Algorithme EM pour les mélanges gaussiens

(X,Y) à valeurs dans $\mathbb{R}^d \times [\![1,K]\!]$

$$\begin{cases} \mathbb{P}\{Y=j\} = \pi_j > 0 & \sum_{j=1}^{K} \pi_j = 1 \\ X \mid Y=j \sim \mathcal{N}(\mu_j, \Sigma_j) \end{cases}$$

$\mathcal{P}_g = \{ G_\theta = \mathcal{L}(X,Y), \theta \in \Theta \}$

$\theta = (\pi, (\mu_1, \Sigma_1) \dots (\mu_K, \Sigma_K))$.

$[0,1]^K \; \mathbb{1}^T \pi = 1$ SDP

Observations $X_0 \dots X_n$ iid de même loi que X.

$\mathcal{P}_M = \{ M_\theta = \mathcal{L}(X), (X,Y) \sim G_\theta \}$

$\mathcal{P}_C = \{ Q_\theta = \mathcal{L}(Y \mid X), (X,Y) \sim G_\theta \}$.

$$p_{\mu,\Sigma}(x) = \frac{1}{\sqrt{2\pi}^d \sqrt{\det \Sigma}} e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)}.$$

$X \mid Y=j$

G_θ a pour densité $g_\theta(x,y) = \pi_y \, p_{\mu_y, \Sigma_y}(x)$. $\quad X, Y=y$

$M_\theta : m_\theta(x) = \sum_{y=1}^{K} g_\theta(x,y)$

$$Q_{\theta,x} : q_{\theta,x}(y) = \frac{g_\theta(x,y)}{m_\theta(x)}.$$

▷ Expectation step.

A l'itération t, $\hat{\theta}_t$

$Z_1^t \dots Z_n^t$ tq $\{(X_1, Z_1^t) \mid \hat{\theta}_t \dots (X_n, Z_n^t) \mid \hat{\theta}_t\}$ iid

$Z_i^t \mid X_1^n \overset{\mathcal{L}}{=} Z_i^t \mid \hat{\theta}_t X_i \sim Q_{\hat{\theta}_t, X_i}$

$$\mathbb{P}\{Z_i^t = j \mid X_1^n\} = \frac{g_{\hat{\theta}_t}(x_i, j)}{m_{\hat{\theta}_t}(x_i)} = p_{ij}^t$$

$$\forall \theta : F(\theta, \hat{\theta}_t) = \mathbb{E}\left[\sum_{i=1}^{n} \log g_\theta(X_i, Z_i^t) \mid X_1^n \right].$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{K} p_{ij}^t \log g_\theta(X_i, j).$$

avec $\log(g_\theta(x_i,j)) = \log \pi_j - \frac{1}{2}(x_i-\mu_j)^T \Sigma_j^{-1}(x_i-\mu_j) - \frac{1}{2}\log(\det(\Sigma_j$

$$p_{ij}^{t} = \frac{\hat{\pi}_j^t \, p_{\hat{\mu}_j,\hat{\Sigma}_j}(x_i)}{\sum_{l=1}^{K} \hat{\pi}_l^t \, p_{\hat{\mu}_l,\hat{\Sigma}_l}(x_i)}$$

▷ <u>Maximisation step</u>.

$$\hat{\theta}_{t+1} \in \underset{\theta\in\Theta}{\operatorname{argmax}}\, F(\theta/\hat{\theta}_t) .$$

$\to \pi$
$\to \mu$
$\to \Sigma$

En dep le calcul, on obtient :

$$\begin{cases} \hat{\theta}_{t+1} = (\hat{\pi}^{t+1}, \hat{\mu}_j^{t+1}, \hat{\Sigma}_j^{t+1}) \\ \hat{\mu}_j^{t+1} = \dfrac{\sum_{i=1}^{n} p_{ij}^t x_i}{\sum_{i=1}^{n} p_{ij}^t} \\ \hat{\Sigma}_j^{t+1} = \dfrac{\sum p_{ij}^t (x_i-\hat{\mu}_j^t)(x_i-\hat{\mu}_j^t)^T}{\sum p_{ij}^t} \\ \hat{\pi}_j^{t+1} = \dfrac{1}{n}\sum_{i=1}^{n} p_{ij}^t \end{cases}$$

