Task 1

Climate researchers are interested in modeling the hurricane trajectories to forecast the wind speed. Let t be time (in hours) since a hurricane began, and For each hurricane i, we denote $Y_i(t)$ be the wind speed of the ith hurricane at time t. The following Bayesian model was suggested.

$$Y_i(t+6) = \beta_{0,i} + \beta_{1,i} Y_i(t) + \beta_{2,i}\Delta_{i,1}(t) + \beta_{3,i}\Delta_{i,2}(t) + \beta_{4,i}\Delta_{i,3}(t) + \mathbf{X}_i{}^{\top}\boldsymbol{\gamma} + \epsilon_i(t)$$

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 $t-6$ and t,

$\mathbf{X}_i = (x_{i,1}, x_{i,2}, x_{i,3})^{\top}$ are covariates with fixed effect $\boldsymbol{\gamma}$ ($\boldsymbol{\gamma}$ is a vector.), where $x_{i,1}$ be the month of year when the i-th hurricane started, $x_{i,2}$ be the calendar year of the i hurricane, and $x_{i,3}$ be the type of the i-th hurricane. (We assume $x_{i,2}$ are numeric variable, $\mathbf{x}_{i,1} = (x_{i,1,4}, \ldots, x_{i,12})^{\top}$ is categorical variable (April-January, January as reference category, 10 categories, 9 dummy variables) and $\mathbf{x}_{i,3} = (x_{i,3,1}, \ldots, x_{i,3,4})^{\top}$ is categorical variable (5 categories, 4 dummy variables). we assume $\boldsymbol{\gamma} = (\gamma_{1,4}, \ldots, \gamma_{1,12}, \gamma_2, \gamma_{3,1}, \ldots, \gamma_{3,4})^{\top}$, where the prior of $\boldsymbol{\gamma} \sim N(\mathbf{0}, 0.05^2 \boldsymbol{I}_{14})$)

and $\epsilon_{i,t}$ follows a normal distributions with mean zero and variance σ^2, independent across t.

In the model, $\boldsymbol{\beta}_i = (\beta_{0,i}, \beta_{1,i}, ..., \beta_{4,i})$ 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 $\boldsymbol{\Sigma}$.

Prior distributions

1. $\boldsymbol{\mu}$: a normal distribution with mean vector $\mathbf{0}$ and variance-covariance matrix $\boldsymbol{V}$, reflecting the prior knowledge that the mean coefficients should be centered around zero but allowing for some variability across hurricanes. The variance-covariance matrix $\boldsymbol{V}$ can be set to a diagonal matrix with large variances on the diagonal and small covariances off-diagonal, reflecting the prior knowledge that the coefficients may have some correlation but are largely independent across hurricanes.

2. $\boldsymbol{\Sigma}$: an inverse-Wishart distribution with degrees of freedom ν and scale matrix $\boldsymbol{S}$, reflecting the prior knowledge that the covariance matrix of the coefficients should be positive definite and have some structure. The degrees of freedom ν can be set to a small value (e.g., 5) to reflect a relatively weak prior, while the scale matrix $\boldsymbol{S}$ can be set to a diagonal matrix with large variances on the diagonal and small covariances off-diagonal, reflecting the prior knowledge that the covariance matrix should be diagonal or nearly diagonal.

3. All the fixed effects $\gamma \sim N(0, 0.05^2)$

4. σ: a half-Cauchy distribution with scale parameter 10, reflecting the prior knowledge that the residual variance should be positive and large enough to account for any unexplained variability in the wind speed data.

Task 1: Let $\mathbf{B} = (\boldsymbol{\beta}_1^\top, ..., \boldsymbol{\beta}_n^\top)^\top$, derive the posterior distribution of the parameters $\Theta = (\mathbf{B}^\top, \boldsymbol{\mu}^\top, \sigma^2, \boldsymbol{\Sigma}, \boldsymbol{\gamma}^\top)$.

