Basile TERVER

TP 2 Computational Statistics

Ex. 1 :

(1) On simule X en utilisant la méthode d'inversion.
Soit F la fonction de répartition de X et F^{-1} son inverse généralisé, i.e : $F^{-1} : u \in [0,1] \mapsto \inf\{x : F(x) \geq u\}$.
Par le théorème de la réciproque, si U est une variable aléatoire de loi $\mathcal{U}[0,1]$, $F^{-1}(U) \sim \mathcal{L}oi(X)$. On a explicitement :
$\forall u \in [0,1], F^{-1}(u) = \sum_{j=1}^{n} x_j \mathbb{1}_{[p_0+\dots+p_{j-1}; p_1+\dots+p_j]}(u)$
où l'on note $p_0 = 0$.
Ainsi, si on peut générer une variable uniforme $U \sim \mathcal{U}[0,1]$, on simule un échantillon $(u_1, \dots, u_k)$ et on pose $y_i = x_j$ si $u_i \in [p_0+\dots+p_{j-1}; p_1+\dots+p_j]$, alors $(y_1, \dots, y_k)$ est un échantillon de la loi $\mathcal{L}(X)$.

Ex. 2

(1) (i) Les paramètres sont $\theta = (p_j, \alpha_j, \Sigma_j)_{j=1}^{p}$
avec $(\alpha_j)_{j=1}^{p} \in \text{Simplexe}(p) = \{x \in \mathbb{R}^p : \sum_{i=1}^{p} x_i = 1\}$, $\mu_j \in \mathbb{R}^d$ et $\Sigma_j \in S_d^{++}(\mathbb{R})$ pour tout $j \in [\![1, p]\!]$ (le groupe des matrices symétriques définies positives, qui est un ouvert de $S_d(\mathbb{R}) \subset \mathcal{M}_d(\mathbb{R})$).

(ii) $p_\theta(x) = \sum_{j=1}^{p} p_\theta(Z=j)\, p_\theta(x \mid Z=j)$ avec :

- $p_\theta(x \mid Z=j) = \frac{1}{(2\pi)^{d/2}\sqrt{\det(\Sigma_j)}} \exp\left(-\frac{(x-\mu_j)^T \Sigma_j^{-1} (x-\mu_j)}{2}\right)$
- $p_\theta(Z=j) = \alpha_j$

Pour un échantillon iid $(x_i)_{i=1}^{n}$ de la loi X on a donc
$p_\theta(x_1, \dots, x_n) = \prod_{i=1}^{n} \sum_{j=1}^{p} \alpha_j p_\theta(x_i \mid Z=j)$, donc :
$\log p_\theta(x_1, \dots, x_n) = \sum_{i=1}^{n} \log\left[\sum_{j=1}^{p} \alpha_j p_\theta(x_i \mid Z=j)\right]$

Scanned with CamScanner

• Rappel:

③ Dans l'algorithme EM, on veut trouver

$\theta^* = \arg\max_\theta p_\theta(x_1, \dots, x_n)$

On a, en toute généralité, pour toute observation x,

$$p_\theta(x) = \int_{\mathcal{Z}} p_\theta(x,z)\,dz = \int_{\mathcal{Z}} \frac{p_\theta(x,z)}{q(z)} q(z)\,dz = E_{Z\sim q}\left[\frac{p_\theta(x,Z)}{q(Z)}\right]$$

où $\mathcal{Z}$ est l'espace des variables latentes. Par Jensen

on a : $\log p_\theta(x) \geq E_{Z\sim q}[\log(p_\theta(x,Z))] - \underbrace{E_{Z\sim q}[\log(q(Z))]}_{H(Z)\text{: entropie}}$

$= F(q,\theta)(x)$

Comme $\arg\max_q F(q,\theta)(\cdot) = p_\theta(\cdot|x)$, on a donc l'algorithme EM : à la fin du pas t, on calcule :

(*) $\theta^{t+1} = \arg\max_\theta E_{Z\sim p_{\theta^t}(\cdot|x)}[\log(p_\theta(x,Z))]$

• Dans notre cas : La formule (*) donne, pour n observations :

$$\theta^{t+1} = \arg\max_\theta E_{(Z_i\sim p_{\theta^t}(\cdot|x_i))_{i=1}^n}\left[\log p_\theta((x_i)_{i=1}^n, (Z_i)_{i=1}^n)\right]$$

(par question ①)

$$= \arg\max_\theta \sum_{i=1}^n \underbrace{E_{Z_i\sim p_{\theta^t}(\cdot|x_i)}[\log p_\theta(x_i, Z_i)]}_{\text{Notons } F_i(\theta) \text{ cette quantité}}$$

Par la formule de Bayes : $\log(p_\theta(x,z)) = \log p_\theta(x|z) + \log p_\theta(z)$.

