Sparse Subspace Clustering (SSC)

Javier A. Salazar Cavazos

University of Michigan

June 15, 2022

Outline

1. Background
   - Motivation, Subspaces, & Clustering
   - Subspace Clustering, Challenges, & Literature Review
2. Sparse Subspace Clustering
   - Affinity Matrix Construction
   - Spectral Clustering
   - Results
3. References & Questions

Motivation

Face Clustering [1]

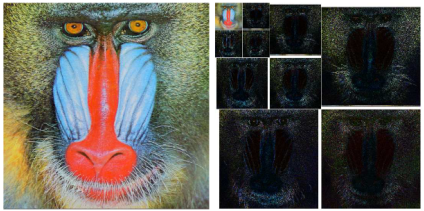


Image Compression [2]

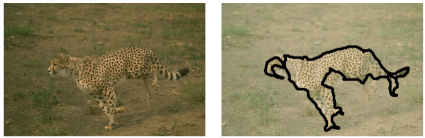
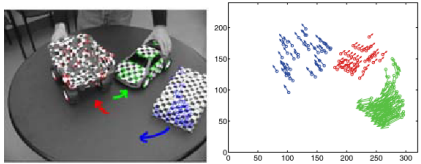


Image Segmentation [3]



Motion Estimation [4]

Image Credit: Respective Papers

Subspaces

For a vector space $\mathcal{V}$ defined on a field $\mathbb{F}$, a nonempty set $\mathcal{S} \subseteq \mathcal{V}$ is called a *linear* subspace iff

- S is closed under vector addition (i.e. for $\boldsymbol{u}, \boldsymbol{v} \in \mathcal{S} \implies \boldsymbol{u} + \boldsymbol{v} \in \mathcal{S}$)
- S is closed under scalar multiplication (i.e. $\boldsymbol{v} \in \mathcal{S} \implies \alpha \boldsymbol{v} \in \mathcal{S}$)

Key Points:

- Every *linear* subspace includes $\mathbf{0}$
- A shifted *linear* subspace is called an affine subspace
- Left singular matrix contains subspace basis

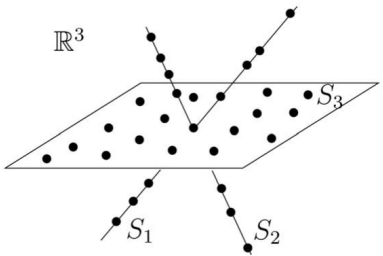


Image Credit: René Vidal @ Johns Hopskins

Clustering

An unsupervised machine learning method to identify and group "similar" unlabeled data points in order to find structure or patterns

By "similar", this can mean different things:

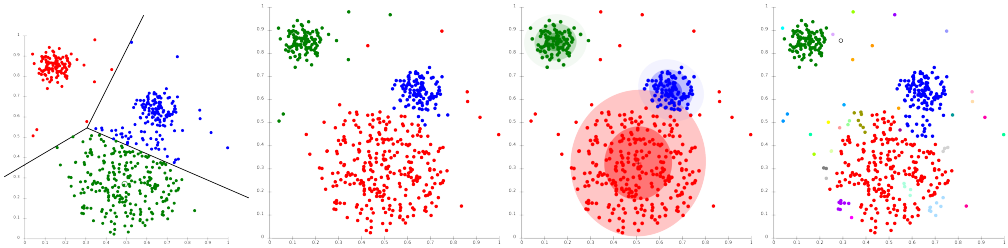
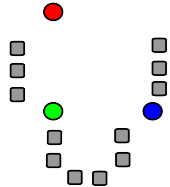
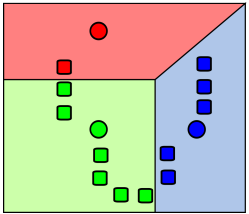


Image Credit: Chire @ Wikipedia, CC BY-SA 3.0

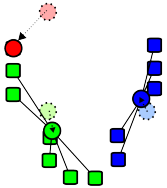
K-Means Clustering



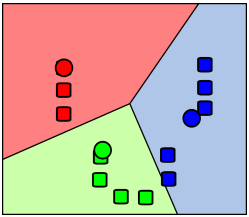
1. Initialize centroids



2. Generate clusters



3. Update centroids



4. Repeat steps 2 & 3

Note!

NP-hard problem $\Longrightarrow$ initialization is extremely important!

