Project 3 (A): Baysian modeling of hurrican trajectories

Hurrican Data

hurrican703.csv collected the track data of 703 hurricanes in the North Atlantic area since 1950. For all the storms, their location (longitude & latitude) and maximum wind speed were recorded every 6 hours. The data includes the following variables

1. **ID**: ID of the hurricans
2. **Season**: In which the hurricane occurred
3. **Month**: In which the hurricane occurred
4. **Nature**: Nature of the hurricane

- ET: Extra Tropical
- DS: Disturbance
- NR: Not Rated
- SS: Sub Tropical
- TS: Tropical Storm

5. **time**: dates and time of the record

6. **Latitude** and **Longitude**: The location of a hurricane check point
7. **Wind.kt** Maximum wind speed (in Knot) at each check point

Load and explore the hurrican data through visulaizations

```
library(ggplot2)
dt= read.csv("data/hurrican703.csv")
ggplot(data=dt, aes(x = Longitude, y = Latitude)) +
  stat_summary_2d(data = dt, aes(x = Longitude, y = Latitude,
  z = dt$Wind.kt), fun = median, binwidth = c(1, 1), show.legend = TRUE)
library(data.table)
dt <- as.data.table(dt)
summary(dt)
```

Overlay the hurrican data in the world map

```
library(maps)
map <- ggplot(data = dt, aes(x = Longitude, y = Latitude)) +
  geom_polygon(data = map_data(map = 'world'), aes(x = long, y = lat, group = group))
map +
  stat_summary_2d(data = dt, aes(x = Longitude, y = Latitude,
  z = dt$Wind.kt), fun = median, binwidth = c(1, 1), show.legend = TRUE, alpha = 0.75) +
  ggtitle(paste0("Atlantic Windstorm mean knot"))
```

Additional Plots

```
map <- ggplot(dt, aes(x = Longitude, y = Latitude, group = ID)) +
  geom_polygon(data = map_data("world"),
               aes(x = long, y = lat, group = group),
               fill = "gray25", colour = "gray10", size = 0.2) +
```

```
  geom_path(data = dt, aes(group = ID, colour = Wind.kt), size = 0.5) +
  xlim(-138, -20) + ylim(3, 55) +
  labs(x = "", y = "", colour = "Wind \n(knots)") +
  theme(panel.background = element_rect(fill = "gray10", colour = "gray30"),
        axis.text.x = element_blank(), axis.text.y = element_blank(),
        axis.ticks = element_blank(), panel.grid.major = element_blank(),
        panel.grid.minor = element_blank())

seasonrange <- paste(range(dt[, Season]), collapse=" - ")

map + ggtitle(paste("Atlantic named Windstorm Trajectories (",
                     seasonrange, ")\n"))
```

Show hurricance tracks by month

```
mapMonth <- map + facet_wrap(~ Month) +
  ggtitle(paste("Atlantic named Windstorm Trajectories by Month (",
                 seasonrange, ")\n"))
mapMonth
```

A Hierarchical Bayesian model for hurricane trajectories.

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

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

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 and $t+6$,

$\mathbf{X}_i = (x_{i,1}, x_{i,2}, x_{i,3})$ are covariates with fixed effect γ, 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.

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

In the model, $\boldsymbol{\beta}_i = (\beta_{0,i}, \beta_{1,i}, ..., \beta_{7,i})$ 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 $\boldsymbol{\beta}$ and covariance matrix Σ.

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.

Your tasks:

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

2. Design and implement a custom MCMC algorithm for the outlined Bayesian hierarchical model. Monitor the convergence of the MCMC chains, using diagnostic plots and summary statistics to check for any issues.

3. Compute posterior summaries and 95% credible intervals of γ, the fixed effects associated with the covariates in the model. Using the estimated Bayesian model, answer the following questions: (1) are there seasonal differences in hurricane wind speeds, and (2) is there evidence to support the claim that hurricane wind speeds have been increasing over the years?

4. With the estimated model parameters and covariate values, you can calculate the predicted wind speed for each time point using the model equation. This way, you can track the hurricane and compare the predicted wind speeds with the actual wind speeds recorded during the hurricane. Please evaluate how well the estimated Bayesian model can track individual hurricanes.

5. Write a report of your findings.