Fair Kernel K-Means: from Single Kernel to Multiple Kernel

NeurIPS2024

Peng Zhou, Rongwen Li, Liang Du

November 13, 2024

An Hui University, Shan Xi University

Background

fairness in clustering (Chierichetti et al.)

For a subset **Y**, the balance of **Y** is defined as

$$\text{balance}(Y) = \min\left(\frac{\#\,\text{RED}(Y)}{\#\,\text{BLUE}(Y)}, \frac{\#\,\text{BLUE}(Y)}{\#\,\text{RED}(Y)}\right) \in [0, 1]$$

The balance of a clustering $\mathcal{C}$ is defined as:

$$\text{balance}(\mathcal{C}) = \min_{C \in \mathcal{C}} \text{balance}\,(C)$$

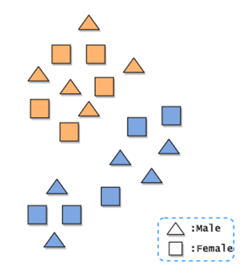

Figure 1: The effect of fair clustering

Contribution

- We design a simple yet effective fair regularization term.
- Based on this regularization term, we first propose fair kernel k-means.
- We also provide a generalization analysis for our method and obtain some interesting conclusions about fairness and clustering performance.

Fair Regularization

- $\mathbf{G} \in 0, 1^{n\times t}$ (protected group indicators)
- $\mathbf{Y} \in 0, 1^{n\times c}$ (cluster indicators)

Theorem 1

Given $\mathbf{G}$ and $\mathbf{Y}$ defined as mentioned before, we can obtain the maximum of fairness by optimizing the following objective function:

$$\min_{\mathbf{Y}\in Ind} \operatorname{Tr}\left(\mathbf{Y}^T\mathbf{G}\mathbf{G}^T\mathbf{Y}\left(\mathbf{Y}^T\mathbf{Y}\right)^{-1}\right).$$

Fair Kernel K-means

fair kernel k-means

$$\min_{\mathbf{Y}\in Ind} \mathrm{Tr}(\mathbf{K}) - \mathrm{Tr}\left(\left(\mathbf{Y}^T\mathbf{Y}\right)^{-\frac{1}{2}}\mathbf{Y}^T\mathbf{K}\mathbf{Y}\left(\mathbf{Y}^T\mathbf{Y}\right)^{-\frac{1}{2}}\right) + \lambda\,\mathrm{Tr}\left(\mathbf{Y}^T\mathbf{G}\mathbf{G}^T\mathbf{Y}\left(\mathbf{Y}^T\mathbf{Y}\right)^{-1}\right)$$

$$\Longleftrightarrow \max_{\mathbf{Y}\in Ind} \mathrm{Tr}\left(\mathbf{Y}^T\left(\mathbf{K} - \lambda\mathbf{G}\mathbf{G}^T\right)\mathbf{Y}\left(\mathbf{Y}^T\mathbf{Y}\right)^{-1}\right)$$

fair multiple kernel k-means

$$\min_{\mathbf{Y},\gamma} \mathrm{Tr}\left(\tilde{\mathbf{K}}^*\left(\mathbf{I} - \mathbf{Y}\left(\mathbf{Y}^T\mathbf{Y}\right)^{-1}\mathbf{Y}^T\right)\right)$$

$$s.t.\ \mathbf{Y} \in Ind,\ \boldsymbol{\gamma}^T\mathbf{1} = 1,\ \gamma_p \geq 0,\ \tilde{\mathbf{K}}^* = \sum_{p=1}^{m} \gamma_p^2 \tilde{\mathbf{K}}^{(p)}$$

p.s.d. guarantee

Lemma 1

Given two real symmetric matrices $\mathbf{A}$ and $\mathbf{B}$ with the same size, where the smallest eigenvalue of $\mathbf{A}$ is σ_A and the largest eigenvalue of $\mathbf{B}$ is σ_B. If $\sigma_A \geq \sigma_B$, then $\mathbf{A} - \mathbf{B}$ is p.s.d.