Image Credit: Weston.pace @ Wikipedia, CC BY-SA 3.0

Subspaces + Clustering

n = # of points, k = # of subspaces, D = dim($\boldsymbol{y}$), d_i = dim($\mathcal{S}_i$)

Points:
$Y = [\, \boldsymbol{y}_1, ..., \boldsymbol{y}_n \,]$ for $\{\, \boldsymbol{y}_j \in \mathbb{R}^D \,\}_{j=1}^{n}$ drawn from $\{\mathcal{S}_i\}_{i=1}^{k}$

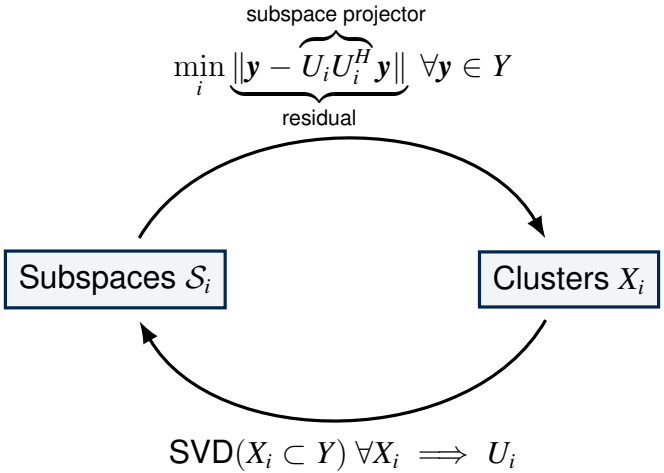
Subspaces:
$\mathcal{S}_i = \{\, \boldsymbol{y} \in \mathbb{R}^D : \boldsymbol{y} = U_i \boldsymbol{z} \,\}$ where $\boldsymbol{z}$ is coordinate in basis U_i

Goal:
Find subspace bases $U_i \in \mathbb{R}^{D \times d_i}$ and segmentation $X_i \subset Y, X_i \in \mathcal{S}_i$

Recall!

The SVD can only return one subspace basis ($k = 1$ case) for Y
How do we solve this when $k > 1$?

The Chicken & Egg Dilemma



Literature Review [5]

Algebraic Methods (e.g. Generalized PCA)

- Geometric and algebraic interpretations

Iterative Methods (e.g. K-Subspaces)

- Alternate between segmentation and subspaces

Statistical Methods (e.g. Mixture of Probabilistic PCA)

- Model assumptions combined with prior beliefs

Affinity-based Methods (e.g. Sparse Subspace Clustering)

- Measures “similarity” between points

Outline

1. Background
   - Motivation, Subspaces, & Clustering
   - Subspace Clustering, Challenges, & Literature Review
2. Sparse Subspace Clustering
   - Affinity Matrix Construction
   - Spectral Clustering
   - Results
3. References & Questions

SSC Affinity Measure

Self-Expressiveness Property

Each data point in a union of subspaces can be efficiently reconstructed by a combination of other points in the dataset.

$$\text{For } \boldsymbol{y}_i \in \bigcup_{l=1}^{n} \mathcal{S}_l, \quad \text{self-expressive: } \boldsymbol{y}_i = \underbrace{Y\boldsymbol{c}_i}_{\text{not unique!}} \quad \text{where } \underbrace{c_{ii} = 0}_{\boldsymbol{y}_i \neq \boldsymbol{y}_i \text{ (trivial)}}$$

self-expressiveness $>>$ other affinity measures

Recall that $\boldsymbol{u}, \boldsymbol{v} \in \mathcal{S} \implies \boldsymbol{u} + \boldsymbol{v} \in \mathcal{S}$
self-expressiveness $=$ natural result of subspace definition

Question: Why is $\boldsymbol{c}_i$ not unique in general?

Sparse Subspace Clustering [6]

Nonconvex formulation (vector case):

$$\min_{\boldsymbol{c}_i} \|\boldsymbol{c}_i\|_0 \quad \text{s.t.} \quad \boldsymbol{y}_i = Y\boldsymbol{c}_i, c_{ii} = 0$$

Convex formulation (matrix case):

$$\min_{C} \underbrace{\|C\|_{1,1}}_{\|C\|_{1,1} \triangleq \|\text{vec}(C)\|_1} \quad \text{s.t.} \quad Y = YC, \text{diag}(C) = \mathbf{0}$$

Convex formulation (matrix case + outliers + noise):

$$\min_{C,E,Z} \overbrace{\|C\|_{1,1}}^{\text{sparse affinity}} + \overbrace{\lambda_e \|E\|_{1,1}}^{\text{sparse outliers}} + \frac{\lambda_z}{2} \overbrace{\|Z\|_F^2}^{\text{noise}} \quad \text{s.t.} \quad Y = YC + E + Z, \overbrace{\text{diag}(C) = \mathbf{0}}^{C \neq I} \quad (1)$$

Optimization - ADMM

Introduce auxiliary variable $A = C - \text{mdiag}(C)$ where $\text{mdiag}(C) = \text{diagm}(\text{diag}(C))$ to decouple terms such that:

$$\min_{C,E,A} \|C\|_{1,1} + \lambda_e \|E\|_{1,1} + \frac{\lambda_z}{2} \| \underbrace{Y - YA - E}_{Z} \|_F^2 \quad \text{s.t.} \quad A = C - \text{mdiag}(C)$$