$$\underset{\substack{\pi\in[0,1]^K \\ \mathbb{1}^T\pi=1}}{\text{maximise}} \sum_{j=1}^{K}\left(\sum_{i=1}^{n} p_{ij}^t\right)\log(\pi_j) \iff \text{minimiser } \underbrace{-\sum_{j=1}^{K}\frac{1}{n}\sum_{i=1}^{n} p_{ij}^t \log\pi_j + \sum_j \frac{1}{n}\sum_i p_{ij}^t \log\left(\frac{1}{n}\sum p_{ij}^t\right)}_{\sum_j \underbrace{\left(\frac{1}{n}\sum p_{ij}^t\right)}_{\square_j} \log\frac{\square_j}{\pi_j}}$$

minimiser $D_{KL}(\square \| \pi)$ $\iff$ (on a $\square_j \in [0,1]$ et $\sum_{j=1}^{K} \square_j = 1$
π vecteur de probabilité

D'où $\hat{\pi}^{t+1} = \square$

<u>EM $\mathcal{N}(\cdot)$ ~ soft k-means</u>

Itérer :
$p_{ij}^t \simeq \mathbb{P}\{Y_i=j / X_i\}$
$\hat{\mu}_j$
$\hat{\Sigma}_j$

Cas général de EM.

$$\begin{cases} (X,Y) \sim G_\theta = g_\theta \cdot (\psi \times d) \\ X \sim M_\theta = m_\theta \cdot \mu \\ Y/X \sim Q_{\theta,X} = q_{\theta,X} \cdot \nu \end{cases} \quad X = (X_1 \dots X_n) \quad Y = (Y_1 \dots Y_n)$$

On cherche $\hat{\theta}_t$ un estimateur de θ, $F(\theta|\hat{\theta}_t) = \mathbb{E}[\log(g_\theta(X,Z_t))/X]$

$Z_t | X \sim Q_{\hat{\theta}_t|x}$.

Lemme

$Q_x = q_x \cdot \nu$ et $Z | X = x \sim Q_x$.

$\forall \theta \in \Theta$, $\log(m_\theta(X)) = \mathbb{E}[\log(g_\theta(X,Z)) | X] + D_{KL}(Q_x \| Q_{\theta,x}) + H(Q_x)$

$D_{KL}(Q_x \| Q_{\theta,x}) = \mathbb{E}\left[\log \dfrac{q_x(Z)}{q_{\theta,x}(Z)} \Big| X\right]$ et $H(Q_x) = \mathbb{E}[\log q_x(Z) | X]$.

En particulier, $Q_x = Q_{\theta',x}$

$$\log m_\theta(X) = F(\theta|\theta') + D_{KL}(\theta' \| \theta) + H(\theta').$$

Preuve

Formule de Bayes, $g_\theta(x,z) = q_{\theta,x}(z)\, m_\theta(x)$.

$$\begin{aligned} \log m_\theta(X) &= \mathbb{E}[\log(m_\theta(X)) | X] \\ &= \mathbb{E}[\log(g_\theta(X,Z)) | X] - \mathbb{E}[\log(q_{\theta,X}(Z)) | X] + \mathbb{E}[\log q_X(Z) / X] \\ &= \mathbb{E}[\log(g_\theta(X,Z)) | X] - \underbrace{\mathbb{E}\left[\log \frac{q_X(Z)}{q_{\theta,X}(Z)} / X\right]}_{D_{KL}(Q_X \| Q_{\theta,X})} - \underbrace{\mathbb{E}[\log q_X(Z) / X]}_{H(Q_X)} \end{aligned}$$

$\square$

Proposition

$P = p \cdot \mu$ et $Q = q \cdot \mu$. Alors : $0 \leq D_{KL}(P \| Q) \in \overline{\mathbb{R}}$.

De plus, si $Q \ll P$:

$$D_{KL}(P \| Q) \iff p = q \quad p\text{-presque partout}$$

Preuve

$D_{KL}(P\|Q) = \mathbb{E}\left[\log \frac{p(z)}{q(z)}\right]$ avec $z \sim P$

Si $!(P \ll Q)$ alors $D_{KL} = +\infty$.