D'où : $F_i(\theta) = E_{Z_i\sim p_{\theta^t}(\cdot|x_i)}[\log(p_\theta(x_i|Z_i)) + \log(p_\theta(Z_i))]$

$$= \sum_{j=1}^p p_{\theta^t}(Z_i=j|x_i)\left[\log(p_\theta(x_i|Z_i=j)) + \log(\underbrace{p_\theta(Z_i=j)}_{=\alpha_j})\right] \quad (E_1)$$

On a :

$$p_{\theta^t}(Z|x) = \frac{p_{\theta^t}(x,Z)}{p_{\theta^t}(x)} = \frac{p_{\theta^t}(x|Z)\,p_{\theta^t}(Z)}{\int p_{\theta^t}(x|\tilde{Z})\,p_{\theta^t}(\tilde{Z})\,d\tilde{Z}}, \text{ donc :}$$

$$p_{\theta^t}(Z_i=j|x_i) = \frac{p_{\theta^t}(x_i|Z_i=j)\,\alpha_j^t}{\underbrace{\sum_{k=1}^p p_{\theta^t}(x_i|Z=k)\,\alpha_k^t}_{\text{notons } \tau_{ij} \text{ cette quantité (indépendante de } \theta)}}.$$

D'après la question ①, on a :

$$\log(p_\theta(x_i|Z_i=j)) = -\frac{d}{2}\log(2\pi) - \frac{1}{2}\log(\det(\Sigma_j)) - \frac{1}{2}(x_i-\mu_j)^T\Sigma_j^{-1}(x_i-\mu_j) \quad (E_2)$$

Scanned with CamScanner

- On note $V = (\mathbb{R}^d \times (\mathbb{R}_+^*) \times S_d^{++}(\mathbb{R}))^p$

- En somme, pour obtenir θ^{t+1} sachant l'étape t terminée, on doit résoudre le problème de maximisation suivant :

(M) $$\max_{\theta=(\mu_j,\alpha_j,A_j)\in V} \sum_{i=1}^{n} F_i(\theta) \quad s.c : \sum_{j=1}^{p} \alpha_j = 1, \quad \alpha_j \geq 0 \ \forall j \in [\![1,p]\!]$$

où l'on note $A_j := \Sigma_j^{-1}$ dans les formules précédentes.

On note $F(\theta) := \sum_{i=1}^{n} F_i(\theta)$. (Remarque : $A_j \in S_d^{++}(\mathbb{R}) \Leftrightarrow \Sigma_j \in S_d^{++}(\mathbb{R})$ d'après E_1)

- Chaque $F_i(\theta)$ est concave en θ comme somme positive (avec coefficients $\tau_{ij} = p_{\theta^t}(Z_i = j | x_i)$ indépendants de θ) de fonctions concaves en θ. En effet, d'après E_2 et E_1, on a :

$$\log(p_\theta(x_i | Z_i = j)) = -\frac{d}{2}\log(2\pi) + \underbrace{\frac{1}{2}\log(\det(A_j))}_{\text{concave en } A_j} - \underbrace{\frac{1}{2}(x_i - \mu_j)^T A_j (x_i - \mu_j)}_{\text{concave en } (\mu_j, A_j)}$$

et trivialement $\theta \mapsto \log(p_\theta(Z_i = j)) = \log \alpha_j$ est concave.

Donc la fonction objectif F du problème de maximisation (M) est concave.

- Les contraintes d'égalité affines "$\sum_{j=1}^{p} \alpha_j = 1$" peuvent s'écrire sous la forme de contraintes d'inégalité affines, donc convexes.

Les contraintes d'inégalité "$\alpha_j \geq 0, \forall j \in [\![1,p]\!]$" sont affines donc convexes.

Les contraintes d'inégalité "$A_j \in S_d^{++}(\mathbb{R}) \ \forall j \in [\![1,p]\!]$" sont convexes bien que non-linéaires, ce qui est le résultat fondamental de la programmation semi-définie (on utilise la norme de Frobenius et le cône des matrices symétriques définies positives).

