Convergence and Recovery Guarantees of the K-Subspaces Method for Subspace Clustering

Peng Wang

Department of Electrical Engineering and Computer Science
University of Michigan, Ann Arbor

(Joint Work with Huikang Liu, Anthony Man-Cho So, and Laura Balzano)

Proceedings of the 39th International Conference on Machine Learning
(ICML 2022)

July 19, 2022

Outline

Introduction

Main Results

K-Subspaces Method (KSS)

- Applications in computer vision, image segmentation, and network analysis...[1]

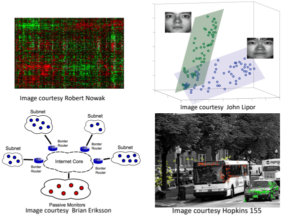
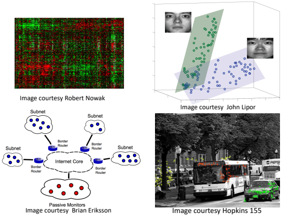


Image courtesy Robert Nowak
Image courtesy John Lipor
Image courtesy Brian Eriksson
Image courtesy Hopkins 155

[1] Source papers for images: (Top right) Lipor, J. and Balzano, L. "Clustering quality metrics for subspace clustering." Pattern Recognition 104 (2020): 107328; (Bottom left) Eriksson, B., Balzano, L., and Nowak, R. "High-rank matrix completion." In Artificial Intelligence and Statistics, pp. 373-381. PMLR, 2012.

K-Subspaces Method (KSS)

- Applications in computer vision, image segmentation, and network analysis...[1]



- KSS Formulation

$$\min_{\mathcal{C}, \boldsymbol{U}} \sum_{k=1}^{K} \sum_{i \in \mathcal{C}_k} \|\boldsymbol{z}_i - \boldsymbol{U}_k \boldsymbol{U}_k^T \boldsymbol{z}_i\|^2, \quad (1)$$

where $\boldsymbol{z}_i \in \mathbb{R}^n$ for all $i \in [N]$ denote data points, $\mathcal{C}_k$ for all $k \in [K]$ denote estimated clusters, and $\boldsymbol{U}_k$ denotes an orthonormal basis of the corresponding cluster.

[1] Source papers for images: (Top right) Lipor, J. and Balzano, L. "Clustering quality metrics for subspace clustering." Pattern Recognition 104 (2020): 107328; (Bottom left) Eriksson, B., Balzano, L., and Nowak, R. "High-rank matrix completion." In Artificial Intelligence and Statistics, pp. 373-381. PMLR, 2012.

KSS Method with TIPS Initialization

- Given the t-th iterate $(\mathcal{C}_1^t, \ldots, \mathcal{C}_K^t, \boldsymbol{U}_1^t, \ldots, \boldsymbol{U}_K^t)$, KSS alternates between
  - **Subspace update step**

$$\boldsymbol{U}_k^{t+1} = \mathrm{PCA}\left(\sum_{i \in \mathcal{C}_k^t} \boldsymbol{z}_i \boldsymbol{z}_i^T, \hat{d}_k\right),$$

where $\hat{d}_k$ for all $k \in [K]$ are candidate subspace dimensions.

KSS Method with TIPS Initialization

- Given the t-th iterate $(\mathcal{C}_1^t, \dots, \mathcal{C}_K^t, \boldsymbol{U}_1^t, \dots, \boldsymbol{U}_K^t)$, KSS alternates between
  - **Subspace update step**

$$\boldsymbol{U}_k^{t+1} = \mathrm{PCA}\left(\sum_{i \in \mathcal{C}_k^t} \boldsymbol{z}_i \boldsymbol{z}_i^T, \hat{d}_k\right),$$

    where $\hat{d}_k$ for all $k \in [K]$ are candidate subspace dimensions.
  - **Cluster assignment step**

$$i \in \mathcal{C}_k^{t+1}, \text{ if } k \in [K] \text{ satisfies } \|{\boldsymbol{U}_k^{t+1}}^T \boldsymbol{z}_i\| \geq \|{\boldsymbol{U}_\ell^{t+1}}^T \boldsymbol{z}_i\|, \ \forall\ \ell \neq k.$$

KSS Method with TIPS Initialization

- Given the t-th iterate $(\mathcal{C}_1^t, \ldots, \mathcal{C}_K^t, \boldsymbol{U}_1^t, \ldots, \boldsymbol{U}_K^t)$, KSS alternates between
  - **Subspace update step**

$$\boldsymbol{U}_k^{t+1} = \mathrm{PCA}\left(\sum_{i \in \mathcal{C}_k^t} \boldsymbol{z}_i \boldsymbol{z}_i^T, \hat{d}_k\right),$$

    where $\hat{d}_k$ for all $k \in [K]$ are candidate subspace dimensions.
  - **Cluster assignment step**

$$i \in \mathcal{C}_k^{t+1}, \text{ if } k \in [K] \text{ satisfies } \|{\boldsymbol{U}_k^{t+1}}^T \boldsymbol{z}_i\| \geq \|{\boldsymbol{U}_\ell^{t+1}}^T \boldsymbol{z}_i\|, \ \forall\ \ell \neq k.$$

- KSS: a generalization of the **k-means method**, handle clusters in subspaces.

