Spectral Clustering

Mikkel N. Schmidt

April 17, 2005

- Patterns

$$\{x_i\}_{i=1}^{N} \tag{1}$$

- Patterns

$$\{x_i\}_{i=1}^{N} \tag{1}$$

- Weights

$$\begin{aligned} \{w_{ij}\}_{i,j=1}^{N} \\ w_{ij} = w_{ji} \\ w_{ii} = 0 \end{aligned} \tag{2}$$

Weight w_{ij} indicates similarity between patterns x_i and x_j.

- Patterns

$$\{x_i\}_{i=1}^{N} \tag{1}$$

- Weights

$$\{w_{ij}\}_{i,j=1}^{N} \tag{2}$$
$$w_{ij} = w_{ji}$$
$$w_{ii} = 0$$

Weight w_{ij} indicates similarity between patterns x_i and x_j.

- Labels

$$\{y_i\}_{i=1}^{N} \tag{3}$$
$$y_i \in \{-\frac{1}{2}, \frac{1}{2}\}$$

- Minimize the weight that spans the two sets

$$\min_{y_i \in \{-\frac{1}{2}, \frac{1}{2}\}} \Sigma_{i,j}(y_i - y_j)^2 w_{ij} \tag{4}$$

- Minimize the weight that spans the two sets

$$\min_{y_i \in \{-\frac{1}{2}, \frac{1}{2}\}} \Sigma_{i,j} (y_i - y_j)^2 w_{ij} \tag{4}$$

- Balance the number of patterns assigned to each cluster

$$\min_{\substack{y_i \in \{-\frac{1}{2}, \frac{1}{2}\} \\ |\mathbf{y}^T \mathbf{1}| \leq \beta}} \Sigma_{i,j} (y_i - y_j)^2 w_{ij} \tag{5}$$

- Relaxation

$$\min_{\substack{\mathbf{y}\in\mathbb{R}^N \\ |\mathbf{y}^T\mathbf{1}|\leq\beta \\ \mathbf{y}^T\mathbf{y}=N/4}} \quad \Sigma_{i,j}(y_i - y_j)^2 w_{ij} \tag{6}$$

Vector of labels $\mathbf{y} \in \mathbb{R}^N$ with constant sum of squared elements.

- Relaxation

$$\min_{\substack{\mathbf{y}\in\mathbb{R}^N \\ |\mathbf{y}^T\mathbf{1}|\leq\beta \\ \mathbf{y}^T\mathbf{y}=N/4}} \quad \Sigma_{i,j}(y_i - y_j)^2 w_{ij} \tag{6}$$

  Vector of labels $\mathbf{y} \in \mathbb{R}^N$ with constant sum of squared elements.

- For convenience, scale each y_i by $2/\sqrt{N}$

$$\min_{\substack{\mathbf{y}\in\mathbb{R}^N \\ |\mathbf{y}^T\mathbf{1}|\leq\alpha \\ \mathbf{y}^T\mathbf{y}=1}} \quad \Sigma_{i,j}(y_i - y_j)^2 w_{ij} \tag{7}$$

  where $\alpha = \frac{2\beta}{\sqrt{N}}$

- Weight Matrix

$$\mathbf{W} \in \mathbb{R}^{N \times N} \qquad (8)$$
$$\mathbf{W}_{ij} = w_{ij}$$

- Weight Matrix

$$\mathbf{W} \in \mathbb{R}^{N \times N} \qquad \mathbf{W}_{ij} = w_{ij} \tag{8}$$

- Diagonal Matrix

$$\mathbf{D} \in \mathbb{R}^{N \times N} \qquad d_i = \Sigma_{j=1}^{N} w_{ij} \tag{9}$$

- The Laplacian

$$\mathbf{L} = \mathbf{D} - \mathbf{W} \tag{10}$$

is symmetric positive semi-definite with smallest eigenvalue 0 and corresponding eigenvector $\mathbf{1}$.

- The Laplacian

$$\mathbf{L} = \mathbf{D} - \mathbf{W} \tag{10}$$

is symmetric positive semi-definite with smallest eigenvalue 0 and corresponding eigenvector $\mathbf{1}$.

