Homework 8

Chris Cameron

11/17/2021

Import data & monte-carlo library

```
swim = read.table('data/swim.dat')
```

```
## Warning in read.table("data/swim.dat"): incomplete final line found by
## readTableHeader on 'data/swim.dat'
```

```
library(MCMCpack)
```

```
## Loading required package: coda

## Loading required package: MASS

## ##
## ## Markov Chain Monte Carlo Package (MCMCpack)

## ## Copyright (C) 2003-2021 Andrew D. Martin, Kevin M. Quinn, and Jong Hee Park

## ##
## ## Support provided by the U.S. National Science Foundation

## ## (Grants SES-0350646 and SES-0350613)
## ##
```

```
library(pracma) #For matrix inversion
```

```
##
## Attaching package: 'pracma'

## The following object is masked from 'package:MCMCpack':
##
##     procrustes
```

1c)

Set the seed, priors, design matrix and number of samples for Gibbs sampler:

```r
set.seed(89780)

a = 0.1
b = 0.1

beta0 = c(23, 0)
sigma0 = rbind(c(5,0), c(0,2))

#Design Matrix
X = cbind(rep(1, 6), seq(1, 11, by = 2))

#Number of samples
num_samples = 5000
```

Run the Gibbs sampler

```r
swim_pred = apply(swim, MARGIN = 1, function(y) { #Row Wise

  BETA = matrix(nrow = num_samples, ncol = length(beta0))
  TAU = numeric(num_samples)

  # Initial values
  beta = beta0
  tau = 0.2

  for (s in 1:num_samples) {

    Sigma.n <- inv(inv(sigma0) + tau * (t(X) %*% X))
    Beta.n <- Sigma.n %*% (inv(sigma0) %*% beta0 + tau * (t(X) %*% y))

    beta = mvrnorm(1, Beta.n, Sigma.n)

    SSR = (t(y) %*% y) - (2 * t(beta) %*% t(X) %*% y) + (t(beta) %*% t(X) %*% X %*% beta)

    tau = rgamma(1, a + 3, (b + SSR) / 2)

    BETA[s,] = beta
    TAU[s] = tau
  }

  return(list(BETA, TAU))
})
```

1d)

```
BETA = purrr::map(swim_pred, 1) #Extract beta values for all swimmers
TAU = purrr::map(swim_pred, 2) #Above, but for Tau

pred_x = c(1, 13)

pred_time = matrix(nrow = num_samples,ncol = nrow(swim))

for (i in 1:nrow(swim)) {
  pred_y = rnorm(num_samples, BETA[[i]] %*% pred_x, 1/sqrt(TAU[[i]]))
  pred_time[,i] = pred_y
}
```

```
slowest = apply(pred_time, MARGIN = 1, which.max)
table(slowest) / length(slowest)
```

```
## slowest
##      1      2      3      4
## 0.0064 0.6804 0.0276 0.2856
```

When computing the probability that any one of the four swimmers has the highest time, swimmer 1 has the lowest predicted probability of having the highest time, at less than 1%. As such, based on these probabilities the coach should select swimmer 1. To see whether swimmer 1 is expected to be better overall or is simply more consistent than their fellow swimmers, we can also check the probability of having the lowest time.

```
fastest = apply(pred_time, MARGIN = 1, which.min)
table(fastest) / length(fastest)
```

```
## fastest
##      1      2      3      4
## 0.7516 0.0118 0.2186 0.0180
```

Checking the probability that a given swimmer only further confirms that the coach should pick swimmer 1, as they have a predicted probability of over .75 of having the lowest time.

Multivariate Methods

Given:

$$y_i \mid \theta, \Sigma \overset{i.i.d.}{\sim} MVN(\theta_{d\times 1}, \Sigma_{d\times d}), \quad i = 1, \cdots, n,$$

and the independent priors

$$\theta_{d\times 1} \sim MVN(\mu_{d\times 1}, T_{d\times d}), \qquad \Sigma_{d\times d} \sim \text{inverseWishart}(\nu, \Psi^{-1}_{d\times d}).$$

a) Show that $(\theta^T T^{-1}\mu)^T = \mu^T T^{-1}\theta$.

$$(\theta^T T^{-1}\mu)^T = \mu^T (T^{-1})^T \theta$$
$$\mu^T (T^{-1})^T \theta = \mu^T (T^T)^{-1}\theta$$

