Bayesian learning Lab 2

Mohamed Ali

2023-05-10

Linear and polynomial regression

To answer this question a conjugate prior for the linear regression model will be used in which: The joint prior for $\beta \ and \ \sigma^2$:

$$\beta|\sigma^2 \sim \ \mathcal{N}(\mu_0, \sigma^2{\Omega_0}^{-1})$$
$$\sigma^2 \sim Inv - \chi^2(v_0, \sigma^2)$$

While the posterior:

$$\beta|\sigma^2, y \sim \ \mathcal{N}(\mu_n, \sigma^2{\Omega_n}^{-1})$$
$$\sigma^2 \sim Inv - \chi^2(v_n, {\sigma^2}_n)$$
$$where:$$
$$\mu_n = (X'X\Omega_0)^{-1} \ (X'X\hat{\beta} + \Omega_0\mu_0)$$
$$\Omega_n = X'X + \Omega_0$$
$$v_n = v_0 + n$$
$$v_n\sigma^2 = v_n\sigma^2 + (y'y + \mu_0'\Omega_0\mu_0 - \mu_n'\Omega_n\mu_n)$$

First we start by reading the files and then we Create the covariate_time vaiable as (the number of days sinvce the beginning of the year / 365)

```
#reading the files
df<-read_xlsx("Linkoping2022.xlsx")
#Creating the covariate_time vaiable as
#(the number of days sinvce the beginning of the year / 365)
a<- df$datetime
#begunning of the year
b<- '2022-01-01'
a<-format.POSIXlt(strptime(a,'%Y-%m-%d'))
b<-format.POSIXlt(strptime(b,'%Y-%m-%d'))
#time diff from the
x<-as.vector(difftime(a,b,units='days'))
df$cov_tm<-x/365
```

A

We assume to use a conjugate prior from the linear regression in Lec 5 , we have been given the prior hyperparamteres as follow:

```
#We assume to use a conjugate prior from
#the linear regression in Lec 5 ,
#we have been given the prior hyperparamteres as follow:
mu_0= as.matrix(c(0,100,-100),ncol=3)
omega_0=0.01*diag(3)
v_0=1
segma2_0=1
n= length(df$temp)
ndraws=10
```

We have the joint prior for beta and segma2 defined as $\beta|\sigma^2 \sim \mathcal{N}(\mu_0, \sigma^2\Omega_0^{-1})$ and $\sigma^2 \sim Inv - \chi^2(v_o, \sigma_0^2)$ follows Inv-Chi(v0,sigma2_0)

First we draw our random samaple from inv-chi2 using the below defined function from Lec 3 slide 5

Now we estimate the betas values using the formula $\beta|\sigma^2 \sim \mathcal{N}(\mu_0, \sigma^2\Omega_0^{-1})$ and we fit the regression based ontemp = beta0 + beta1 time + beta 2 time^2 + erorr

*Note we have our error follows the normal distrbution by 0 and σ^2

After we done with the draws now for every draw compute the regression curve