MCMC Algorithm Derivation

Yifei Wang

Given XZ, k, the log likelihood and the composite likelihood are

$$
\begin{aligned}
\ell\left(\tau_k^{XZ^2}, \mu_k^{XZ}\right) &= -\frac{1}{2} \sum_{i \in N_{XZ}^k} \left[\log\left(s_{k,i}^{XZ^2} + \tau_k^{XZ^2}\right) + \frac{\left(y_{k,i}^{XZ} - \mu_k^{XZ}\right)^2}{s_{k,i}^{XZ^2} + \tau_k^{XZ^2}}\right] \\
&= \sum_{i \in N_{XZ}^k} \left[\log \frac{1}{\sqrt{s_{k,i}^{XZ^2} + \tau_k^{XZ^2}}} - \frac{\left(y_{k,i}^{XZ} - \mu_k^{XZ}\right)^2}{2\left(s_{k,i}^{XZ^2} + \tau_k^{XZ^2}\right)}\right] \\
L\left(\tau_k^{XZ^2}, \mu_k^{XZ}\right) &= \left[\prod_{i \in N_{XZ}^k} \left(\frac{1}{\sqrt{s_{k,i}^{XZ^2} + \tau_k^{XZ^2}}}\right)\right] \cdot \exp\left(\sum_{i \in N_{x_2}^k} -\frac{\left(y_{k,i}^{XZ} - \mu_k^{XZ}\right)^2}{2\left(s_{k,i}^{xz^2} + \tau_k^{XZ^2}\right)}\right) \\
&= \left[\prod_{i \in N_{XZ}^k} \left(\frac{1}{\sqrt{s_{k,i}^{XZ^2} + \tau_k^{XZ^2}}}\right)\right] \cdot \exp\left(\sum_{i \in N_{XZ}^k} -\frac{\left(y_{k,i}^{XZ} - \bar{y}_k^{XZ}\right)^2}{2\left(s_{k,i}^{XZ^2} + \tau_k^{XZ^2}\right)} - \sum_{i \in N_{kz}^k} \frac{\left(\bar{y}_k^{XZ} - \mu_k^{XZ}\right)^2}{2\left(s_{k,i}^{XZ^2} + \tau_k^{XZ^2}\right)}\right) \\
&= \left[\prod_{i \in N_{XZ}^k} \left(\frac{1}{\sqrt{s_{k,i}^{XZ^2} + \tau_k^{XZ^2}}}\right)\right] \cdot \exp\left(-\sum_{i \in N_{XZ}} \frac{\left(y_{k,i}^{XZ} - \bar{y}_k^{XZ}\right)^2}{2\left(s_{k,i}^{XZ^2} + \tau_k^{XZ}\right)}\right) \exp\left(-\sum_{i \in N_{XZ}^k} \frac{\left(\bar{y}_k^{XZ} - \mu_k^{XZ}\right)^2}{2\left(s_{k,i}^{XZ^2} + \tau_k^{XZ^2}\right)}\right) \\
&= \left[\prod_{i \in N_{XZ}^k} \frac{1}{\sqrt{s_{k,i}^{XZ^2} + \tau_k^{XZ^2}}}\right] \cdot \exp\left[\sum_{i \in N_{kz}^k} -\frac{\left(y_{k,i}^{XZ} - \bar{y}_k^{XZ}\right)^2}{2\left(s_{k,i}^{XZ} + \tau_k^{XZ^2}\right)}\right] \cdot \exp\left[\left(\bar{y}_k^{XZ} - \mu_k^{XZ}\right)^2 \sum_{i \in N_{XZ}} \frac{1}{2\left(s_{k,i}^{XZ^2} + \tau_k^{XZ^2}\right)}\right]
\end{aligned}
$$

where

$$
\begin{aligned}
n_k^{xz} &= \sum_{i \in N_{XZ}^k} \nVdash_{i \in N_{XZ}^k} \\
\bar{y}_k^{xz} &= \frac{1}{n_{xz}^k} \sum_{i \in N_{XZ}^k} y_{k,i}^{xz}
\end{aligned}
$$

Let

$$
\sum_{i \in N_{XZ}^k} \frac{1}{\left(s_{k,i}^{XZ} + \tau_k^{XZ^2}\right)} = \frac{n_{xz}^k}{\sigma_k^{XZ^2}}
$$

