TP 3 : Hasting - Metropolis and Gibbs Samplers

①

Kevin

Wardakhan

Exercise 1 : 1.A - A population model for longitudinal data

1)

We consider $\theta = (\bar{t}_0, \bar{v}_0, \sigma_\xi, \sigma_\tau, \sigma)$ and $z := ((z_i)_{i \in [\![1,N]\!]}, z_{pop}) \in \mathbb{R}^{2N+2}$

where $N \in \mathbb{N}^*$.

The complete data log-likelihood is :

$$\log(q(y, z, \theta)) = \underbrace{\log(q(y|z,\theta))}_{(1)} + \underbrace{\log(q(z|\theta))}_{(2)} + \underbrace{\log(q(\theta))}_{(3)}$$

Let's compute each term separately.

$$(3) = \log(q(\theta)) = \log(q(\bar{t}_0, \bar{v}_0, \sigma_\xi, \sigma_\tau, \sigma))$$

(priors are supposed independent ⓘ)

$$= \log(q(\bar{t}_0)) + \log(q(\bar{v}_0)) + \sum_{\tilde{\sigma} \in \{\sigma_\xi, \sigma_\tau, \sigma\}} \log(q(\tilde{\sigma}))$$

$$= -\frac{1}{2s_{t_0}^2}(\bar{t}_0 - \bar{\bar{t}}_0)^2 - \frac{1}{2s_{v_0}^2}(\bar{v}_0 - \bar{\bar{v}}_0)^2 + \sum_{\tilde{\sigma} \in \{\sigma_\xi, \sigma_\tau, \sigma\}} \log(q(\tilde{\sigma})) + C_1$$

where $C_1 = -\frac{1}{2}\left[\log(2\pi s_{t_0}^2) + \log(2\pi s_{v_0}^2)\right] \in \mathbb{R}$

Moreover,

$$\sum_{\tilde{\sigma}\in\{\sigma,\sigma_\xi,\sigma_\tau\}} \log(q(\tilde{\sigma})) = -\log(\sigma^2) - \log(\sigma_\tau^2) - \log(\sigma_\xi^2)$$

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

$$- \frac{v^2}{2\sigma^2} + c_2$$

where $c_2 = -\log\left(\Gamma\left(\frac{m_\xi}{2}\right)\right) - \log\left(\Gamma\left(\frac{m_\tau}{2}\right)\right) - \log\left(\Gamma\left(\frac{m}{2}\right)\right)$

$$+ m_\xi \log\left(\frac{1}{\sqrt{2}}\right) + m_\tau \log\left(\frac{1}{\sqrt{2}}\right) + m \log\left(\frac{1}{\sqrt{2}}\right)$$

Thus, $\log(q(\theta)) = -\frac{1}{2\delta_{r_0}^2}(\bar{r}_0 - \bar{\bar{r}}_0)^2 - \frac{1}{2\delta_{v_0}^2}(\bar{v}_0 - \bar{\bar{v}}_0)^2$

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

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

$$+ m \log\left(\frac{v}{\sigma}\right) - \log(\sigma^2) - \frac{v^2}{2\sigma^2} + c_3$$

where $c_3 = c_1 + c_2$.

Exercise 1:

1) We now compute (2).

$$(2) = \log(q(z|\theta)) = \left[\sum_{i=1}^{N} \log(q(z_i|\theta))\right] + \log(q(z_{pop}|\theta))$$

First, $\log(q(z_{pop}|\theta)) = \log(q(t_0, v_0 | \bar{t}_0, \bar{v}_0))$

$$\begin{cases} t_0 \sim \mathcal{N}(\bar{t}_0, \sigma_{t_0}^2) \\ v_0 \sim \mathcal{N}(\bar{v}_0, \sigma_{v_0}^2) \\ t_0 \perp\!\!\!\perp v_0 \end{cases}$$

$$\overset{\Downarrow}{=} \log(q(t_0|\bar{t}_0)) + \log(q(v_0|\bar{v}_0))$$

$$= -\frac{1}{2\sigma_{t_0}^2}(t_0 - \bar{t}_0)^2 - \frac{1}{2\sigma_{v_0}^2}(v_0 - \bar{v}_0)^2 + C_4$$

where $C_4 = -\frac{1}{2}\left[\log(2\pi\sigma_{t_0}^2) + \log(2\pi\sigma_{v_0}^2)\right] \in \mathbb{R}$

