Basile Terver
Copie 1

Computational Statistics TP 3

Exercice 1

① Les paramètres du modèles sont $\theta = (\underbrace{\bar{t}_0, \bar{v}_0}_{\theta_{pop}}, \underbrace{\sigma_\xi, \sigma_\tau, \sigma}_{\theta_{ind}})$

Les variables latentes sont:

$z_{pop} = (t_0, v_0)$, $\{z_i\}_i = \{(\alpha_i, \tau_i)\}_i$.

Donc, la log-vraisemblance du modèle pour les observations $\{y_{ij}\}_{i,j}$ est par la formule de Bayes:

$$\log p(\{y_{ij}\}_{i,j}, \{z_i\}_i, z_{pop}, \theta) = \log p(\{y_{ij}\}_{ij} \mid \{z_i\}_i, z_{pop}, \theta) \quad (1)$$
$$+ \log p(\{z_i\}_i, z_{pop} \mid \theta) \quad (2)$$
$$+ \log p(\theta) \quad (3)$$

On se concentre successivement sur chacun de ces 3 termes

(1) Comme $y_{ij} \overset{iid}{\sim} \mathcal{N}(d_i(t_{i,j}), \sigma^2)$, on a

$$\log p(\{y_{ij}\} \mid \{z_i\}_i, z_{pop}, \theta) = \sum_{i=1}^{N} \sum_{j=1}^{K_i} \log p(y_{ij} \mid \{z_i\}, z_{pop}, \theta)$$
$$= \sum_{i=1}^{N} \sum_{j=1}^{K_i} \left[-\frac{1}{2} \log(2\pi\sigma^2) - \frac{(y_{ij} - d_i(t_j))^2}{2\sigma^2} \right]$$
$$= -\frac{1}{2} \sum_{i=1}^{N} \left[K_i \log(2\pi\sigma^2) + \sum_{j=1}^{K_i} \frac{(y_{ij} - d_i(t_{ij}))^2}{\sigma^2} \right]$$
$$= -\frac{1}{2} \sum_{i=1}^{N} \left[K_i \log(\sigma^2) + \sum_{j=1}^{K_i} \frac{(y_{ij} - d_i(t_{ij}))^2}{\sigma^2} \right] + \underbrace{C_1}_{\text{constante}}$$

où $d_i(t_{ij}) = p_0 + v_0 \alpha_i (t_{ij} - t_0 - \tau_i)$

(2) Par la formule de Bayes on a puisque les $\{z_i\}$ sont iid:

$$\log p(\{z_i\}, z_{pop} \mid \theta) = \log p(z_{pop} \mid \theta) + \sum_{i=1}^{N} \log p(z_i \mid \theta)$$

- $z_{pop} = (t_0, v_0)$ avec $t_0 \sim \mathcal{N}(\bar{t}_0, \sigma_{t_0}^2)$ et $v_0 \sim \mathcal{N}(\bar{v}_0, \sigma_{v_0}^2)$

donc $\log p(z_{pop} \mid \theta) = -\frac{1}{2} \left[\frac{(t_0 - \bar{t}_0)^2}{\sigma_{t_0}^2} - \frac{(v_0 - \bar{v}_0)^2}{\sigma_{v_0}^2} \right] + \underbrace{C_2}_{\text{constante}}$

- $z_i = (\alpha_i, \tau_i)$ avec $\alpha_i = \exp(\xi_i)$ et $\xi_i \overset{iid}{\sim} \mathcal{N}(0, \sigma_\xi^2)$

donc $\alpha_i \overset{iid}{\sim} \text{Log-}\mathcal{N}(0, \sigma_\xi^2)$

$\tau_i \overset{iid}{\sim} \mathcal{N}(0, \sigma_\tau^2)$

Scanned with CamScanner

Donc : $\log p(\{z_i\}_i|\theta) = \sum_{i=1}^{N}\left[\log p(\tau_i|\theta) + \log p(\xi_i|\theta)\right]$

$$= \sum_{i=1}^{N}\left[-\log(\sigma_\xi \sigma_\tau) - \frac{1}{2\sigma_\xi^2}\xi_i^2 - \frac{1}{2\sigma_\tau^2}\tau_i^2\right] + \underbrace{C_3}_{\text{constante}}$$

