We have a dataset for the level of ozone pollution in Los Angeles in the year 1976. The dataset has been split into training and test datasets, namely $Xtraining$, $Xtest$, $Ytraining$ and $Ytest$. The X datasets contain the 90 predictors such as wind speed, humidity, pressure, etc. The Y datasets contain the response variables. Both X and Y datasets have been centered and normalised. Here we are fitting a linear model, $y = X * \beta + \epsilon$. In our model, y is the response

variable, X is a predictor matrix and β is the coefficient vector. ϵ is the error term normal distributed, $N(0, v)$ where v is an unknown variance hyperparameter. Each observation is assumed to be independent from each other. We would like to carry out Bayesian inference on β using Gibbs Sampling.

We are adopting the Generalised Double Pareto model whereby

$$\begin{aligned} \beta_j &\sim N(0, \sigma^2 \tau_j) \\ \tau_j &\sim \exp\left(\frac{\lambda_j^2}{2}\right) \\ \lambda_j &\sim \Gamma(\alpha, \nu) \\ \pi_0(v) &\propto \frac{1}{v} \end{aligned} \tag{1}$$

In our example, we will be assuming that $\alpha = 10$ and $\nu = 1$. We run the 10000 iterations and burn in 1000 of them.