For dual variable Λ, we have augmented Lagrangian function as follows:

$$\mathcal{L}_\mu(C, E, A, \Lambda) = \|C\|_{1,1} + \lambda_e \|E\|_{1,1} + \frac{\lambda_z}{2} \|Y - YA - E\|_F^2 + \langle \Lambda, A - C + \text{mdiag}(C) \rangle + \frac{\mu}{2} \|A - C + \text{mdiag}(C)\|_F^2$$

Optimization - More ADMM

$$E = \min_E \mathcal{L}_\mu(\cdot) = \min_E \|E\|_{1,1} + \frac{1}{2(\lambda_e \lambda_z^{-1})}\|E - (Y - YA)\|_F^2 = \text{prox}_{\lambda_e \lambda_z^{-1}\|\cdot\|_{1,1}}(Y - YX)$$

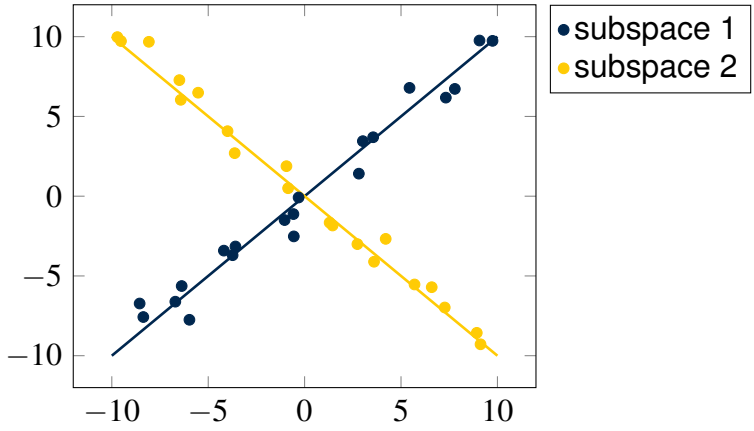
$$\begin{aligned} A &= \min_A \mathcal{L}_\mu(\cdot) = \min_A \frac{\lambda_z}{2}\|Y - YA - E\|_F^2 + \frac{\mu}{2}\|A - C + \text{mdiag}(C) + \frac{1}{\mu}\Lambda\|_F^2 \\ &= (\lambda_z Y^T Y + \mu I)^{-1}[\lambda_z Y^T (Y - E) + \mu(C - \text{mdiag}(C)) - \Lambda] \end{aligned}$$

$$\begin{aligned} C &= \min_C \mathcal{L}_\mu(\cdot) = \min_C \|C\|_{1,1} + \frac{\mu}{2}\|A - C + \text{mdiag}(C) + \frac{1}{\mu}\Lambda\|_F^2 \\ &= J - \text{mdiag}(J) \text{ where } J = \text{prox}_{\mu^{-1}\|\cdot\|_{1,1}}(A + \frac{1}{\mu}\Lambda) \end{aligned}$$

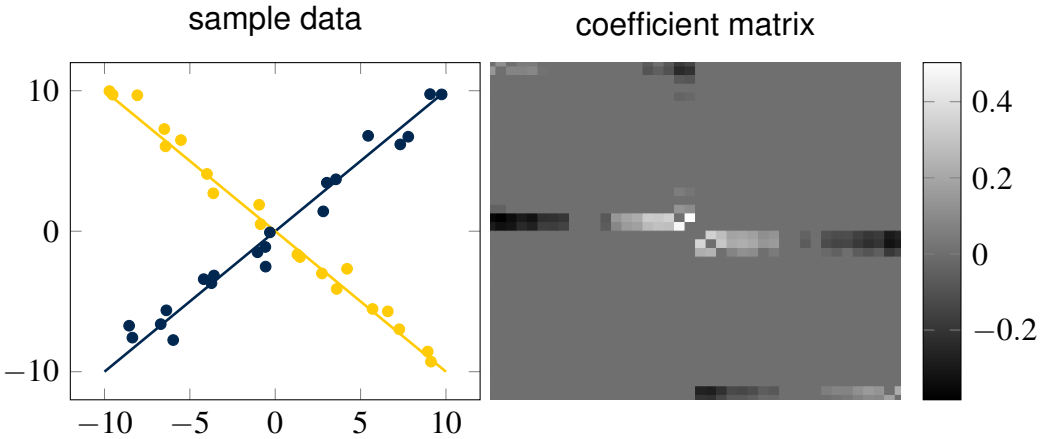
$$\Lambda = \Lambda + \mu(A - C + \text{mdiag}(C))$$

Example Problem

Let $Y = [\underbrace{\boldsymbol{y}_1^{(1)}, ..., \boldsymbol{y}_t^{(1)}}_{\text{subspace 1}}, \underbrace{\boldsymbol{y}_{t+1}^{(2)}, ..., \boldsymbol{y}_n^{(2)}}_{\text{subspace 2}}]$ where $\boldsymbol{y}_i \in \mathbb{R}^2 \ \ \forall i, n = 40$ pts



