Hierarchical Multiple Kernel Clustering

Jiyuan Liu

liujiyuan13@nudt.edu.cn

January 8, 2021

Why Multiple Kernel Clustering ?

Problems of single kernel clustering methods:

1. how to choose kernel mapping and corresponding parameters;
2. rapidly increasing multi-view data.

Naive approaches:

1. searching;
2. concatenating data of all views directly.

Multiple Kernel Clustering (MKC)

1. generate kernels with various mapping functions and parameters;
2. generate kernels corresponding to each view.

MKC Categorization

Early-fusion approaches

1. directly learn a consensus kernel or graph from multiple ones;
2. afterwards generate the final partition.

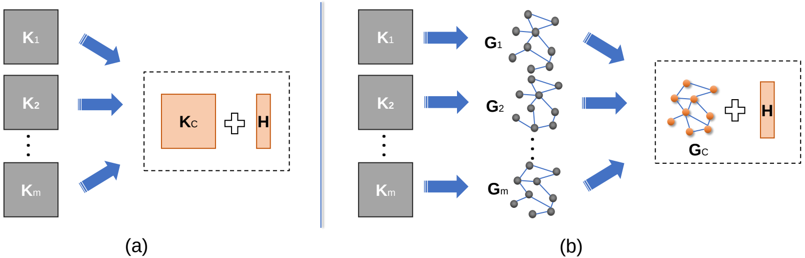

Figure 1: A brief introduction to early-fusion MKC methods.

MKC Categorization

Late-fusion approaches

1. firstly obtain multiple partitions from each kernel independently;
2. then construct the final partition on them.

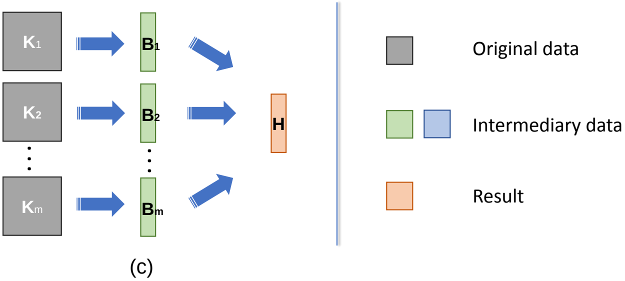

Figure 2: A brief introduction to late-fusion MKC methods.

MKC Problem

Problem

Both of them try to encode the clustering information from kernels or graphs with size $\mathcal{R}^{n\times n}$ to partition matrices with size $\mathcal{R}^{n\times k}$, where n and k are the number of samples and clusters, respectively. This **sudden drop of dimension** would result in the **loss of advantageous details** for clustering.

Visualization of Detail Loss

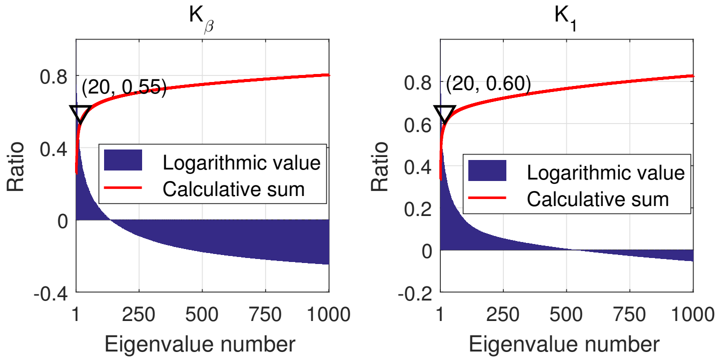

Figure 3: Eigenvalue distributions of the first kernel $\mathbf{K}_1$ and consensus kernel $\mathbf{K}_\beta$ obtained by MKKM on CCV. The eigenvalues are sorted from large to small, and only the first 1000 out of 6773 ones are plotted. The bar plot shows the times of each logarithmic eigenvalue to the first one. Meanwhile, the curve presents the calculative sum of the sorted eigenvalues.

Visualization of HMKC

We gradually group samples into $\{c_t\}_{t=1}^{s}$ clusters, together with generating a sequence of intermediary matrices with size $\mathcal{R}^{n\times c_t}$, in which $n > c_1 > \cdots > c_s > k$. A consensus partition with size $\mathcal{R}^{n\times k}$ is simultaneously learned and conversely guides the construction of intermediary matrices.

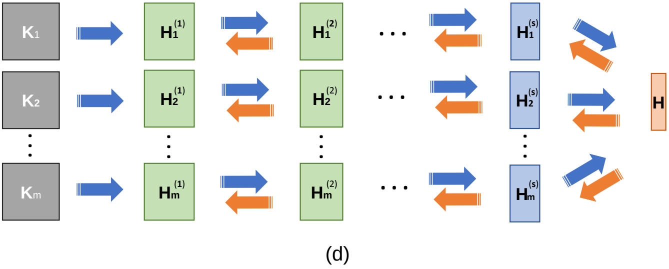

Figure 4: A brief introduction to the proposed HMKC.