$$\mathrm{Tr}\left(\mathbf{Y}^T\left(\mathbf{K}-\lambda\mathbf{G}\mathbf{G}^T\right)\mathbf{Y}\left(\mathbf{Y}^T\mathbf{Y}\right)^{-1}\right)+\alpha\,\mathrm{Tr}(\mathbf{I}) = \mathrm{Tr}\left(\mathbf{Y}^T\left(\mathbf{K}+\alpha\mathbf{I}-\lambda\mathbf{G}\mathbf{G}^T\right)\mathbf{Y}\left(\mathbf{Y}^T\mathbf{Y}\right)^{-1}\right)$$

Pseudo-codes of our method

Algorithm 1 Fair Multiple Kernel K-means

Input: Kernel matrices $\{\mathbf{K}^{(p)}\}_{p=1}^{m}$, protected groups $\mathcal{G}_1, \cdots, \mathcal{G}_t$, fairness hyper-parameter λ.

1: Construct protected group indicator matrix $\mathbf{G}$ and calculate α as $\alpha = |\mathcal{G}_{max}| * \lambda$.

2: Construct the corresponding fair kernel by $\tilde{\mathbf{K}}^{(p)} = \mathbf{K}^{(p)} + \alpha\mathbf{I} - \lambda\mathbf{G}\mathbf{G}^T$ for each base kernel matrix $\mathbf{K}^{(p)}$.

3: Initialize $\boldsymbol{\gamma} = \frac{1}{m}$ and $\mathbf{Y}$ by running standard kernel k-means on $\sum_{p=1}^{m} \gamma_p^2 \mathbf{K}^{(p)}$.

4: **repeat**

5: Update $\mathbf{Y}$ row by row by solving $\max_{\mathbf{Y} \in Ind} \mathrm{Tr}\left(\mathbf{Y}^T \tilde{\mathbf{K}}^{*} \mathbf{Y} \left(\mathbf{Y}^T \mathbf{Y}\right)^{-1}\right)$.

6: Update $\boldsymbol{\gamma}$ by $\gamma_p = \frac{h_p^{-1}}{\sum_{j=1}^{m} h_j^{-1}}$.

7: **until** Converges

Output: The final partition matrix $\mathbf{Y}$.

Assumption

Assumption 1

Each $\tilde{\mathbf{K}}^{(p)} = \mathbf{K}^{(p)} + \alpha\mathbf{I} - \lambda\mathbf{G}\mathbf{G}^T$ is a valid kernel matrix, i.e., $\tilde{\mathbf{K}}^{(p)}$ is symmetric and p.s.d.

Assumption 2

All $\mathbf{K}^{(p)}$ are upper bounded. We denote b as the maximum of elements in all $\mathbf{K}^{(p)}$.

Generalization Analysis

- function class of our method:

$$\mathcal{F} = \left\{ f : \mathbf{x} \mapsto \min_{\mathbf{y} \in \{\mathbf{e}_1, \ldots, \mathbf{e}_c\}} \|\Phi_{\boldsymbol{\gamma}}(\mathbf{x}) - \mathbf{M}\mathbf{y}\|_{\mathcal{H}}^2 \,\middle|\, \boldsymbol{\gamma}^T \mathbf{1} = 1, \gamma_p \geq 0, \mathbf{m}_k \in \mathcal{H} \right\}.$$

Theorem 2

any $\delta \geq 0$, with probability at least $1 - \delta$, the following inequality holds:

$$\mathbb{E}[f(\mathbf{x})] \leq \frac{1}{n} \sum_{i=1}^{n} f(\mathbf{x}_i) + \frac{2\sqrt{2\pi}}{\sqrt{n}} \left[(1 + c^2)(b + \alpha) - (1 + \frac{c^2}{t})\lambda + c\sqrt{2(b + \alpha - \lambda)\left(b + \alpha - \frac{\lambda}{t}\right)} \right]$$
$$+ \left(4(b + \alpha) - 2\left(1 + \frac{1}{t}\right)\lambda\right)\sqrt{\frac{\log(1/\delta)}{2n}}$$

