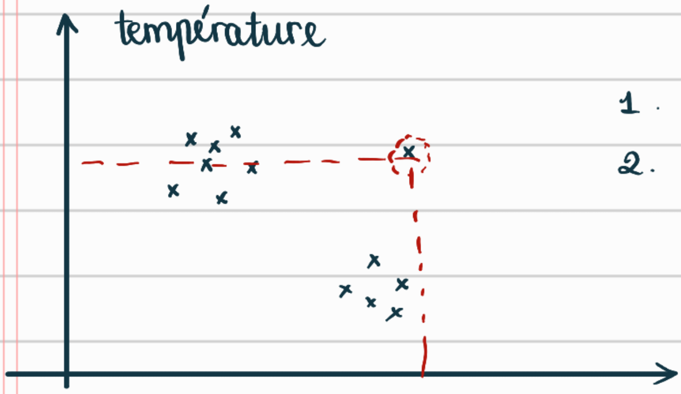
III. Algorithme EM appliqué aux mélanges gaussiens.

Exemple d'application de l'estimation de modèle :



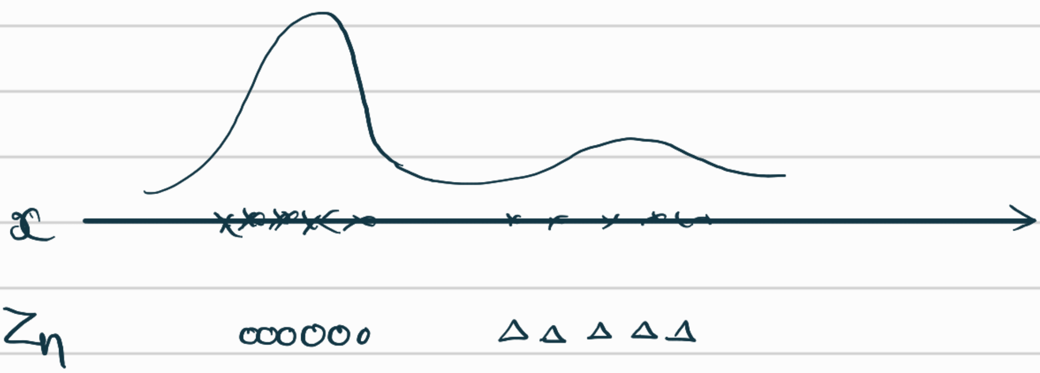
1. On estime un modèle $p(x)$
2. On veut un règle de décision
   $p(x_0) < \varepsilon \rightarrow$ anomalie

Exemple introductif "simple" :

Estimation par MV d'une gaussien simple $(K = 1)$
$X = (x_1 \_\_ x_N)^T$ iid $x_N \sim \mathcal{N}(\mu^*, \Sigma^*)$ $\forall n \in [\![1, N]\!]$.
On cherche $\theta_{ML} = (\mu_{ML}, \Sigma_{ML})$ qui maximise
log-vraisemblance :

$$\ell(\theta) = \ln(p(X|\theta)) = \ln\left(\prod_{n=1}^{N} p(x_n|\theta)\right)$$

$$= \sum_{n=1}^{N} \ln(p(x_n|\theta))$$



On introduit la log-vraisemblance "complétée" :

$$\ell'(\theta) = \ln(p(X|\theta))$$

Idée :

$$\ell(\theta) = \ln(p(x|\theta)) = \ln(p(x,z|\theta)) - \ln(p(z|x,\theta))$$

$$= \ln\left(\frac{p(x,z|\theta)}{q(z)}\right) - \ln\left(\frac{p(z|x,\theta)}{q(z)}\right)$$

$\times q(z)$

$$\sum_z = \underbrace{\sum_z q(z)\cdot\ln\left(\frac{p(x,z|\theta)}{(z)}\right)}_{\mathcal{L}(q,\theta)} + \underbrace{\sum_z -q(z)\cdot\ln\left(\frac{p(z|x,\theta)}{q(z)}\right)}_{= KL(q \,\|\, p(z|x,\theta))}$$

Rappel sur le Kullback-Leibler :

Si P et Q sont 2 var de densités p et q, alors $KL(P \| Q) = \int_{-\infty}^{+\infty} p(x)\cdot\ln\left(\frac{p(x)}{q(x)}\right) dx$

"KL de P per % à Q"

Rmqs : • meme de l'information

EM algorithm : Expectation Maximization algo