derive posterior distribution

Renjie Wei

2022-05-02

The suggested Bayesian model

$$Y_i(t+6) = \beta_{0,i} + \beta_{1,i}Y_i(t) + \beta_{2,i}\Delta_{i,1}(t) + \beta_{3,i}\Delta_{i,2}(t) + \beta_{4,i}\Delta_{i,3}(t) + \epsilon_i(t)$$

where $Y_i(t)$ the wind speed at time t (i.e. 6 hours earlier), $\Delta_{i,1}(t)$, $\Delta_{i,2}(t)$ and $\Delta_{i,3}(t)$ are the changes of latitude, longitude and wind speed between t and $t+6$, and $\epsilon_{i,t}$ follows a normal distributions with mean zero and variance σ^2, independent across t.

In the model, $\boldsymbol{\beta}_i = (\beta_{0,i}, \beta_{1,i}, ..., \beta_{5,i})$ are the random coefficients associated the ith hurricane, we assume that

$$\boldsymbol{\beta}_i \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

follows a multivariate normal distributions with mean $\boldsymbol{\mu}$ and covariance matrix Σ.

We assume the following non-informative or weak prior distributions for σ^2, $\boldsymbol{\beta}$ and Σ.

$$P(\sigma^2) \propto \frac{1}{\sigma^2}; \quad P(\boldsymbol{\mu}) \propto 1; \quad P(\Sigma^{-1}) \propto |\Sigma|^{-(d+1)} \exp(-\frac{1}{2}\Sigma^{-1})$$

d is dimension of β.

1

Let $\mathbf{B} = (\boldsymbol{\beta}_1^\top, ..., \boldsymbol{\beta}_n^\top)^\top$, derive the posterior distribution of the parameters $\Theta = (\mathbf{B}^\top, \boldsymbol{\mu}^\top, \sigma^2, \Sigma)$.

Note from given Bayesian model, let

$$\epsilon_i(t) = Y_i(t+6) - \Big(\beta_{0,i} + \beta_{1,i}Y_i(t) + \beta_{2,i}\Delta_{i,1}(t) + \beta_{3,i}\Delta_{i,2}(t) + \beta_{4,i}\Delta_{i,3}(t)\Big) \overset{i.i.d}{\sim} N(0, \sigma^2)$$

or

$$Y_i(t+6) \sim N(\boldsymbol{X}_i(t)\boldsymbol{\beta}_i^\top, \sigma^2)$$

where $\boldsymbol{X}_i(t) = (1, Y_i(t), \Delta_{i,1}(t), \Delta_{i,2}(t), \Delta_{i,3}(t))$, and $\boldsymbol{\beta}_i = (\beta_{0,i}, \beta_{1,i}, \beta_{2,i}, \beta_{3,i}, \beta_{4,i})$. Therefore, the wind speed of i^{th} hurricane at time t follows the normal distribution with the pdf below

$$f_{Y_i(t+6)}(y_i(t+6) \mid \boldsymbol{X}_i(t), \boldsymbol{\beta}_i, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \exp\Big\{ -\frac{1}{2\sigma^2}\Big(y_i(t+6) - \boldsymbol{X}_i(t)\boldsymbol{\beta}_i^\top\Big)^2\Big\}$$

Therefore, the conditional distribution of Y_i, the wind speed of i^{th} hurricane follows the multivariate normal distribution below, (since $Y_i(t)$'s are independent across t)

$$(\boldsymbol{Y}_i \mid \boldsymbol{X}_i, \boldsymbol{\beta}_i, \sigma^2) \sim \mathcal{N}(\boldsymbol{X}_i\boldsymbol{\beta}_i^\top, \sigma^2 I)$$

where Y_i is an m_i-dimensional vector and $\boldsymbol{X}_i$ is a $m_i \times d$ matrix.

Hence, the joint likelihood function of all $i's$ hurricanes can be expresses as

$$L_Y(\mathbf{B}^\top, \sigma^2) = \prod_{i=1}^{n} \Big\{ \det(2\pi\sigma^2 I)^{-\frac{1}{2}} \exp\Big(-\frac{1}{2}(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i^\top)^\top(\sigma^2 I)^{-1}(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i^\top)\Big)\Big\}$$

where I is an identical matrix with dimension consistent with Y_i.

