

## ST 758 Project 2

Jessica Miller

Nov. 10, 2016

For this assignment, I did not have a lot of time to complete it around recent and sudden events in the last two weeks. However, I would like to take this opportunity to express some thoughts I had about how I might approach it.

One of the first approaches that comes to mind is to reduce the number of locations. The more locations, the larger the distance matrix and the longer it takes to compute the log-likelihood. To explore this, I would plot the locations and identify a grid of knots. This grid would represent and approximate the spatial domain without using all the locations. If there is clustering in the locations, I would weight the knots by location population in that area. In order to address accuracy, I would explore the number of knots to choose by increasing their density, under the assumption that more knots comes closer to the truth, while tracking the speed of the function computation. By balancing the two, I would get an approximation that is accurate and yet computes much more quickly than if I used every single location.

Something I would also explore would be Fourier transforms. I know there is a one-to-one map from a covariance function to a spectral density and that using spectral densities are significantly faster to compute. Since the covariance functions in the assignment are similar to the common exponential and Matern functions, I would find their corresponding spectral densities and compare the timing of the Gaussian log-likelihood using Fourier computations to the timing using covariance functions.

Finally, I would focus on the covariance matrix and likelihood functions themselves. The diagonals of each are equal to 1, which implies the matrix could be almost Toeplitz. Exploiting such characteristics would improve the calculation of the inverse and determinant. There are a variety of methods to approximate the likelihoods, such as pairing the data into bivariate Normals instead of multivariate. One could also divide the locations up into blocks, thus computing smaller distance and covariance matrices, and then multiply all the likelihoods together.

These are all methods I am very interested in exploring. It is unfortunate that I was not able to.