# Small Problem 1: Bayesian Linear Regression

## Summary

**Given:**

A set of training points where and

Prior probability distribution over the and parameters of . Priors will be diagonal Gaussian for and Gamma for

**Find:**

Query 1: The posterior distribution of

**Metrics:**

Metric 1: Expected squared Euclidean distance between the predicted mean and the true mean , where the expectation is taken with respect to the posterior distribution.

Metric 2: Total variation distance between the computed posterior and the correct posterior over .

## Details

The file “problem-1-generator.R” contains R code to generate the true regression coefficients and the input training data. The model is

The file contains 500 training examples generated from a single run of the R code. There are four covariates generated uniformly from . The values of the variables that generated the data are

see “problem-1-prior.Sigma.csv”

see “problem-1-data.csv”

Queries/Metrics:

1. Let be the posterior distribution of the estimated weight vector. One metric is the expected squared error under this distribution.
2. We have provided samples generated from the true posterior distribution . We can estimate the total variation distance between the true distribution and your estimate using the samples generated by your estimated distribution: