Kernel Normalized Cut: a Theoretical Revisit

*Yoshikazu Terada*\(^1,3\) & Michio Yamamoto\(^2,3\)

\(^1\)Graduate School of Engineering Science, Osaka University
\(^2\)Graduate School of Environmental and Life Science, Okayama University
\(^3\)RIKEN Center for Advanced Intelligence Project (AIP)

Unsupervised Learning (Room 103)
12:05 - 12:10, Jun 13, 2019 (Thu)
ICML2019@Long Beach
What is Normalized cut?

- Normalized cut (Ncut; Shi and Malik, 2000)
  - **Ncut** = Graph partitioning method
  - **Goal** = To find “clusters” in the graph:
    - Many edges inside the cluster
    - Fewer edges between different clusters
  - **Ncut** = Balanced cut
    - Each cluster is “reasonably large”!
    - Cut between different clusters is small.

- **Objective function of Ncut** (Number of clusters = 2)
  - \( K := (k_{ij})_{n \times n} : \) Similarity matrix, \( d_i := \sum_{i=1}^{n} k_{ij}, \) \( \text{vol}(A) := \sum_{i \in A} d_i, \)
  - Min cut: \( \text{Mcut}(A, B) := \sum_{i \in A} \sum_{j \in B} k_{ij} \)

\[
\text{Ncut}(A, B) = \text{Mcut}(A, B) \left\{ \frac{1}{\text{vol}(A)} + \frac{1}{\text{vol}(B)} \right\}
\]
Normalized cut and its related methods

- Normalized cut, Spectral clustering, Weighted kernel $k$-means
  - Ncut is an NP hard problem $\Rightarrow$ Normalized Spectral clustering (SC) $\Rightarrow$ Continuous relaxation of Ncut

- Ncut and Weighted Kernel $K$-Means (WKKM) (Dhillon et al., 2007)
  - WKKM with kernel $h$ and weight $w_i : H = (h_{ij})_{n \times n}$, $W = \text{diag}(w_1, \ldots, w_n)$
    \[ \sum_{i=1}^{n} w_i \min_m \| \psi_h(X_i) - \mu_m \|_h^2 = \text{Const.} - \text{tr}(\tilde{U}^T W^{1/2} H W^{1/2} \tilde{U}) \]
  - Ncut $\Rightarrow$ WKKM with $H = D^{-1} K D^{-1}$ and $W = D$ ($D = \text{diag}(d_1, \ldots, d_n)$)

- Setting

- Data points
- Similarity matrix
- Clustering result!
Overview of this study

We study theoretical properties of clustering based on Ncut!

- Theoretical properties of Ncut
  - Weighted KM in $n$-dim. space
  - Ncut for data points
  - Norm. SC for data points
  - Norm. graph Laplacian (eigenvector)

- Empirical
  - Dhillon et al. (2007, IEEE PAMI)
  - Shi and Malik (2000, IEEE PAMI)

- This study
  - W. KM in $n$-dim. space
  - Ncut for data points
  - Norm. SC for data points
  - von Luxburg et al. (2008, AoS)

- Population
  - Weighted KM in RKHS
  - Ncut for population distribution
  - Optimality of the partition is not clear
  - Limit operator in func. space (eigenfunction)

- We also derive the fast rate of convergence of the normalized cut!
Numerical experiments

Note that we used the same tuning parameter in both Ncut and SC!