Example Solution



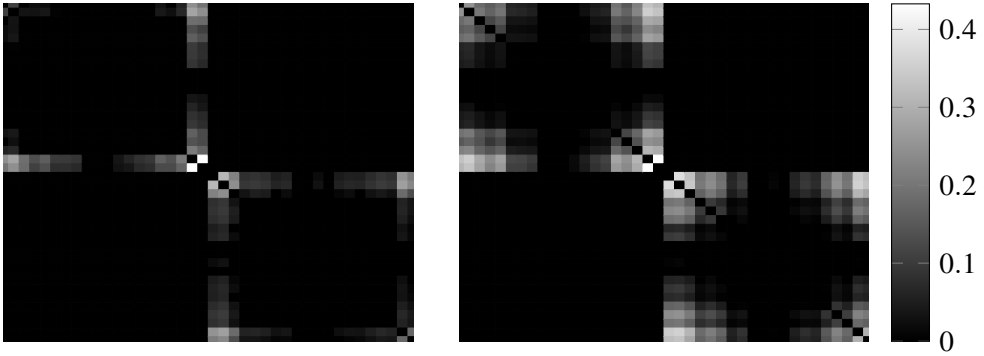
Question: What points did SSC select? What happens near the origin?

Example Solution

Note: Plots shown are for $W = |C| + |C|^T$ for visualization convenience

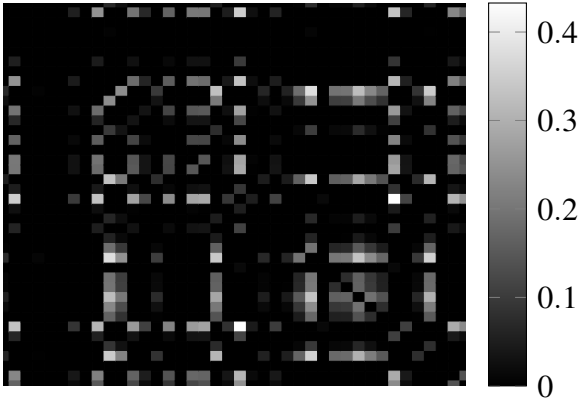
high sparsity

low sparsity



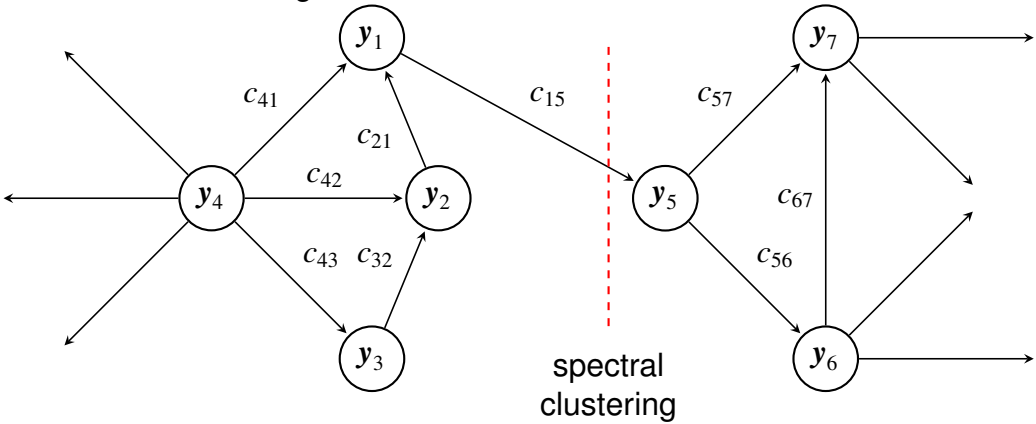
The Problem

$$Y = [\, \boldsymbol{y}_1^{(1)}, \boldsymbol{y}_2^{(2)}, \boldsymbol{y}_3^{(2)}, \boldsymbol{y}_4^{(1)}, \boldsymbol{y}_5^{(1)}, \boldsymbol{y}_6^{(2)}, \dots \,] \implies C \text{ below}$$



Spectral Clustering [7]

Data points $\Longrightarrow$ nodes/vertices
coefficients $\Longrightarrow$ edges



Graph Notation

$\mathcal{G}(V, C) \implies$ directed graph
Let $W = |C| + |C|^T$ s.t. $\mathcal{G}(V, W) \implies$ undirected graph
W is not self-expressive ($Y \neq YW$) but this is fine

Direction does not matter!

- $\boldsymbol{y}_i$ may use $\boldsymbol{y}_j$ but the other way is not guaranteed for C
- W ensures both nodes are connected together
- connections tell the big picture! (w.r.t. clustering)

Let $D_W = \text{diag}(d_1, ..., d_n)$ where $d_i = \sum_{j=1}^{n} w_{ij}$ (degree of nodes)

- If $\boldsymbol{y}_i$ shows up many times $\implies$ high degree node

Graph Laplacians

Let $L = D - W$ be the unnormalized Laplacian matrix

Proposition 1

1. For $\boldsymbol{f} \in \mathbb{R}^n$, we have that:
$$\boldsymbol{f}^T L \boldsymbol{f} = \frac{1}{2} \sum_{i,j=1}^{n} w_{ij}(f_i - f_j)^2$$
2. L is symmetric and positive semi-definite
3. $\lambda_{\text{min}}(L) = 0$ and $\nu_{\text{min}} = \mathbf{1}$
4. $0 = \lambda_1 \leq \lambda_2 \leq ... \leq \lambda_n$

Graph Theory Primer

Let $A_1, ..., A_k$ form a partition on $\mathcal{G}(V, E)$
Let $W(A, B) = \sum_{i \in A, j \in B} w_{ij}$

Then, our problem becomes the following:

$$\min_{A_1,...,A_k} \text{cut}(A_1, ..., A_k) = \frac{1}{2} \sum_{i=1}^{k} W(A_i, \bar{A}_i)$$

To make things "balanced", consider this:

$$\min_{A_1,...,A_k} \text{RatioCut}(A_1, ..., A_k) = \frac{1}{2} \sum_{i=1}^{k} \frac{W(A_i, \bar{A}_i)}{|A_i|}$$

RatioCut

1. Define indicator vectors as $h_{i,j} = \{1/\sqrt{|A_j|}$ if $V_i \in A_j, 0$ if $V_i \notin A_j\}$
2. Collect vectors into matrix $H \in \mathbb{R}^{n \times k}$ (k clusters)
3. Observe that vectors are orthonormal (i.e. $H^T H = I$)
4. Without proof, it can be shown that $h_i^T L h_i = (H^T L H)_{ii} = \text{cut}(A_i, \bar{A}_i)/|A_i|$

Thus all together, RatioCut can be written as follows:

$$\text{RatioCut}(A_1, ..., A_k) = \sum_{i=1}^{k} \frac{\text{cut}(A_i, \bar{A}_i)}{|A_i|} = \sum_{i=1}^{k} (H^T L H)_{ii} = \text{Tr}(H^T L H)$$

Note: For simplicity, this is assuming unnormalized Laplacian.

More RatioCut

Our optimization problem now becomes this:

$$\min_{A_1,...,A_k} \text{RatioCut}(A_1, ..., A_k) = \text{Tr}(H^T L H) \text{ s.t. } H^T H = I$$

We make this convex by allowing arbitrary real values for H:

$$\min_{H \in \mathbb{R}^{n \times k}} \text{Tr}(H^T L H) \text{ s.t. } H^T H = I$$

The solution is $\hat{H} = [\, \nu_1(L), ..., \nu_k(L) \,]$ assuming $\lambda_1 \leq ... \leq \lambda_n$

Big Picture!

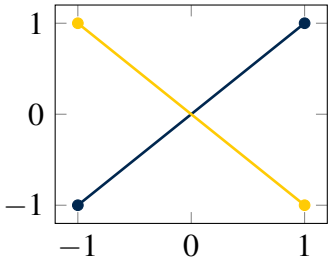
Spectral clustering is a relaxation to various graph cutting problems

Spectral Embedding

For simplicity, assume $n = 4$ and $k = 2$.

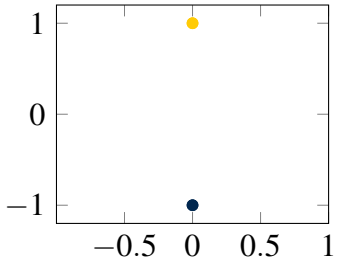
Recall $h_i = \{\sqrt{\frac{|\bar{A}|}{|A|}}$ if $V_i \in A, -\sqrt{\frac{|A|}{|\bar{A}|}}$ if $V_i \in \bar{A}\} \implies h^T L h = \text{cut}(A, \bar{A})/|A|$

sample data



$$h = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} \implies$$

spectral embedding



Question: Why is it difficult to cluster sample data? Is the right plot easier?

Spectral Clustering

Algorithm SpectralClustering(W, k)

Require: $W \in \mathbb{R}^{n \times n}, k \in \mathbb{N}$ ▷ W = affinity matrix, k = # of subspaces

