835	0.8835	**0.9005**	0.4940	**0.0460**	0.0480	0.1000	0.0480	0.0480	0.0472
DS-577 (k=3)	0	0	0	0.7988	0.5427	0.8019	**0.8042**	0.4183	0.0995	0.1496	0.0462	0.1618	**0.0242**	0.0319
2d-4c-no0 (k=4)	0	0	0	0.7719	0.0201	0.8114	**0.8141**	0.3127	**0.0049**	0.0649	0.0274	0.2226	0.0052	0.0103
2d-4c-no1 (k=4)	0	0	0	0.7780	0	0.8018	**0.8019**	0.2705	**0.0012**	0.0701	0.0427	0.2341	0.0014	0.0125
2d-4c-no4 (k=4)	0	0	0	0.7369	0.3045	0.7388	**0.7566**	0.1723	0.0759	0.0799	0.0459	0.1117	**0.0076**	0.0139
Adult (k=10)	0.4362	0	0.5599	0.7995	0.8996	0.9051	**0.9180**	0.0573	0.0428	0.0559	0.0484	0.0128	0.0082	**0.0074**
Bank (k=6)	0.2929	0.3368	0.5306	0.7958	0.8076	0.8090	**0.8137**	0.0691	0.0357	0.0358	0.0334	0.0251	**0.0203**	0.0210
Census1990 (k=5)	0.5129	0.6964	0.7418	0.7984	0.9119	0.9151	**0.9169**	0.0671	0.0511	0.0470	0.0554	0.0174	**0.0141**	**0.0141**
CreditCard (k=10)	0.7390	0	0.8851	0.7999	0.7353	**0.8900**	**0.8900**	0.0363	**0.0034**	0.0188	0.0340	0.0376	0.0170	0.0194
Diabetic (k=10)	0.8376	0	0.8239	0.8239	0.8728	0.8744	**0.8777**	0.0357	**0.0235**	0.0285	0.0289	0.0243	0.0242	0.0260

The bold entities indicate the best results.

TABLE IV
CLUSTERING QUALITY

Dataset	DI ↑							SSE ↓						
	Lloyd	FSCUN	FSCN	FAC	VFC	FrKM	BFKM	Lloyd	FSCUN	FSCN	FAC	VFC	FrKM	BFKM
Elliptical (k=2)	0.0644	0.0118	**0.0716**	0.0010	**0.0716**	**0.0716**	0.0040	**206.2982**	344.4216	343.9640	446.1201	343.9640	343.9640	351.1090
DS-577 (k=3)	**0.0079**	0.0005	0.0068	0.0005	0.0005	0	0	**71.0134**	449.0292	361.4299	501.0075	377.3101	518.1536	516.0655
2d-4c-no0 (k=4)	**0.0076**	0.0032	0.0002	0.0002	0.0002	0	0	**114.4834**	1.5283E+03	1.3558E+03	1.4101E+03	419.5988	1.4789E+03	1.4555E+03
2d-4c-no1 (k=4)	0.0017	**0.0112**	0.0001	0	0	0	0	**82.3122**	1.6150E+03	1.2718E+03	1.4734E+03	287.2313	1.5835E+03	1.5397E+03
2d-4c-no4 (k=4)	**0.0057**	0.0002	0.0001	0.0001	0.0001	0	0	**104.0023**	705.9991	666.3117	676.0752	420.1287	714.9866	704.5242
Adult (k=10)	**0.0006**	0	0.0004	0	0	0	0	**9.5083E+03**	1.4380E+04	1.0251E+04	9.7740E+03	1.0028E+04	1.0277E+04	1.0583E+04
Bank (k=6)	**0.0190**	0	0.0018	0	0.0013	0.0001	0	**1.23134E+03**	1.7858E+03	1.2543E+03	1.2533E+03	1.2774E+03	1.3698E+03	1.3298E+03
Census1990 (k=5)	0.0507	0.0580	0.0258	**0.0859**	0.0527	0.0433	0.0640	**1.7604E+03**	1.8207E+03	1.8219E+03	1.8405E+03	1.8807E+03	1.8526E+03	1.8520E+03
CreditCard (k=10)	0.0094	**0.0275**	0	0	0.0093	0	0	8.1998E+03	1.8420E+04	9.3441E+03	8.2005E+03	**8.1834E+03**	8.2827E+03	8.2267E+03
Diabetic (k=10)	**0.0406**	0	0	0.0374	0	0	0	243.2634	3.2616E+03	235.2583	**231.9906**	254.8696	327.1158	298.4838

The bold entities indicate the best results.

Experiments

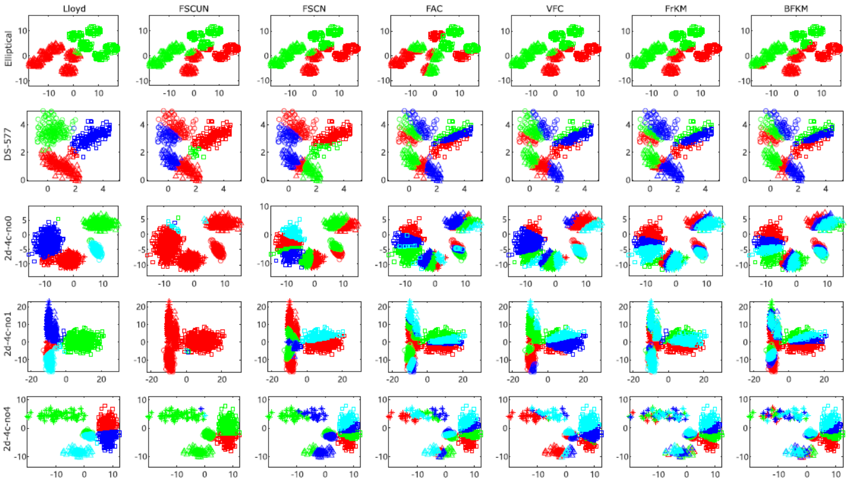

Fig. 6. Clustering results on the synthetic datasets. The clustering result is represented by different colors and shapes.

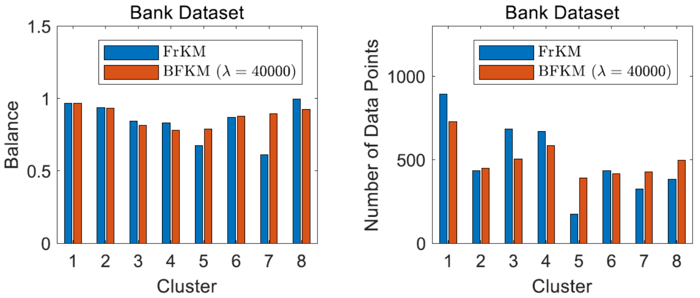

Fig. 9. FR and sample distribution of FrKM and BFKM in clusters (kw $= 8$). The results are shown for the Bank dataset.