Since T is a covariance matrix, $T = T^T$

$$\mu^T (T^T)^{-1}\theta = \mu^T (T)^{-1}\theta$$

b) Use (a) to show that $p(\theta) \propto e^{-\frac{1}{2}(\theta^T T^{-1}\theta - 2\theta^T T^{-1}\mu)}$

$$p(\theta) \propto det(T)^{-\frac{n}{2}} e^{-\frac{1}{2}\sum_i (\theta_i-\mu)^T T^{-1}(\theta_i-\mu)}$$
$$p(\theta) \propto e^{-\frac{1}{2}\sum_i (\theta_i-\mu)^T T^{-1}(\theta_i-\mu)}$$
$$p(\theta) \propto e^{-\frac{1}{2}(\theta-\mu)^T T^{-1}(\theta-\mu)}$$
$$p(\theta) \propto e^{-\frac{1}{2}(\theta^T-\mu^T) T^{-1}(\theta-\mu)}$$
$$p(\theta) \propto e^{-\frac{1}{2}((\theta^T-\mu^T)(T^{-1}\theta - T^{-1}\mu))}$$
$$p(\theta) \propto e^{-\frac{1}{2}(\theta^T T^{-1}\theta - \theta^T T^{-1}\mu - \mu^T T^{-1}\theta + \mu^T T^{-1}\mu)}$$
$$p(\theta) \propto e^{-\frac{1}{2}(\theta^T T^{-1}\theta)} e^{-\frac{1}{2}(-\theta^T T^{-1}\mu)} e^{-\frac{1}{2}(\mu^T T^{-1}\theta)} e^{-\frac{1}{2}(\mu^T T^{-1}\mu)}$$
$$p(\theta) \propto e^{-\frac{1}{2}(\theta^T T^{-1}\theta)} e^{-\frac{1}{2}(-\theta^T T^{-1}\mu)} e^{-\frac{1}{2}(\mu^T T^{-1}\theta)}$$
$$p(\theta) \propto e^{-\frac{1}{2}(\theta^T T^{-1}\theta - \theta^T T^{-1}\mu - \mu^T T^{-1}\theta)}$$
$$p(\theta) \propto e^{-\frac{1}{2}(\theta^T T^{-1}\theta - 2\theta^T T^{-1}\mu)} \quad (part\ a)$$

c) Use (b) to show that:

$$p(\theta \mid \Sigma, y) \sim MVN\left\{\mu^*(\Sigma) := T^*(n\boldsymbol{\Sigma}^{-1}\bar{y} + T^{-1}\mu), T^* := (n\boldsymbol{\Sigma}^{-1} + T^{-1})^{-1}\right\}.$$

$$p(y|\theta,\Sigma) = \prod_{i=1}^{n} p(y_i|\theta,\Sigma) \sim \prod_{i=1}^{n} MVN(\theta,\Sigma)$$
$$p(y|\theta,\Sigma) = \prod_{i=1}^{n} (2\pi)^{\frac{-p}{2}} det(\Sigma)^{\frac{-1}{2}} e^{-\frac{1}{2}(y_i-\theta)^T\Sigma^{-1}(y_i-\theta)}$$

$$\propto e^{-\frac{1}{2}\sum_i \left((y_i-\theta)^T\Sigma^{-1}(y_i-\theta)\right)}$$

$$\propto e^{-\frac{1}{2}\left(\sum_i y_i^T\Sigma^{-1}y_i - 2\sum_i \theta_i^T\Sigma^{-1}y_i + \Sigma_i\theta^T\Sigma^{-1}\theta\right)}$$

$$\propto e^{-\frac{1}{2}(-2\theta^T\Sigma^{-1}n\overline{y}+n\theta^T\Sigma^{-1}\theta)}$$

$$p(\theta|\Sigma, y) \propto p(y|\theta,\Sigma)p(\theta)$$

$$\propto e^{-\frac{1}{2}(-2\theta^T\Sigma^{-1}n\overline{y}+n\theta^T\Sigma^{-1}\theta)}e^{-\frac{1}{2}(\theta^T T^{-1}\theta - 2\theta^T T^{-1}\mu)}$$

$$\propto e^{-\frac{1}{2}(-2\theta^T\Sigma^{-1}n\overline{y}+n\theta^T\Sigma^{-1}\theta+\theta^T T^{-1}\theta-2\theta^T T^{-1}\mu)}$$

$$\propto e^{\theta^T(\Sigma^{-1}n\overline{y}+T^{-1}\mu)-\frac{1}{2}\theta^T(n\Sigma^{-1}T^{-1})\theta}$$

$$\propto e^{\theta^T(n\Sigma^{-1}\overline{y}+T^{-1}\mu)-\frac{1}{2}\theta^T(n\Sigma^{-1}T^{-1})\theta}$$

$$p(\theta|\Sigma, y) \sim MVN(n\Sigma^{-1}T^{-1})^{-1}(n\Sigma^{-1}\overline{y}+T^{-1}\mu), (n\Sigma^{-1}T^{-1})^{-1})$$

d) Show that:

$$\operatorname{tr}(\Psi\boldsymbol{\Sigma}^{-1})+\sum_{i=1}^{n}(y_i-\boldsymbol{\theta})^T\boldsymbol{\Sigma}^{-1}(y_i-\boldsymbol{\theta}) = \operatorname{tr}\left\{\left(\Psi+\sum_{i=1}^{n}(y_i-\boldsymbol{\theta})(y_i-\boldsymbol{\theta})^T\right)\boldsymbol{\Sigma}^{-1}\right\}.$$

$$\operatorname{tr}(\Psi\boldsymbol{\Sigma}^{-1})+\sum_{i=1}^{n}(y_i-\boldsymbol{\theta})^T\boldsymbol{\Sigma}^{-1}(y_i-\boldsymbol{\theta}) = \operatorname{tr}(\Psi\boldsymbol{\Sigma}^{-1})+\sum_{i=1}^{n}\operatorname{tr}((y_i-\boldsymbol{\theta})^T\boldsymbol{\Sigma}^{-1}(y_i-\boldsymbol{\theta}))$$

$$\operatorname{tr}(\Psi\boldsymbol{\Sigma}^{-1})+\sum_{i=1}^{n}\operatorname{tr}((y_i-\boldsymbol{\theta})^T(y_i-\boldsymbol{\theta})\Sigma^{-1}) \quad \text{Lemma 2}$$

$$\operatorname{tr}(\Psi\boldsymbol{\Sigma}^{-1})+\sum_{i=1}^{n}\operatorname{tr}((y_i-\boldsymbol{\theta})^T(y_i-\boldsymbol{\theta})\Sigma^{-1})$$

$$\operatorname{tr}(\Psi\boldsymbol{\Sigma}^{-1})+\operatorname{tr}\left(\sum_{i=1}^{n}(y_i-\boldsymbol{\theta})^T(y_i-\boldsymbol{\theta})\Sigma^{-1}\right)$$

$$\operatorname{tr}\left(\Psi\boldsymbol{\Sigma}^{-1}+\sum_{i=1}^{n}(y_i-\boldsymbol{\theta})^T(y_i-\boldsymbol{\theta})\Sigma^{-1}\right)$$

$$\operatorname{tr}\left(\left(\Psi+\sum_{i=1}^{n}(y_i-\boldsymbol{\theta})^T(y_i-\boldsymbol{\theta})\right)\Sigma^{-1}\right)$$

e) Use d to show that $[\Sigma|\theta, \text{data}] \sim IW(\nu^*, \Psi^*(\theta))$ where:

$$\nu^* = n + v \text{ and } \Psi^*(\theta) = (\Psi + \sum_{i=1}^{n}(y_i - \boldsymbol{\theta})(y_i - \boldsymbol{\theta})^T)^{-1}.$$

$$- -$$

$$[\Sigma|\theta, data] \sim p(\Sigma|\theta, y)$$

$$p(y|\theta,\Sigma) \propto det(\Sigma)^{-\frac{n}{2}} e^{-\Sigma_i (y_i-\theta)^T \Sigma^{-1}(y_i-\theta)/2}$$

$$\propto det(\Sigma)^{-\frac{n}{2}} e^{-\Sigma_i (y_i-\theta)^T (y_i-\theta)\Sigma^{-1}/2}$$

$$\propto det(\Sigma)^{-\frac{n}{2}} e^{-tr(\Sigma_i (y_i-\theta)^T (y_i-\theta)\Sigma^{-1}/2)}$$

$$p(\Sigma|\theta, y) = p(\Sigma)p(y|\theta,\Sigma)$$

$$\propto det(\Sigma)^{-(v+p+1)/2} e^{-tr(\psi^{-1}\Sigma^{-1})/2} * det(\Sigma)^{-\frac{n}{2}} e^{-tr(\Sigma_i (y_i-\theta)^T (y_i-\theta)\Sigma^{-1})/2}$$

$$\propto det(\Sigma)^{-(v+p+1+n)/2} e^{-tr(\psi^{-1}\Sigma^{-1})/2 - tr(\Sigma_i (y_i-\theta)^T (y_i-\theta)\Sigma^{-1})/2}$$

$$\propto det(\Sigma)^{-(v+p+1+n)/2} e^{-tr((\psi + \Sigma_i (y_i-\theta)(y_i-\theta)^T)\Sigma^{-1})/2} \quad \text{(part d)}$$

$$p(\Sigma|\theta, y) \sim IW(n+v, (\psi + \Sigma_{i=1}^{n}(y_i-\theta)(y_i-\theta)^T)^{-1}$$

$$[\Sigma|\theta, data] \sim IW(n+v, (\psi + \Sigma_{i=1}^{n}(y_i-\theta)(y_i-\theta)^T)^{-1}$$