Si $P \ll Q$, $D_{KL}(P\|Q) = \mathbb{E}\left[-\log \frac{q(z)}{p(z)}\right]$ car $p(z) > 0$ ps.

$\geq -\log\big(\underbrace{\mathbb{E}_{p(z)}[q(z)/p(z)]}_{1}\big)$ par l'inégalité de Jensen.

≥ 0

Par stricte convexité de la fonction $-\log$, on a :

$D_{KL}(P\|Q) = 0$ ssi $\frac{q(z)}{p(z)} = \text{cste}$ ps.

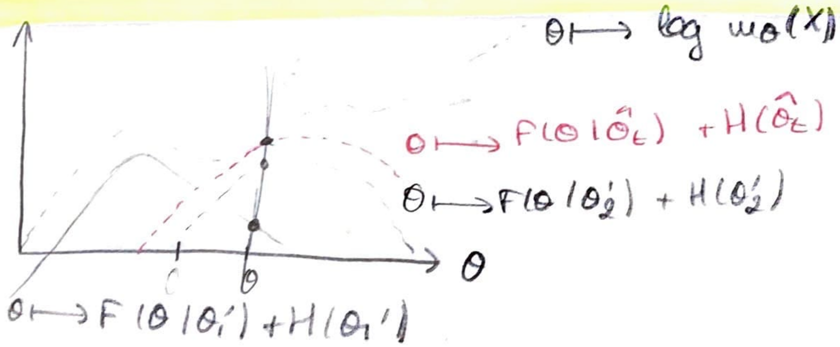
ie $\frac{q}{p} = \text{cst}$ $\mathbb{P}$-ps. car $Q \ll P$.
Q-ps

Donc par intégration, $\int q = \text{cste} \int p$ et cste $= 1$ et $p = q$ $\mathbb{P}$-ps. □

Proposition

$\forall \theta, \theta'$, $\log w_\theta(X) \geq F(\theta, \theta') + H(\theta')$

et $\log w_\theta(X) = F(\theta, \theta') + H(\theta')$ pour $\theta = \theta'$.



E step $\hat{\theta}_t' \in \arg\max F(\hat{\theta}_t | \theta) + H(\theta)$.

$\hat{\theta}_t' = \hat{\theta}_t$

M step: $\hat{\theta}_{t+1} \in \arg\max_{\theta} F(\theta | \hat{\theta}_t) + H(\hat{\theta}_t)$.

Théorème

$\hat{\theta}_{t+1} \in \underset{\theta \in \Theta}{\arg\max}\ F(\theta \mid \hat{\theta}_t)$, alors la suite $\log(m_{\hat{\theta}_t}(X))$ est croissante

Preuve

$$\log(m_{\hat{\theta}_t}(X)) = F(\hat{\theta}_t \mid \hat{\theta}_t) + H(\hat{\theta}_t)$$
$$\underset{\text{def max}}{\leq} F(\hat{\theta}_{t+1} \mid \hat{\theta}_t) + H(\hat{\theta}_t) \leq \log(m_{\hat{\theta}_{t+1}}(X)).$$

II. Minimisation de coût / risque.

Objectif : Partition $(C_1 \dots C_K)$ de $\mathbb{R}^d$ tel que $(C_1 \dots C_K)$ minimise un risque empirique.

→ coûts centrés (k-means)
→ coût distribué (clustering spectral).

d : dissimilarité, $d(x,y) \gg 1$ pour $x \neq y$.

Pour $(C_1 \dots C_K)$ un partitionnement, $\mu(C_j) \in \underset{\mu \in \mathbb{R}^d}{\arg\min}\ \mathbb{E}[d(X,\mu)\mathbb{1}_{X \in C_j}]$.

Coûts centrés

$$\underset{(C_1 \dots C_K) \in \mathcal{P}(\mathbb{R}^d)}{\text{minimiser}}\ \underbrace{\mathbb{E}[d(X, \mu(C_j))\mathbb{1}_{X \in C_j}]}_{D(C_1 \dots C_K)}.$$

Quantification (cf Biau)

$\mathbb{R}^d$, $q(x) = \sum_{j=1}^{K} \mu(C_j)\mathbb{1}_{x \in C_j}$

$D(C_1 \dots C_K) = \mathbb{E}[d(X, q(X))]$

① Coût des k-moyennes

$d(x,y) = \|x-y\|_2^2$.

