Biostats 270: HW 3

Andrew Holbrook

Spring 2024

For the following, please include code and output in a single pdf file. Assignment due by the beginning of class on Thursday 5/9.

1. Use a kernel that is the sum of a Matérn kernel ($\nu = 1/2$ and $\rho = 10$) and nugget $\delta_{ij} = 0.00001$ to simulate a single GP realization over values $\{1, \dots, 100\}$, i.e., $(y(1), \dots, y(100))$. Conditioning on $(y(1), \dots, y(50))$ and using the true model parameters, simulate 100 independent draws from the conditional distribution of $(y(51), \dots, y(100))$. **Plot** the truth, the predicted paths and 95% predictive intervals.

2. Use a kernel that is the sum of a Matérn kernel ($\nu = 1/2$ and $\rho = 4$) multiplied by variance term $\sigma^2 = 2$ and nugget $\delta_{ij} = 0.00001$ to simulate a single GP realization over 10 random values $x_i \in (0, 100)$.
   (a) Placing priors of your choosing on σ^2 and ρ, use Metropolis-Hastings to generate posterior samples conditioning on $y(x_i)$, $i \in \{1, \dots, 10\}$. Run the chain long enough such that, after removing a reasonable burn-in, the effective sample size for both parameters is above 1000. **Plot** both posterior densities.
   (b) Thin your sample so you only have 1000 stored MCMC states. Use these samples to simulate posterior predictive curves over a grid ranging from 0 to 100 with 300 points. **Plot** the original observations, the posterior predictive curves and 95% predictive intervals.

3. Letting $a = 1$, $b = 2$ and $\mu = 0.1$, use Ogata's modified thinning algorithm to simulate one sample path of an exponential Hawkes process on the interval $(0, 100)$. How many events do you observe? Conditioning on this realization and assuming we know $\mu = 0.1$, perform posterior inference on a and b after specifying appropriate priors. Make sure to achieve an effective sample size over 1000 for both. **Plot** posterior densities. Are these reasonable results?

4. Consider the same situation as in Problem 3 but let $\mu = 2$. Again draw a single sample path. How many events do you observe?
   (a) Perform posterior inference on a and b using the naive likelihood. How long does it take to generate 100k samples?
   (b) Perform posterior inference on a and b using the linear-time likelihood. How long does it take to generate 100k samples?