SDS 385, Big Data
Fall 2016
Giorgio Paulon & Jennifer Starling
giorgio.paulon@utexas.edu, jstarling@utexas.edu

Spatial Smoothing of Continuous Gamma Radiation Data via Graph Fused Lasso and Fixed Rank Kriging

The radiation data set consists of 814,458 gamma radiation level readings observed on the University of Texas at Austin campus. These data were gathered over the period of a few months, and include the latitude, longitude, altitude, temperature at time of reading, and the gamma radiation level.

The general purpose of spatial smoothing in this setting is to interpolate the response variable over areas of the grid without readings. In the radiation data case, the data was gathered using measurements obtained from a device attached to a police car, so readings are limited to latitudes and longitudes where a vehicle travels. The goal is to interpolate the radiation levels over areas of the campus not covered in the police car's travels.

We propose two potential approaches. The first is to discretize the data and apply the Graph Fused Lasso technique, with an optimized ADMM approach to decrease computation time. Potential approaches are as described in Exercise 08, including the Tansey et al. (2014) paper. The second is to work with the data in continuous form, and implement Fixed Rank Kriging. This approach is preferable to discretizing the data, as it preserves specific location information instead of splitting the campus into a grid. Fixed Rank Kriging is also superior to traditional Kriging, as traditional Kriging is computationally intensive and slow.

We will select a final approach after reviewing sources, including Katzfuss and Cressie (2011), Tutorial on Fixed Rank Kriging (FRK) of CO_2 Data. Our primary goal is to implement Fixed Rank Kriging.