732A73 Bayesian Learning

Computer lab 2

Johannes Hedström & Mikael Montén

STIMA
Department of Computer and Information Science
Linköpings universitet

2024-04-16

Contents

```
library(coda)
library(readxl)
```

```
# loading the data
lin <- read_excel("Linkoping2022.xlsx")

women <- read.table('WomenAtWork.dat', header=TRUE)
```

1 1 - Linear and polynomial regression

The dataset Linkoping2022.xlsx contains daily average temperatures (in degree Celcius) in Linköping over the course of the year 2022. Use the function read_xlsx(), which is included in the R package readxl (install.packages("readxl")), to import the dataset in R. The response variable is temp and the covariate time that you need to create yourself is defned by

$$time = \frac{\text{the number of days since the beginning of the year}}{365}$$

A Bayesian analysis of the following quadratic regression model is to be performed:

$$temp = \beta_0 + \beta_1 \cdot time + \beta_2 \cdot time^2 + \epsilon, \epsilon \overset{iid}{\sim} N(0, \sigma^2)$$

1.1 a

Use the conjugate prior for the linear regression model. The prior hyperparameters μ_0, Ω_0, v_0 and σ_0^2 shall be set to sensible values. Start with $\mu_0 = (0, 100, -100)T, \Omega_0 = 0.01 \cdot I_3, v_0 = 1$ and $\sigma_0^2 = 1$. Check if this prior agrees with your prior opinions by simulating draws from the joint prior of all parameters and for every draw compute the regression curve. This gives a collection of regression curves; one for each draw from the prior. Does the collection of curves look reasonable? If not, change the prior hyperparameters until the collection of prior regression curves agrees with your prior beliefs about the regression curve. [Hint: R package mvtnorm can be used and your $Inv \sim \chi^2$ simulator of random draws from Lab 1.]

1.2 b

1.3 c

1.4 d

2 2 - Posterior approximation for classification with logistic regressino

2.1 a

2.2 b

2.3 c