- On écrit le Lagrangien associé à (M) pour tout $\lambda \in \mathbb{R}$ et $\theta \in V$, $\mathcal{L}(\theta, \lambda) = \sum_{i=1}^{n} F_i(\theta) + \lambda(1 - \sum_{j=1}^{p} \alpha_j)$

Scanned with CamScanner

Enfin, tout point $\theta \in V$ est un point où la contrainte $\sum_{j=1}^{p} \alpha_j = 1$ est qualifiée.

- Le théorème de Kuhn-Tucker donne donc que $\theta \in V$ est un maximum global de F sur V si, et seulement si il existe $\lambda \in \mathbb{R}$ tel que (θ, λ) soit un point-selle du lagrangien $\mathcal{L}$ sur $V \times \mathbb{R}$, i.e tel que :

$$\theta \in V,\ \sum_{j=1}^{p} \alpha_j = 1, \quad \begin{cases} \dfrac{\partial \mathcal{L}(\theta,\lambda)}{\partial \alpha_j} = 0 & \forall j \in [\![1,p]\!] \quad (L_1) \\ \dfrac{\partial \mathcal{L}(\theta,\lambda)}{\partial \mu_j} = 0 & \forall j \in [\![1,p]\!] \quad (L_2) \\ \dfrac{\partial \mathcal{L}(\theta,\lambda)}{\partial (A_j)} = 0 & \forall j \in [\![1,p]\!] \quad (L_3) \end{cases}$$

On rappelle que, d'après (E_1) et (E_2) :

$$\mathcal{L}(\lambda,\theta) = \sum_{i=1}^{n} \sum_{j=1}^{p} \tau_{ij} \left[-\frac{d}{2}\log(2\pi) + \frac{1}{2}\log(\det A_j) - \frac{1}{2}(x_i-\mu_j)^T A_j (x_i-\mu_j) + \log(\alpha_j) \right] + \lambda\left(1 - \sum_{k=1}^{p} \alpha_k\right)$$

Notons $\tau_{ij} = \dfrac{\tilde{\tau}_{ij}}{\sum_{k=1}^{p} \tilde{\tau}_{ik}}$ soit $\tilde{\tau}_{ij} = p_{\theta^t}(x_i \mid Z_i = j)\, \alpha_j^t$ donc $\sum_{j=1}^{p} \tau_{ij} = 1$ et $\sum_{i=1}^{n} \tau_{ij} =$

(L_1) donne : $\sum_{i=1}^{n} \tau_{ij} \dfrac{1}{\alpha_j} - \lambda = 0 \quad \forall j \in [\![1,p]\!]$ et comme $\sum_{j=1}^{p} \alpha_j = 1$ et $\sum_{j=1}^{p} \tau_{ij} = 1$ on obtient $\lambda \sum_{j=1}^{p} \alpha_j = \sum_{i=1}^{n} \sum_{j=1}^{p} \tau_{ij} = n$ soit $\lambda = n$

(L_2) $\forall j \in [\![1,p]\!]$, $\sum_{i=1}^{n} \tau_{ij} A_j (x_i - \mu_j) = 0$ soit

(L_3) donne : $\sum_{i=1}^{n} \tau_{ij} \left[\dfrac{\Sigma_j}{2} - \dfrac{1}{2}(x_i-\mu_j)(x_i-\mu_j)^T \right] = 0$ car pour une matrice symétrique définie positive M : $\partial_M \log\det M = M^{-1}$.

On obtient donc :

$$\boxed{\alpha_j^{t+1} = \frac{\sum_{i=1}^{n} \tau_{ij}}{n}}$$

donc $\sum_{i=1}^{n} \tau_{ij} = n\alpha_j^{t+1}$ et donc

(L_2) nous donne $n\alpha_j^{t+1} A_j^{t+1} \mu_j^{t+1} = \sum_{i=1}^{n} \tau_{ij} A_j^{t+1} x_i$ donc $n\alpha_j^{t+1} \mu_j^{t+1} = \sum_{i=1}^{n} \tau_{ij} x_i$

Soit $\boxed{\mu_j^{t+1} = \dfrac{\sum_{i=1}^{n} \tau_{ij} x_i}{\sum_{i=1}^{n} \tau_{ij}}}$ car A_j est inversible pour tout $j \in [\![1,p]\!]$

