Exploring Bayes: Bayesian Linear Regression

In this basic practical we will try and implement some of the ideas supplied in the first lecture. This is partly a warm-up and logistics exercise as we get used to the Theory Lecture/ Practical Lecture format. For those unfamiliar with the jargon then this will be a crash course.

1. Consider $y = x + \epsilon$ with $\epsilon \sim \mathcal{N}(\mu, \sigma^2)$ and $\mu = 0$, $\sigma = 1$. Simulate the output of this model.
2. Write down the likelihood $P(D|\alpha)$ for the linear model $y = \alpha x + \epsilon$ with noise (fix $\sigma = 1$ and $\mu = 0$).
3. Use your simulated data to return an α that maximizes this likelihood (fix $\sigma = 1$ and $\mu = 0$).
4. What priors should you consider?
5. What are conjugate priors?
6. If we fix $\sigma = 1$ and $\mu = 0$, what is the conjugate prior?
7. Take draws from original model 1 and estimate the maximum a-posteriori (MAP) value for α.
8. What is the posterior on α if its prior is conjugate?
9. What is the affect of conjugate prior hyperparameter choice on the ratio between the MAP estimate and the ML estimate as the amount of data goes large?
10. Re-use the posteriors as priors but now use these to make inferences with a new model (e.g. $y = 3x + \epsilon$). Compare the posteriors with the posteriors you'd have if you had simply used the naive priors in 6
11. More Advanced: What is the conjugate prior when σ and μ are not fixed as constants but are also uncertain?

This is Bayesian Linear Regression - you'll find it in lots of standard texts.