Plus, $\forall i \in [\![1, N]\!]$,

$$\log(q(z_i|\theta)) \overset{(\forall i, \alpha_i \perp\!\!\!\perp \tau_i)}{=} \log(q(\alpha_i|\theta)) + \log(q(\tau_i|\theta))$$

$$\begin{cases} \alpha_i \sim \text{Log}\mathcal{N}(0, \sigma_\xi^2) \\ \tau_i \overset{iid}{\sim} \mathcal{N}(0, \sigma_\tau^2) \end{cases} = -\log(\alpha_i) - \frac{\log(\alpha_i)^2}{2\sigma_\xi^2} - \frac{1}{2\sigma_\tau^2}\tau_i^2 - \log(\sigma_\xi \sigma_\tau) + c_5$$

where $c_5 = -\log(2\pi)$

Thus,

$$q(z|\theta) = -\left[\sum_{i=1}^{N} \log(d_i) + \frac{\log(d_i)^2}{2\sigma_\xi^2} + \frac{\tau_i^2}{2\sigma_\tau^2}\right] - N\log(\sigma_\xi \sigma_\tau)$$

$$-\frac{1}{2\sigma_{t_0}^2}(t_0 - \bar{t}_0)^2 - \frac{1}{2\sigma_{v_0}^2}(v_0 - \bar{v}_0)^2 + C_6$$

where $C_6 = C_4 + NC_5$.

Finally, we compute (1)

$$(1) = \log(q(y|z,\theta)) = \sum_{i=1}^{N} \log(q(y_i|z,\theta))$$

$$= \sum_{i=1}^{N} \sum_{j=1}^{k_i} \log(q(y_{ij}|z,\theta))$$

$\forall i \in [\![1, N]\!]$, $\forall j \in [\![1, k_i]\!]$, $y_{ij} \sim \mathcal{N}(d_i(t_{ij}), \sigma^2)$, so

$$\log(q(y_{ij}|z,\theta)) = -\frac{1}{2\sigma^2}(y_{ij} - d_i(t_{ij}))^2 - \frac{1}{2}\log(\sigma^2) + C_7$$

where $C_7 = -\frac{1}{2}\log(2\pi)$

So, $$\log(q(y|z,\theta)) = \sum_{i=1}^{N} \sum_{j=1}^{k_i} \left[-\frac{1}{2\sigma^2}(y_{ij} - d_i(t_{ij}))^2 - \frac{1}{2}\log(\sigma^2)\right] + C_8$$

where $C_8 = \sum_{i=1}^{N} \sum_{j=1}^{k_i} C_7 = C_7 \sum_{i=1}^{N} k_i$

Exercise 1:

1) Finally, we obtain that:

$$\log(q(y,z;\theta)) = \sum_{i=1}^{N} \sum_{j=1}^{k_i} \left[-\frac{1}{2\sigma^2}(y_{ij} - d_i(t_{ij}))^2 - \frac{1}{2}\log(\sigma^2)\right]$$

$$+ \sum_{i=1}^{N}\left[-\log(d_i) - \frac{\log(d_i)^2}{2\sigma_\xi^2} - \frac{\tau_i^2}{2\sigma_\tau^2}\right] - N\log(\sigma_\xi \sigma_\tau)$$

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

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

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

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

$$+ m \log\left(\frac{v}{\sigma}\right) - \log(\sigma^2) - \frac{v^2}{2\sigma^2} + C$$

where $C = C_8 + C_6 + C_3$

We can write the log-likelihood as

$$-\phi(\theta) + S(y,z)^T \psi(\theta) + \log(h(y,z)) \text{ by setting:}$$

- $$S(y,z) = \begin{pmatrix} \frac{1}{N_b}\sum_{i=1}^{N}\sum_{j=1}^{b_i}(y_{ij} - d_i(t_{ij}))^2 \\ \frac{1}{N}\sum_{i=1}^{N}\log(\alpha_i)^2 \\ \frac{1}{N}\sum_{i=1}^{N}\tau_i^2 \\ t_0 \\ v_0 \end{pmatrix}$$ (here $N_b = \sum_{i=1}^{N} b_i$)

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

- $$\phi(\theta) = \frac{\bar{t}_0^2}{2\sigma_{t_0}^2} + \frac{\bar{v}_0^2}{2\sigma_{v_0}^2} + \frac{\bar{\bar{t}}_0^2}{2s_{t_0}^2} + \frac{\bar{\bar{v}}_0^2}{2s_{v_0}^2} - \frac{\bar{t}_0\bar{\bar{t}}_0}{s_{t_0}^2} - \frac{\bar{v}_0\bar{\bar{v}}_0}{s_{v_0}^2}$$
$$- \frac{v_\xi^2}{2\sigma_\xi^2} - \frac{v_\tau^2}{2\sigma_\tau^2} - \frac{v^2}{2\sigma^2} + (N_b + m + 2)\log(\sigma) + (N_b + m_\xi + 2)\log(\sigma_\xi)$$
$$+ (N_b + m_\tau + 2)\log(\sigma_\tau)$$

