Paco GOZE
11/11/2024
Compstats

TP2 : Algorithme EM - Importance sampling

I-

① $X = \{x_1, \dots, x_n\}$ avec x_i distincts
$\forall i \in [\![1,n]\!],\ p_i > 0$ et $\sum_{i=1}^{n} p_i = 1$

Pour générer X variable aléatoire discrète sur $\{x_1, \dots, x_n\}$ telle que : $\forall i \in [\![1,n]\!],\ \mathbb{P}(X = x_i) = p_i$: méthode d'inversion.

$F(t) = \mathbb{P}(X \leq t) = \sum_{i=1}^{n} p_i \mathbb{1}_{x_i \leq t}$

On pose $\boxed{G :]0,1[\to [\![1,n]\!],\ \alpha \mapsto \inf\{ j \in [\![1,n]\!] ;\ \sum_{i=1}^{j-1} p_i < \alpha \leq \sum_{i=1}^{j} p_i \}}$

Alors si $U \sim \mathcal{U}_{[0,1]}$, on pose $X = G(U)$ et X suit la distribution recherchée.

II-

① À l'étape (n) de l'algorithme Population Monte-Carlo, on cherche

$$\max_{\alpha, \mu, \Sigma} \sum_{i=1}^{n} \tilde{\omega}_i^{(0)} \log\left(\sum_{j=1}^{m} \alpha_j \varphi(x_i^{(0)} ; \theta_j) \right)$$

puis on réécrit :

$$\theta^{(1)} \in \underset{\theta}{\operatorname{argmax}} \sum_{i=1}^{n} \tilde{\omega}_i^{(0)} \log(q_\theta(x_i^{(0)}))$$

avec $x_i^{(0)} \sim q_{\theta^{(0)}}$ et $\tilde{\omega}_i^{(0)} = \frac{\nu(x_i^{(0)})}{q_{\theta^{(0)}}(x_i^{(0)})}$

~~On a~~ $\cancel{\tilde{w}_i^{(0)} \log\left(q_\theta(x_i^{(0)})\right) \geq \tilde{w}_i^{(0)} \mathbb{E}_{q_{\theta^{(0)}}(\cdot|x_i^{(0)})} \log\left[q_\theta(x_i^{(0)}, z_i^{(0)})\right]}$

On a $\tilde{w}_i^{(0)} \log\left(q_\theta(x_i^{(0)})\right) \geq \tilde{w}_i^{(0)} \mathbb{E}_{q_{\theta^{(0)}}(\cdot|x_i^{(0)})}\left[\log\left(\dfrac{q_\theta(x_i^{(0)}, z_i^{(0)})}{q_{\theta^{(0)}}(z_i^{(0)}|x_i^{(0)})}\right)\right]$

Rque: j'utilise un indice (0) ici, mais en fait de même $\forall t$.

On veut donc maximiser en θ :

$$\sum_{i=1}^{n} \tilde{w}_i^{(0)} \mathbb{E}_{q_{\theta^{(0)}}(\cdot|x_i^{(0)})} \log\left(q_\theta(x_i^{(0)}, z_i^{(0)})\right)$$

avec $x_i^{(0)} \sim q_{\theta^{(0)}}$

On procède avec l'algorithme EM :

E-Step : calculer $\displaystyle\sum_{i=1}^{n} \tilde{w}_i^{(t)} \mathbb{E}_{q_{\theta^{(t)}}(\cdot|x_i^{(t)})} \log\left(q_\theta(x_i^{(t)}, z_i^{(t)})\right)$ $(*)$

avec $x_i^{(t)} \sim q_{\theta^{(t)}}$

et $\tilde{w}_i^{(t)} = w_i^{(t)} / \sum_{i=1}^{n} w_i^{(t)}$ où $w_i^{(t)} = \dfrac{\nu(x_i^{(t)})}{q_{\theta^{(t)}}(x_i^{(t)})}$

M-Step : $\theta^{t+1} = \arg\max_\theta (*)$.

On est dans la même situation que dans le cas vu en cours. Seulement, il y a les facteurs $\tilde{w}_i^{(t)}$ à prendre en compte.

On veut donc maximiser (avec $x_1^{(t)}, \dots, x_n^{(t)}$) :

$$\sum_{i=1}^{n} \tilde{w}_i^{(t)} \left[\sum_{j=1}^{M} \alpha_j^{t+1}(x_i^{(t)}) \log(\alpha_j^{(t)}) + \alpha_j^{(t+1)}(x_i^{(t)}) \log \varphi_{\mu_j^{(t)}, \sigma_j^{2(t)}}(x_i^{(t)})\right]$$

sc $\displaystyle\sum_{j=1}^{M} \alpha_j^{(t)} = 1$

où $\alpha_j^{(t+1)}(x_i^{(t)}) = \varphi_{\mu_j^{(t)}, \sigma_j^{(t)2}}(x_i^{(t)}) \lambda_j^{(t)} \Big/ \displaystyle\sum_{j=1}^{M} \varphi_{\mu_j^{(t)}, \sigma_j^{(t)2}}(x_i^{(t)}) \lambda_j^{(t)}$

On obtient (optimisation avec Lagrangien) :

$$\boxed{\begin{aligned} \mu_j^{(t+1)} &= \sum_{i=1}^{n} \tilde{\omega}_i^{(t)} \alpha_j^{(t+1)}(x_i^{(t)})\, x_i^{(t)} \Big/ \sum_{i=1}^{n} \tilde{\omega}_i^{(t)} \alpha_j^{(t+1)}(x_i^{(t)}) \\ \alpha_j^{(t+1)} &= \frac{\sum_{i=1}^{n} \tilde{\omega}_i^{(t)} \alpha_j^{(t+1)}(x_i^{(t)})}{\sum_{j=1}^{M} \sum_{i=1}^{n} \tilde{\omega}_i^{(t)} \alpha_j^{(t+1)}(x_i^{(t)})} \\ \Sigma_j^{(t+1)} &= \frac{\sum_{i=1}^{n} \tilde{\omega}_i^{(t)} \alpha_j^{(t+1)}(x_i^{(t)}) \left(x_i^{(t)} - \mu_j^{(t+1)}\right)\left(x_i^{(t)} - \mu_j^{(t+1)}\right)^T}{\sum_{i=1}^{n} \tilde{\omega}_i^{(t)} \alpha_j^{(t+1)}(x_i^{(t)})} \end{aligned}}$$