Intro to Spatial Point Patterns Project

Problem Setup

This assignment is a continuation of the heat-related 911 calls example. The file `Spatial911 PtPtrn.RData` contains the following information:

- `calls` - a 1389×7 data frame containing demographic information on 1389 911 calls for heat-related health problems.
- `houston` - a shape-file data from with information on Houston city boundaries. Try:

  `> plot(houston)`

- `pp.grid` - a 1428×2 data frame of lat-long points that form a grid covering the city of Houston. Try:

  `> plot(houston)`

  `> points(pp.grid$Longitude,pp.grid$Latitude`

- `hrldas.grid` - a 15625×2 data frame of lat-long points that form a complete grid covering the city of Houston. Try:

  `> plot(houston)`

  `> points(hrldas.grid$Longitude,hrldas.grid$Latitude`

- `kp.gp` - a length 15625 logical vector such that `pp.grid = hrldas.grid[kp.gp,]`.

For this assignment we wish to consider only the spatial locations of the 911 calls. Let $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_N$ where $N = 1389$ be the latitude-longitude coordinates of the 911 calls in `calls` and $\mathcal{S}$ denote the set of locations corresponding to `houston`. Furthermore let $\boldsymbol{s}_1^\star, \ldots, \boldsymbol{s}_K^\star$ be the $K = 1428$ spatial locations in `pp.grid`.

For spatial point patterns, we wish to model the intensity surface $\Lambda(\boldsymbol{s})$ as a smooth function of $\boldsymbol{s}$. Unlike age from the previous assignment, $\boldsymbol{s}$ indexes a continuous spatial location. Hence, $\Lambda(\boldsymbol{s})$ is a continuous (rather than discrete) intensity surface. Therefore, to make the modeling easier define we discretize $\Lambda(\boldsymbol{s})$ as,

$$\Lambda(\boldsymbol{s}) = \delta \sum_{k=1}^{K} \frac{\lambda_k}{|\mathcal{G}_k|} \mathbb{1}\{\boldsymbol{s} \in \mathcal{G}_k\} \tag{1}$$

where $\mathcal{G}_k$ corresponds to the set of points that are closest to grid point $\boldsymbol{s}_k^\star$ and $|\mathcal{G}_k| = \int_{\mathcal{G}_k} d\boldsymbol{s}$ is the "area" of grid box $\mathcal{G}_k$. Furthermore, we assume that each $\lambda_k \in (0,1)$ and $\sum_{k=1}^{K} \lambda_k = 1$. Effectively, (1) discretizes the spatial domain $\mathcal{S}$ into K values $\{\lambda_1, \ldots, \lambda_K\}$. The associated likelihood function is

$$L(\Lambda) = \exp\{-\delta\} \, \delta^N \prod_{i=1}^{N} \frac{\lambda_{\{k:\boldsymbol{s}_i \in \mathcal{G}_k\}}}{|\mathcal{G}_k|} \tag{2}$$

Interpretation Questions

1. Show that $\int_{\mathcal{S}} \Lambda(\boldsymbol{s}) d\boldsymbol{s} = \delta$ and interpret δ in context.

2. Show that $\mathbb{P}\text{r}(\boldsymbol{s} \in \mathcal{G}_k) = \lambda_k$ and interpret λ_k in context.

Modeling Questions

Note that we could assume $\lambda_1, \ldots, \lambda_K \sim$ Dirichlet as we did in the age of 911 caller analysis to get a conjugate analysis but this would ignore spatial smoothness in the intensity surface Λ. Instead let, $\lambda_k \propto \exp\{\lambda_k^\star\}$ where λ_k follows a Gaussian process with constant mean μ and covariance given by the Matérn covariance function. In other words,

$$\boldsymbol{\lambda}^\star = \begin{pmatrix} \lambda_1^\star \\ \vdots \\ \lambda_K^\star \end{pmatrix} \sim \mathcal{N}\left(\boldsymbol{0}, \sigma^2 \boldsymbol{M}\right) \tag{3}$$

where $\boldsymbol{M}$ is a $K \times K$ correlation matrix with ij^{th} element $M_\nu(\|\boldsymbol{s}_i^\star - \boldsymbol{s}_j^\star\| \mid \phi)$ where $M_\nu(d \mid \phi)$ is the Matérn correlation function with smoothness ν and decay ϕ. It the follows that λ_k is the normalized version of $\exp\{\lambda_k^\star\}$ such that $\lambda_k = \exp\{\lambda_k^\star\}/\sum_{j=1}^K \exp\{\lambda_j^\star\}$. Therefore, $\{\lambda_k^\star\}_k$ are the true parameters that need to be estimated.

1. Fix $\nu = 3.5$ to enforce smoothness and let $\phi_k = 343$.

2. Let the prior for δ be a gamma distribution with shape a_δ and rate b_δ (you choose a_δ and b_δ).

3. Let the prior for σ^2 be inverse-gamma with shape a_σ and rate b_σ (you choose a_σ and b_σ).

4. Write a Metropolis within Gibbs sampler to generate draws of $\boldsymbol{\lambda}^\star$, δ and σ^2. Hint: the complete conditional distributions for μ and σ^2 will be Normal and Inverse-gamma, respectively.