Self-Tuning Spectral Clustering

Jing Guo(jg3527)

COLUMBIA | ENGINEERING
The Fu Foundation School of Engineering and Applied Science

Motivation

$$\boxed{A_{ij} = \exp(-||s_i - s_j||^2/2\sigma^2)}$$

- **Algorithm for Spectral Clustering:**

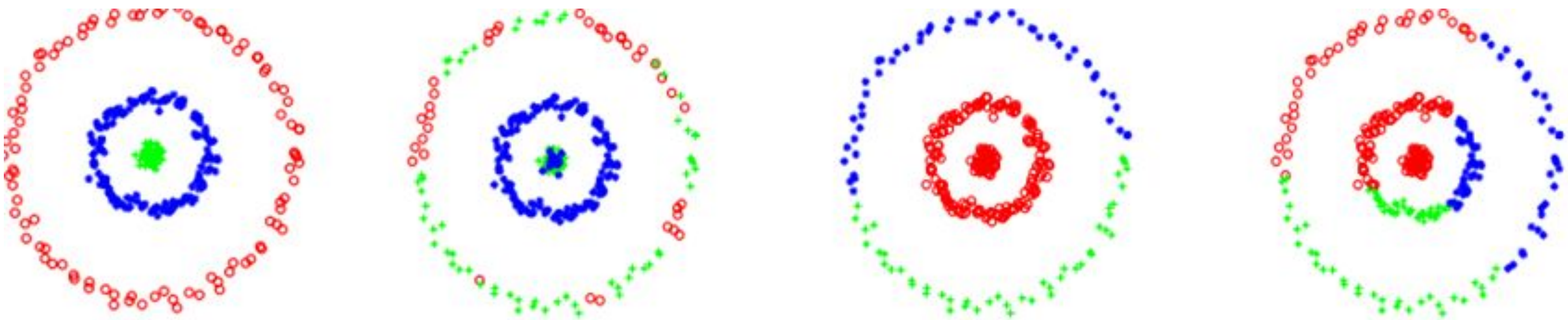
Given a set of points $S = \{s_1, \ldots, s_n\}$ in $\mathbb{R}^l$ that we want to cluster into k subsets:

1. Form the affinity matrix $A \in \mathbb{R}^{n \times n}$ defined by $\boxed{A_{ij} = \exp(-||s_i - s_j||^2/2\sigma^2)}$ if $i \neq j$, and $A_{ii} = 0$.
2. Define D to be the diagonal matrix whose (i, i)-element is the sum of A's i-th row, and construct the matrix $L = D^{-1/2}AD^{-1/2}$.[1]
3. Find $x_1, x_2, \ldots, x_k$, the k largest eigenvectors of L (chosen to be orthogonal to each other in the case of repeated eigenvalues), and form the matrix $X = [x_1 x_2 \ldots x_k] \in \mathbb{R}^{n \times k}$ by stacking the eigenvectors in columns.
4. Form the matrix Y from X by renormalizing each of X's rows to have unit length (i.e. $Y_{ij} = X_{ij}/(\sum_j X_{ij}^2)^{1/2}$).
5. Treating each row of Y as a point in $\mathbb{R}^k$, cluster them into k clusters via K-means or any other algorithm (that attempts to minimize distortion).
6. Finally, assign the original point s_i to cluster j if and only if row i of the matrix Y was assigned to cluster j.

- **The optimal σ does not exist if the input data includes clusters with different local scales.**
- **Want to estimate the Number of Clusters automatically.**

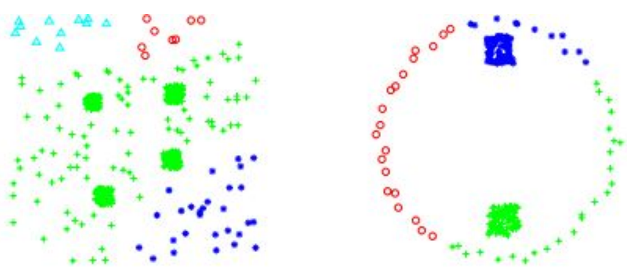
- σ has high impact on clustering quality.

σ=0.041235 σ=0.015625 σ=0.35255 σ=1



- Spectral clustering without local scaling.