Then the composite likelihood and marginal likelihood can also

$$
\begin{aligned}
L\left(\tau_k^{XZ^2}, \mu^{XZ}\right) &= \left[\prod_{i \in N_{XZ}^k} \frac{1}{\sqrt{s_{k,i}^{XZ^2} + \tau_k^{XZ^2}}}\right] \exp\left[\sum_{i \in N_{Nk}^k} -\frac{\left(y_{k,i}^{XZ} - \bar{y}_k^{XZ}\right)^2}{2\left(s_{k,i}^{xz^2} + \tau_k^{XZ^2}\right)}\right] \exp\left[-\frac{n_{xz}^k}{2\sigma_k^{xz^2}}\left(\bar{y}_{xz}^k - \mu_k^{xz}\right)^2\right] \\
L\left(\tau_k^{XZ^2}\right) &= \int_{-\infty}^{\infty} L\left(\tau_k^{XZ^z}, \mu_k^{XZ}\right) d\mu_k^{XZ} \\
&= \left[\prod_{i \in N_{XZ}^k} \frac{1}{\sqrt{s_{k,i}^{XZ^2} + \tau_k^{XZ^2}}}\right] \exp\left[\sum_{i \in N_{Nk}^k} -\frac{\left(y_{k,i}^{XZ} - \bar{y}_k^{XZ}\right)^2}{2\left(s_{k,i}^{xz^2} + \tau_k^{XZ^2}\right)}\right] \sqrt{\frac{2\pi\sigma_k^{xz^2}}{n_{xz}^k}} \int_{-\infty}^{\infty} \sqrt{\frac{n_{xz}^k}{2\pi\sigma_k^{xz^2}}} \exp\left[-\frac{n_{xz}^k}{2\sigma_k^{xz^2}}\left(\mu_k^{xz} - \bar{y}_x^k\right.\right. \\
&= \left[\prod_{i \in N_{XZ}^k} \frac{1}{\sqrt{s_{k,i}^{XZ^2} + \tau_k^{XZ^2}}}\right] \exp\left[\sum_{i \in N_{Nk}^k} -\frac{\left(y_{k,i}^{XZ} - \bar{y}_k^{XZ}\right)^2}{2\left(s_{k,i}^{xz^2} + \tau_k^{XZ^2}\right)}\right] \sqrt{\frac{2\pi\sigma_k^{xz^2}}{n_{xz}^k}}
\end{aligned}
$$

Conditioning on $\tau_k^{XZ^2}$ and y, the posterior distribution of μ_k^{XZ} follows

$$
\begin{aligned}
p(\mu_k^{XZ} | \tau_k^{XZ^2}, y) &= \frac{P\left(\mu_k^{XZ}, \tau_k^{XZ^2}, y\right)}{P\left(\tau_k^{XZ^2}, y\right)} \\
&= \frac{p\left(y \mid \tau_k^{XZ^2}, \mu_k^{XZ}\right) \cdot P\left(\tau_k^{XZ^2}, \mu_k^{XZ}\right)}{p\left(\tau_k^{XZ^2}, y\right)} \\
&= \frac{P\left(y \mid \tau_k^{XZ^2}, \mu_k^{XZ}\right) P\left(\tau_k^{XZ^2}, \mu_k^{XZ}\right)}{P\left(y \mid \tau_k^{XZ^2}\right) P\left(\tau_k^{XZ^2}\right)} \\
&\propto \frac{L\left(\tau_k^{XZ^2}, \mu_k^{XZ}\right)}{L\left(\tau_k^{XZ^2}\right)} \\
&= \sqrt{\frac{n_{XZ}^k}{2\pi\sigma_k^{xz^2}}} \exp\left[-\frac{n_{xz}^k}{2\sigma_k^{xz^2}}\left(\bar{y}_{XZ}^k - \mu_k^{XZ}\right)^2\right]
\end{aligned}
$$