Exercise 1:

4) We want to compute for $k \in \mathbb{N}$ fixed:

$$\theta^{(k)} = \underset{\theta \in \Theta}{\operatorname{argmax}} \ \vartheta_k(\theta) := -\phi(\theta) + S_k(y,z)^T \psi(\theta)$$

where $S_k(y,z) = (S_k^{(i)}(y,z))_{i \in [\![1,5]\!]}$.
Differentiating $\theta \longmapsto \vartheta_k(\theta)$ gives:

$$\nabla \vartheta_k(\theta) = -\nabla \phi(\theta) + J_\psi(\theta)^T S_k(y,z)$$

Moreover:

$$\bullet \ -\nabla \phi(\theta) = \begin{pmatrix} \dfrac{\bar{t}_0}{s_{t_0}^2} - \bar{t}_0 \left(\dfrac{1}{s_{t_0}^2} + \dfrac{1}{\sigma_{t_0}^2} \right) \\ \dfrac{\bar{v}_0}{s_{v_0}^2} - \bar{v}_0 \left(\dfrac{1}{s_{v_0}^2} + \dfrac{1}{\sigma_{v_0}^2} \right) \\ \dfrac{v_\xi^2}{\sigma_\xi^3} - \dfrac{(N_k + m_\xi + 2)}{\sigma_\xi} \\ \dfrac{v_\tau^2}{\sigma_\tau^3} - \dfrac{(N_k + m_\tau + 2)}{\sigma_\tau} \\ \dfrac{v^2}{\sigma^3} - \dfrac{(N_k + m + 2)}{\sigma} \end{pmatrix}$$

and $J_\psi(\theta)^T = \begin{pmatrix} 0 & 0 & 0 & \sigma_{r_0}^{-2} & 0 \\ 0 & 0 & 0 & 0 & \sigma_{v_0}^{-2} \\ 0 & \frac{N}{\sigma_\xi^3} & 0 & 0 & 0 \\ 0 & 0 & \frac{N}{\sigma_\tau^3} & 0 & 0 \\ \frac{N_b}{\sigma^3} & 0 & 0 & 0 & 0 \end{pmatrix}$

Thus, $\nabla v_b(\theta) = 0$

$$\Leftrightarrow \theta_b := \begin{pmatrix} \bar{r}_0 \\ \bar{v}_0 \\ \sigma_\xi \\ \sigma_\tau \\ \sigma \end{pmatrix} = \begin{pmatrix} \left(\frac{\bar{\bar{r}}_0}{s_{r_0}^2} + \frac{S_b^{(4)}}{\sigma_{r_0}^2}\right) \frac{s_{r_0}^2 \sigma_{r_0}^2}{s_{r_0}^2 + \sigma_{r_0}^2} \\ \left(\frac{\bar{\bar{v}}_0}{s_{v_0}^2} + \frac{S_b^{(5)}}{\sigma_{v_0}^2}\right) \frac{s_{v_0}^2 \sigma_{v_0}^2}{s_{v_0}^2 + \sigma_{v_0}^2} \\ \sqrt{\frac{S_b^{(2)} N + v_\xi^2}{N_b + m_\xi + 2}} \\ \sqrt{\frac{S_b^{(3)} N + v_\tau^2}{N_b + m_\tau + 2}} \\ \sqrt{\frac{S_b^{(1)} N_b + v^2}{N_b + m + 2}} \end{pmatrix}$$

Exercise 3: Multiplicative Hasting-Metropolis

Let f be a density function on $]-1,1[$, a random variable $\mathcal{E}$ having f for distribution and independent of $B \sim \mathrm{Ber}(\frac{1}{2})$. Let X be a random variable such that we have mutual independence of $(X, B, \mathcal{E})$.

1) Let's compute the density of the jumping distribution $Y \sim q(X,Y)$.