Prior distribution

Let n denote the number of hurricanes in the training data. (we need to partition all hurricanes into training & test.)

$$
\begin{aligned}
\pi(\Theta) &= \pi(\mathbf{B}^\top, \boldsymbol{\mu}^\top, \boldsymbol{\Sigma}, \boldsymbol{\gamma}^\top, \sigma) \\
&= \pi(\mathbf{B}^\top \mid \boldsymbol{\mu}^\top, \boldsymbol{\Sigma}) \pi(\boldsymbol{\mu}^\top, \boldsymbol{\Sigma}) \pi(\boldsymbol{\gamma}) \pi(\sigma) \\
&= \left(\prod_{i=1}^{n} \pi(\boldsymbol{\beta}_i^\top \mid \boldsymbol{\mu}^\top, \boldsymbol{\Sigma}) \right) \pi(\boldsymbol{\mu}) \pi(\boldsymbol{\Sigma}) \pi(\boldsymbol{\gamma}) \pi(\sigma) \\
&= \prod_{i=1}^{n} \left[(2\pi)^{-5/2} |\boldsymbol{\Sigma}|^{-1/2} \exp\left(-\frac{1}{2} (\boldsymbol{\beta}_i - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{\beta}_i - \boldsymbol{\mu}) \right) \right] && \boldsymbol{\beta}_i \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \\
&\quad \times (2\pi)^{-5/2} |\boldsymbol{V}|^{-1/2} \exp\left(-\frac{1}{2} \boldsymbol{\mu}^\top \boldsymbol{V}^{-1} \boldsymbol{\mu} \right) && \boldsymbol{\mu} \sim N(\mathbf{0}, \boldsymbol{V}) \\
&\quad \times \frac{|\boldsymbol{S}|^{\nu/2}}{2^{5\nu/2} \Gamma_5(\nu/2)} |\boldsymbol{\Sigma}|^{-(v+6)/2} \exp\left(-\frac{1}{2} \mathrm{tr}(\boldsymbol{S} \boldsymbol{\Sigma}^{-1}) \right) && \boldsymbol{\Sigma} \sim \mathcal{W}^{-1}(\boldsymbol{S}, \nu) \\
&\quad \times (2\pi)^{-14/2} \cdot 20^{14} \cdot \exp\left(-\frac{1}{2} \cdot 20^2 \cdot \boldsymbol{\gamma}^\top \boldsymbol{\gamma} \right) && \boldsymbol{\gamma} \sim N(\mathbf{0}, 0.05^2 \boldsymbol{I}_{14}) \\
&\quad \times \frac{2}{10\pi} \cdot \frac{I(\sigma > 0)}{1 + (\sigma/10)^2} && \sigma \sim \text{half-Cauchy}(0, 10 \\
&\propto |\boldsymbol{\Sigma}|^{-n/2} \exp\left(-\frac{1}{2} \sum_{i=1}^{n} (\boldsymbol{\beta}_i - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{\beta}_i - \boldsymbol{\mu}) \right) \times \exp\left(-\frac{1}{2} \boldsymbol{\mu}^\top \boldsymbol{V}^{-1} \boldsymbol{\mu} \right) \\
&\quad \times |\boldsymbol{\Sigma}|^{-(v+6)/2} \exp\left(-\frac{1}{2} \mathrm{tr}(\boldsymbol{S} \boldsymbol{\Sigma}^{-1}) \right) \times \exp\left(-\frac{1}{2} \cdot 400 \boldsymbol{\gamma}^\top \boldsymbol{\gamma} \right) \times \frac{I(\sigma > 0)}{1 + (\sigma/10)^2} \\
&\propto \frac{|\boldsymbol{\Sigma}|^{-(n+v+6)/2}}{1 + (\sigma/10)^2} \exp\left[-\frac{1}{2} \left(\sum_{i=1}^{n} (\boldsymbol{\beta}_i - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{\beta}_i - \boldsymbol{\mu}) + \boldsymbol{\mu}^\top \boldsymbol{V}^{-1} \boldsymbol{\mu} + \mathrm{tr}(\boldsymbol{S} \boldsymbol{\Sigma}^{-1}) + 400 \|\gamma\|_2^2 \right) \right], \quad \sigma > 0
\end{aligned}
$$

Likelihood (or density?)

