

# Bayesian Modeling of Hurricane Trajectories

Hongjie Liu, Xicheng Xie, Jiajun Tao, Zijian Xu, Shaohan Chen

May 1st, 2023

# Outline

- ▶ Background
- ▶ Task 1
- ▶ Task 2
- ▶ Task 3
- ▶ Task 4
- ▶ Discussions
- ▶ Reference
- ▶ Q&A

# Background

## **Bayesian Modeling of Hurricane Trajectories:**

- ▶ In this project we are interested in modeling the hurricane trajectories to forecast the wind speed.

# Background

## **Data Source:**

- ▶ “hurricane703.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.

# Task 1

## Objective:

Let  $B = (\beta_1^T, \dots, \beta_n^T)$ , derive the posterior distribution of the parameters  $\Theta = (\mathbf{B}^T, \boldsymbol{\mu}^T, \sigma^2, \Sigma)$

## Task 2

### **Objective:**

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.

## Task 2 - MCMC Algorithm

---

### Algorithm 1 MCMC Algorithm (Part 1)

---

**Require:**  $\mathbf{Y}; \beta_0, \mu_0, \Sigma_0, \sigma_0, \gamma_0$

**Ensure:**  $\hat{\beta}, \hat{\mu}, \hat{\Sigma}, \hat{\sigma}, \hat{\gamma} \approx \beta, \mu, \Sigma, \sigma, \gamma$

$i \leftarrow 0$ , where  $i$  is the current number of iterations

**while** iteration times is not met **do**

$i \leftarrow i + 1$

Gibble sampling for  $\beta$

$$\beta^{(k)} \sim \mathcal{N}(\beta | \mathbf{Y}, \mu^{(k-1)}, \Sigma^{(k-1)}, \sigma^{(k-1)}, \gamma^{(k-1)})$$

Gibble sampling for  $\mu$

$$\mu^{(k)} \sim \mathcal{N}(\mu | \mathbf{Y}, \beta^{(k)}, \Sigma^{(k-1)}, \sigma^{(k-1)}, \gamma^{(k-1)})$$

Gibble sampling for  $\Sigma$

$$\Sigma^{(k)} \sim \mathcal{W} \langle -\nabla \sqcup (\Sigma | \mathbf{Y}, \beta^{(k)}, \mu^{(k)}, \sigma^{(k-1)}, \gamma^{(k-1)}) \rangle$$

Gibble sampling for  $\gamma$   $\gamma^{(k)} \sim p(\Sigma | \mathbf{Y}, \beta^{(k)}, \mu^{(k)}, \Sigma^{(k)}, \sigma^{(k-1)})$

## Task 2 - MCMC Algorithm

---

**Algorithm 2** MCMC Algorithm (Part 2)

---

Metropolis-Hastings for  $\sigma$

Propose a new value  $\sigma^*$  from a normal distribution with mean  $\sigma^{(k-1)}$  and a small variance.

Compute the acceptance ratio

$$\gamma = \frac{p(\gamma^{(k)}|\mathbf{Y}, \beta^{(k)}, \boldsymbol{\mu}^{(k)}, \boldsymbol{\Sigma}^{(k)}, \sigma^{(k)})q(\gamma^{(k-1)}|\gamma^{(k)})}{p(\gamma^{(k-1)}|\mathbf{Y}, \beta^{(k)}, \boldsymbol{\mu}^{(k)}, \boldsymbol{\Sigma}^{(k)}, \sigma^{(k)})q(\gamma^{(k)}|\gamma^{(k-1)})}$$

Generate a random number  $u$  from a uniform distribution between 0 and 1.

If  $u \leq \gamma$ , set  $\sigma^{(k)} = \sigma^*$ , otherwise set  $\sigma^{(k)} = \sigma^{(k-1)}$ .

**end while**

---



# Task 2 - MCMC Algorithm

## Initial Values:

- ▶  $\beta_i$ : This can be obtained through the random effects term in the lmm model. The random effects term can be added to the fixed effects term to obtain  $\beta_i^{(0)}$ .
- ▶  $\mu$ : This can be obtained through the fixed effects term `lmmcoefficients[1]`.
- ▶  $\sigma^2$ : This can be obtained through the model residual `sigma0`.
- ▶  $\Sigma^{-1}$ : This can be obtained through the `VarCorr(lmm)` function which returns the covariance matrix of the random effects in the model. The inverse of this matrix can be taken to obtain  $\Sigma^{-1(0)}$ .

Table 1: Initial Values

Parameter	Value
$\mu$	(24.25, 0.94, -0.02, -0.24, 0.47)
$\gamma$	(-0.01, 0.35, 0.28, 0.37, 0.12, 0.08)
$\Sigma$	$\begin{pmatrix} 0.36 & -0.01 & 0.04 & 0.12 & 0.03 \\ -0.01 & 0.00 & -0.00 & -0.00 & 0.00 \\ 0.04 & -0.00 & 0.04 & 0.03 & -0.02 \\ 0.12 & -0.00 & 0.03 & 0.07 & 0.00 \\ 0.03 & 0.00 & -0.02 & 0.00 & 0.02 \end{pmatrix}$
$\sigma^2$	5.27