Conditioning on μ_k^{XZ} and y, the postier distribution of τ_k^{XZ} follows

$$
\begin{aligned}
p(\tau_k^{XZ^2} | \mu_k^{XZ}, y) &\propto P(y | \tau_k^{XZ^2}, \mu_k^{XZ}) P(\tau_k^{XZ^2}, \mu_k^{XZ}) \\
&\propto L\left(\tau_k^{XZ^2}, \mu_k^{XZ}\right)\left(\tau_k^{XZ}\right)^{-2} \\
&\propto \left[\prod_{i \in N_{XZ}^k}\left(\frac{1}{\sqrt{s_{k,i}^{XZ^2} + \tau_k^{XZ^2}}}\right)\right] \cdot \exp\left(\sum_{i \in N_{x_2}^k} -\frac{\left(y_{k,i}^{XZ} - \mu_k^{XZ}\right)^2}{2\left(s_{k,i}^{xz^2} + \tau_k^{XZ^2}\right)}\right)\left(\tau_k^{XZ}\right)^{-2}
\end{aligned}
$$

where

$$
\begin{gathered}
P\left(\tau_k^{XZ^2}, \mu_k^{XZ}\right) \propto \left(\tau_k^{XZ^2}\right)^{-1} \\
\frac{P\left(\tau_k^{XZ^2}, \mu_k^{XZ}\right)}{P\left(\tau_k^{XZ^2}\right)} \propto 1
\end{gathered}
$$

We use Gibbs Sampler to sample μ_k^{XZ} and $\tau_k^{XZ^2}$ from the posterior distribution

Input: data y

Output: A realization of length N from a Markov Chain for $\left(\tau_k^{XZ}, \mu_k^{XZ}\right)$

for $k \in \{1, 2\}$
for $XZ \in \{AB, AC, BC\}$
for $t = 1 \cdots N$:
sample $\mu_k^{XZ(p)} \sim N\left(\bar{y}_{XZ}^k, \frac{\left(\sigma_k^{XZ}\right)^2}{n_{xz}^k}\right)$ where $\frac{\left(\sigma_k^{XZ}\right)^2}{n_{xz}^k} = \frac{1}{\sum_{i \in N_{XZ}^k} \frac{1}{\left(s_{k,i}^{XZ} + \tau_k^{XZ^2}\right)}}$ and $\bar{y}_{xz}^k = \frac{1}{n_{xz}^k} \sum_{i \in N_{XZ}^k} y_{k,i}^{xz}$
sample $\tau_k^{XZ(p)} \sim p\left(\tau_k^{XZ^2} | y, \mu_k^{XZ(p)}\right)$
$\tau_k^{XZ^2(t)} = \tau_k^{XZ^2(p)}$
$\mu_k^{XZ(t)} = \mu_k^{XZ(p)}$

We use Metropolis-Hasting algorithm to sample τ_k^{XZ} from $p(\tau_k^{XZ^2} | y, \mu_k^{XZ})$.

Input: data y

Output: A realization of length N from a Markov Chain for τ_k^{XZ} from $p(\tau_k^{XZ^2} | y)$

Propose $q\left(\tau_{k,t}^{XZ^2} | \tau_{k,t-1}^{XZ^2}\right) = N(\tau_{k,t-1}^{XZ^2}, 0.1)$

$\tau_{k,0}^{XZ^2} = 0.3$
for t from 1 to N:
sample $\tau_{k,*}^{XZ^2} \sim N(\tau_{k,t-1}^{XZ^2}, 0.1)$
$a = \frac{P\left(\tau_{k,*}^{XZ^2} | y, \mu_k^{XZ}\right) / q\left(\tau_{k,*}^{XZ^2} | \tau_{k,t-1}^{XZ^2}\right)}{P\left(\tau_{k,t-1}^{XZ^2} | y, \mu_k^{XZ}\right) / q\left(\tau_{k,t-1}^{XZ^2} | \tau_{k,*}^{XZ^2},\right)}$
if $runif(1) < min(a, 1)$:
$\tau_{k,t}^{XZ^2} = \tau_{k,*}^{XZ^2}$
else:
$\tau_{k,t}^{XZ^2} = \tau_{k,t-1}^{XZ^2}$

To include the magnitude adjustment, we first derive the relationship between the magnitude adjusted likelihood and the composite likelihood without adjustment

$$\begin{aligned}
\ell_{magn}(\theta; y) &= k\ell_c(\theta, y) \\
\exp\left[\ell_{magn}(\theta; y)\right] &= \exp\left(k\ell_c(\theta; y)\right) \\
L_{magn}(\theta; y) &= \left\{\exp\left[\ell_c(\theta; y)\right]\right\}^k \\
L_{magn}(\theta; y) &= \left[L_c(\theta; y)\right]^k
\end{aligned}$$

where $k = p / \sum_{i=1}^p \lambda_i$ and $\lambda_1, \ldots, \lambda_p$ are eigenvalues from $H(\theta_0)^{-1} J(\theta_0)$.