$D = 0 \in \mathbb{R}^{n \times n}$

$D_{i,i} \leftarrow \sum_{j=1}^{n} w_{ij} \ \forall i$ ▷ node degree matrix

$L \leftarrow I - D^{-1}W$ ▷ normalized laplacian

$E \leftarrow$ EVD(L).vectors ▷ eigenvalue decomposition

$\nu \leftarrow E_{:,1:k}$ ▷ assuming smallest to largest λ

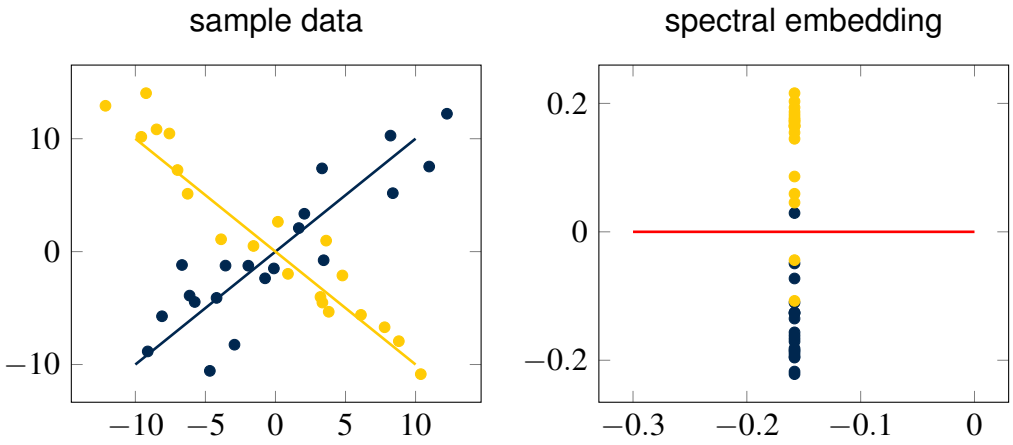
$X_1, ..., X_k \leftarrow$ kmeans(ν^T, k) ▷ cluster assignments $\forall$ pts

return $X_1, ..., X_k \subset Y, \ \ X_i \in \mathcal{S}_i$

Note!

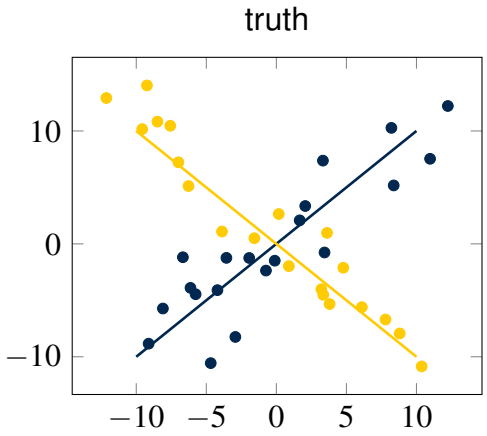
Use partial decomposition methods (e.g. power method) for big data

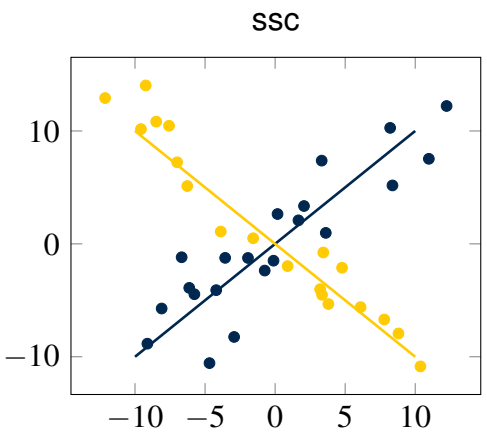
Example Problem



Question: Conceptually, which points will fail to be classified correctly?

Example Problem





SSC Overview

Algorithm SSC(Y, k)

Require: $Y \in \mathbb{R}^{d \times n}, k \in \mathbb{N}$ ▷ Y = data matrix, k = # of subspaces

$C \leftarrow$ ADMM$(Y, \boldsymbol{\lambda}, \mu)$ ▷ solve optimization EQ1

$C_{:,i} \leftarrow C_{:,i} / \|C_{:,i}\|_\infty \; \forall i$ ▷ normalize edge weights

$W \leftarrow |C| + |C|^T$ ▷ undirected graph matrix

$\{X_1, ..., X_k\} \leftarrow$ SpectralClustering(W, k) ▷ see pseudocode 1

for $i = 1, ..., k$ **do** ▷ for all clusters

  $U_i \leftarrow$ PCA$(X_i \subset Y)$ ▷ use SVD, weighted PCA, etc...

end for

return $\{X_1, ..., X_k\}, \{U_1, ..., U_k\}$ ▷ return clustered points and subspaces

Results

SSC: $\|E\|_{1,1} + \|C\|_{1,1} + \|Z\|_F^2 \quad \text{s.t.} \quad Y = YC + E + Z$

LRR: $\|E\|_{1,1} + \|C\|_* + \|Z\|_F^2 \quad \text{s.t.} \quad Y = YC + E + Z$

LRSC: $\|E\|_{1,1} + \|C\|_* + \|X - Y\|_F^2 + \|Z\|_F^2 \quad \text{s.t.} \quad X = XC + E + Z$

LSA: Compute k-NN for $\boldsymbol{y}_j$, apply PCA to get $\hat{\mathcal{S}}_j$, get affinity by doing $A_{jk} = \exp\left[-\sum_{m=1}^{\min(d_j,d_k)} (\sin\theta_{jk}^m)^2\right]$ where θ is principal angle

SCC: Randomly take points,form volume of the simplex formed by the points, and get affinity by taking advantage of polar curvature

Dataset:
Extended Yale B dataset $\rightarrow$ face clustering

Results

Clustering Error (%)					
Mean Value	Algorithms				
	LSA	SCC	LRR	LRSC	SSC
2 subjects	32.80	16.62	9.52	5.32	**1.86**
3 subjects	52.29	38.16	19.52	8.47	**3.10**
5 subjects	58.02	58.90	34.16	6.90	**4.31**
8 subjects	59.19	66.11	41.19	23.72	**5.85**
10 Subjects	60.42	73.02	38.85	30.36	**10.94**

Note: No pre-processing of the data for these results

End

Thanks for listening!