(3) $p(\theta)$ est la distribution a priori des paramètres.

$$\log p(\theta) = \log p(\bar{t}_0, \bar{v}_0, \sigma_\xi, \sigma_\tau, \sigma)$$

$$= -\frac{1}{2s_{t_0}^2}(\bar{t}_0 - \bar{\bar{t}}_0)^2 - \frac{1}{2s_{v_0}^2}(\bar{v}_0 - \bar{\bar{v}}_0)^2$$

$$+ m_\xi \log\left(\frac{v_\xi}{\sigma_\xi}\right) - \log(\sigma_\xi^2) - \frac{1}{2\sigma_\xi^2}v_\xi^2$$

$$+ m_\tau \log\left(\frac{v_\tau}{\sigma_\tau}\right) - \log(\sigma_\tau^2) - \frac{1}{2\sigma_\tau^2}v_\tau^2$$

$$+ m \log\left(\frac{v}{\sigma}\right) - \log(\sigma^2) - \frac{1}{2\sigma^2}v^2 + \underbrace{C_4}_{\text{constante}}$$

Car, en prenant le log de la densité $W^{-1}(n, m)$, $\log\left(\frac{1}{\Gamma(\frac{m}{2})}\right)$ est un terme indépendant de $(\sigma_\xi, \sigma_\tau, \sigma)$.

En conclusion :

$$\log p(y, z, \theta) = -\frac{1}{2}\sum_{i=1}^{N}\left[K_i \log(\sigma^2) - \sum_{j=1}^{K_i}\frac{(y_{ij} - d_i(t_{ij}))^2}{\sigma^2}\right] \quad \log p(\{y_{ij}\}|\{z_i\}_i, z_{pp}, \theta)$$

$$+ \sum_{i=1}^{N}\left[-\log(\sigma_\xi \sigma_\tau) - \frac{\xi_i^2}{2\sigma_\xi^2} - \frac{\tau_i^2}{2\sigma_\tau^2}\right] \quad \log p(\{z_i\}_i|\theta)$$

$$-\frac{1}{2\sigma_{t_0}^2}(\bar{t}_0 - \bar{\bar{t}}_0)^2 - \frac{1}{2\sigma_{v_0}^2}(\bar{v}_0 - \bar{\bar{v}}_0)^2 \quad \log p(z_{pp}|\theta)$$

$\log p(\bar{t}_0, \bar{v}_0)$

$$-\frac{1}{2}\left[\frac{(\bar{t}_0 - \bar{\bar{t}}_0)^2}{s_{t_0}^2} + \frac{(\bar{v}_0 - \bar{\bar{v}}_0)^2}{s_{v_0}^2}\right] + \underbrace{C}_{\text{constante}} \quad C = C_1 + C_2 + C_3 + C_4$$

$\log p(\sigma_\xi, \sigma_\tau, \sigma)$

$$+ m_\xi \log\left(\frac{v_\xi}{\sigma_\xi}\right) - \log(\sigma_\xi^2) - \frac{v_\xi^2}{2\sigma_\xi^2} + m_\tau \log\left(\frac{v_\tau}{\sigma_\tau}\right) - \log(\sigma_\tau^2) - \frac{v_\tau^2}{2\sigma_\tau^2} + m\log\left(\frac{v}{\sigma}\right) - \log(\sigma^2) - \frac{v^2}{2\sigma^2}$$

On peut donc écrire $p(y, z, \theta)$ comme appartenant à la famille exponentielle courbe avec

$$S(y, z) = \begin{pmatrix} \frac{1}{\sum K_i}\sum_{i=1}^{N}\sum_{j=1}^{K_i}(y_{ij} - d_i(t_{ij}))^2 \\ \frac{1}{N}\sum_{i=1}^{N}\xi_i^2 \\ \frac{1}{N}\sum_{i=1}^{N}\tau_i^2 \\ v_0 \\ t_0 \end{pmatrix}$$