Then we update the likelihood part in the $p(\tau_k^{XZ^2} | y)$ from Metropolis-Hasting algorithm as below

Input: data y

Output: A realization of length N from a Markov Chain for τ_k^{XZ} from $p(\tau_k^{XZ^2} | y)$

Propose $q\left(\tau_{k,t}^{XZ^2} | \tau_{k,t-1}^{XZ^2}\right) = N(\tau_{k,t-1}^{XZ^2}, 0.1)$

Set initial value $\tau_{k,0}^{XZ^2} = 0.3$
For t from 1 to N:
Sample $\tau_{k,*}^{XZ^2} \sim N(\tau_{k,t-1}^{XZ^2}, 0.1)$
Calculate $H(\tau_{k,t-1}, \mu_k)$ and $J(\tau_{k,t-1}, \mu_k)$
Get $k = p/\text{Trace}\left[H(\tau_{k,t-1}, \mu_k)^{-1} J(\tau_{k,t-1}, \mu_k)\right]$ where p is the number of columns of $\left[H(\tau_{k,t-1}, \mu_k)^{-1} J(\tau_{k,t-1}, \mu_k)\right]$

$a = \frac{P\left(\tau_{k,*}^{XZ^2}|y,\mu_k^{XZ}\right)^k / q\left(\tau_{k,*}^{XZ^2}|\tau_{k,t-1}^{XZ^2}\right)}{P\left(\tau_{k,t-1}^{XZ^2}|y,\mu_k^{XZ}\right)^k / q\left(\tau_{k,t-1}^{XZ^2}|\tau_{k,*}^{XZ^2},\right)}$
If $runif(1) < min(a, 1)$:
    $\tau_{k,t}^{XZ^2} = \tau_{k,*}^{XZ^2}$
Else:
    $\tau_{k,t}^{XZ^2} = \tau_{k,t-1}^{XZ^2}$

To adjust the likelihood on sampling μ, we inherent the derivation and adjust the likelhood part

$$p(\mu_k^{XZ}|\tau_k^{XZ^2}, y) \propto \frac{L\left(\tau_k^{XZ^2}, \mu_k^{XZ}\right)^k}{L\left(\tau_k^{XZ^2}\right)^k}$$
$$= \left\{\sqrt{\frac{n_{XZ}^k}{2\pi\sigma_k^{xz^2}}} \exp\left[-\frac{n_{xz}^k}{2\sigma_k^{xz^2}}\left(\bar{y}_{XZ}^k - \mu_k^{XZ}\right)^2\right]\right\}^k$$

Since above form is not a known distribution anymore, we intend to use the Metropolis-Hasting algorithm again to simulate the μ_k^{XZ} as below

Propose $q\left(\mu_{k,t}^{XZ}|\mu_{k,t-1}^{XZ}\right) = N(\mu_{k,t-1}^{XZ}, 0.1)$

Set initial value $\mu_{k,0}^{XZ^2} = 0.3$
For t from 1 to N:
    Sample $\mu_{k,*}^{XZ^2} \sim N(\mu_{k,t-1}^{XZ}, 0.1)$
    Calculate $H\left(\tau_{k,t-1}, \mu_{k,t-1}\right)$ and $J\left(\tau_{k,t-1}, \mu_{k,t-1}\right)$
    Get $k = p/\text{Trace}\left[H\left(\tau_{k,t-1}, \mu_{k,t-1}\right)^{-1} J\left(\tau_{k,t-1}, \mu_{k,t-1}\right)\right]$ where p is the number of columns of $\left[H\left(\tau_{k,t-1}, \mu_{k,t-1}\right)^{-1} J\left(\tau_{k,t-1}, \mu_{k,t-1}\right)\right]$
    $a = \frac{P\left(\mu_{k,*}^{XZ}|y,\tau_k^{XZ^2}\right)^k / q\left(\mu_{k,*}^{XZ}|\mu_{k,t-1}^{XZ^2}\right)}{P\left(\mu_{k,t-1}^{XZ^2}|y,\tau_k^{XZ^2}\right)^k / q\left(\mu_{k,t-1}^{XZ}|\mu_{k,*}^{XZ}\right)}$
    If $runif(1) < min(a, 1)$:
        $\mu_{k,t}^{XZ} = \mu_{k,*}^{XZ}$
    Else:
        $\mu_{k,t}^{XZ} = \mu_{k,t-1}^{XZ}$