II Mélanges gaussiens (introduction)

Densité d'un mélange de K gaussiennes de paramètres
(pdf) $(\pi_k)_{1\leq k\leq K}$, $(\mu_k)_{1\leq k\leq K}$ et $(\Sigma_k)_{1\leq k\leq K}$:

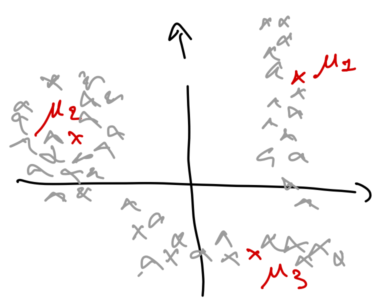
$$p(x) = \sum_{k=1}^{K} \pi_k \mathcal{N}(x \mid \mu_k, \Sigma_k)$$

Notations :

- $x_1, \dots, x_N$: observations $\in \mathbb{R}^D$
- $z_1, \dots, z_N$: variable latente / non observée $\in [\![1, K]\!]$
  (z_n correspond à la gaussienne "qui a généré x_n")
- $(\pi_k)_{1\leq k\leq K}$: coefficients de mixage : $\pi_k = p(z=k)$
- $(\mu_k)_{1\leq k\leq K}$: moyennes des K gaussiennes ($\mu_k \in \mathbb{R}^D \ \forall k$)
  $\mu_k = \mathbb{E}(x \mid z=k)$
- $(\Sigma_k)_k$: matrices de covariance des K gaussiennes

Probabilités a priori : $\pi_k = p(z=k)$

Probabilités a posteriori : $\gamma_{nk} = p(z=k \mid x=x_n)$ $[1\leq n\leq N, 1\leq k\leq K]$
(responsabilité de la $k^{\text{ème}}$ gaussienne dans l'explicativité de x_n)



$$\gamma_{nk} = p(z=k \mid x=x_n)$$

$$\underset{\text{Bayes}}{=} \frac{p(z=k)\, p(x=x_n \mid k)}{p(x=x_n)}$$

$$\underset{\text{Proba totales}}{=} \frac{p(z=k)\, p(x=x_n \mid k)}{\sum_{j=1}^{K} \pi_j \mathcal{N}(x_n \mid \mu_j, \Sigma_j)}$$

<u>Avec :</u> $\mathcal{N}(x \mid \mu, \Sigma) = \frac{1}{(2\pi)^{D/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\right)$

pdf d'une gaussienne simple dans $\mathbb{R}^D$,

de paramètres μ et Σ.

<u>Log-vraisemblance des observations sachant un modèle :</u>

$$l(X \mid \theta) = \ln\left(p(X \mid \theta)\right) = \ln\left(\prod_{n=1}^{N} p(x_n \mid \theta)\right) = \sum_{n=1}^{N} \ln\left(p(x_n \mid \theta)\right)$$

ens. des observat°

paramètres d'un modèle