Let m_i denote the number of observations for the i-th hurricane, excluding the first and second observations. We denote the wind speed data for the i-th hurricane (excluding the first and second observations) as $\boldsymbol{Y}_i = (Y_{i,1}, \ldots Y_{i,m_i})^\top$, where $Y_{i,k}$ represents the wind speed at time $t_{0,i} + 6k$, and $t_{0,i}$ is the time of the second observation for the i-th hurricane. We denote the wind speed data for all hurricanes as $\boldsymbol{Y} = (\boldsymbol{Y}_1^\top, \boldsymbol{Y}_2^\top, \ldots, \boldsymbol{Y}_n^\top)^\top$. We denote $\Delta_{i,k,1}$, $\Delta_{i,k,2}$ and $\Delta_{i,k,3}$ as $\Delta_{i,1}(t_{0,i} + 6k)$, $\Delta_{i,2}(t_{0,i} + 6k)$ and $\Delta_{i,3}(t_{0,i} + 6k)$, and denote $\mathbf{Z}_{i,k} = (1, Y_{i,k}, \Delta_{i,k,1}, \Delta_{i,k,2}, \Delta_{i,k,3})^\top$. ($\mathbf{Z}_{i,k}$ contains $Y_{i,k}$, the previous observation of wind speed)

Given that

$$
Y_{i,j} \mid (Y_{i,j-1}, \boldsymbol{\beta}_i^\top, \boldsymbol{\gamma}^\top, \sigma) \sim N(\mathbf{Z}_{i,j-1}^\top \boldsymbol{\beta}_i + \mathbf{X}_i^\top \boldsymbol{\gamma}, \sigma^2), \quad \forall j = 1, \cdots, m_i
$$

we have

$$\begin{aligned}
f(\boldsymbol{Y}^\top \mid \Theta) &= \prod_{i=1}^{n} f(\boldsymbol{Y}_i^\top \mid \mathbf{B}^\top, \boldsymbol{\mu}^\top, \boldsymbol{\Sigma}, \boldsymbol{\gamma}^\top, \sigma) \\
&= \prod_{i=1}^{n}\prod_{j=1}^{m_i} f(Y_{i,j} \mid Y_{i,j-1}, \boldsymbol{\beta}_i^\top, \boldsymbol{\gamma}^\top, \sigma) \\
&= \prod_{i=1}^{n}\prod_{j=1}^{m_i} \left[\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(Y_{i,j} - \mathbf{Z}_{i,j-1}^\top \boldsymbol{\beta}_i - \mathbf{X}_i^\top \boldsymbol{\gamma})^2}{2\sigma^2}\right)\right] \\
&= \left(\prod_{i=1}^{n}(\sqrt{2\pi}\sigma)^{-m_i}\right)\cdot \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\sum_{j=1}^{m_i}(Y_{i,j} - \mathbf{Z}_{i,j-1}^\top \boldsymbol{\beta}_i - \mathbf{X}_i^\top \boldsymbol{\gamma})^2\right).
\end{aligned}$$

Posterior distribution

$$\begin{aligned}
\pi(\Theta \mid \boldsymbol{Y}^\top) &\propto f(\boldsymbol{Y}^\top \mid \Theta)\pi(\Theta) \\
&\propto \frac{|\boldsymbol{\Sigma}|^{-(n+v+6)/2}}{1+(\sigma/10)^2}\exp\left[-\frac{1}{2}\left(\sum_{i=1}^{n}(\boldsymbol{\beta}_i - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{\beta}_i - \boldsymbol{\mu}) + \boldsymbol{\mu}^\top \boldsymbol{V}^{-1}\boldsymbol{\mu} + \mathrm{tr}(\boldsymbol{S}\boldsymbol{\Sigma}^{-1}) + 400\|\boldsymbol{\gamma}\|_2^2\right)\right] \\
&\quad \times I(\sigma > 0)\sigma^{-\sum_{i=1}^{n} m_i}\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\sum_{j=1}^{m_i}(Y_{i,j} - \mathbf{Z}_{i,j-1}^\top \boldsymbol{\beta}_i - \mathbf{X}_i^\top \boldsymbol{\gamma})^2\right).
\end{aligned}$$