From Bayesian theorem, the posterior distribution for Θ is

$$\pi(\boldsymbol{\Theta}|\boldsymbol{Y}) = \pi(\mathbf{B}^\top, \boldsymbol{\mu}^\top, \sigma^2, \Sigma \mid \boldsymbol{Y}) \propto L_Y(\mathbf{B}^\top, \sigma^2) \times \pi(\mathbf{B} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \times \pi(\boldsymbol{\mu}) \times \pi(\sigma^2) \times \pi(\boldsymbol{\Sigma}),$$

where $\pi(\mathbf{B} \mid \boldsymbol{\beta}, \Sigma)$ is the joint multivariate normal density of β, since

$$\boldsymbol{\beta}_i \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

Therefore

$$\pi(\mathbf{B} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \prod_{i=1}^{n} \Big\{ \det(2\pi\boldsymbol{\Sigma})^{-\frac{1}{2}} \exp(-\frac{1}{2}(\boldsymbol{\beta}_i - \boldsymbol{\mu})\boldsymbol{\Sigma}^{-1}(\boldsymbol{\beta}_i - \boldsymbol{\mu})^\top)\Big\}.$$

So we have the following posterior distribution:

$$\begin{aligned}\pi(\mathbf{B}^\top, \boldsymbol{\mu}^\top, \sigma^2, \Sigma \mid Y) \propto & \prod_{i=1}^{n} \Big\{ (2\pi\sigma^2)^{-m_i/2} \exp\big\{ -\frac{1}{2}(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i^\top)^\top(\sigma^2 I)^{-1}(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i^\top)\big\}\Big\} \\ & \times \prod_{i=1}^{n} \Big\{ \det(2\pi\boldsymbol{\Sigma})^{-\frac{1}{2}} \exp\big\{ -\frac{1}{2}(\boldsymbol{\beta}_i - \boldsymbol{\mu})\boldsymbol{\Sigma}^{-1}(\boldsymbol{\beta}_i - \boldsymbol{\mu})^\top\big\}\Big\} \times \frac{1}{\sigma^2} \times \det(\boldsymbol{\Sigma})^{-(d+1)} \exp\big\{ -\frac{1}{2}\boldsymbol{\Sigma}^{-1}\big\}\end{aligned}$$

To apply MCMC, we need to find conditional posterior distribution of each parameter.

1. For $\pi(\mathbf{B}|.)$

$$
\begin{aligned}
\pi(\mathbf{B}|.) &\propto L_Y(\mathbf{B}^\top,\sigma^2)\times\pi(\mathbf{B}\mid\boldsymbol{\mu},\boldsymbol{\Sigma})\\
&\propto \prod_{i=1}^{n}\Big\{(2\pi\sigma^2)^{-m_i/2}\exp\Big(-\frac{1}{2}(\boldsymbol{Y}_i-\boldsymbol{X}_i\boldsymbol{\beta}_i^\top)^\top(\sigma^2 I)^{-1}(\boldsymbol{Y}_i-\boldsymbol{X}_i\boldsymbol{\beta}_i^\top)\Big)\Big\}\\
&\times\prod_{i=1}^{n}\Big\{\det(2\pi\boldsymbol{\Sigma})^{-\frac{1}{2}}\exp(-\frac{1}{2}(\boldsymbol{\beta}_i-\boldsymbol{\mu})\boldsymbol{\Sigma}^{-1}(\boldsymbol{\beta}_i-\boldsymbol{\mu})^\top)\Big\}\\
&\propto\prod_{i=1}^{n}\exp\{-\frac{1}{2}\Big((\boldsymbol{Y}_i-\boldsymbol{X}_i\boldsymbol{\beta}_i^\top)^\top(\sigma^2 I)^{-1}(\boldsymbol{Y}_i-\boldsymbol{X}_i\boldsymbol{\beta}_i^\top)+(\boldsymbol{\beta}_i-\boldsymbol{\mu})\boldsymbol{\Sigma}^{-1}(\boldsymbol{\beta}_i-\boldsymbol{\mu})^\top)\Big)\}
\end{aligned}
$$

Considering the exponential term in each component in the product,

$$
\begin{aligned}
&(\boldsymbol{Y}_i-\boldsymbol{X}_i\boldsymbol{\beta}_i^\top)^\top(\sigma^2 I)^{-1}(\boldsymbol{Y}_i-\boldsymbol{X}_i\boldsymbol{\beta}_i^\top)+(\boldsymbol{\beta}_i-\boldsymbol{\mu})\boldsymbol{\Sigma}^{-1}(\boldsymbol{\beta}_i-\boldsymbol{\mu})^\top)\\
&=\boldsymbol{Y}_i^\top(\sigma^2 I)^{-1}\boldsymbol{Y}_i+\boldsymbol{\beta}_i\boldsymbol{X}_i^\top(\sigma^2 I)^{-1}\boldsymbol{X}_i\boldsymbol{\beta}_i^\top-2\boldsymbol{Y}_i(\sigma^2 I)^{-1}\boldsymbol{X}_i\boldsymbol{\beta}_i^\top\\
&+\boldsymbol{\beta}_i\boldsymbol{\Sigma}^{-1}\boldsymbol{\beta}_i^\top+\boldsymbol{\mu}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}^\top-2\boldsymbol{\mu}\boldsymbol{\Sigma}^{-1}\boldsymbol{\beta}_i^\top\\
&=\boldsymbol{\beta}_i\boldsymbol{V}\boldsymbol{\beta}_i^\top-2\boldsymbol{M}\boldsymbol{\beta}_i^\top+\boldsymbol{R}
\end{aligned}
$$