Donc (L_3) donne $\boxed{\Sigma_j^{t+1} = \dfrac{\sum_{i=1}^{n} (x_i-\mu_j)(x_i-\mu_j)^T \tau_{ij}}{\sum_{i=1}^{n} \tau_{ij}}}$

Ainsi, le pseudo-code de l'algorithme EM appliqué à un modèle de GMM peut s'écrire :

Scanned with CamScanner

Basile TERVER Copie 2

Init : $\theta = (\alpha, \Sigma, \mu)$

Repeat until convergence :

$$\tau_{ij} \leftarrow \frac{\mathcal{N}(\mu_j, \Sigma_j)(x_i)\,\alpha_j}{\sum_{k=1}^{p} \mathcal{N}(\mu_k, \Sigma_k)(x_i)\,\alpha_k} \qquad \text{(E-step)}$$

$$\alpha_j \leftarrow \frac{\sum_{i=1}^{n} \tau_{ij}}{n} \qquad \text{(M-step)}$$

$$\mu_j \leftarrow \frac{\sum_{i=1}^{n} \tau_{ij} x_i}{\sum_{i=1}^{n} \tau_{ij}}$$

$$\Sigma_j \leftarrow \frac{\sum_{i=1}^{n} (x_i - \mu_j)(x_i - \mu_j)^T \tau_{ij}}{\sum_{i=1}^{n} \tau_{ij}}$$

Exercice 3

(4) Dans l'étape (t+1) de Population Monte-Carlo, il s'agit de trouver, à l'étape (iii) :

$$\text{PM}(t+1)\ \theta^{(t+1)} \in \underset{\theta}{\operatorname{argmax}} \sum_{i=1}^{n} \tilde{\omega}_i^{(t)} \log(q_\theta(X_i^{(t)})) \quad \text{avec } \tilde{\omega}_i^{(t)} \propto \frac{\pi(X_i^{(t)})}{q_{\theta^{(t)}}(X_i^{(t)})}$$

De même que pour l'Exercice 2, on a, sachant que $\forall i \in [\![1, n]\!]$, $\tilde{\omega}_i^{(t)} \geq 0$, que pour tout $i \in [\![1, n]\!]$:

$$\tilde{\omega}_i^{(t)} \log(q_\theta(X_i^{(t)})) \geq \tilde{\omega}_i^{(t)} E_{Z_i^{(t)} \sim q_{\theta^{(t)}}(\cdot | X_i^{(t)})} \left[\log\left(\frac{p_\theta(X_i^{(t)}, Z_i^{(t)})}{q_{\theta^{(t)}}(Z_i^{(t)} | X_i^{(t)})} \right) \right]$$

On cherche toujours à maximiser en θ la borne inférieure (à droite dans l'équation ci-dessus), dont le seul terme dépendant de θ est : $\tilde{\omega}_i^{(t)} E_{Z_i^{(t)} \sim q_{\theta^{(t)}}(\cdot | X_i^{(t)})} \left[\log(p_\theta(X_i^{(t)}, Z_i^{(t)})) \right]$

Ainsi, on peut appliquer l'algorithme pour résoudre PM(t+1) avec pour étape (M) :

$$\theta_{(k+1)}^{(t+1)} \in \underset{\theta}{\operatorname{argmax}} \sum_{i=1}^{n} \tilde{\omega}_i^{(t)} \underbrace{E_{Z_i^{(t)} \sim q_{\theta_{(k)}^{(t+1)}}(\cdot | X_i^{(t)})} \left[\log(p_\theta(X_i^{(t)}, Z_i^{(t)})) \right]}_{\text{mêmes } F_i(\theta) \text{ qu'en Exercice 2 mais les observations sont désormais } X_i^{(t)}, \text{ générées sous } q_{\theta^{(t)}}}$$

