Bayesian Linear Models

Andy Shen, Devin Francom

LANL: CCS-6

Task

GIT TEST

Let $\boldsymbol{y} = (y_1, \ldots, y_n)$ be the regression response and $\boldsymbol{X}$ be a $n \times p$ matrix of covariates. Here is the traditional linear model likelihood:

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim N(\mathbf{0}, \sigma^2 \boldsymbol{I})$$

The Bayesian version just needs priors for the unknowns, which are $\boldsymbol{\beta}$ and σ^2, and then you just "turn the Bayesian crank", which means you multiply the likelihood and priors, get the full conditionals, and sample the posterior with MCMC. Use $\boldsymbol{\beta} \sim N(\mathbf{0}, \tau^2 \boldsymbol{I})$ and $\sigma^2 \sim IG(a, b)$ as priors, and derive the following:

$$\begin{aligned}
p(\boldsymbol{\beta}, \sigma^2 | \boldsymbol{X}, \boldsymbol{y}) &\propto ?? \\
p(\boldsymbol{\beta} | \sigma^2, \boldsymbol{X}, \boldsymbol{y}) &\propto ?? \\
p(\sigma^2 | \boldsymbol{\beta}, \boldsymbol{X}, \boldsymbol{y}) &\propto ??
\end{aligned}$$

Hint: the full conditionals will be recognizable distributions (conjugate).

1 Posterior Distributions

Prior Distributions

We know that $\boldsymbol{\beta} \sim N(\mathbf{0}, \tau^2\mathbf{I})$ and $\sigma^2 \sim IG(a, b)$.

Moreover,

$$p(\sigma^2) = \frac{\beta^\alpha}{\Gamma(\alpha)}(\sigma^2)^{-\alpha-1} exp\bigg[\frac{\beta}{\sigma^2}\bigg]$$
$$\propto (\sigma^2)^{-\alpha-1} exp\bigg[\frac{\beta}{\sigma^2}\bigg]$$

and

$$p(\boldsymbol{\beta}) = (2\pi\tau^2)^{-\frac{p}{2}} exp\bigg[-\frac{1}{2\tau^2}\boldsymbol{\beta}'\boldsymbol{\beta}\bigg] \propto (\tau^2)^{-\frac{p}{2}} exp\bigg[-\frac{1}{2\tau^2}\boldsymbol{\beta}'\boldsymbol{\beta}\bigg]$$

Likelihood Function

We know that $y_i \sim N(0, \sigma^2)$ and $\boldsymbol{y} \sim N(\boldsymbol{X\beta}, \sigma^2\boldsymbol{I})$.

Then,

$$p(\boldsymbol{\beta} \mid \cdot) \propto N(\mathbf{0}, \tau^2\ \mathbf{I})\ L(\boldsymbol{\beta}, \sigma^2 \mid \boldsymbol{y})$$

and

$$p(\sigma^2 \mid \cdot) \propto IG(a, b)\ L(\boldsymbol{\beta}, \sigma^2 \mid \boldsymbol{y})$$

Leveraging the fact that $\mathbb{E}(\boldsymbol{y}) = \boldsymbol{X\beta}$, the derivation of the likelihood is as follows (shoutout to Stats 100C Lec 3, Spring 2020):

$$L(\boldsymbol{\beta}, \sigma^2|\boldsymbol{y}) = \prod_{i=1}^{N} p(y_i|\boldsymbol{\beta}, \sigma^2)$$

$$= \prod_{i=1}^{N}(2\pi\sigma^2) exp\bigg[-\frac{1}{2\sigma^2}(y_i - \mu_i)^2\bigg]$$

$$= (2\pi\sigma^2)^{-\frac{n}{2}} exp\bigg[-\frac{1}{2\sigma^2}\sum_{i=1}^{N}(y_i - \mu_i)^2\bigg]$$

$$\propto (\sigma^2)^{-\frac{n}{2}} exp\bigg[-\frac{1}{2\sigma^2}\sum_{i=1}^{N}(y_i - \mu_i)^2\bigg]$$

$$\propto (\sigma^2)^{-\frac{n}{2}} exp\bigg[-\frac{1}{2\sigma^2}(\boldsymbol{y} - \boldsymbol{X\beta})'(\boldsymbol{y} - \boldsymbol{X\beta})\bigg]$$