Scanned with CamScanner

et $\psi(\theta) = \begin{pmatrix} -\frac{\sum_{i=1}^{N} K_i}{2\sigma^2} \\ -\frac{N}{2\sigma_\xi^2} \\ -\frac{N}{2\sigma_\tau^2} \\ \frac{\bar{v}_0}{\sigma_{v_0}^2} \\ \frac{\bar{t}_0}{\sigma_{t_0}^2} \end{pmatrix}$

Les termes restants appartiennent à $\varphi(\theta)$:

$$\varphi(\theta) = \left(\sum_{i=1}^{N} K_i + m + 2\right) \log \sigma$$
$$+ (N + m_\xi + 2) \log \sigma_\xi$$
$$+ (N + m_\tau + 2) \log \sigma_\tau$$
$$+ \bar{v}_0^2 \left(\frac{1}{2\sigma_{v_0}^2} + \frac{1}{2s_{v_0}^2}\right) - \frac{\bar{v}_0 \cdot \bar{\bar{v}}_0}{s_{v_0}^2}$$
$$+ \bar{t}_0^2 \left(\frac{1}{2\sigma_{t_0}^2} + \frac{1}{2s_{t_0}^2}\right) - \frac{\bar{t}_0 \cdot \bar{\bar{t}}_0}{s_{t_0}^2} + \underbrace{C'}_{\text{constante indépendante de } \theta}$$

On a bien $\log p(y, z, \theta) = -\varphi(\theta) + \langle S(y,z) | \psi(\theta) \rangle$ à une constante indépendante de θ près.

③ $p(z|y,\theta) = \frac{p(z,y,\theta)}{p(y,\theta)} = p(y|z,\theta)\frac{p(z,\theta)}{p(y,\theta)} = p(y|z,\theta)\, p(z|\theta)\frac{p(\theta)}{p(y,\theta)}$
$= \frac{p(y|z,\theta)\, p(z|\theta)}{p(y|\theta)}$ donc $\log p(z|y,\theta) = \log p(y|z,\theta) + \log p(z|\theta) + \underbrace{C''}_{\text{constante indépendante de } z}$

④ On note $f_K : \theta \in \Theta \mapsto -\varphi(\theta) + \langle S^{(K)} | \psi(\theta) \rangle$ la fonction objectif à l'étape K de l'algorithme MCMC-SAEM. Cette fonction est concave et différentiable. L'estimateur MAP correspond à $\theta^{(K)}$, qui est tel que : $\nabla_\theta f_K(\theta)|_{\theta^{(K)}} = 0$.

On a : $\nabla_\theta \varphi(\theta) = \begin{pmatrix} \bar{t}_0\left(\frac{1}{\sigma_{t_0}^2} + \frac{1}{s_{t_0}^2}\right) - \frac{\bar{\bar{t}}_0}{s_{t_0}^2} \\ \bar{v}_0\left(\frac{1}{\sigma_{v_0}^2} + \frac{1}{s_{v_0}^2}\right) - \frac{\bar{\bar{v}}_0}{s_{v_0}^2} \\ (N + m_\xi + 2)\frac{1}{\sigma_\xi} \\ (N + m_\tau + 2)\frac{1}{\sigma_\tau} \\ \left(\sum_{i=1}^{N} K_i + m + 2\right)\frac{1}{\sigma} \end{pmatrix}$

Scanned with CamScanner

Et : $\nabla_\theta \langle S^{(k)}, \psi(\theta)\rangle = \begin{pmatrix} \frac{t_0^{(k)}}{\sigma_{t_0}^2} \\ \frac{v_0^{(k)}}{\sigma_{v_0}^2} \\ \frac{1}{2\sigma_\xi^3}\left[\sum_{i=1}^N (\xi_i^{(k)})^2\right] \\ \frac{1}{\sigma_\tau^3}\left[\sum_{i=1}^N \sum_{j=1}^{k_i} (y_{ij} - d_i(t_{ij}))^2\right] \end{pmatrix}$

