P8160 - Baysian modeling of hurrican trajectories

Jiahao Fan, Tianwei Zhao, Yijin Wang, Youlan Shen, Yujia Li

2023-04-29

Contents

Background and Objectives

As hurricanes affect people, the ability to forecast hurricanes is essential for minimizing the risks in suffered areas. Hereby, a hierarchical Bayesian strategy for modeling North Atlantic hurricane counts since 1950 is illustrated. Model integration would be expected to achieve through a Markov chain Monte Carlo algorithm. Contingent on the availability of values for the covariates, the models can be used to explore the seasonality among track data and make predictive inferences on hurricanes.

The data given has 703 observations and following features are recorded for each hurricanes in the North Atlantic:

- ID: ID of the hurricanes
- Season: In which the hurricane occurred
- Month: In which the hurricane occurred
- Nature: Nature of the hurricane (ET-Extra tropical;DS-Disturbance;NR-Not rated;SS-Sub tropical; TS-Tropical storm)
- time: dates and time of the record
- Latitude and Longitude: The location of a hurricane check point
- Wind.kt: Maximum wind speed (in Knot) at each check point

Methods

Baysian Model

MCMC Algorithm

Given the fact that we cannot directly sample from the joint posterior distribution calculated above. We choose to use Gibbs sampling. Gibbs sampling is a sequential sampler that samples from the conditional distribtuion of each parameter so that the final collection of samples forms a Markov chain whose stationary distribution is the joint distribution that we are interested in.

In our case, conditional posteriors are more manageable than the joint posterior. Conditonal posteriors for each parameter are calculated as the following:

!!! ADD CONDITONAL POSTERIOR

Still, we could not find a regular distribution that has the density function that matches with the conditional posterior of σ^2. Therefore, we incorporated Metropolis Hasting Random Walk into Gibbs sampling. Metropolis Hasting Random Walk finds approximated values of σ^2 by accepting and rejecting proposed value with an acceptance probability that helps retain balance of the system. In our algorithm, we work with log-scaled posterior values instead of the original-scaled posterior because $\frac{1}{\sigma^N}$ dominates the entire posterior value where N is the total number of observations in the data set.

In our investigation of hurricane wind speed, we used Gibbs sampling and as Metropolis Hasting Random Walk as the following:

Algorithm 1 Gibbs Sampling

function GIBBS(starting_values, V,S,X,D,Y, nitr,ν,H, hurricane_idx)
  create posts results lists to collect results for B, μ, A, γ, σ^2
  $B^1 \leftarrow$ f(B|start_μ, start_σ^2, start_A, start_γ, Y)
  $\mu^1 \leftarrow$ f(μ|B^1, start_σ^2, start_A, start_γ, Y)
  $A^1 \leftarrow$ f(A|B^1, μ^1, start_σ^2, start_γ, Y)
  $\gamma^1 \leftarrow$ f(γ|B^1, μ^1, start_σ^2, A^1, Y)
  σ^2_chain $\leftarrow$ metropolis_hasting_sigma_squared(start_σ^2, iter = 500, a = 0.8,Y,D,B^1,X,γ^1)
  $(\sigma^2)^1 \leftarrow$ mean(σ^2_chain[401:500])
  for i in 2:nitr **do**
    $B^i \leftarrow$ f(B|μ^{i-1}, $(\sigma^2)^{i-1}$, A^{i-1}, γ^{i-1}, Y)
    $\mu^i \leftarrow$ f(μ|B^i, $(\sigma^2)^{i-1}$, A^{i-1}, γ^{i-1}, Y)
    $A^i \leftarrow$ f(A|B^i, μ^i, $(\sigma^2)^{i-1}$, γ^{i-1}, Y)
    $\gamma^i \leftarrow$ f(γ|B^i, μ^i, $(\sigma^2)^{i-1}$, A^i, Y)
    σ^2_chain $\leftarrow$ metropolis_hasting_sigma_squared($(\sigma^2)^{i-1}$, iter = 500, a = 0.8,Y,D,B^i,X,γ^i)
    $(\sigma^2)^i \leftarrow$ mean(σ^2_chain[401:500])
  end for
  return posts results lists that collect results for B, μ, A, γ, σ^2
end function

Figure 1: Gibbs Sampling

Algorithm 2 Metropolis Hasting sigma squared

function METROPOLIS HASTING SIGMA SQUARED(start_sigma_squared, nitr, a, B, X,D,Y, H, hurricane_idx)
  Initialize chain
  chain[1] $\leftarrow$ start_sigma_squared
  for i in 2:itr **do**
    curr $\leftarrow$ *chain*$[i-1]$
    prop $\leftarrow$ a sample from Unif(curr-a, curr+a)
    prop_posterior $\leftarrow$ sigma_squared_log_posterior(prop,...)
    curr_posterior $\leftarrow$ sigma_squared_log_posterior(curr,...)
    posterior_ratio $\leftarrow$ exp(prop_posterior - curr_posterior)
    $\alpha \leftarrow$ min(posterior_ratio, 1)   ▷ acceptance probability
    $u \leftarrow runif(1)$
    if u $<$ α **then**
      chain$[i] \leftarrow$ *prop*   ▷ accept proposal
    else
      chain$[i] \leftarrow$ *curr*   ▷ reject proposal
    end if
  end for
  return *chain*
end function

Figure 2: Metropolis Hasting Random Walk

We decide starting values of each parameters by fitting a linear mixed effects model that is the same as the proposed model in a frequentist way and then extract its corresponding coefficients. Starting values for γ are the fixed effects coefficients in the linear mixed effects model.

Table 1: Starting Values for Gibbs Sampling

Terms	Starting Value
γ	Fixed effect coefficients in the fitted Linear Mixed Effects Model
$\beta_{i_1,\ldots,i_n}$	Random effect coefficients in the fitted Linear Mixed Effects Model
σ^2	Residual Variance in the fitted Linear Mixed Effects Model
β_0	Sum of random effect and fixed effect intercepts in the fitted Linear Mixed Effects Model
μ	Mean of the $\beta_{i_1,\ldots,i_n}$ for each predictor
$A = \Sigma^{-1}$	Mean of the Wishart distribution with $\nu = 5$ and scale matrix with 0.7 on the main diagnoal and 0.2 on off-diagnoal

Results

Seasonal Analysis

Prediction

Discussion

Group Contributions

References

Appendices

Figures and Tables