$\mu(C_j) = \frac{1}{\mathbb{P}\{X \in C_j\}} \mathbb{E}[X\mathbb{1}_{X \in C_j}]$.

Empiriquement : critère $D(C_1 \dots C_K) = \frac{1}{n}\sum_{j=1}^{K}\sum_{X \in \hat{C}_j} \|x - \mu(\hat{C}_j)$

$\hat{C}_j = C_j \cap (X_1 \dots X_n)$.

$\mu(\hat{C}_j) = \frac{1}{\#\{X_i \in \hat{C}_j\}} \sum_{X \in \hat{C}_j} X$

② k-medoids

Même critère

$\mu(\hat{g}) = x_t$ avec $t \in \underset{1 \leq i \leq n}{\mathrm{argmin}} \sum_{x \in \hat{g}} \| x - x_i \|_2^2$

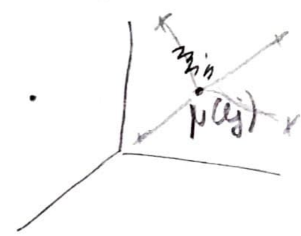
un point plutôt que la moyenne

③ k-median

$D(g_1 \dots g_K) = \frac{1}{n} \sum \sum \| \quad \|_2^{\circ}$

$\mu(\hat{g}) \in \{x_1 \dots x_n\}$

Dispersion intra-classe

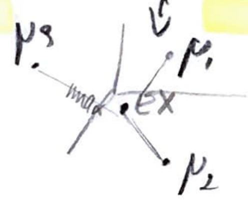


Cas des k-moyennes

$D(g_1 \dots g_K) = \mathbb{E}\left[\| X - \mu(g_j) \|_2^2 \mathbb{1}_{X \in g_j} \right]$

Propriété

$D(g_1 \dots g_K) + \sum_{j=1}^{K} \mathbb{P}\{X \in g_j\} \| \mathbb{E}(X - \mu(g_j)) \|^2 = \mathbb{E}\left[\| X - \mathbb{E}[X] \|^2 \right]$.



Preuve

$\mu(g_j) = \frac{1}{\mathbb{P}\{X \in g_j\}} \mathbb{E}[X \mathbb{1}_{X \in g_j}]$.

TODO ("c'est marrant")

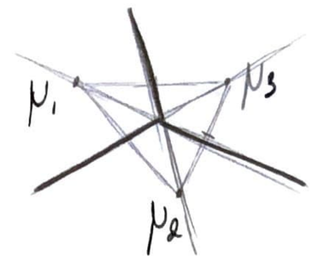
Algorithme

$$\hat{D}(C_1 \cdots C_K) = \frac{1}{n} \sum_{j=1}^{K} \sum_{x \in \hat{C}_j} \| x - \mu(\hat{C}_j) \|^2 .$$

solu° ①: tester ttes les partitions possibles de $(X_1 \longrightarrow X_n) \rightarrow$ tp coûteux

$$\min_{C_1 \cdots C_K \in \mathcal{P}(\mathbb{R}^d)} \hat{D}(C_1 \cdots C_K)$$

$$\min_{\substack{(C_1 \cdots C_K) \in \mathcal{P}(\mathbb{R}^d) \\ (\mu_1 \cdots \mu_K) \in \mathbb{R}^d}} \frac{1}{n} \sum \sum \| x - \mu_j \|^2 \qquad \mu_j = \mu(\hat{C}_j)$$



Algorithme des k-means

$\mu_1 \cdots \mu_K$ initialisé

itérer : $(C_1^{t+1} \cdots C_K^{t+1})$ = Partitionnement de Voronoï pour les centres $(\mu_1 \cdots \mu_K)$.