- Alternative: Normalized Laplacian

$$\mathbf{L}_{norm} = D^{-\frac{1}{2}} \left(\mathbf{D} - \mathbf{W}\right) D^{-\frac{1}{2}} \tag{11}$$

- Eigenvalues of the Laplacian

$$0 = \lambda_1 < \lambda_2 \leq \cdots \leq \lambda_N \tag{12}$$

- Eigenvalues of the Laplacian

$$0 = \lambda_1 < \lambda_2 \leq \cdots \leq \lambda_N \tag{12}$$

- (Orthonormal) Eigenvectors

$$\mathbf{v}^{[1]}, \mathbf{v}^{[2]}, \ldots, \mathbf{v}^{[N]} \tag{13}$$

- Eigenvalues of the Laplacian

$$0 = \lambda_1 < \lambda_2 \leq \cdots \leq \lambda_N \tag{12}$$

- (Orthonormal) Eigenvectors

$$\mathbf{v}^{[1]}, \mathbf{v}^{[2]}, \ldots, \mathbf{v}^{[N]} \tag{13}$$

- $\mathbf{v}^{[1]} = \mathbf{1}/\sqrt{N}$

- Eigenvalues of the Laplacian

$$0 = \lambda_1 < \lambda_2 \leq \cdots \leq \lambda_N \tag{12}$$

- (Orthonormal) Eigenvectors

$$\mathbf{v}^{[1]}, \mathbf{v}^{[2]}, \ldots, \mathbf{v}^{[N]} \tag{13}$$

- $\mathbf{v}^{[1]} = \mathbf{1}/\sqrt{N}$
- $\mathbf{v}^{[2]}$ is referred to as the *Fiedler vector.*

- Out minimization problem

$$\min_{\substack{\mathbf{y}\in\mathbb{R}^N \\ |\mathbf{y}^T\mathbf{1}|\leq\alpha \\ \mathbf{y}^T\mathbf{y}=1}} \Sigma_{i,j}(y_i - y_j)^2 w_{ij} \tag{14}$$

- Out minimization problem

$$\min_{\substack{\mathbf{y}\in\mathbb{R}^N \\ |\mathbf{y}^T\mathbf{1}|\leq\alpha \\ \mathbf{y}^T\mathbf{y}=1}} \Sigma_{i,j}(y_i - y_j)^2 w_{ij} \tag{14}$$

- Matrix-vector formulation

$$\min_{\substack{\mathbf{y}\in\mathbb{R}^N \\ |\mathbf{y}^T\mathbf{1}|\leq\alpha \\ \mathbf{y}^T\mathbf{y}=1}} \mathbf{y}^T\mathbf{L}\mathbf{y} \tag{15}$$

- Out minimization problem

$$\min_{\substack{\mathbf{y}\in\mathbb{R}^N \\ |\mathbf{y}^T\mathbf{1}|\leq\alpha \\ \mathbf{y}^T\mathbf{y}=1}} \quad \Sigma_{i,j}(y_i - y_j)^2 w_{ij} \tag{14}$$

- Matrix-vector formulation

$$\min_{\substack{\mathbf{y}\in\mathbb{R}^N \\ |\mathbf{y}^T\mathbf{1}|\leq\alpha \\ \mathbf{y}^T\mathbf{y}=1}} \quad \mathbf{y}^T\mathbf{L}\mathbf{y} \tag{15}$$

- Solution

$$\mathbf{y} = \alpha\mathbf{v}^{[1]} + \sqrt{1-\alpha^2}\mathbf{v}^{[2]} \tag{16}$$

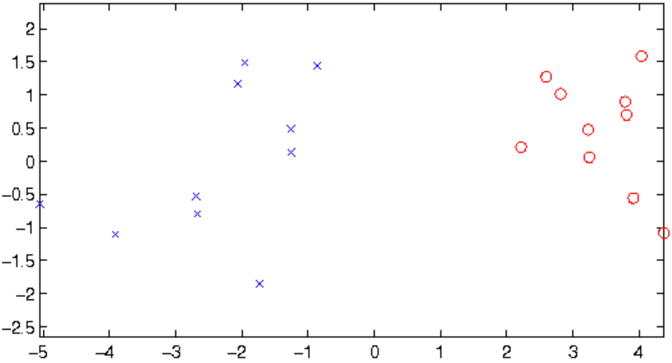
(Proof given in [Higham 2004])

