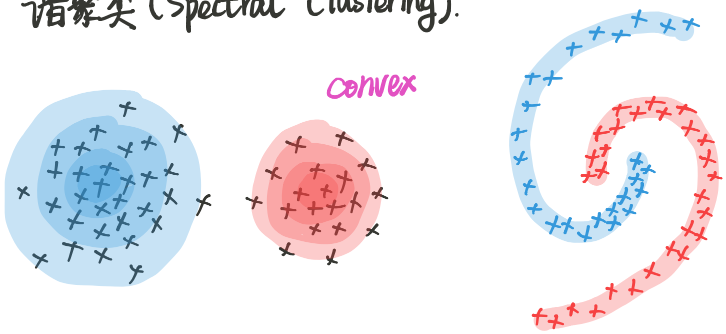
22. 谱聚类 Spectral Clustering

P1 谱聚类(Spectral Clustering).

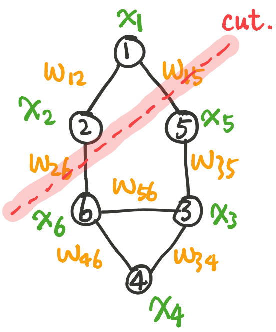


(convex)

$\begin{cases} \text{compactness: K-Means、GMM} \\ \text{connectivity: spectral clustering} \end{cases}$

P2

Graph-based(带权重的无向图)



$G=\{V,E\} \longrightarrow \begin{cases} V=\bigcup_{k=1}^{K} A_k \\ A_i \cap A_j = \emptyset, \forall i,j \in \{1,2,\cdots,K\} \end{cases}$

$V=\{1,2,3,\cdots,N\}$

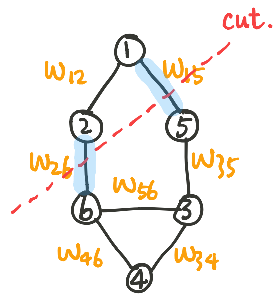
$X=(x_1\ x_2 \cdots x_N)=\begin{bmatrix} x_1^T \\ x_2^T \\ \vdots \\ x_N^T \end{bmatrix}_{N\times p}$

$E: W=[w_{ij}],\ 1\le i,j\le N$

W: similarity matrix (affinity matrix).

其中 $w_{ij}=\begin{cases} K(x_i,x_j)=\exp\left\{-\frac{\|x_i-x_j\|^2}{2\sigma^2}\right\}, & (i,j)\in E \\ 0 & (i,j)\notin E \end{cases}$ 高斯核

定义: $A\subset V,\ B\subset V,\ A\cap B=\emptyset,\ W(A,B)=\sum_{\substack{i\in A\\ j\in B}} w_{ij}$ (cost).



对于图中的cut, $A=\{1,2\}$, $B=\{3,4,5,6\}$, $W(A,B)=w_{15}+w_{26}$.

假如: 一共有K个类别.

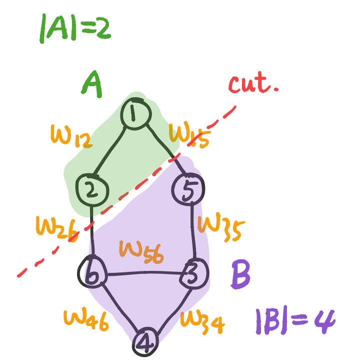
$cut(V)=cut(A_1,A_2,\cdots,A_K)=\sum_{k=1}^{K} W(A_k,\bar{A}_k)$ ↳ $V-A_k$.

$=\sum_{k=1}^{K} W(A_k,V)-\sum_{k=1}^{K} W(A_k,A_k)$.

目标: $\min_{\{A_k\}_{k=1}^K} cut(V)$ $(V\to A_1,A_2,\cdots,A_K)$ 将V中的节点分成K类.

问题: 需要标准化. (A、B集合元素个数不同)

$cut=\sum_{k=1}^{K}\frac{W(A_k,\bar{A}_k)}{D(A_k)}$ $\quad D(A_k)=\sum_{i\in A_k} d_i,\ d_i=\sum_{j=1}^{n} w_{ij}$



$d_i=\sum_{i=1}^{N} w_{ij}$

目标: $\min_{\{A_k\}_{k=1}^K} Ncut(V)$ $\quad Ncut(V)=\sum_{k=1}^{K}\frac{W(A_k,\bar{A}_k)}{\sum_{i\in A_k} d_i},\ d_i=\sum_{j=1}^{n} w_{ij}$

P3. $N_{cut}(V) = N_{cut}(A_1, A_2, \cdots A_K) = \sum_{k=1}^{K} \frac{W(A_k,\bar{A}_k)}{\sum_{i\in A_k} d_i} = \sum_{k=1}^{K} \frac{W(A_k,V)-W(A_k,A_k)}{\sum_{i\in A_k}\sum_{j=1}^{N} w_{ij}}$

Model: $\{A_k\}_{k=1}^{K} = \underset{\{A_k\}_{k=1}^{K}}{\arg\min}\ N_{cut}(V)$

indicator vector:

$\begin{cases} y_i \in \{0,1\}^K \\ \sum_{j=1}^{K} y_{ij}=1 \end{cases}$ $\quad y_i = \begin{bmatrix} y_{i1}\\ y_{i2}\\ \vdots \\ y_{iK}\end{bmatrix}$ $\quad 1\le i\le N$, $1\le j\le K$