Let $\varphi : \mathbb{R} \to \mathbb{R}$ a Borel bounded function.

$$\mathbb{E}[\varphi(Y) | X] = \mathbb{E}[\varphi(Y) (\mathbb{1}_{\{B=1\}} + \mathbb{1}_{\{B=0\}}) | X]$$

$$= \mathbb{E}[\varphi(Y) \mathbb{1}_{\{B=1\}} | X] + \mathbb{E}[\varphi(X) \mathbb{1}_{\{B=0\}} | X]$$

$$= g(X) + h(X)$$

where
$$\begin{cases} g(x) = \mathbb{E}[\varphi(Y) \mathbb{1}_{\{B=1\}} | X = x] & i) \\ h(x) = \mathbb{E}[\varphi(Y) \mathbb{1}_{\{B=0\}} | X = x] & ii) \end{cases}$$

i) Moreover, $\forall x \neq 0$ (if $x = 0$, then $Y = 0$ a.s)

$$g(x) = \mathbb{E}\left[\varphi(Y)\, \mathbb{1}_{\{B=1\}} \mid X = x\right]$$

$\downarrow$ $B = 1$ a.s

$$= \mathbb{E}\left[\varphi(\varepsilon X)\, \mathbb{1}_{\{B=1\}} \mid X = x\right]$$

$\downarrow$ $(\varepsilon, B) \perp\!\!\!\perp X$

$$= \mathbb{E}\left[\varphi(\varepsilon x)\, \mathbb{1}_{\{B=1\}}\right]$$

$\downarrow$ $\varepsilon \perp\!\!\!\perp B$

$$= \frac{1}{2}\, \mathbb{E}\left[\varphi(\varepsilon x)\right]$$

$$= \frac{1}{2} \int_{-1}^{1} \varphi(zx)\, f(z)\, dz = \frac{1}{2} \int_{-|x|}^{|x|} \varphi(y)\, f\left(\frac{y}{x}\right) \frac{1}{|x|}\, dy$$

by doing the change of variable $\psi : \begin{cases}]-1, 1[\longrightarrow]-|x|, |x|[\\ z \longmapsto zx =: y \end{cases}$

ii) By doing the same for h, we get: $\forall x \neq 0$,

$$h(x) = \frac{1}{2}\, \mathbb{E}\left[\varphi\left(\frac{x}{\varepsilon}\right)\right] = \frac{1}{2} \int_{\mathbb{R}} \varphi(y)\, f\left(\frac{x}{y}\right) \frac{|x|}{y^2}\, \mathbb{1}_{|x| < |y| < 1}\, dy$$

by letting the change of variable:

$$\psi' : \begin{cases}]-1, 1[\longrightarrow \{y \in \mathbb{R} \mid |x| < |y| < 1\} \\ z \longmapsto \frac{x}{z} =: y \end{cases}$$

Exercise 3:

Finally, we get:

$$\mathbb{E}[\ell(Y) \mid X] = g(x) + h(x)$$

$$= \int_{\mathbb{R}} \ell(y) \left[\frac{1}{2} \left(f\left(\frac{y}{x}\right) \frac{1}{|x|} \mathbb{1}_{|y|<|x|}(y) + f\left(\frac{x}{y}\right) \frac{|x|}{y^2} \mathbb{1}_{|x|<|y|<1}(y) \right) \right] dy$$

So, $\forall x \neq 0$, $\forall y \in \mathbb{R}$,

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

Remark:

If $\begin{cases} y = 0, \\ x \neq 0, \end{cases}$ we set the convention $f\left(\frac{x}{y}\right) \frac{|x|}{y^2} \mathbb{1}_{|x|<|y|<1} = 0$

so that $q(x,0) = \frac{1}{2} f(0) \frac{1}{|x|}$

2) We know that the classical Hasting-Metropolis algorithm is:

$$\alpha(X,Y) = \begin{cases} \min\left(1, \dfrac{\pi(Y)\, q(X,Y)}{\pi(X)\, q(X,Y)}\right) & \text{if } \pi(X)\, q(X,Y) > 0 \\ 1 & \text{otherwise} \end{cases}$$

where π is a fixed invariant probability measure for the chain.

Let x, y s.t $q(x,y) > 0$, then:

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

$$= \left|\frac{y}{x}\right| \mathbb{1}_{|x|<1} \mathbb{1}_{|y|<1} = \left|\frac{y}{x}\right|$$

because the chain lives in $]-1, 1[$.

So, for a fixed π,

$$\alpha(x,y) = \begin{cases} \min\left(1, \frac{\pi(y)}{\pi(x)} \left|\frac{y}{x}\right|\right) & \text{if } x \neq 0 \text{ and } \pi(x) > 0 \\ 1 & \text{otherwise} \end{cases}$$