Attention : ici l'indice $(k+1)$ dans $\theta_{(k+1)}^{(t+1)}$ indique l'itération d'EM $(k+1)$ de l'itération $(t+1)$ de Population Monte-Carlo à ne pas confondre avec le nombre de clusters K. Comme $q_{\theta_{(k)}^{(t+1)}}$ est un mélange gaussien avec modèle hiérarchique où Z est la variable catégorique du tirage du cluster comme dans l'Exercice 2, on peut appliquer le même calcul des $F_i(\theta)$ que dans l'exercice 2 :

$$F_i(\theta) = \sum_{j=1}^{K} \tau_{ij} \left[\log(\alpha_j) - \frac{1}{2} \log(\det(\Sigma_j)) - \frac{1}{2} (X_i^{(t)} - \mu_j)^T \Sigma_j^{-1} (X_i^{(t)} - \mu_j) \right]$$

avec $\theta = (\alpha_j, \mu_j, \Sigma_j)_{j=1}^{K}$ et $\boxed{\tau_{ij}^{(k)} = \frac{q_{\theta_{(k)}^{(t+1)}}(X_i^{(t)} | Z_i^{(t)} = j)\,\alpha_j^{(t)}}{\sum_{\ell=1}^{K} q_{\theta_{(k)}^{(t+1)}}(X_i^{(t)} | Z_i^{(t)} = \ell)\,\alpha_\ell^{(t)}}}$

Scanned with CamScanner

Donc pour résoudre (PM), (E+) avec l'EM on cherche à résoudre

$$\theta_{(k+1)}^{(t+1)} \in \underset{\theta}{\operatorname{argmax}} \sum_{j=1}^{K} \left\{ \log(\alpha_j)(\tilde{\omega}\tau)_j^{(k)} - \frac{1}{2}(\tilde{\omega}\tau)_j^{(k)} \log(\det(\Sigma_j)) - \frac{1}{2}\sum_{i=1}^{n} \tilde{\omega}_i^{(t)} \tau_{ij}^{(k)} (X_i^{(t)} - \mu_j)^T \Sigma_j^{-1} (X_i^{(t)} - \mu_j) \right\}$$

s.c : $(\alpha_j > 0$ et $\sum_{j=1}^{K} \alpha_j = 1$

où l'on a noté $(\tilde{\omega}\tau)_j^{(k)} = \sum_{i=1}^{n} \tilde{\omega}_i^{(t)}$

On a toujours un problème concave donc on détermine un maximum par Kuhn-Tucker en trouvant des points critiques du lagrangien :

- $\nabla_\alpha \mathcal{L}(\theta, \lambda) = 0 \iff \lambda = \frac{(\omega\tau)_j^{(k)}}{\alpha_j}$

  donc $\sum_{j=1}^{K} (\omega\tau)_{j,(k)} = \lambda \sum_{j=1}^{K} \alpha_j = \lambda$

  et : $\sum_j (\omega\tau)_j = \sum_j \sum_i \tilde{\omega}_i^{(t)} \tau_{ij} = \sum_{i=1}^{n} \tilde{\omega}_i^{(t)} := \tilde{\omega}^{(t)}$

  donc $\lambda = \tilde{\omega}^{(t)}$ et $\boxed{(\alpha_j)_{(k+1)}^{(t+1)} = \frac{(\omega\tau)_j^{(k)}}{\tilde{\omega}^{(t)}}}$ $\forall j \in \{1, \dots, K\}$

- $\nabla_{\mu_j} \mathcal{L}(\theta, \lambda) = 0 \iff \sum_{i=1}^{n} \tau_{ij}^{(k)} \tilde{\omega}_i^{(t)} \Sigma_j^{-1} (X_i^{(t)} - \mu_j) = 0$

  $\iff \boxed{(\mu_j)_{(k+1)}^{(t+1)} = \frac{\sum_{i=1}^{n} \tilde{\omega}_i^{(t)} \tau_{ij}^{(k)} X_i^{(t)}}{(\omega\tau)_j^{(k)}}}$ $\forall j \in \{1, \dots, K\}$

- $\nabla_{(\Sigma_j)^{-1}} \mathcal{L}(\theta, \lambda) = 0 \iff (\omega\tau)_j^{(k)} \frac{\Sigma_j}{2} - \frac{1}{2}\sum_{i=1}^{n} (X_i^{(t)} - \mu_j)(X_i^{(t)} - \mu_j)^T = 0$

  $$\iff \boxed{(\Sigma_j)_{(k+1)}^{(t+1)} = \frac{1}{(\omega\tau)_j^{(k)}} \sum_{i=1}^{n} \tilde{\omega}_i^{(t)} \tau_{ij}^{(k)} (X_i^{(t)} - \underbrace{\mu_j}_{(\mu_j)_{(k+1)}^{(t+1)}})(X_i^{(t)} - \underbrace{\mu_j}_{(\mu_j)_{(k+1)}^{(t+1)}})^T}$$

Scanned with CamScanner