Maj de $\mu_1^{t+1} \cdots \mu_K^{t+1}$ moyennes arithmétiques de chaque cellule $\hat{C}_j = C_j \cap \{X_1 \cdots X_n\}$

$$C_1 = \{ x \in \mathbb{R}^d : \| x - \mu_1 \| \leq \| x - \mu_\ell \| , \forall \ell \in [\![1, K]\!] \} .$$

$$C_{k+1} = \{ x \in \mathbb{R}^d : \| x - \mu_{k+1} \| \leq \| x - \mu_\ell \| \ \forall \ell \in [\![1, K]\!] \} \setminus \left(\bigcup_{\ell=1}^{k} C_\ell \right) .$$

Proposition

$$D(C_1^t \cdots C_K^t) = \frac{1}{n} \sum_{j=1}^{K} \sum_{x \in \hat{C}_j} \| x - \mu(\hat{C}_j) \|^2 \text{ est décroissante avec } t .$$

Preuve

$$\hat{D}(C_1^t \cdots C_K^t) = \frac{1}{n} \sum_{j=1}^{K} \sum_{x \in \hat{C}_j^t} \| x - \mu(\hat{C}_j^t) \|^2 \overset{\text{Voronoï}}{\geq} \frac{1}{n} \sum_j \sum_{x \in C_j^{t+1}} \| x - \mu_j^{t+1} \|^2$$

$$\frac{1}{n} \sum_i \sum_j \| X_i - \mu_j^t \| \, \mathbb{1}_{X_i \in C_j^{t+1}}$$

$$\overset{\text{maj}}{\geq} \frac{1}{n} \sum_j \sum_{x \in C_j^{t+1}} \| x - \mu_j^{t+1} \|^2 = \hat{D}(C_1^{t+1} \cdots C_K^{t+1}) .$$

① $\forall x \in C_j^{t+1}$, $\| x - \mu_j^t \|^2 \leq \| x - \mu_\ell^t \|^2 \quad \forall \ell \in [\![1, K]\!]$ vrai pour ℓ tq $x \in C_\ell^t$

$\forall x \in \mathbb{R}^d$, $\sum_{j=1}^{K} \| x - \mu_j^t \|^2 \mathbb{1}_{x \in C_j^{t+1}} \leq \sum_{j=1}^{K} \| x - \mu_j \|^2 \mathbb{1}_{x \in C_j^t}$

$$\sum_{i=1}^{n} \sum \| X_i - \quad \|^2 \mathbb{1} \qquad \leq \sum_{i=1}^{n} \sum_j \| X_i - \mu_j \|$$

② $\mu_j^{t+1} = \frac{1}{card(G_j^{t+1})} \sum_{i=1}^{n} x_i \mathbb{1}_{x_i \in G_j^{t+1}} \in \underset{\mu \in \mathbb{R}^d}{\arg\min} \sum_{i=1}^{n} \|x_i - \mu\| \mathbb{1}_{x_i \in G_j}$

Donc $\sum_{j=1}^{K} \sum_{i=1}^{n} \|x_i - \mu_j^{t+1}\|^2 \mathbb{1}_{x_i \in G_j^{t+1}} \leq \sum_i \sum_j \|x_i - \mu_j^t\|^2 \mathbb{1}$ ___

□

k-means

① Voronoï $\mathbb{1}_{x_i \in G}$

② Maj centroid

Soft-k-means

① calcul $p_{ij} \sim \mathbb{P}(Y=j / X_i) \longrightarrow 0, 1$

② Maj $\hat{\mu}_j$

$\hat{\Sigma}_j = \sigma^2 Id$, $\sigma^2 \to 0$

$\hat{\pi}_j = \frac{1}{K}$

$\Leftrightarrow$

Coûts distribués

s : similarité, $s(x,y)$ grand si $x \simeq y$

$\hat{D}(G_1 \dots G_K) = \sum_{j=1}^{K} \sum_{\substack{X \in \hat{G}_j \\ Y \notin \hat{G}_j}} s(X,Y)$ minimiser

W : matrice d'adjacence $n \times n$: $W_{ij} = s(X_i, X_j)$ s est symétrique, W est symétrique

Partitionnement des indices $[\![1 \dots n]\!]$, $(I_1 \dots I_K)$:

$\hat{D}(I_1 \dots I_K) = \sum_{j=1}^{K} \sum_{\substack{i \in I_j \\ l \notin I_j}} W_{il}$

