Graph clustering

Suppose we have a graph $G=(V,E)$

constructed from $s_{ij} \propto \exp(-d_{ij}^2/\tau)$

when $d_{ij} = \|x_i - x_j\|_2$ for example.

Ex: Correlation networks $d_{ij} = 1/\rho_{ij}$

Graph Laplacian

$f(x_1,x_2)$ on $\mathbb{R}^2$ $\Delta f = \frac{d^2}{dx_1^2} f + \frac{d^2}{dx_2^2} f$

On a graph: $w_{ij} = \begin{cases} s_{ij} & \text{if top } K \text{ neighbor} \\ 0 & \end{cases}$

$D = \text{diag}\left(\sum_{j \in N(i)} w_{ij}\right)$ $W = (w_{ij})_{ij}^N$

$L = D - W.$

Consider the set of operations

$D_1 f(i_1,i_2) = f(i_1+1, i_2) - f(i_1,i_2)$
$D_2 f(i_1,i_2) = f(i_1, i_2+1) - f(i_1,i_2)$

$D_1^2 f(i_1,i_2) = f(i_1+2,i_2) - f(i_1+1,i_2) - D_1 f(i_1,i_2)$

$\Rightarrow f(i_1+1,i_2) + f(i_1-1,i_2) - 2f(i_1,i_2)$

$$D_2^2 f(i_1, i_2) = f(i_1, i_2+1) + f(i_1, i_2-1) - 2 f(i_1, i_2)$$

$$-(D_1^2 + D_2^2) f(i_1, i_2) = 4 f(i_1, i_2) - \sum_{(j,k) \in N(i_1, i_2)} f(j,k)$$

Spectral Clustering

$L = Z \Sigma Z^T$ let $Z = (z_1 | \cdots | z_n)$ be ordered by eigenvalues

$0 = \lambda_n \leq \lambda_2 \leq \cdots \leq \lambda_1$

Then do K-means on the rows of Z. (may want to truncate).

Factor Analysis

Suppose $\underset{n \times p}{X} = \underset{p \times p}{U} \underset{p \times p}{\Sigma} V^T$

Define $S = \sqrt{n}\, U$ $A^T = \frac{D V^T}{\sqrt{N}}$

$U = U \Sigma V^T = U \frac{\sqrt{n}}{\sqrt{n}} D V^T = \underbrace{S}_{\substack{\text{new} \\ \text{coordinate} \\ \text{system}}} \underbrace{A^T}_{\text{new operator}}$

So $x_i = A s_i$ which gives a change of basis.

Choose A such that columns of S are uncorrelated

We consider $\{S_1, S_2, \ldots, S_p\}$ as factor loadings