Donc $\nabla_\theta \ell_k(\theta) = 0$ si, et seulement si $\theta = \theta^{(k)} = (\bar{t}_0^{(k)}, \bar{v}_0^{(k)}, \sigma_\xi^{(k)}, \sigma_\tau^{(k)}, \sigma^{(k)})$

où

$$\begin{pmatrix} \bar{t}_0^{(k)} \\ \bar{v}_0^{(k)} \\ (\sigma_\xi^2)^{(k)} \\ (\sigma_\tau^2)^{(k)} \\ (\sigma^2)^{(k)} \end{pmatrix} = \begin{pmatrix} \left(\frac{t_0^{(k)}}{\sigma_{t_0}^2} + \frac{\bar{\bar{t}}_0}{s_{t_0}^2}\right)\left(\frac{1}{\sigma_{t_0}^2} + \frac{1}{s_{t_0}^2}\right)^{-1} \\ \left(\frac{v_0^{(k)}}{\sigma_{v_0}^2} + \frac{\bar{\bar{v}}_0}{s_{v_0}^2}\right)\left(\frac{1}{\sigma_{v_0}^2} + \frac{1}{s_{v_0}^2}\right)^{-1} \\ \frac{v_\xi^2 + \sum_{i=1}^N (\xi_i^{(k)})^2}{N + m_\xi + 2} \\ \frac{v_\tau^2 + \sum_{i=1}^N (\tau_i^{(k)})^2}{N + m_\tau + 2} \\ \frac{v^2 + \sum_{i=1}^N \sum_{j=1}^{k_i} ((y_{ij} + d_i(t_{ij}))^{(k)})^2}{\sum_{i=1}^N k_i + m + 2} \end{pmatrix} = \begin{pmatrix} \frac{s_{t_0}^2 t_0^{(k)} + \sigma_{t_0}^2 \bar{\bar{t}}_0}{\sigma_{t_0}^2 + s_{t_0}^2} \\ \frac{s_{v_0}^2 v_0^{(k)} + \sigma_{v_0}^2 \bar{\bar{v}}_0}{\sigma_{v_0}^2 + s_{v_0}^2} \\ - - - \\ - - - \\ - - - \end{pmatrix}$$

⑤ • $p(z_i | z_{pop}, y, \theta) = \frac{p(z_i, z_{pop}, y_i, \theta)}{p(z_{pop}, y, \theta)} = p(y | z_i, z_{pop}, \theta) \cdot \frac{p(z_i, z_{pop}, \theta)}{p(z_{pop}, y, \theta)}$

$= p(y | z_i, z_{pop}, \theta)\, p(z_i, z_{pop} | \theta) \frac{p(\theta)}{p(z_{pop}, y, \theta)} = \frac{p(y | z_i, z_{pop}, y, \theta)}{p(z_{pop}, y | \theta)} p(z_i, z_{pop} | \theta)$

Donc :

$\log p(z_i | z_{pop}, y, \theta) = \log p(y | z_i, z_{pop}, y, \theta) + \log p(z_i, z_{pop} | \theta) + \underbrace{\tilde{C}}_{\text{constante indépendante de } z_i}$

Et similairement pour $\log p(z_{pop} | \{z_i\}, y, \theta)$.

⑧ Block Gibbs sampling, compared to "one-at-a-time" approach, accelerates convergence by exploiting dependencies amongst parameters within a block, resulting in more efficient updates. This method is particularly relevant for models with strong correlations between parameters, such as hierarchical models. The reduced autocorrelation between samples enhances the Markov chain's mixing properties, facilitating a more thorough exploration of the posterior distribution.

Scanned with CamScanner

Exercice 2 :

Copie 2 ① Par l'étape (i), $\varepsilon \sim \pi_f$ et $B \sim Ber(\frac{1}{2})$.

