Consignes :
- Sont interdits : Documents, calculettes, téléphones, écouteurs, ordinateurs, tablettes.
- Il est interdit de composer avec un crayon.
- Votre feuille double d'examen doit porter, à l'emplacement réservé, vos nom, prénom, et signature.
- Cette zone réservée doit être cachée par collage.
- Vos feuilles intercalaires doivent être toutes numérotées.
- Le barème est donné à titre indicatif.

Segmentation de séries temporelles : On considère une série temporelle $(Y_1, \ldots, Y_n)$, avec $Y_t \in \mathbb{R}$, régie par un processus latent $(Z_1, \ldots, Z_n)$ à deux états où Z_t représente l'état (binaire) à l'instant $t\,(t = 1, \ldots, n)$. $Z_t \in \{0, 1\}$. On dispose d'une série observée $(y_1, \ldots, y_n)$. L'objectif est de segmenter la série sur la base du modèle défini par la densité :

$$\begin{aligned} f(y_t; \boldsymbol{\theta}) &= \pi(t; \boldsymbol{w}) f(y_t | Z_t = 1; \boldsymbol{\alpha}, \sigma^2) + (1 - \pi(t; \boldsymbol{w})) f(y_t | Z_t = 0; \boldsymbol{\beta}, \sigma^2) \\ &= \pi(t; \boldsymbol{w}) \mathcal{N}(y_t; \boldsymbol{x}_t^\top \boldsymbol{\alpha}, \sigma^2) + (1 - \pi(t; \boldsymbol{w})) \mathcal{N}(y_t; \boldsymbol{x}_t^\top \boldsymbol{\beta}, \sigma^2) \end{aligned}$$

où $\pi(t; \boldsymbol{w})$ est la probabilité de l'état 1 à l'instant t définie par :

$$\pi(t; \boldsymbol{w}) = \mathbb{P}(Z_t = 1; \boldsymbol{w}) = \frac{\exp(\boldsymbol{w}^\top \boldsymbol{x}_t)}{1 + \exp(\boldsymbol{w}^\top \boldsymbol{x}_t)}$$

où $\boldsymbol{x}_t = (1, t)^\top$ et $\boldsymbol{\theta} = (\boldsymbol{w}^\top, \boldsymbol{\alpha}^\top, \boldsymbol{\beta}^\top, \sigma^2)^\top$ est le vecteur paramètre inconnu du modèle.

On estime $\boldsymbol{\theta}$ à partir des données en maximisant la log-vraisemblance $L(\boldsymbol{\theta}) = \sum_{t=1}^n \ln f(y_t; \boldsymbol{\theta})$ par l'algorithme Espérance-Maximisation (EM). On montre que pour ce modèle l'algorithme EM consiste à partir d'un modèle initial de paramètre $\boldsymbol{\theta}^{(0)}$ et alterner à chaque itération q entre les deux étapes E- et M- suivantes jusqu'à la convergence :

Étape E : Calculer la fonction $Q(\boldsymbol{\theta}; \boldsymbol{\theta}^{(q)})$ définie par :

$$Q(\boldsymbol{\theta}; \boldsymbol{\theta}^{(q)}) = \sum_{t=1}^n \Big\{ \tau_t(\boldsymbol{\theta}^{(q)}) \ln \big[\pi(t; \boldsymbol{w}) \mathcal{N}(y_t; \boldsymbol{x}_t^\top \boldsymbol{\alpha}, \sigma^2)\big] + (1 - \tau_t(\boldsymbol{\theta}^{(q)})) \ln \big[(1 - \pi(t; \boldsymbol{w})) \mathcal{N}(y_t; \boldsymbol{x}_t^\top \boldsymbol{\beta}, \sigma^2)\big] \Big\}$$

où $\tau_t(\boldsymbol{\theta}^{(q)}) = \mathbb{P}(Z_t = 1 | y_t, \boldsymbol{\theta}^{(q)})$ est la probabilité a posteriori de l'état 1 à l'instant t.

Étape M : Mettre à jour les paramètres du modèle en calculant $\boldsymbol{\theta}^{(q+1)}$ définie par :

$$\boldsymbol{\theta}^{(q+1)} = \arg\max_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}; \boldsymbol{\theta}^{(q)}).$$

1. Montrer que les probabilités a posteriori calculées à l'étape E- sont définies par :

$$\tau_t(\boldsymbol{\theta}^{(q)}) = \mathbb{P}(Z_t = 1 | y_t; \boldsymbol{\theta}^{(q)}) = \frac{\pi(t; \boldsymbol{w}^{(q)}) \, \mathcal{N}(y_t; \boldsymbol{x}_t^\top \boldsymbol{\alpha}^{(q)}, \sigma^{2(q)})}{f(y_t; \boldsymbol{\theta}^{(q)})} \quad \forall t = 1, \ldots, n \tag{1}$$

2. On rappelle la densité d'une v.a. X de densité normale : $\mathcal{N}(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x - \mu)^2\right)$.
Montrer que l'étape M- est donnée par les équations de mise à jour suivantes :

$$\begin{aligned} \boldsymbol{\alpha}^{(q+1)} &= \Big[\sum_{t=1}^n \tau_t(\boldsymbol{\theta}^{(q)}) \boldsymbol{x}_t \boldsymbol{x}_t^\top\Big]^{-1} \sum_{t=1}^n \tau_t(\boldsymbol{\theta}^{(q)}) y_t \boldsymbol{x}_t & (2) \\ \boldsymbol{\beta}^{(q+1)} &= \Big[\sum_{t=1}^n (1 - \tau_t(\boldsymbol{\theta}^{(q)})) \boldsymbol{x}_t \boldsymbol{x}_t^\top\Big]^{-1} \sum_{t=1}^n (1 - \tau_t(\boldsymbol{\theta}^{(q)})) y_t \boldsymbol{x}_t & (3) \\ \sigma^{2(q+1)} &= \frac{1}{n} \sum_{t=1}^n \Big\{ \tau_t(\boldsymbol{\theta}^{(q)}) (y_t - \boldsymbol{x}_t^\top \boldsymbol{\alpha}^{(q+1)})^2 + (1 - \tau_t(\boldsymbol{\theta}^{(q)})) (y_t - \boldsymbol{x}_t^\top \boldsymbol{\beta}^{(q+1)})^2 \Big\} & (4) \\ \boldsymbol{w}^{(q+1)} &= \boldsymbol{w}^{(q)} - \left(\frac{\partial^2 Q(\boldsymbol{\theta}; \boldsymbol{\theta}^{(q)})}{\partial \boldsymbol{w} \partial \boldsymbol{w}^\top}\right)^{-1}_{\boldsymbol{w} = \boldsymbol{w}^{(q)}} \left(\frac{\partial Q(\boldsymbol{\theta}; \boldsymbol{\theta}^{(q)})}{\partial \boldsymbol{w}}\right)_{\boldsymbol{w} = \boldsymbol{w}^{(q)}} & \\ &= \boldsymbol{w}^{(q)} + \left(\sum_{t=1}^n \boldsymbol{x}_t \boldsymbol{x}_t^\top \pi(\boldsymbol{x}_t; \boldsymbol{w}^{(q)}) (1 - \pi(\boldsymbol{x}_t; \boldsymbol{w}^{(q)}))\right)^{-1} \sum_{t=1}^n \boldsymbol{x}_t (\tau_t(\boldsymbol{\theta}^{(q)}) - \pi(\boldsymbol{x}_t; \boldsymbol{w}^{(q)})) & (5) \end{aligned}$$

3. Qu'appelle-t-on l'algorithme de mise à jour itérative de l'équation (5).