where,

$$
\begin{aligned}
\boldsymbol{V}&=\boldsymbol{\Sigma}^{-1}+\boldsymbol{X}_i^\top(\sigma^2 I)^{-1}\boldsymbol{X}_i\\
\boldsymbol{M}&=\boldsymbol{Y}_i(\sigma^2 I)^{-1}\boldsymbol{X}_i+\boldsymbol{\mu}\boldsymbol{\Sigma}^{-1}\\
\boldsymbol{R}&=\boldsymbol{Y}_i^\top(\sigma^2 I)^{-1}\boldsymbol{Y}_i+\boldsymbol{\mu}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}^\top
\end{aligned}
$$

re-writing the conditional posterior distribution, and ignoring some constant terms

$$
\pi(\mathbf{B}|.)\propto\prod_{i=1}^{n}\exp\{(\boldsymbol{\beta}_i^\top-\boldsymbol{V}^{-1}\boldsymbol{M})^\top\boldsymbol{V}(\boldsymbol{\beta}_i^\top-\boldsymbol{V}^{-1}\boldsymbol{M})\}
$$

Hence, each $\boldsymbol{\beta}_i$ has a conditional posterior multivariate normal distribution

$$
\pi(\boldsymbol{\beta}_i|.)\sim N(\boldsymbol{V}^{-1}\boldsymbol{M},\boldsymbol{V}^{-1})
$$

2. For $\pi(\sigma^2|.)$

$$
\begin{aligned}
\pi(\sigma^2|.) &\propto L_Y(\mathbf{B}^\top,\sigma^2)\times\pi(\boldsymbol{\sigma}^2)\\
&\propto\frac{1}{\sigma^2}\prod_{i=1}^{n}\Big\{(2\pi\sigma^2)^{-m_i/2}\exp\Big(-\frac{1}{2}(\boldsymbol{Y}_i-\boldsymbol{X}_i\boldsymbol{\beta}_i^\top)^\top(\sigma^2 I)^{-1}(\boldsymbol{Y}_i-\boldsymbol{X}_i\boldsymbol{\beta}_i^\top)\Big)\Big\}\\
&\propto\frac{1}{\sigma^2}^{(\frac{\sum_{i=1}^{n}m_i}{2}+1)}\exp\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(\boldsymbol{Y}_i-\boldsymbol{X}_i\boldsymbol{\beta}_i^\top)^\top(\boldsymbol{Y}_i-\boldsymbol{X}_i\boldsymbol{\beta}_i^\top)\}
\end{aligned}
$$

which follows the form of pdf of inverse gamma distribution

$$
f(x;\alpha,\beta)=\frac{\beta^\alpha}{\Gamma(\alpha)}\frac{1}{x}\exp\{-\frac{\beta}{x}\}
$$

in this case, x is replaced by σ^2, α is replaced by $\frac{1}{2}\sum\limits_{i=1}^{n} m_i$, β is replaced by $\frac{1}{2}\sum\limits_{i=1}^{n}(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i^\top)^\top(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i^\top)$

i.e.

$$\pi(\sigma^2|.) \sim IG(\frac{1}{2}\sum_{i=1}^{n} m_i, \frac{1}{2}\sum_{i=1}^{n}(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i^\top)^\top(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}_i^\top))$$

```
# betavec: 1*5
# sigma2: 1*1
# mu: 1*5
# Sigma: 5*5

# d = 5
# n = 702


loglik_sigma2 <- function(sigma2){
    if (sigma2 <= 0) {
        return(-Inf)
    } else {
        return(log(1/sigma2^2))
    }

}

loglik_Sigma <- function(Sigma, d){
    if (min(eigen(Sigma)$values) <= 0) {
        return(-Inf)
    } else {
        loglik = log(det(Sigma)^(-(d+1))*exp(-(1/2)*solve(Sigma)))
        return(loglik)
    }
}

# log likelihood for y given a specific hurricane's data given sigma2 and betavec
loglik_Y <- function(y, X, betavec, sigma2){
    m = length(y)
    loglik = (2*pi*sigma2^2)^(-(m/2))*exp(-(1/2)*t(X %*% t(betavec)) %*% solve(diag(sigma2, m)) %*% (X 
    loglik = log(loglik)
    return(loglik)

}

# log likelihood for betavec given a specific hurricane's mu and Sigma

loglik_betavec <- function(betavec, mu, Sigma){
    if (min(eigen(Sigma)$values) <= 0) {
        return(-Inf)
    } else {
        loglik = log(det(2*pi*Sigma)^(-(1/2))*exp(-(1/2)*(betavec - mu) %*% solve(Sigma) %*% t(betavec 
        return(loglik)
    }
}
```