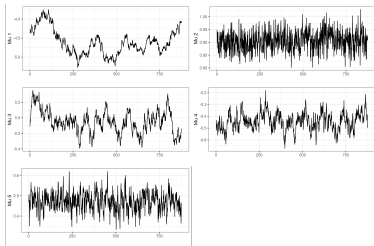
## Task 2 - MCMC Algorithm R code

```
# Gimble sampling algorithm
B_sample <- function(mu, Sigma, gamma, sigma) {
  Sigma.inv <- solve(Sigma)
  B_mean_cov <- function(i) {
    cov <- solve(Sigma.inv + 1/sigma^2 * t(Z[[i]]) %*% Z[[i]])
    mean <- cov %*% (Sigma.inv %*% mu + 1/sigma^2 * colSums((Y[[i]] - (X[i,] %*% gamma)[,]) * Z[[i]]))
    list(mean = mean, cov = cov)
  }
  mean_cov_list <- lapply(1:n, B_mean_cov)
  B <- sapply(mean_cov_list, function(x) {mvrnorm(mu = x$mean, Sigma = x$cov)})
  return(B)
}

# MH algorithm (random walk)
sigma_sample <- function(sigma, B, gamma, a) {
  sigma_new <- sigma + (runif(1) - 0.5) * 2 * a # candidate sigma
  if (sigma_new <= 0) {
    return(sigma)
  }
  RSS <- sum(sapply(1:n, function(i) sum((Y[[i]] - Z[[i]] %*% B[,i] - (X[i,] %*% gamma)[,])^2)))
  log_kernal_ratio <- -sum(m) * log(sigma_new/sigma) +
    log(1 + (sigma/10)^2) - log(1 + (sigma_new/10)^2) -
    0.5 * (1/sigma_new^2 - 1/sigma^2) * RSS
  log_prob <- min(0, log_kernal_ratio)
  sigma <- ifelse(log_prob > log(runif(1)), sigma_new, sigma)
  return(sigma)
}
```

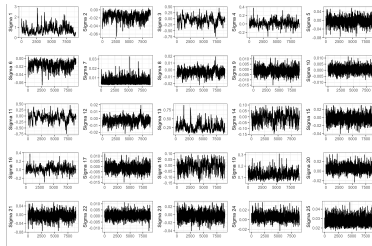
# Task 3 - Results Presentation

Time Series Plot of Mu



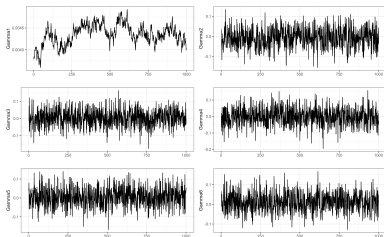
(a) Mu

Time Series Plot of Sigma

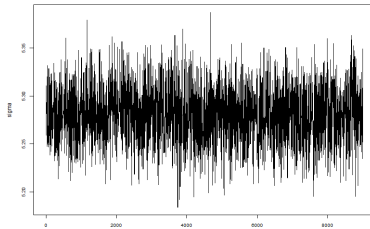


(b) Sigma

Time Series Plot of Gamma



(c) Gamma



(d) sigma

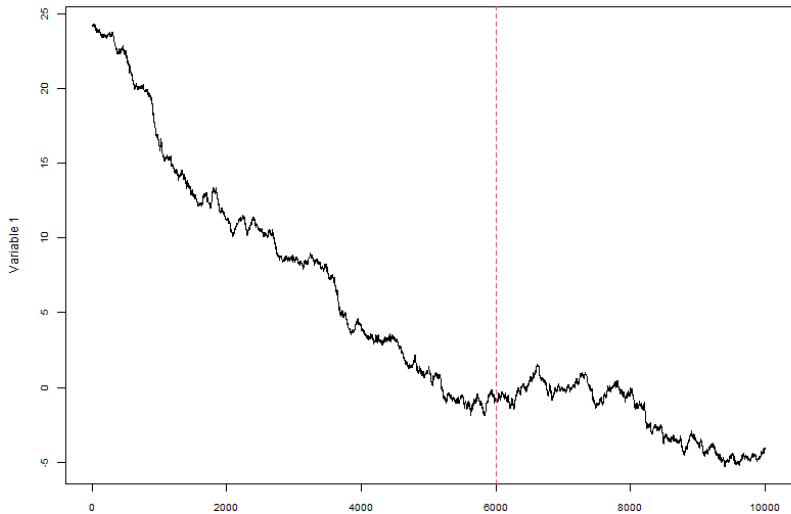
## Task 3

### **Objective:**

Compute posterior summaries and 95% credible intervals of  $\gamma$ , 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?
- (2) Is there evidence to support the claim that hurricane wind speeds have been increasing over the years?

## Task 3 - Parameters Convergence



## Task 4 - Objective

### **Objective:**

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.

## Task 4

**Prediction:** Using the parameters after burn-in, we can obtain the predicted value for each hurricane.

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

**Performance evaluation:** For each hurricane, we can evaluate the estimated Bayesian model performance by calculating

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}}$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

## Task 4

Table 2: Summary of RMSE and R-squared for selected hurricanes

ID	Year	RMSE	R-squared
ABBY.1960	1960	8.8804	0.7700
ABBY.1964	1964	9.6430	0.3033
ABBY.1968	1968	3.5043	0.9360
ABLE.1950	1950	3.6755	0.9813
ABLE.1951	1951	3.4802	0.9767
ABLE.1952	1952	4.5183	0.9583
AGNES.1972	1972	5.2483	0.8881
ALBERTO.1982	1982	8.0473	0.7499
ALBERTO.1988	1988	2.6121	0.7420
ALBERTO.1994	1994	4.3941	0.8807
ALBERTO.2000	2000	3.7896	0.9625
ALBERTO.2006	2006	4.3591	0.7882
ALBERTO.2012	2012	3.2193	0.8036
ALEX.1998	1998	2.9351	0.7289
ALEX.2004	2004	5.4552	0.9539



# Discussions

- ▶ Parameters Convergence Problem

# Reference

Reference

# Q&A

- ▶ Thanks for listening!