Full Conditionals

Multiplying the priors by the likelihood function, we get that:

$$p(\sigma^2 \mid \cdot) \propto (\sigma^2)^{-\alpha-1} exp\bigg[-\frac{b}{\sigma^2}\bigg](\sigma^2)^{-\frac{n}{2}} exp\bigg[-\frac{1}{2\sigma^2}(\boldsymbol{y}-\boldsymbol{X\beta})'(\boldsymbol{y}-\boldsymbol{X\beta})\bigg]$$

$$\propto (\sigma^2)^{-\alpha-1-\frac{n}{2}} exp\bigg[-\frac{b}{\sigma^2}-\frac{1}{2\sigma^2}(\boldsymbol{y}-\boldsymbol{X\beta})'(\boldsymbol{y}-\boldsymbol{X\beta})\bigg]$$

$$= (\sigma^2)^{-\alpha-1-\frac{n}{2}} exp\bigg[-\frac{1}{\sigma^2}\bigg(\frac{2b+(\boldsymbol{y}-\boldsymbol{X\beta})'(\boldsymbol{y}-\boldsymbol{X\beta})}{2}\bigg)\bigg]$$

and

$$p(\boldsymbol{\beta}|\cdot) \propto exp\bigg[-\frac{1}{2\tau^2}\boldsymbol{\beta}'\boldsymbol{\beta}-\frac{1}{2\sigma^2}(\boldsymbol{y}-\boldsymbol{X\beta})'(\boldsymbol{y}-\boldsymbol{X\beta})\bigg]$$

$$\propto exp\bigg[-\frac{1}{2\tau^2\sigma^2}\big[\sigma^2\boldsymbol{\beta}'\boldsymbol{\beta}+\tau^2(\boldsymbol{y}-\boldsymbol{X\beta})'(\boldsymbol{y}-\boldsymbol{X\beta})\big]\bigg]$$

$$\propto exp\bigg[-\frac{1}{2\tau^2\sigma^2}\big[\sigma^2\boldsymbol{\beta}'\boldsymbol{\beta}+\tau^2(\boldsymbol{y}'\boldsymbol{y}-2\boldsymbol{\beta}'\mathbf{X}'\boldsymbol{y}+\boldsymbol{\beta}'\mathbf{X}'\boldsymbol{X\beta})\big]\bigg]$$

and

$$p(\boldsymbol{\beta},\sigma^2 \mid \boldsymbol{y},\boldsymbol{X}) \propto p(\sigma^2)\ p(\boldsymbol{\beta})\ L(\boldsymbol{\beta},\sigma^2|\boldsymbol{y})$$

$$\propto (\sigma^2)^{-\alpha-1} exp\bigg[-\frac{b}{\sigma^2}\bigg](\tau^2)^{-\frac{p}{2}} exp\bigg[-\frac{1}{2\tau^2}\boldsymbol{\beta}'\boldsymbol{\beta}\bigg](\sigma^2)^{-\frac{n}{2}} exp\bigg[-\frac{1}{2\sigma^2}(\boldsymbol{y}-\boldsymbol{X\beta})'(\boldsymbol{y}-\boldsymbol{X\beta})\bigg]$$

$$\propto (\sigma^2)^{-(\frac{n}{2}+\alpha+1)}(\tau^2)^{-\frac{p}{2}} exp\bigg[-\frac{b}{\sigma^2}\bigg] exp\bigg[-\frac{1}{2\tau^2}\boldsymbol{\beta}'\boldsymbol{\beta}\bigg] exp\bigg[-\frac{1}{2\sigma^2}(\boldsymbol{y}-\boldsymbol{X\beta})'(\boldsymbol{y}-\boldsymbol{X\beta})\bigg]$$

$$p(\boldsymbol{\beta},\sigma^2 \mid \boldsymbol{y},\boldsymbol{X}) \propto (\sigma^2)^{-(\frac{n}{2}+\alpha+1)}(\tau^2)^{-\frac{p}{2}} exp\bigg[-\frac{b}{\sigma^2}\bigg] exp\bigg[-\frac{1}{2\tau^2}\boldsymbol{\beta}'\boldsymbol{\beta}\bigg] exp\bigg[-\frac{1}{2\sigma^2}(\boldsymbol{y}-\boldsymbol{X\beta})'(\boldsymbol{y}-\boldsymbol{X\beta})\bigg]$$

The joint posterior distribution is effectively an inverse gamma distribution multiplied by two multivariate normal distributions.