Par l'étape (ii), $Y = \mathbb{1}_{\{B=0\}} \frac{X}{\varepsilon} + \mathbb{1}_{\{B=1\}} \varepsilon X$.

Soit g continue bornée. On a, par indépendance de ε et Y, X,

$$E[g(Y)|X] = \frac{1}{2}\int_{-1}^{1} g\left(\frac{X}{z}\right) f(z)\,dz + \frac{1}{2}\int_{-1}^{1} g(zX) f(z)\,dz$$

$$= \frac{1}{2}\int_{-1}^{0} g\left(\frac{X}{z}\right) f(z)\,dz + \frac{1}{2}\int_{0}^{1} g\left(\frac{X}{z}\right) f(z)\,dz + \frac{1}{2}\int_{-1}^{1} g(zX) f(z)\,dz$$

- Si $X = 0$: $E[g(Y)|X] = g(0)$.
- Si $X > 0$: $E[g(Y)|X] = \frac{1}{2}\int_{-\infty}^{-X} g(y) f\left(\frac{X}{y}\right)\frac{X}{y^2}dy + \frac{1}{2}\int_{X}^{+\infty} g(y) f\left(\frac{X}{y}\right)\frac{X}{y^2}dy + \frac{1}{2}\int_{-X}^{X} g(y) f\left(\frac{y}{X}\right)\frac{dy}{X}$

$\frac{X}{z} = y$

$dz = -\frac{X}{y^2}dy$

et :

$zX = y$

$dzX = dy$

$$\left.\begin{array}{l} h_+ :]-1,0[\ni z \mapsto \frac{X}{z} \in]-\infty, -X[\\ h_- :]0,1[\ni z \mapsto \frac{X}{z} \in]X, +\infty[\\ h_2 :]-1,1[\ni z \mapsto zX \in]-X, X[\end{array}\right\} C^1\text{-difféomorphismes}$$

- Si $X < 0$: $E[g(Y)|X] = -\frac{1}{2}\int_{-\infty}^{X} g(y) f\left(\frac{X}{y}\right)\frac{X}{y^2}dy - \frac{1}{2}\int_{-X}^{+\infty} g(y) f\left(\frac{X}{y}\right)\frac{X}{y^2}dy - \frac{1}{2}\int_{X}^{-X} g(y) f\left(\frac{y}{X}\right)\frac{dy}{X}$

En conclusion on a :

$$q(x, dy) = \mathbb{1}_{\{|y| \geq |x|\}} f\left(\frac{x}{y}\right)\frac{|x|}{2y^2}dy + \mathbb{1}_{\{|y| < |x|\}} f\left(\frac{y}{x}\right)\frac{dy}{2|x|}$$

② Le ratio d'acceptation d'Hastings-Metropolis est :

$\alpha(x,y) = 1 \wedge \frac{\pi(y)q(y,x)}{\pi(x)q(x,y)}$ pour que la chaîne ait une distribution invariante π. On a donc :

$$\alpha(x,y) = 1 \wedge \frac{\pi(y)\left[\mathbb{1}_{\{|x| \geq |y|\}} f\left(\frac{y}{x}\right)\frac{|y|}{2|x|^2} + \mathbb{1}_{\{|x| < |y|\}} f\left(\frac{x}{y}\right)\frac{1}{2|y|}\right]}{\pi(x)\left[\mathbb{1}_{\{|y| \geq |x|\}} f\left(\frac{x}{y}\right)\frac{|x|}{2y^2} + \mathbb{1}_{\{|y| < |x|\}} f\left(\frac{y}{x}\right)\frac{1}{2|x|}\right]}$$

$$= 1 \wedge \frac{\pi(y)}{\pi(x)}\left(\mathbb{1}_{\{|x| \geq |y|\}}\frac{|y|}{|x|} + \mathbb{1}_{\{|x| < |y|\}}\frac{|y|}{|x|}\right)$$

$$= 1 \wedge \frac{\pi(y)|y|}{\pi(x)|x|}$$

Scanned with CamScanner