Milestone 1

Diego Arevalo Fernandez, Charlie Ngo, Diego Osborn, Sardor Sobirov

Problem Description

We want to study hurricane track prediction using a probabilistic state-space model. Where the hurricane's actual physical state (e.g., its location, velocity, intensity, etc.) is treated as a latent state. With this framework, hurricane forecasting essentially becomes a sequential Bayesian inference problem. Meaning, as we are provided with new observations, we can update our belief about the hurricane's current state and predict where it may move next. To do this, we want to use the Kalman filter (a similar use case can be found in Zulfi, Hasan, & Purnomo, 2018). Once we have updated estimates of the storms' current position and motion from the Kalman filter, we can create repeated simulations to generate a collection of possible future tracks, where each simulation allows for some randomness in the process. This yields "spaghetti" plots, which visualize the possible paths of the hurricane, as well as our confidence in these forecasts.

Dataset Source

We will use the NOAA National Hurricane Center HURDAT2 "best track" database, which can be accessed via the NHC data archive (IBTrACS, 2021). This dataset provides the hurricane's position and intensity throughout its lifetime every six hours. These reported positions will serve as the observational data in our model. This dataset also enables us to evaluate our forecasts by comparing our predicted tracks to the positions provided.

Methodology

Although the Kalman filter was not explicitly covered in the course, it is essentially a continuous-state version of the forward algorithm for a HMM, in which the hidden state evolves linearly, rather than discretely, and all randomness is Gaussian. For each hurricane, our model is defined by a transition matrix A, an observation matrix H, and two noise covariances, Q, which captures the inherent uncertainty regarding the storm's movement, and R, which captures the inherent uncertainty regarding the reported position data. Given these parameters, the Kalman filter combines the state-space model's prediction of the state at time t with the new observation y_t to produce an updated estimate of the current state, $P(x_t \mid y_{1:t})$, and also generates a one-step-ahead predictive distribution, $P(x_{t+1} \mid y_{1:t})$, which serves as the model's forecast for the next six-hour interval. To evaluate the model, we'll compute the forecast error at each time step by comparing the Kalman filter's one-step-ahead predicted position to the next observed best-track position.

References

International Best Track Archive for Climate Stewardship (IBTrACS). (2021, June 17). National Centers for Environmental Information (NCEI). https://www.ncei.noaa.gov/products/international-best-track-archive

Zulfi, M., Hasan, M., & Purnomo, K. (2018). The development of rainfall forecasting using Kalman filter. Journal of Physics: Conference Series, 1008, 012006. https://doi.org/10.1088/1742-6596/1008/1/012006