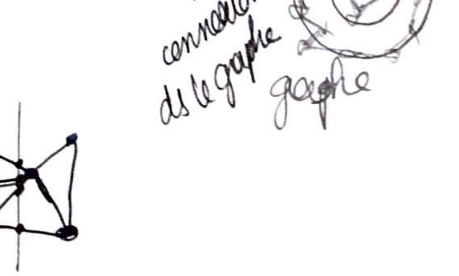
Graphe cut problem

19/10/2021

$s : \|x-y\|_2$

↳ s : ε-neighborhood similarity :

$s(x,y) = \begin{cases} 1 & \text{si } \|x-y\|_2 \leq \varepsilon \\ 0 & \text{sinon} \end{cases}$



$(I_1 \dots I_K)$ partition de $[\![1;n]\!]$ telle que :

$\sum_{j=1}^{K} \sum_{\substack{i \in I_j \\ l \notin I_j}} W_{il}$ ↓ Graph cut.

peu pertinente en pratique

Ratio cut :

$\rightarrow \mathcal{D}_r(\mathcal{I}) = \sum_{j=1}^{n} \frac{1}{|I_j|} \sum_{\substack{i \in I_j \\ l \notin I_j}} W_{il}$ avec $|I_j| = \text{card}(I_j)$

Normalized cut

$\rightarrow \mathcal{D}_n(\mathcal{I}) = \sum_{j=1}^{n} \frac{1}{\text{vol}(I_j)} \sum_{\substack{i \in I_j \\ l \notin I_j}} W_{il}$ avec $\text{Vol}(I_j) = \sum_{i \in I_j} d_i$,

$d_i = \sum_{l=1}^{n} W_{il}$ (degré du $i^{\text{ème}}$ sommet)

Définition

Graphe $W \in \mathbb{R}^{n \times n}$ symétrique + la matrice des degrés $D = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix}$.

Laplacien du graphe : $L = D - W$. $\quad d_i = \sum_{l=1}^{n} W_{il}$

Propriété

$(I_1 \dots I_K)$ partition de $\{1; n\}$,

$\mathcal{D}_r(\mathcal{I}) = \text{tr}(H^T L H)$ avec $H \in \mathbb{R}^{n \times k}$, $H_{ij} = \frac{1}{\sqrt{|I_j|}} 1_{i \in I_j}$

De plus, $H^T H = I_K$.

Preuve

- $H^T H = I_K$ calcul direct.
- $\text{tr}(H^T L H) = \text{tr}((L^{1/2} H)^T L^{1/2} H) = \sum_{j=1}^{K} \underbrace{(L^{1/2} h_j)^T (L^{1/2} h_j)}_{h_j^T L h_j}$, $H = [h_1 | \dots | h_K]$

$\forall u \in \mathbb{R}^n$, $u^T L u = u^T D u - u^T W u = \sum_{i=1}^{n} d_i u_i^2 - \sum_{i,j} W_{ij} u_i u_j$
D diagonale

$= \frac{1}{2} \left(\sum_i d_i u_i^2 + \sum_l d_l u_l^2 - 2 \sum_{i,j} W_{ij} u_i u_j \right)$ W symétrique

$= \frac{1}{2} \left(\sum_{i,j} W_{ij} u_i^2 + \sum_j W_{ij} u_j^2 - 2 \text{———} \right)$?

$= \frac{1}{2} \left(\sum W_{ij} (u_i - u_j)^2 \right)$

Donc $h_j^T L h_j = \frac{1}{2} \sum_{il} w_{il} \left[(h_j)_i - (h_j)_l \right]^2$

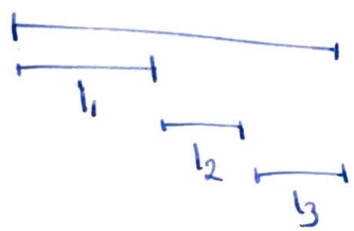
$(h_j)_i = H_{ij} = \frac{1}{\sqrt{|I_j|}} \mathbb{1}_{i \in I_j}$ $\qquad (h_j)_l = H_{lj} = \frac{1}{\sqrt{|I_j|}} \mathbb{1}_{l \in I_j}$

$= \frac{1}{2} \sum_j \frac{1}{|I_j|} \sum_{il} \underbrace{(\mathbb{1}_{i \in I_j} - \mathbb{1}_{l \in I_j})}_{\substack{= 1 \text{ si } i \in I_j \text{ et } l \notin I_j \\ i \notin I_j \text{ et } l \in I_j}} w_{il}$.