KSS Method with TIPS Initialization

- Given the t-th iterate $(\mathcal{C}_1^t, \ldots, \mathcal{C}_K^t, \boldsymbol{U}_1^t, \ldots, \boldsymbol{U}_K^t)$, KSS alternates between
  - **Subspace update step**

$$\boldsymbol{U}_k^{t+1} = \mathrm{PCA}\left(\sum_{i \in \mathcal{C}_k^t} \boldsymbol{z}_i \boldsymbol{z}_i^T, \hat{d}_k\right),$$

    where $\hat{d}_k$ for all $k \in [K]$ are candidate subspace dimensions.
  - **Cluster assignment step**

$$i \in \mathcal{C}_k^{t+1}, \text{ if } k \in [K] \text{ satisfies } \|\boldsymbol{U}_k^{t+1^T} \boldsymbol{z}_i\| \geq \|\boldsymbol{U}_\ell^{t+1^T} \boldsymbol{z}_i\|, \; \forall\, \ell \neq k.$$

- KSS: a generalization of the **k-means method**, handle clusters in subspaces.
- **Initialization method:** Thresholding inner-product based spectral (TIPS) method

Outline

Master Theorems (Informal)

- **Theorem 1.** Suppose that the following hold:
  - (i) (Data input) Let $\{\boldsymbol{z}_i\}_{i=1}^N$ be generated by the UoS model, i.e., $\boldsymbol{z}_i = \boldsymbol{U}_k^* \boldsymbol{a}_i$ for $i \in \mathcal{C}_k^*$, where $\boldsymbol{a}_i \overset{i.i.d.}{\sim} \mathrm{Unif}(\mathbb{S}^{d_k - 1})$ for all $k \in [K]$.
  - (ii) (Affinity requirement) $\mathrm{aff}(S_k^*, S_\ell^*)/\min\{\sqrt{d_k}, \sqrt{d_\ell}\} \leq 1/2$ for all $1 \leq k \neq \ell \leq K$, where $\mathrm{aff}(S_k^*, S_\ell^*) := \|{\boldsymbol{U}_k^*}^T \boldsymbol{U}_\ell^*\|_F$.
  - (iii) (Sampling requirement) $N_{\min} \gtrsim d_k \gtrsim \log N$ for all $k \in [K]$.
  - (iv) (Initial requirement) The initial assignment $\boldsymbol{H}^0 \in \mathcal{M}^{N \times K}$ satisfies

$$d_F(\boldsymbol{H}^0, \boldsymbol{H}^*) \lesssim \frac{N_{\min}}{\sqrt{N}}. \tag{2}$$

  The following statement holds with probability at least $1 - N^{-\Omega(1)}$: The KSS method converges superlinearly and finds the true partition within $\Theta(\log \log N)$ iterations.

Master Theorems (Informal)

- **Theorem 1.** Suppose that the following hold:
  (i) (Data input) Let $\{\boldsymbol{z}_i\}_{i=1}^N$ be generated by the UoS model, i.e., $\boldsymbol{z}_i = \boldsymbol{U}_k^* \boldsymbol{a}_i$ for $i \in \mathcal{C}_k^*$, where $\boldsymbol{a}_i \overset{i.i.d.}{\sim} \mathrm{Unif}(\mathbb{S}^{d_k-1})$ for all $k \in [K]$.
  (ii) (Affinity requirement) $\mathrm{aff}(S_k^*, S_\ell^*)/\min\{\sqrt{d_k}, \sqrt{d_\ell}\} \le 1/2$ for all $1 \le k \ne \ell \le K$, where $\mathrm{aff}(S_k^*, S_\ell^*) := \|\boldsymbol{U}_k^{*T} \boldsymbol{U}_\ell^*\|_F$.
  (iii) (Sampling requirement) $N_{\min} \gtrsim d_k \gtrsim \log N$ for all $k \in [K]$.
  (iv) (Initial requirement) The initial assignment $\boldsymbol{H}^0 \in \mathcal{M}^{N\times K}$ satisfies

$$d_F(\boldsymbol{H}^0, \boldsymbol{H}^*) \lesssim \frac{N_{\min}}{\sqrt{N}}. \quad (2)$$

  The following statement holds with probability at least $1 - N^{-\Omega(1)}$: The KSS method converges superlinearly and finds the true partition within $\Theta(\log\log N)$ iterations.

- **Theorem 2.** It holds with probability at least $1 - N^{-\Omega(1)}$ that the TIPS method can return a qualified initial point that satisfies (2).

Comments on the Master Theorems

- While the KSS formulation is NP-hard in the worst case, the assumption that $\{\boldsymbol{z}_i\}_{i=1}^N$ arises from the semi-random UoS model allows us to conduct an average-case analysis.

Comments on the Master Theorems

- While the KSS formulation is NP-hard in the worst case, the assumption that $\{\boldsymbol{z}_i\}_{i=1}^N$ arises from the semi-random UoS model allows us to conduct an average-case analysis.
- A neighborhood of size $O\left(\frac{N_{\min}}{\sqrt{N}}\right)$ around each true cluster forms a *basin of attraction* in the UoS model, in which the KSS method converges superlinearly.

Comments on the Master Theorems

- While the KSS formulation is NP-hard in the worst case, the assumption that $\{z_i\}_{i=1}^N$ arises from the semi-random UoS model allows us to conduct an average-case analysis.
- A neighborhood of size $O\left(\frac{N_{\min}}{\sqrt{N}}\right)$ around each true cluster forms a *basin of attraction* in the UoS model, in which the KSS method converges superlinearly.
- Any method that can return a point satisfying (2) is qualified as an initialization scheme for the KSS method. In this work, we design a TIPS method that can provably generate a qualified point.

Thank You!