- A closer look at the solution

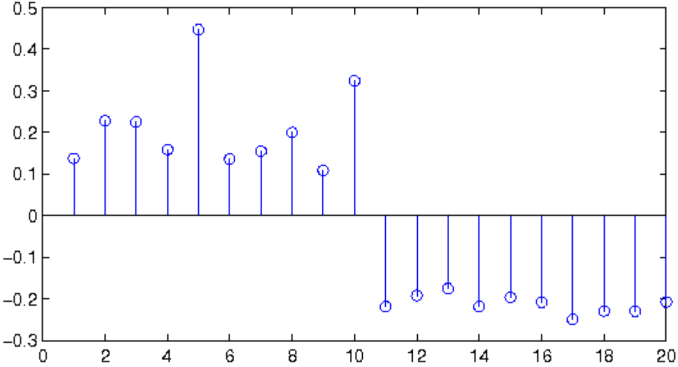
$$\mathbf{y} = \underbrace{\alpha \mathbf{v}^{[1]}}_{Multiple\ of\ \mathbf{1}} + \sqrt{1-\alpha^2}\mathbf{v}^{[2]} \tag{17}$$

Clustering only depends on $\mathbf{v}^{[2]}$.
Solution is independent of the balancing constant, α.

Data



Second eigenvector, $\mathbf{x}^{[2]}$



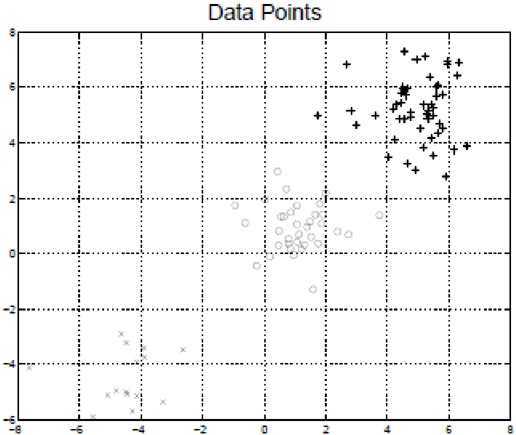


Figure 1: Data points used for Figure 2. Pairwise similarity weight is taken to be the reciprocal of the Euclidean distance.

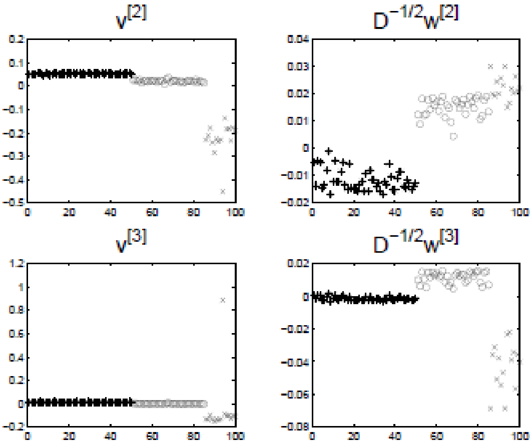
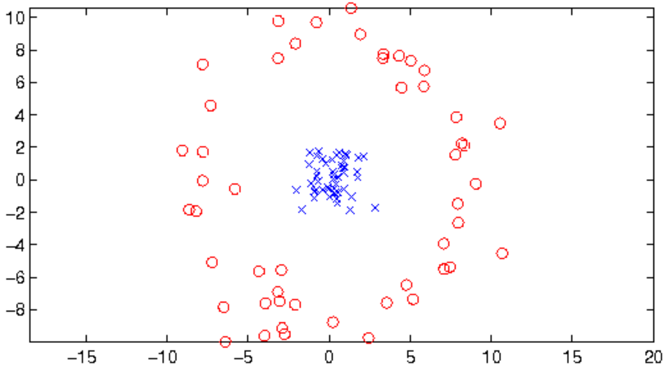


Figure 2: Components of the second and third eigenvectors for the data from Figure 1. Left unnormalized. Right normalized.

Data



Second eigenvector, $\mathbf{x}^{[2]}$