$= \frac{1}{2} \sum_j \frac{1}{|I_j|} \sum_{\substack{i \in I_j \\ l \notin I_j}} w_{il} = D_r(I)$. $\square$

__exemple :__

" ≐ one-hot-encoding "



$$H_r = \begin{pmatrix} \Big| \, 1/|I_1| & & 0 \\ & \Big| \, \frac{1}{|I_2|} & \\ 0 & & \Big| \, \frac{1}{|I_3|} \end{pmatrix}$$

__Ratio cut :__

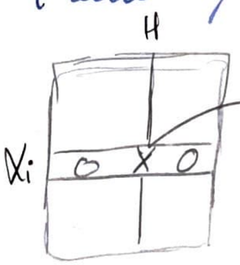
$\begin{cases} \text{minimise} & D_r(I) \\ I \in \mathcal{P}([\![1, n]\!]) & \end{cases}$ $\qquad \begin{cases} \text{minimise}_{H \in \mathbb{R}^{n \times K}} & \mathrm{tr}(H^T L H) \\ & H^T H = I_K \\ & H_{ij} \in \{0, \frac{1}{\sqrt{|I_j|}}\} \end{cases}$

inconnu ! → retirer cette contrainte.

$L = \sum_{i=1}^{n} \lambda_i v_i v_i^T$ élts propres de L (réelle symétrique)

avec $\underbrace{\boxed{\lambda_1 \leq \ldots}}_{K} \leq \lambda_n$.

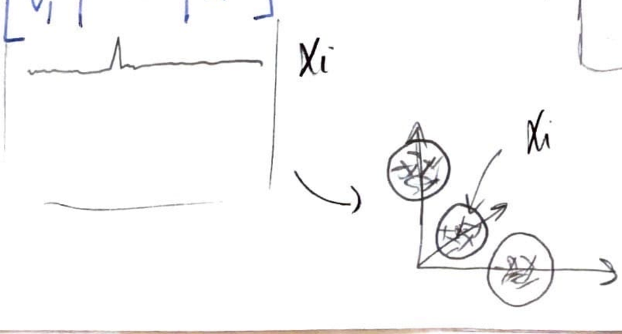
$H = [v_1 | \ldots | v_K]$.

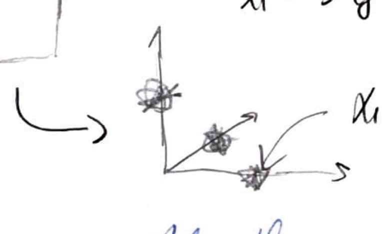


$\mathbb{1}_{i \in I_j}$

$X_i \rightarrow I_j$ tq $H_{ij} \neq 0$.

$j \in \mathrm{argmax}_{1 \leq l \leq K} H_{il}$





Algorithme : cf notes de cours

k-moyennes

Algo itératif et simple

↳ solution approché du pb exact

Clustering Spectral

Problème relâché

↳ solution exacte = solution approchée du problème d'origine

mat avec 1 seule col ≡ quotient de Rayleight

↳ cf Fisher généralisé et ss

Normalized cut

$$\mathcal{D}_n(\mathcal{Z}) = \sum_{j=1}^{K} \frac{1}{\mathrm{Vol}(\mathcal{Z}_j)} \sum_{il} w_{il}$$

D : matrice des degrés

Propriété

$$\mathcal{D}_n(\mathcal{Z}) = \mathrm{tr}(H^T L H) \quad \text{avec } H_{ij} = \frac{1}{\sqrt{\mathrm{Vol}(\mathcal{Z}_j)}} \quad u_i \in \mathcal{Z}_j$$

$$H^T D H = I_K.$$

Si D non inversible, un point du graphe est isolé

$L_w = D^{-1} L$ si D est inversible.

$= \sum \lambda_i v_i v_i^T$, $d_1 \leq \dots \leq d_n$.

$H = [v_1 | \dots | v_K]$.

22/10/2022