Model Building

1. **From $\mathbf{K}_p$ to $\mathbf{H}_p^{(1)}$.**

We firstly group data samples into c_1 clusters, where $n > c_1 > k$, to extract the clustering details from kernels.

$$\begin{aligned} \max_{\mathbf{H}_p^{(1)}} \quad & \mathrm{Tr}\left(\mathbf{K}_p \mathbf{H}_p^{(1)} \mathbf{H}_p^{(1)\top}\right) \\ s.t.\ & \mathbf{H}_p^{(1)\top} \mathbf{H}_p^{(1)} = \mathbf{I}_{c_1},\ \mathbf{H}_p^{(1)} \in \mathbb{R}^{n \times c_1}. \end{aligned} \tag{1}$$

Model Building

2. **From $\mathbf{H}_p^{(t-1)}$ to $\mathbf{H}_p^{(t)}$.**

Without loss of generality, we formulate the relation between $\mathbf{H}_p^{(t-1)}$ and $\mathbf{H}_p^{(t)}$ as

$$\begin{aligned} &\max_{\mathbf{H}_p^{(t)}} \ \mathrm{Tr}\left(\mathbf{H}_p^{(t-1)}\mathbf{H}_p^{(t-1)\top}\mathbf{H}_p^{(t)}\mathbf{H}_p^{(t)\top}\right) \\ &s.t.\ \mathbf{H}_p^{(t)\top}\mathbf{H}_p^{(t)} = \mathbf{I}_{c_t},\ \mathbf{H}_p^{(t)} \in \mathbb{R}^{n\times c_t}. \end{aligned} \tag{2}$$

Model Building

3. **From $\mathbf{H}_p^{(s)}$ to $\mathbf{H}$.**

We also perform kernel k-means on each one to obtain a consensus partition and employ an vector $\boldsymbol{\beta} \in \mathbb{R}^m$ to adjust their weights, which can be formulated as

$$\begin{aligned} \max_{\mathbf{H}, \boldsymbol{\beta}} \quad & \sum_{p=1}^{m} \beta_p \operatorname{Tr}\left(\mathbf{H}_p^{(s)} \mathbf{H}_p^{(s)\top} \mathbf{H} \mathbf{H}^\top\right) \\ \text{s.t.} \quad & \boldsymbol{\beta}^\top \boldsymbol{\beta} = 1, \mathbf{H}^\top \mathbf{H} = \mathbf{I}_k, \mathbf{H} \in \mathbb{R}^{n \times k}, \end{aligned} \tag{3}$$

Objective

Utilizing the aforementioned three steps into an unified objective, we can obtain

$$
\begin{aligned}
\max_{\mathbf{H},\{\mathbf{H}_p\}_{p=1}^{m},\beta,\gamma} \ & \sum_{p=1}^{m} \gamma_p^{(1)} \mathrm{Tr}\left(\mathbf{K}_p \mathbf{H}_p^{(1)} \mathbf{H}_p^{(1)\top}\right) \\
& + \sum_{t=2}^{s} \sum_{p=1}^{m} \gamma_p^{(t)} \mathrm{Tr}\left(\mathbf{H}_p^{(t-1)} \mathbf{H}_p^{(t-1)\top} \mathbf{H}_p^{(t)} \mathbf{H}_p^{(t)\top}\right) \\
& + \sum_{p=1}^{m} \beta_p \mathrm{Tr}\left(\mathbf{H}_p^{(s)} \mathbf{H}_p^{(s)\top} \mathbf{H}\mathbf{H}\right) \\
s.t.\ & \mathbf{H}^\top \mathbf{H} = \mathbf{I}_k,\ \mathbf{H} \in \mathbb{R}^{n\times k},\ \mathbf{H}_p^{(t)\top} \mathbf{H}_p^{(t)} = \mathbf{I}_{c_t}, \\
& \mathbf{H}_p^{(t)} \in \mathbb{R}^{n\times c_t},\ n > c_1 > \cdots > c_s > k,\ \boldsymbol{\beta}^\top \boldsymbol{\beta} = 1, \\
& \beta_p \geq 0,\ \boldsymbol{\beta} \in \mathbb{R}^m,\ \boldsymbol{\gamma}^{(t)\top} \boldsymbol{\gamma}^{(t)} = 1,\ \gamma_p^{(t)} \geq 0,\ \boldsymbol{\gamma}^{(t)} \in \mathbb{R}^m,
\end{aligned}
\tag{4}
$$

in which $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}^{(t)}$ are the weights corresponding to each kernel.

Experiment Settings

Datasets

Table 1: Specifications of the used datasets.

Dataset	Number of Samples	Number of Kernels	Number of Clusters
AR10P	130	6	10
BBCSport	544	6	5
CCV	6773	3	20
Flower17	1360	7	17
Flower102	8189	4	102
Heart	270	13	2
Ionosphere	351	33	2
Plant	940	69	4

Experiment Setting

Comparative Methods

For comparative methods, we adopt the source codes, which are publicly available. Meanwhile, we open the source code of HMKC on `https://github.com/liujiyuan13/HMKC-code_release`.