Questions?

References

[1] J. Ho, M.-H. Yang, J. Lim, K.-C. Lee, and D. Kriegman, "Clustering appearances of objects under varying illumination conditions," in *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings.*, IEEE, vol. 1, 2003, pp. I–I.

[2] W. Hong, J. Wright, K. Huang, and Y. Ma, "Multiscale hybrid linear models for lossy image representation," *IEEE Transactions on Image Processing*, vol. 15, no. 12, pp. 3655–3671, 2006.

[3] A. Y. Yang, J. Wright, Y. Ma, and S. S. Sastry, "Unsupervised segmentation of natural images via lossy data compression," *Computer Vision and Image Understanding*, vol. 110, no. 2, pp. 212–225, 2008.

[4] R. Vidal, R. Tron, and R. Hartley, "Multiframe motion segmentation with missing data using powerfactorization and gpca," *International Journal of Computer Vision*, vol. 79, no. 1, pp. 85–105, 2008.

[5] R. Vidal, "Subspace clustering," *IEEE Signal Processing Magazine*, vol. 28, no. 2, pp. 52–68, 2011.

[6] E. Elhamifar and R. Vidal, "Sparse subspace clustering: Algorithm, theory, and applications," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 11, pp. 2765–2781, 2013.

[7] U. Von Luxburg, "A tutorial on spectral clustering," *Statistics and computing*, vol. 17, no. 4, pp. 395–416, 2007.

Proposition 2

Definition - Connected Components

For $\mathcal{G}(V, E)$, nonempty set $A \subset V$ is a connected component iff

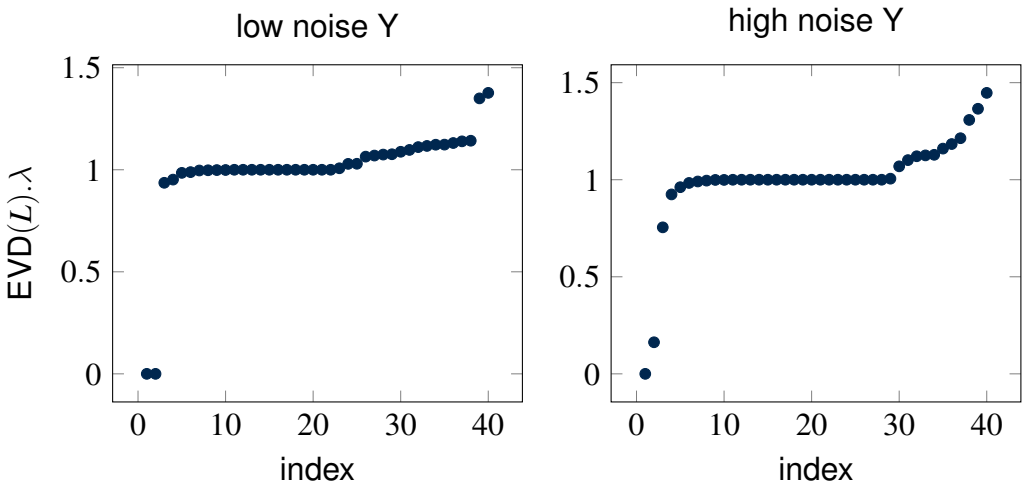
1. Any two vertices in A can be joined s.t. all intermediate points also lie in A
2. No connections between vertices in A and $\bar{A}$

Proposition 2 [7]

Let $\mathcal{G}(\cdot)$ be an undirected graph with non-negative weights. For L,

1. Multiplicity k of $\lambda = 0$ is # of connected components
2. $\nu_{\lambda=0}(L) = \{\mathbb{1}_{A_1}, ..., \mathbb{1}_{A_k}\}$

Spectral Gap Heuristic



Elbow Method Heuristic

Consider K-means perspective instead of graph theory

$\boldsymbol{\mu}_j$: sample mean for cluster j
c_j: cluster j
k: number of subspaces
$V^T_{i,:}$: spectral embedding point i

Define within cluster sum of squares as the following:

$$\text{WCSS}(k) = \sum_{j=1}^{k} \sum_{V^T_{i,:} \in c_j} \|V^T_{i,:} - \boldsymbol{\mu}_j\|_2^2$$

Goal: Plot $WCSS(k)$ and find the "elbow" or "knee" of the plot

Optimization - Proximal Approach

For simplicity, assume no outliers exist. Then,

$$\min_C \|C\|_{1,1} + \frac{\lambda_z}{2}\|Y - YC\|_F^2 \quad \text{s.t.} \quad \text{diag}(C) = \mathbf{0}$$