From these results, we get that:

$$\sigma^2 \mid \cdot \sim IG\bigg(\alpha+\frac{n}{2},\quad \frac{2b+(\boldsymbol{y}-\boldsymbol{X\beta})'(\boldsymbol{y}-\boldsymbol{X\beta})}{2}\bigg)$$

and

$$\boldsymbol{\beta} \mid \cdot \sim \mathcal{N}\bigg(\tau^2(\sigma^2\boldsymbol{I}+\tau^2\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y},\ (\sigma^2\boldsymbol{I}+\tau^2\boldsymbol{X}'\boldsymbol{X})^{-1}\bigg)$$

2 MCMC Sampler

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
## The following objects are masked from 'package:stats':
##
##     filter, lag
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
dat <- read.csv("data.csv")
dat2 <- dat[,-1]
y <- dat[,"y"]
X <- dat[,-c(1,2)] %>% as.matrix()
X <- cbind(1, X)
nrow(X) # n
```

```
## [1] 1000
```

```
ncol(X) # p = 11, p + 1 = 12
```

```
## [1] 12
```

Sigma

$$p(\sigma^2 \mid \cdot) \propto (\sigma^2)^{-\alpha-1-\frac{n}{2}} exp\Bigg[-\frac{1}{\sigma^2}\bigg(\frac{2b + (\boldsymbol{y} - \boldsymbol{X\beta})'(\boldsymbol{y} - \boldsymbol{X\beta})}{2}\bigg)\Bigg]$$

$$\sigma^2 \mid \cdot \sim IG\bigg(\alpha + \frac{n}{2}, \quad \frac{2b + (\boldsymbol{y} - \boldsymbol{X\beta})'(\boldsymbol{y} - \boldsymbol{X\beta})}{2}\bigg)$$

We code the full conditional of σ^2 as follows:

```
p_sig <- function(a = 1, b = 1, n = nrow(X), beta) {
  a_term <- a + (n/2)
  b_term <- 0.5 * (2*b + (t(y - (X %*% beta)) %*% (y - (X %*% beta)))) #y, X defined above
  1 / rgamma(1, shape = a_term, rate = 1/b_term)
}
```

Beta

The full conditional distribution of $\boldsymbol{\beta}$ is effectively the product of two multivariate normal distributions.

$$p(\boldsymbol{\beta}|\cdot) \propto exp\Bigg[-\frac{1}{2\tau^2}\boldsymbol{\beta}'\boldsymbol{\beta} - \frac{1}{2\sigma^2}(\boldsymbol{y} - \boldsymbol{X\beta})'(\boldsymbol{y} - \boldsymbol{X\beta})\Bigg]$$

$$\propto (\tau^2)^{-\frac{p}{2}} exp\Bigg[-\frac{1}{2\tau^2}\boldsymbol{\beta}'\boldsymbol{\beta}\Bigg](\sigma^2)^{-\frac{n}{2}} exp\Bigg[-\frac{1}{2\sigma^2}(\boldsymbol{y} - \boldsymbol{X\beta})'(\boldsymbol{y} - \boldsymbol{X\beta})\Bigg]$$

$$\propto exp((\boldsymbol{\beta} - \boldsymbol{\mu_\beta})' \ \boldsymbol{\Sigma_\beta^{-1}} \ (\boldsymbol{\beta} - \boldsymbol{\mu_\beta}))$$

After solving for the values of μ_β and $\boldsymbol{\Sigma}_\beta$, we get that

$$\boldsymbol{\beta} \mid \cdot \sim \mathcal{N}\Big(\tau^2(\sigma^2\boldsymbol{I} + \tau^2\boldsymbol{X'X})^{-1}\boldsymbol{X'y}, \ (\sigma^2\boldsymbol{I} + \tau^2\boldsymbol{X'X})^{-1}\Big)$$

This is a p-variate normal distribution with mean $\tau^2(\sigma^2\boldsymbol{I} + \tau^2\boldsymbol{X'X})^{-1}\boldsymbol{X'y}$ and covaraince matrix $(\sigma^2\boldsymbol{I} + \tau^2\boldsymbol{X'X})^{-1}$.

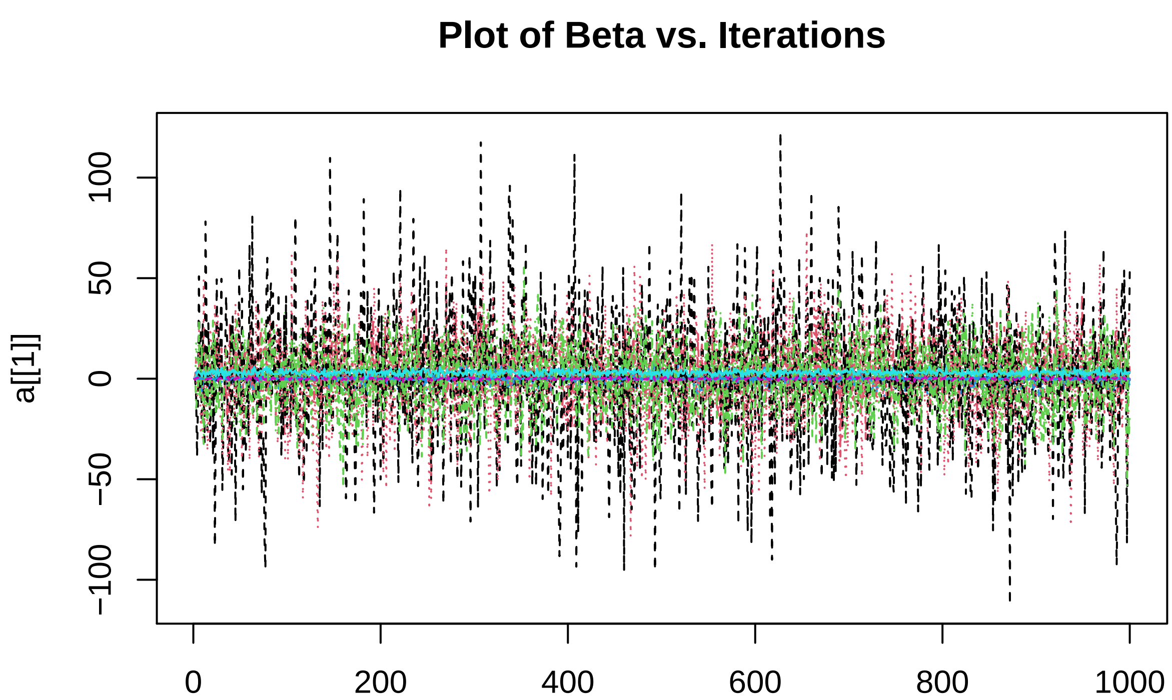
The code for the full conditional distribution of $\boldsymbol{\beta}$ is as follows:

```
library(mvtnorm)
p_beta <- function(sig_sq, tau_sq = 1, p = ncol(X)) {
  sig <- solve( (sig_sq * diag(p) ) + (tau_sq * t(X) %*% X))
  mu <- tau_sq * (sig) %*% t(X) %*% y
  #browser()
  rmvnorm(1, mean = mu, sigma = sig)
}
```

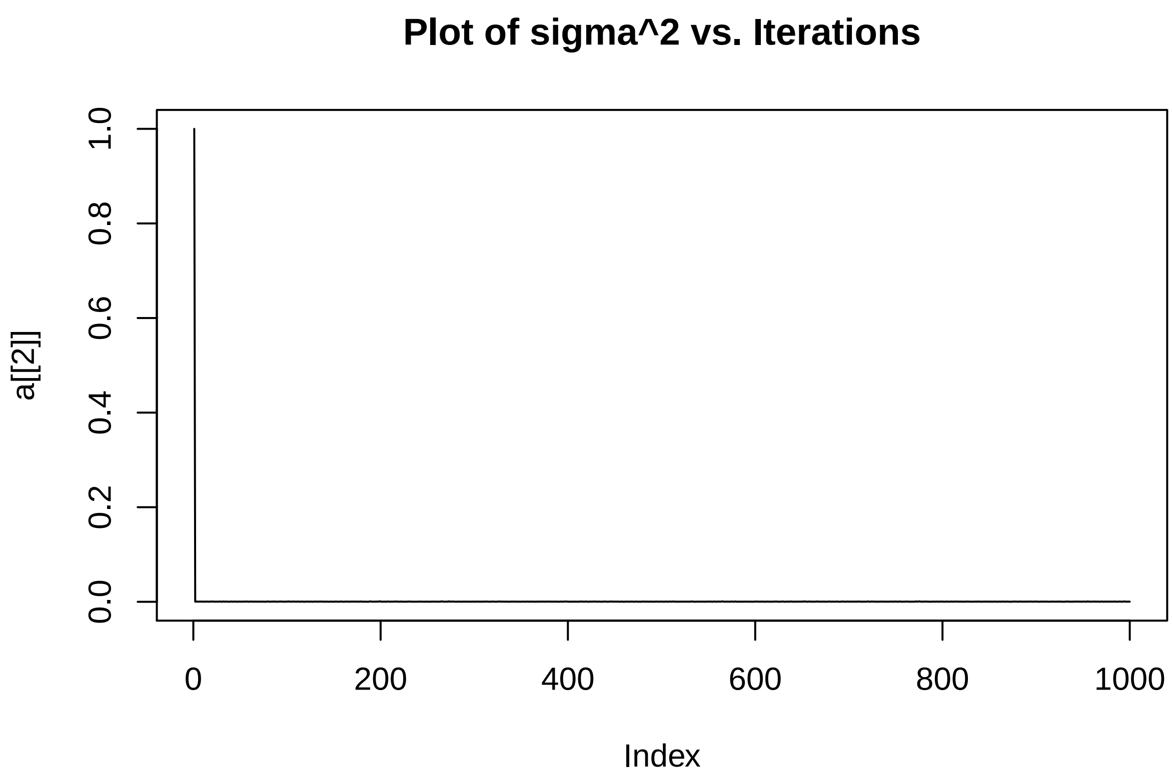
Gibbs Sampler

```
set.seed(12)
gibbs <- function(its, p = ncol(X)) {
  mat_beta <- matrix(NA, its, p)
  mat_sig <- rep(NA, its)
  mat_sig[1] <- 1
  mat_beta[1,] <- rep(0, p)
  for(i in 2:its) {
    mat_beta[i,] <- p_beta(sig_sq = mat_sig[i-1])
    #browser()
    mat_sig[i] <- p_sig(beta = mat_beta[i,])
  }
  list(mat_beta, mat_sig)
}
its <- 1000
a <- gibbs(its = its)
```

```
matplot(a[[1]], type = "l", main = "Plot of Beta vs. Iterations")
```




```r
plot(a[[2]], type = "l", main = "Plot of sigma^2 vs. Iterations")
```




```r
#matplot(a[[1]][,c(2,3,11)], type = "l")
tail(a[[2]])
```

```
## [1] 0.0004697128 0.0003748579 0.0002217820 0.0002698336 0.0003628946
## [6] 0.0002579338
```

We know that:

$$s^2 = \frac{S(\hat{\boldsymbol{\beta}})}{n - p - 1}$$

```
mod <- lm(y ~ ., data = dat2)
res <- mod$residuals
S_beta <- t(res) %*% res
s_sq <- S_beta / (nrow(X) - (ncol(X) - 1) - 1) #or just 1000 - 12
s_sq #same ting from summary table
```

```
##              [,1]
## [1,] 0.0001086339
```

```
summary(mod)$sigma^2 #from summary table
```

```
## [1] 0.0001086339
```

```
mean(a[[2]])
```

```
## [1] 0.00134174
```

```
colMeans(a[[1]])
```

```
##  [1]  0.011452345  4.593248506  4.270065949 -0.062665978 -0.002023421
##  [6] -0.001355014  1.234894935 -0.734097955 -0.281359595  0.011302019
## [11]  3.044805609  0.020362648
```

```
mod$coefficients
```

```
##   (Intercept)             a             b             c           x_n
## -2.436116e-03  5.048074e+00  4.054884e+00 -4.004919e-02  2.562824e-04
##           x_m          vel1          vel2          vel3            G1
## -3.616514e-06  7.584315e-01 -3.210139e-01 -5.395596e-02  4.487446e-03
##        delta2        delta3
##  3.003534e+00 -4.994286e-03
```