Bayesian Models for Service Quality Analysis of the Boston City 311

Jane Huang
Harvard University

Isadora Nun
Harvard University

Weiwei Pan
Harvard University

Francisco Rivera
Harvard University

Abstract

Fill in later.

1 Introduction

In 2015, Mayor Martin J. Walsh launched Boston 311, a platform to better enable residents of Boston to report non-emergency issues to the City, such as graffiti and broken street lights, and access City services. The program is similar to the 311 program that has been in place in New York City since 2003, and is aimed to streamline City services for Boston's residents. There is now a number of on-going and active data-analytics research projects based on the database of 311 service requests.

We suspect that analysis of response times will show complex patterns of heterogeneity. In particular, we are interested in the significance of the impact of factors such as geography or neighborhood demographics. We will approach this in two ways:

1. We propose to model latent classes in response time to service requests in the city of Boston. City officials describe service response time as roughlly depending on the type of request and the volume of requests.

2. We will model the service response time as a function of the type as well as the location of the request.

2 Background

2.1 Bayesian Gaussian Mixture Models

We will assume that our data is given by $Y = (\mathbf{x}_1, \ldots, \mathbf{x}_N)$, where $\mathbf{x}_n \in \mathbb{R}^M$ is a vector encoding

the reported location of the problem (in terms of longitude and latitude) and the request type. We will assume that Y has a pdf that is a mixture of K number of Gaussian components with parameters $\mu = \{\mu_k : 1 \leq k \leq K\}$ and $\Sigma = \{\Sigma_k : 1 \leq k \leq K\}$. Let $\pi \in \mathbb{R}^k$ denote the set of mixture coefficients for this model. Then, the likelihood of observing y_i, given θ, π is

$$L(\mathbf{x}_n) = p(\mathbf{x}_n|\mu, \Sigma, \pi) = \sum_{k=1}^{K} \pi_k \mathcal{N}(\mathbf{x}_n|\mu_k, \Sigma_k) \quad (1)$$

where π_k is the mixture coefficient for the k-th component.

Alternatively, we can model the component membership information as missing data. That is, we model the component membership indicators as a set of latent variables, $Z = (\mathbf{z}_1, \ldots, \mathbf{z}_N)$, where $\mathbf{z}_n \in \mathbb{R}^K$ is a binary vector encoding of the membership of the n-th data in the k-th component. That is,

$$z_{nk} = \begin{cases} 1, & \mathbf{x}_n \text{ belongs to the } k\text{-th component} \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

We will assume that the probability of observing $\mathbf{z}_n$, given the mixture coefficients, as

$$p(\mathbf{z}_n|\pi) = \prod_{k=1}^{K} \pi_k^{z_{nk}}. \quad (3)$$

Now, we can rewrite the likelihood of our data set Y as

$$L(Y) = \prod_{n=1}^{N} \prod_{k=1}^{K} \mathcal{N}(x_n|\mu_k, \Sigma_k^{-1})^{z_{nk}}. \quad (4)$$

Finally, a basic generative model for our mixture can

be specified by the following:

$$\pi \sim Dir(\alpha_0) \tag{5}$$

$$\Sigma_k^{-1} \sim Wish(W_0, \nu_0) \tag{6}$$

$$\mu_k | \Sigma_k^{-1} \sim \mathcal{N}(\eta_0, (\beta_0 \Sigma_k^{-1})^{-1}) \tag{7}$$

$$z_n | \pi \sim \prod_{k=1}^{K} \pi_k^{z_{nk}} \tag{8}$$

$$x_n | Z, \mu, \Sigma^{-1} \sim \prod_{k=1}^{K} \mathcal{N}(\mu_k, \Sigma_k^{-1})^{z_{nk}} \tag{9}$$

Note that the parameters of our model is $\theta = (\alpha_0, W_0, \nu_0, \eta_0, \beta_0)$.

2.2 Hamiltonian Monte Carlo within Gibbs Sampling

2.3 Variational Inference

3 Bayesian GMM with Demographic Information as Covariates