PhD Course: Advanced Signal Processing (02901)
Exercise: Introduction to Bayesian Inference.

by Mikkel N. Schmidt

Exercises for "Introduction to Bayesian Inference".
Questions marked ***Q** are optional.

1 Bayesian modeling

Suppose that a medical test for a disease is correct 95% percent of the time. If a tested patient has the disease, the test returns a positive result with prbability .95, and if a tested patient does not have the disease, a positive result is returned with probability .05. Suppose that only 0.1% of the population has that disease, so that a randomly selected patient has a 0.001 prior probability of having the disease.

Q 1.1 *If a random person takes the test, and the result is positive, what is the probabilty that she has the disease?*
- *Use your intuition to make a guess.*
- *Compute the probability (using Bayes' rule)*

Q 1.2 *Discussion: If someone suspects that she might have the disease, and takes the test with positive result, what is the probability that she has the disease?*

Conditional probability: Suppose you're on a game show, and you're given the choice of three doors: Behind one door is a car; behind the others, goats. You pick a door, say No. 1, and the host, who knows what's behind the doors, opens another door, say No. 3, which has a goat. He then says to you, "Do you want to pick door No. 2?"

Q 1.3 *Is it to your advantage to switch your choice?*

2 Ordinary linear regression

We consider the normal ordinary linear regression model, where the distribution of $\boldsymbol{y}$ given $\boldsymbol{X}$ is a normal whose mean is a linear function of $\boldsymbol{X}$:

$$\boldsymbol{y}|\boldsymbol{w}, \boldsymbol{X} \sim \mathcal{N}(\boldsymbol{y}|\boldsymbol{X}\boldsymbol{w}, \sigma^2\boldsymbol{I}), \tag{1}$$

where $\mathcal{N}(\cdot|\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes the multivariate normal density with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, $\boldsymbol{y}$ is an N-dimensional vector of observations, $\boldsymbol{X}$ is an $N \times K$ matrix of explanatory variables (K variables for each observation), $\boldsymbol{w}$ is a K-dimensional vector of weights, σ^2 is a noise variance parameters, and $\boldsymbol{I}$ is the $N \times N$ identity matrix.

Q 2.1 *Look up the definition of the multivariate normal density to get familiar with the notation in Eq. (1).*

We choose the convenient standard noninformative improper prior over the parameters

$$\boldsymbol{w}, \sigma^2 | \boldsymbol{X} \sim \sigma^{-2} \tag{2}$$

Q 2.2 *Use Bayes' rule to derive the posterior distribution of the parameters.*

Q 2.3 *Show that the conditional posterior distribution of $\boldsymbol{w}$ given σ^2 is given by*

$$\boldsymbol{w} | \sigma^2, y \sim \mathcal{N}(\boldsymbol{w} | \bar{\boldsymbol{\mu}}, \bar{\boldsymbol{\Sigma}}), \tag{3}$$

where

$$\bar{\boldsymbol{\mu}} = (\boldsymbol{X}^\top \boldsymbol{X})^{-1} \boldsymbol{X}^\top \boldsymbol{y}, \tag{4}$$

$$\bar{\boldsymbol{\Sigma}} = (\boldsymbol{X}^\top \boldsymbol{X})^{-1} \tag{5}$$

Q 2.4 *How can we generate random numbers from this normal distribution in* MATLAB*?*

***Q 2.5** *Show that the marginal posterior distribution of σ^2 is given by the following scaled-inverse-χ^2 distribution*

$$\sigma^2 | \boldsymbol{y} \sim \frac{p(\boldsymbol{w}, \sigma^2 | \boldsymbol{y})}{p(\boldsymbol{w} | \sigma^2, \boldsymbol{y})} = \mathrm{Inv}{-}\chi^2(\sigma^2 | N-K, \bar{s}^2), \tag{6}$$

where

$$\bar{s}^2 = \frac{1}{N-K} (\boldsymbol{y} - \boldsymbol{X}\bar{\boldsymbol{w}})^\top (\boldsymbol{y} - \boldsymbol{X}\bar{\boldsymbol{w}}) \tag{7}$$

Q 2.6 *How can we generate random numbers from this scaled-inverse-χ^2 distribution in* MATLAB*?*

Suppose now that we observe a new matrix of explanatory variables, $\tilde{\boldsymbol{X}}$, and wish to predict the corresponding observations, $\tilde{\boldsymbol{y}}$.

Q 2.7 *Write down the expression for the posterior predictive distribution, $\tilde{\boldsymbol{y}} | \boldsymbol{y}, \boldsymbol{X}, \tilde{\boldsymbol{X}}$ (as an integral over the parameters). How can we generate samples from this distribution?*

3 Graphical models

Q 3.1 *Draw a graphical model for the normal ordinary regression model.*

Q 3.2 *How would the graphical model change if*

- *We extend the model hierarchically to include hyper-priors?*
- *We extend the model to multi-dimensional outputs?*

4 Markov chain Monte Carlo

Q 4.1 *Gibbs sampling can be seen as a special case of Metropolis-Hasings, where the proposal distributions are given by the conditional distributions. What is the accept rate? (For simplicity, consider a two-dimensional problem where x_1 is updated and x_2 is fixed.)*

Q 4.2 *Which posterior conditional distributions are required to implement Gibbs sampling in the normal ordinary regression model?*

Q 4.3 *Implement a Gibbs sampler for normal ordinary linear regression in* Matlab

- *Examine the supplied data set and discuss what you expect from regression analysis. Make a crude guess about the values of the parameters.*
- *Apply the algorithm to the supplied data set and examine the posterior distribution of the parameters—does it agree with your guess?*
- *Generate samples from the posterior predictive distribution for the new inputs $\tilde{\boldsymbol{X}}$ and plot them.*

5 Variational inference

The following exercise can be used as mini-project for presentation at the end of the course.

***Q 5.1** *Implement a variational inference procedure for normal ordinary linear regression, and compare results with Gibbs sampling.*