$y_{ij}=1 \Leftrightarrow$ 第i个样本属于第j个类别

$Y=(y_1\ y_2 \cdots y_N)^T_{N\times K}$ $\Rightarrow$ $\hat{Y} = \underset{Y}{\arg\max}\ N_{cut}(V)$

$$N_{cut}(V) = \sum_{k=1}^{K} \frac{W(A_k,\bar{A}_k)}{\sum_{i\in A_k} d_i} = tr\, OP^{-1}$$

$$= tr \begin{bmatrix} \frac{W(A_1,\bar{A}_1)}{\sum_{i\in A_1} d_i} & 0 & \cdots & 0 \\ 0 & \frac{W(A_1,\bar{A}_1)}{\sum_{i\in A_2} d_i} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots\ 0 & \cdots & \frac{W(A_1,\bar{A}_1)}{\sum_{i\in A_k} d_i} \end{bmatrix}$$

$$= tr \underbrace{\begin{bmatrix} W(A_1,\bar{A}_1) & & & \\ & W(A_2,\bar{A}_2) & & \\ & & \ddots & \\ & & & W(A_K,\bar{A}_K) \end{bmatrix}}_{O_{K\times K}} \cdot \underbrace{\begin{bmatrix} \sum_{i\in A_1} d_i & & & \\ & \sum_{i\in A_2} d_i & & \\ & & \ddots & \\ & & & \sum_{i\in A_k} d_i \end{bmatrix}^{-1}}_{P_{K\times K}}$$

P4 已知: W, Y, 求 O、P

$$\underbrace{Y^T Y_{N\times K}}_{K\times K} = (y_1, \cdots, y_N)\begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix} = \sum_{i=1}^{N} y_i y_i^T = \begin{bmatrix} N_1 & & & \\ & N_2 & & \\ & & \ddots & \\ & & & N_K \end{bmatrix}_{K\times K} = \begin{bmatrix} \sum_{i\in A_1} 1 & & & \\ & \sum_{i\in A_2} 1 & & \\ & & \ddots & \\ & & & \sum_{i\in A_k} 1 \end{bmatrix}$$

$y_1y_1^T + \cdots + y_Ny_N^T$

$$y_i = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \begin{matrix} \\ \\ \\ \rightarrow k \\ \\ \\ \end{matrix} \qquad y_i y_i^T = \begin{array}{c} k \\ \left[\begin{array}{c|c} 0 & 0 \\ \hline & 1 \\ 0 & 0 \end{array}\right] k \end{array}$$

如果$y_{1k}=1$, $y_{2k}=1$, 则样本$x_1, x_2 \in A_k$

N_k: 在N个样本中属于类别k的样本个数.

$\sum_{k=1}^{K} N_k = N$, $\quad N_k = |A_k| = \sum_{i\in A_k} 1$

$$\sum_{i=1}^{N} y_i \underset{\substack{\uparrow \\ \text{一维实数}}}{d_i} y_i^T \Rightarrow P = \begin{bmatrix} \sum_{i\in A_1} d_i & & & \\ & \sum_{i\in A_2} d_i & & \\ & & \ddots & \\ & & & \sum_{i\in A_k} d_i \end{bmatrix} = Y^T D Y \qquad D = \begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_N \end{bmatrix} = diag(W\cdot 1_N)$$

$$\begin{bmatrix} w_{11} & \cdots & \\ \vdots & \ddots & \\ & & w_{NN} \end{bmatrix}\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} d_1 \\ \vdots \\ d_N \end{bmatrix}$$

P5

$$W(A_k,\bar{A}_k) = \underbrace{W(A_k,V)}_{\sum_{i\in A_k} d_i} - \underbrace{W(A_k,A_k)}_{\sum_{i\in A_k}\sum_{j\in A_k} w_{ij}}.$$

$$O = \begin{pmatrix} \sum_{i\in A_1} d_i & & \\ & \ddots & \\ & & \sum_{i\in A_k} d_i \end{pmatrix} - \begin{pmatrix} W(A_1,A_1) & & \\ & \ddots & \\ & & W(A_k,A_k) \end{pmatrix} = Y^T D Y - Y^T W Y$$

$$Y^T W Y = (y_1 \cdots y_N)\begin{bmatrix} w_{11} & \cdots & w_{1N} \\ \vdots & \ddots & \vdots \\ w_{N1} & \cdots & w_{NN} \end{bmatrix}\begin{bmatrix} y_1^T \\ \vdots \\ y_N^T \end{bmatrix} = \left(\sum_{i=1}^{N} y_i w_{i1} \quad \cdots \quad \sum_{i=1}^{N} y_i w_{iN}\right)\begin{pmatrix} y_1^T \\ \vdots \\ y_N^T \end{pmatrix}$$

$$= \sum_{i=1}^{N}\sum_{j=1}^{N} y_i y_j^T w_{ij}$$

$$= \begin{bmatrix} \sum_{i\in A_1}\sum_{j\in A_1} w_{ij} & \cdots & \sum_{i\in A_1}\sum_{j\in A_k} w_{ij} \\ \vdots & & \vdots \\ \sum_{i\in A_k}\sum_{j\in A_1} w_{ij} & \cdots & \sum_{i\in A_k}\sum_{j\in A_k} w_{ij} \end{bmatrix}$$

$$\hat{Y} = \underset{Y}{\operatorname{argmin}}\ \operatorname{tr}\left(Y^T(D-W)Y\cdot(Y^T D Y)^{-1}\right)$$

$L = D - W$

↳ Laplacian Matrix