σ=0.035897 σ=0.03125



Clustering with local scaling

- Instead of selecting a single scaling parameter σ we propose to calculate a local scaling parameter σ_i for each data point s_i.
- The affinity between a pair of points can thus be written as:

$$\hat{A}_{ij} = \exp\left(\frac{-d^2(s_i, s_j)}{\sigma_i \sigma_j}\right)$$

Instead of

$$A_{ij} = \exp\left(\frac{-d^2(s_i, s_j)}{\sigma^2}\right)$$

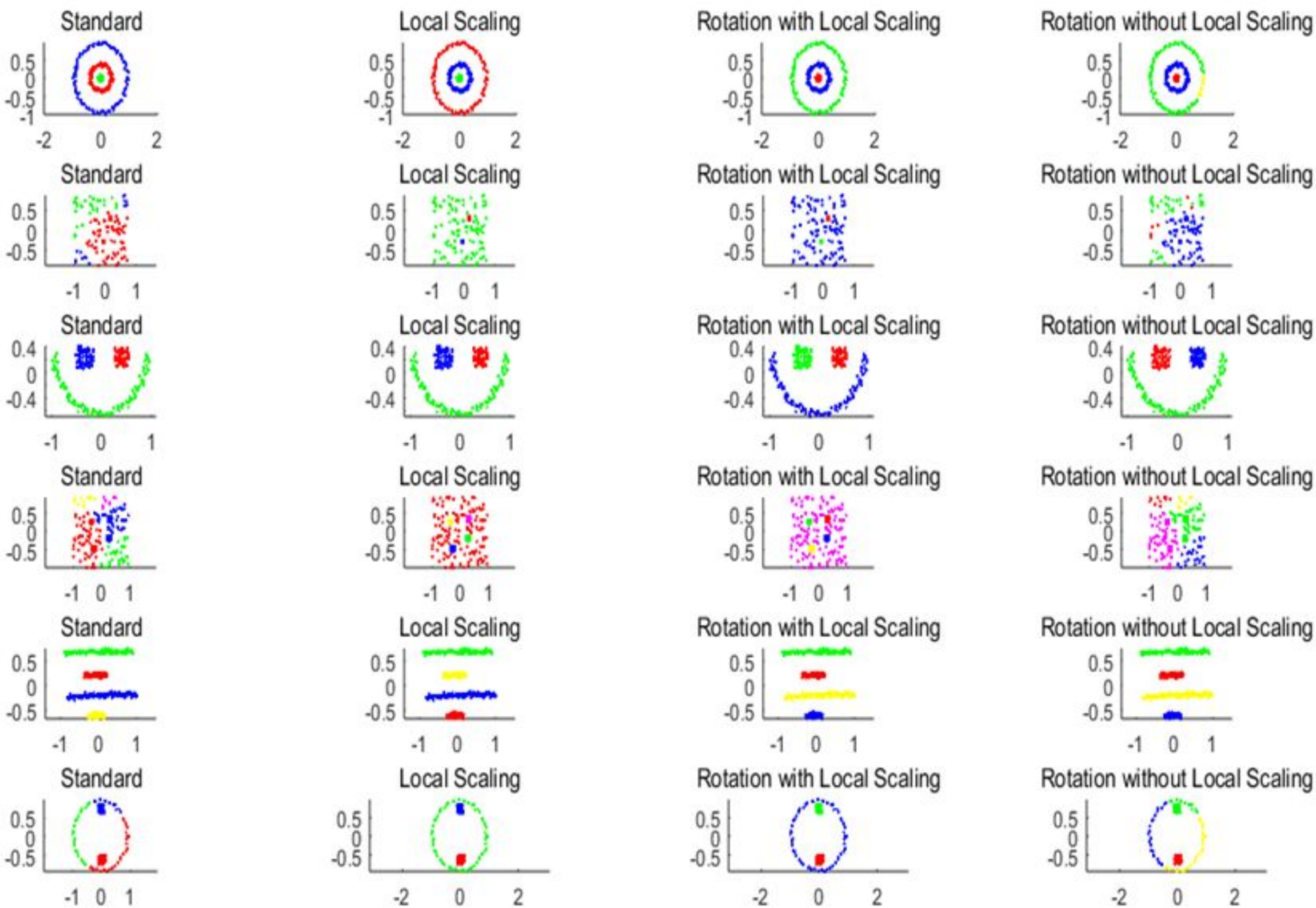
Estimating the Number of Clusters Automatically.

- For each possible group number C we recover the rotation which best aligns X's columns with the canonical coordinate system. Let $Z \in R^{n \times C}$ be the matrix obtained after rotating the eigenvector matrix X, i.e., $Z = XR$ and denote $M_i = \max_j Z_{ij}$. We wish to recover the rotation R for which in every row in Z there will be at most one non-zero entry. We thus define a cost function:

$$J = \sum_{i=1}^{n} \sum_{j=1}^{C} \frac{Z_{ij}^2}{M_i^2}$$

- Minimizing this cost function over all possible rotations will provide the best alignment with the canonical coordinate system.
- Set the final group number C_{best} to be the largest group number with minimal alignment cost.

Experiments



Future Work

- Try to find better substitution of σ^2 to improve local scaling.

- Test this algorithm with images and audio files.

Reference

- Ng A Y, Jordan M I, Weiss Y. On spectral clustering: Analysis and an algorithm[J]. Advances in neural information processing systems, 2002, 2: 849-856.
- Zelnik-Manor L, Perona P. Self-tuning spectral clustering[C]//Advances in neural information processing systems. 2004: 1601-1608.
- http://www.vision.caltech.edu/lihi/Demos/SelfTuningClustering.html