Introduce binary mask $\Omega = \mathbf{1}'\mathbf{1} - I$ and let $\mathcal{P}(X) = (\Omega \odot X)$ s.t.

$$\min_C \|\mathcal{P}(C)\|_{1,1} + \frac{\lambda_z}{2}\|Y - Y\mathcal{P}(C)\|_F^2$$

Assuming $C_0 = \{C : \text{diag}(C) = \mathbf{0}\}$, then the solution is as follows:

$$\left.\begin{array}{l} \nabla(C) = \mathcal{P}(-Y^T(Y - YC)) \\ \text{prox} = \text{soft}(C, \lambda_z^{-1}) \end{array}\right\} \quad \text{Note: } \nabla_C(\tfrac{1}{2}\|C \odot \Omega\|_F^2) = \Omega \odot C \odot \Omega$$

Graph Laplacians Part 2

From proposition 1, we see that:

$$\boldsymbol{f}^T L \boldsymbol{f} = \frac{1}{2} \sum_{i,j=1}^{n} w_{ij} (f_i - f_j)^2$$

Assume our similarity measure $w_{ij} = 1/d_{ij}^2$
where $d_{ij} = \text{dist}(\boldsymbol{y}_i, \boldsymbol{y}_j) = \|\boldsymbol{y}_i - \boldsymbol{y}_j\|_2$

Then, we have the following:

$$\boldsymbol{f}^T L \boldsymbol{f} = \frac{1}{2} \sum_{i,j=1}^{n} w_{ij} (f_i - f_j)^2 \approx \frac{1}{2} \sum_{i,j=1}^{n} \left(\frac{f_i - f_j}{d_{ij}} \right)^2 \approx \int |\nabla g|^2 \, \text{dx}$$

Laplacian & RatioCut Relationship

Assume $k = 2$ and $f_i = \{\sqrt{\frac{|\bar{A}|}{|A|}}$ if $\nu_i \in A$, $-\sqrt{\frac{|A|}{|\bar{A}|}}$ if $\nu_i \in \bar{A}\}$

$$\begin{aligned}
\boldsymbol{f}^T L \boldsymbol{f} &= \frac{1}{2}\sum_{i,j=1}^{n} w_{ij}(f_i - f_j)^2 \\
&= \frac{1}{2}\sum_{i\in A, j\in \bar{A}} w_{ij}\left(\sqrt{\frac{|\bar{A}|}{|A|}} + \sqrt{\frac{|A|}{|\bar{A}|}}\right)^2 + \frac{1}{2}\sum_{i\in \bar{A}, j\in A} w_{ij}\left(\sqrt{\frac{|\bar{A}|}{|A|}} + \sqrt{\frac{|A|}{|\bar{A}|}}\right)^2 \\
&= \text{cut}(A,\bar{A})\left(\frac{|\bar{A}|}{|A|} + \frac{|A|}{|\bar{A}|} + 2\right) = \text{cut}(A,\bar{A})\left(\frac{|A| + |\bar{A}|}{|A|} + \frac{|A| + |\bar{A}|}{|\bar{A}|}\right) \\
&= (|A| + |\bar{A}|)\left(\frac{\text{cut}(A,\bar{A})}{|A|} + \frac{\text{cut}(A,\bar{A})}{|\bar{A}|}\right) = |V| \cdot \text{RatioCut}(A,\bar{A})
\end{aligned}$$

Note: For simplicity, this is assuming unnormalized Laplacian.

Proposition 1 Proof

Given that $L = D - W$, we have that:

$$\begin{aligned} \boldsymbol{f}^T L \boldsymbol{f} &= \boldsymbol{f}^T D \boldsymbol{f} - \boldsymbol{f}^T W \boldsymbol{f} = \sum_{i=1}^{n} d_i f_i^2 - \sum_{i,j=1}^{n} f_i f_j \, w_{ij} \\ &= \frac{1}{2} \left(\sum_{i=1}^{n} d_i f_i^2 - 2 \sum_{i,j=1}^{n} f_i f_j \, w_{ij} + \sum_{j=1}^{n} d_j f_j^2 \right) \\ &= \frac{1}{2} \sum_{i,j=1}^{n} w_{ij} \, (f_i - f_j)^2 \end{aligned}$$

Affine Subspace Extension

1. $k = 1$
   Get sample mean value, de-mean data, & take SVD
2. $k > 1$
   Let $\boldsymbol{x}_1, ..., \boldsymbol{x}_n \in A = \{\boldsymbol{d} + \boldsymbol{z} : \boldsymbol{z} \in \mathcal{S}\}$ where $\boldsymbol{d}$ is displacement vector
   Let $\lambda_1, ..., \lambda_n \in \mathbb{R}$ be scalar values

$$\sum_i^n \lambda_i = 1 \implies \lambda_1 \boldsymbol{x}_1 + ... + \lambda_n \boldsymbol{x}_n \in A$$

For example, if $\lambda_i = \frac{1}{n}$, then we get the centroid location

Question: Given this, how would one generalize for SSC?