Generalization Analysis

$$\frac{2\sqrt{2\pi}}{\sqrt{n}}\left[(1+c^2)\,(b+\alpha)-(1+\frac{c^2}{t})\lambda+c\sqrt{2\,(b+\alpha-\lambda)\left(b+\alpha-\frac{\lambda}{t}\right)}\right]$$
$$+\left(4\,(b+\alpha)-2\left(1+\frac{1}{t}\right)\lambda\right)\sqrt{\frac{\log(1/\delta)}{2n}}$$
$$\geq\frac{2\sqrt{2\pi}}{\sqrt{n}}\left[(1+c^2)b+\left(|\mathcal{G}_{max}|-1+\frac{c^2(|\mathcal{G}_{max}|t-1)}{t}\right)\lambda+c\sqrt{2(b+(|\mathcal{G}_{max}|-1)\lambda)\left(b+\frac{|\mathcal{G}_{max}|t-1}{t}\lambda\right)}\right]$$
$$+(4b+4(|\mathcal{G}_{max}|-1)\lambda)\sqrt{\frac{\log(1/\delta)}{2n}}$$

Experiments

Table 1: Comparison results on the single kernel setting. The best and second best results are denoted in **bold** and underlined, respectively.

Data sets		K-means	KKM	SC	FairSC	VFC	FFC	FKKM-f	FKKM
D&S	ACC	0.555	0.552	0.558	0.433	0.539	0.521	**0.648**	0.636
	NMI	0.650	0.602	0.652	0.575	0.617	0.583	**0.724**	0.683
	Bal	0	0	0	0	0.186	0.100	0	**0.559**
	MNCE	0.156	0.531	0.023	0	0.923	0.712	0.477	**0.991**
HAR	ACC	0.524	0.620	0.680	0.742	0.600	0.602	0.689	**0.771**
	NMI	0.596	0.609	0.618	0.703	0.654	0.490	0.625	**0.710**
	Bal	0	0	0	0	0.200	0.007	0	**0.250**
	MNCE	0.933	0.930	0.914	0	0.983	0.953	0.920	**0.989**
MNIST-USPS	ACC	0.363	0.396	0.406	**0.458**	0.360	0.437	0.403	0.432
	NMI	0.423	0.421	**0.435**	0.429	0.306	0.412	0.426	0.380
	Bal	0	0	0	0	0.142	0.217	0	**0.847**
	MNCE	0	0.003	0	0	0.544	0.684	0	**0.997**
Jaffe	ACC	0.927	0.948	0.901	0.957	0.981	0.901	0.954	**1**
	NMI	0.914	0.922	0.889	0.943	0.969	0.918	0.930	**1**
	Bal	0	0	0	0	0.400	0.250	0	**0.500**
	MNCE	0.808	0.900	0.765	0.827	0.983	0.924	0.897	**0.989**
Credit Card	ACC	0.362	0.381	0.311	0.351	0.381	0.364	0.400	**0.404**
	NMI	0.139	0.140	0.126	0.123	0.142	0.139	0.145	**0.148**
	Bal	0.510	0.550	0.567	0.603	0.586	0.550	0.536	**0.624**
	MNCE	0.953	0.961	0.967	0.973	0.970	0.969	0.956	**0.985**
K1b	ACC	0.742	0.669	0.667	**0.853**	0.778	0.663	0.826	0.809
	NMI	0.589	0.537	0.536	**0.666**	0.553	0.503	0.628	0.591
	Bal	0.666	0.775	0.763	0.667	0.794	0.773	0.703	**0.800**
	MNCE	0.971	0.989	0.987	0.971	0.990	0.989	0.978	**0.991**

Visualization

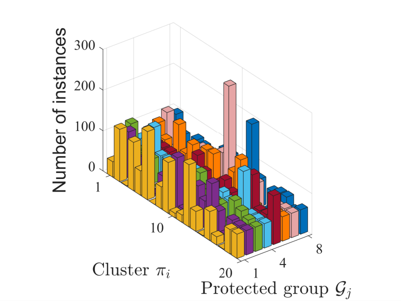

Figure 2: standard kernel

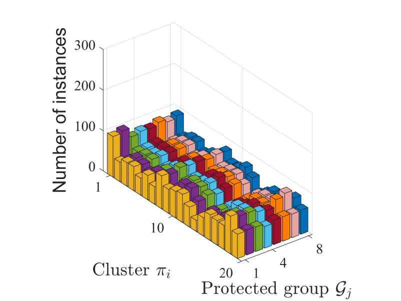

Figure 3: our fair kernel

Last

Thank you!