Parameter Settings

We consider two experimental settings. The first is called HMKC-1 with employing 1-layer of intermediary matrices, $\{\mathbf{H}_p^{(1)}\}_{p=1}^m \in \mathbb{R}^{n\times c}$, where c is searched from $[2k, 3k, \cdots, 20k]$. While the second one is named HMKC-2 with employing 2-layer of intermediary matrices, $\{\mathbf{H}_p^{(1)}\}_{p=1}^m \in \mathbb{R}^{n\times c_1}$ and $\{\mathbf{H}_p^{(2)}\}_{p=1}^m \in \mathbb{R}^{n\times c_2}$, in which $c_1 \geq c_2$ and c_1, c_2 are grid searched from $[2k, 3k, \cdots, 20k]$.

Experiment Results

Effectiveness of Intermediary Matrices

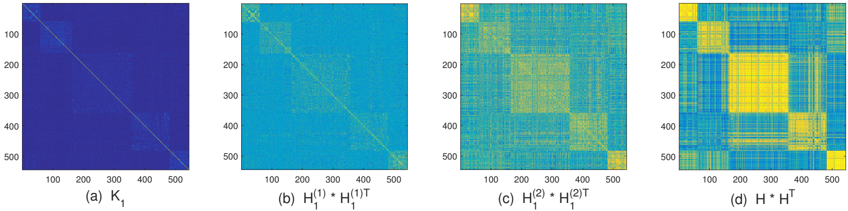

Figure 5: Visualization of kernel matrix, the proposed intermediary matrices and final partition of in HMKC-2 model.

Experiment Results

Performance Comparison

Table 2: Accuracy comparison of MSC algorithms in recent literature.

	Algorithm	AR10P	BBCSport	CCV	Flower17	Flower102	Heart	Ionosphere	Plant
	A-MKKM	38.46	65.99	19.74	51.03	27.29	82.22	61.25	61.70
	SB-KKM	49.23	76.65	20.08	42.06	33.13	76.30	70.09	51.60
	CSRC	36.92	67.65	23.40	51.69	35.19	80.37	75.78	54.57
	MKKM	37.69	66.36	18.01	45.37	21.96	53.33	63.82	56.38
	RMSC	34.62	86.03	16.29	52.57	32.97	83.33	84.62	55.53
	RMKC	43.85	66.36	19.74	52.35	33.55	82.22	66.10	61.70
	RMKKM	32.31	53.13	17.11	53.16	29.61	76.30	65.81	55.32
ACC	MKKM-MR	43.08	66.36	22.47	60.00	40.27	83.33	61.25	62.87
	ONKC	48.46	67.65	23.12	59.85	41.26	83.70	63.82	64.89
	LFAM	46.92	72.61	26.66	59.63	44.61	82.22	68.09	62.77
	SPC	38.46	83.09	-	57.50	-	75.93	71.23	60.43
	MVCC	43.85	74.26	22.10	51.47	37.23	82.96	55.56	55.64
	SPMKC	54.62	40.81	13.54	35.81	-	57.04	72.93	53.94
	HMKC-1	56.15	89.52	36.57	66.91	46.84	**86.67**	86.32	64.68
	HMKC-2	**60.00**	**90.99**	**37.37**	**71.18**	**50.32**	86.30	**86.89**	**67.02**

Experiment Results

Performance Comparison

Table 3: NMI comparison of MSC algorithms in recent literature.

	Algorithm	AR10P	BBCSport	CCV	Flower17	Flower102	Heart	Ionosphere	Plant
	A-MKKM	34.57	53.92	17.16	50.19	46.32	32.40	3.29	26.82
	SB-KKM	51.44	59.39	17.73	45.14	48.99	20.49	10.35	17.18
	CSRC	36.90	55.41	20.96	52.63	53.73	28.53	15.99	21.82
	MKKM	38.92	54.67	15.52	45.35	42.30	0.02	5.12	20.02
	RMSC	29.70	73.89	13.79	56.35	53.36	34.51	36.52	23.83
	RMKC	44.38	54.30	17.16	50.42	49.74	32.40	7.02	26.82
NMI	RMKKM	27.96	28.48	12.54	53.31	48.55	20.49	9.33	19.71
	MKKM-MR	42.62	54.67	18.62	57.11	57.38	35.22	3.29	28.29
	ONKC	51.86	54.74	19.19	56.85	57.39	36.22	31.13	31.13
	LFAM	47.11	54.99	19.85	57.83	57.58	32.40	1.15	27.64
	SPC	39.02	65.49	-	61.32	-	19.67	0.35	29.65
	MVCC	38.56	51.52	15.97	56.26	52.35	34.54	0.63	20.85
	SPMKC	58.39	7.31	8.11	40.78	-	10.95	1.64	19.28
	HMKC-1	54.46	72.73	31.64	63.37	60.40	42.98	39.11	34.11
	HMKC-2	**60.32**	**76.72**	**32.71**	**65.45**	**62.47**	**43.40**	**40.55**	**36.43**

Overview

Thanks !