

STA 602. HW09

Yicheng Shen

11/5/2022

[1] PH 7.1

- (a) If we look into the prior distributions for θ and Σ separately, we can notice that the marginal density with respect to θ is just a constant, whose integral over the real line is infinite. If we take the integral and it does not integrate to 1, it is not a valid probability distribution.
- (b) Full conditional posteriors are shown below:

Joint Posterior: $p_J(\theta, \Sigma | y_1, \dots, y_n) \propto p_J(\theta, \Sigma) p_J(y_1, \dots, y_n | \theta, \Sigma)$

$$\propto |\Sigma|^{-(p+2)/2} |\Sigma|^{-n/2} \exp\left[-\frac{1}{2} \sum_{i=1}^n (y_i - \theta)^\top \Sigma^{-1} (y_i - \theta)\right]$$

$$\begin{aligned} p_J(\theta | \Sigma, y_1, \dots, y_n) &\propto \exp\left[-\frac{1}{2} \sum_{i=1}^n (y_i - \theta)^\top \Sigma^{-1} (y_i - \theta)\right] \\ &\propto \exp\left[-\frac{n}{2} (\bar{\mathbf{y}} - \theta)^\top \Sigma^{-1} (\bar{\mathbf{y}} - \theta)\right] \\ &\sim N(\bar{\mathbf{y}}, \Sigma/n) \end{aligned}$$

$$\begin{aligned} p_J(\Sigma | \theta, y_1, \dots, y_n) &\propto |\Sigma|^{-(p+2+n)/2} \exp\left[-\frac{1}{2} \sum_{i=1}^n (y_i - \theta)^\top \Sigma^{-1} (y_i - \theta)\right] \\ &\sim \text{Inverse-Wishart}(n+1, S_\theta) \end{aligned}$$

$$\begin{aligned} p_J(\Sigma | y_1, \dots, y_n) &\propto |\Sigma|^{-(p+2+n)/2} \int \exp\left[-\frac{1}{2} \sum_{i=1}^n (y_i - \theta)^\top \Sigma^{-1} (y_i - \theta)\right] d\theta \\ &\sim \text{Inverse-Wishart}\left(n, \sum_{i=1}^n (y_i y_i^\top - \bar{y} \bar{y}^\top)\right) \end{aligned}$$

[2] PH 7.2

(a) Derive the resulting log likelihood:

$$\begin{aligned}
l(\theta, \Psi | \mathbf{Y}) &= \log p(y_1 \cdots y_n | \theta, \Psi) \\
&= \sum_{i=1}^n \log \{ (2\pi)^{-p/2} |\Psi|^{1/2} \exp[-\frac{1}{2} (y_i - \theta)^\top \Psi (y_i - \theta)] \} \\
&= C + \frac{n}{2} \log(\Psi) + \sum_{i=1}^n [-\frac{1}{2} (y_i - \theta)^\top \Psi (y_i - \theta)] \\
&= C + \frac{n}{2} \log(\Psi) - \frac{1}{2} \sum_{i=1}^n \text{Tr}[(y_i - \theta)(y_i - \theta)^\top \Psi] \\
&= C + \frac{n}{2} \log(\Psi) - \frac{1}{2} \sum_{i=1}^n \text{Tr}[(y_i - \bar{y} + \bar{y} - \theta)(y_i - \bar{y} + \bar{y} - \theta)^\top \Psi] \\
&= C + \frac{n}{2} \log(\Psi) - \frac{1}{2} \text{Tr}[\sum_{i=1}^n (y_i - \bar{y})(y_i - \bar{y})^\top \Psi] - \frac{1}{2} \text{Tr}[(\bar{y} - \theta)(\bar{y} - \theta)^\top \Psi] \\
&\quad - \frac{1}{2} \text{Tr}[\sum_{i=1}^n (y_i - \bar{y})(\bar{y} - \theta)^\top \Psi] - \frac{1}{2} \text{Tr}[\sum_{i=1}^n (\bar{y} - \theta)(y_i - \bar{y})^\top \Psi] \\
&= C + \frac{n}{2} \log(\Psi) - \frac{1}{2} \text{Tr}[\sum_{i=1}^n (y_i - \bar{y})(y_i - \bar{y})^\top \Psi] - \frac{1}{2} \text{Tr}[(\bar{y} - \theta)(\bar{y} - \theta)^\top \Psi] \\
\text{Let } S &= \sum_{i=1}^n (y_i - \bar{y})(y_i - \bar{y})^\top \text{ by definition} \\
&= C + \frac{n}{2} \log(\Psi) - \frac{1}{2} \text{Tr}[S\Psi] - \frac{1}{2} \text{Tr}[(\bar{y} - \theta)(\bar{y} - \theta)^\top \Psi] \\
\log p(\theta, \Psi) &= l(\theta, \Psi | \mathbf{Y}) / n + c \\
p(\theta, \Psi) &\propto |\Psi|^{\frac{1}{2}} \exp\{-\frac{1}{2} [\text{Tr}(S\Psi) + \text{Tr}((\bar{y} - \theta)(\bar{y} - \theta)^\top \Psi)]\} \\
&\propto |\Psi|^{\frac{1}{2}} \exp\{-\frac{1}{2} \text{Tr}(S\Psi)\} \exp\{-\frac{1}{2} (\bar{y} - \theta)^\top \Psi (\bar{y} - \theta)\} \\
&\propto \text{IW}(p+1, S^{-1}) \times \text{MVN}(\bar{y}, \Psi^{-1}) \\
p_U(\theta, \Psi) &= p_U(\Psi) \times p_U(\theta | \Psi)
\end{aligned}$$

(b)

$$\begin{aligned}
p(\Psi) &= IW(p+1, S^{-1}); p(\Sigma) = \text{Wishart}(p+1, S^{-1}) \\
p(\theta, \Sigma | Y) &\propto p(\theta | \Sigma) \times p(\Sigma) \times p(y_1 \dots y_n | \theta, \Sigma) \\
&\propto |\Sigma|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\bar{y} - \theta)^\top \Sigma^{-1}(\bar{y} - \theta)\right\} \times |\Sigma|^{-\frac{p+1+p+1}{2}} \exp\left\{-\frac{1}{2}[\text{Tr}(S\Sigma^{-\frac{1}{2}})]\right\} \\
&\quad \times |\Sigma|^{-n/2} \exp\left[-\frac{1}{2} \sum_{i=1}^n (y_i - \theta)^\top \Sigma^{-1}(y_i - \theta)\right] \\
&= |\Sigma|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\bar{y} - \theta)^\top \Sigma^{-1}(\bar{y} - \theta)\right\} \times |\Sigma|^{-\frac{p+1+p+1}{2}} \exp\left\{-\frac{1}{2}[\text{Tr}(S\Sigma^{-\frac{1}{2}})]\right\} \\
\text{shown in (a).} \quad &\times |\Sigma|^{-n/2} \exp\left[-\frac{1}{2} \sum_{i=1}^n (y_i - \bar{y})^\top \Sigma^{-1}(y_i - \bar{y}) - \frac{1}{2} \sum_{i=1}^n (\bar{y} - \theta)^\top \Sigma^{-1}(\bar{y} - \theta)\right] \\
&= |\Sigma|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\bar{y} - \theta)^\top \Sigma^{-1}(\bar{y} - \theta)\right\} \times |\Sigma|^{-\frac{p+1+p+1}{2}} \exp\left\{-\frac{1}{2}[\text{Tr}(S\Sigma^{-\frac{1}{2}})]\right\} \\
&\quad \times |\Sigma|^{-n/2} \exp\left[-\frac{1}{2} \sum_{i=1}^n (y_i - \bar{y})^\top \Sigma^{-1}(y_i - \bar{y})\right] \exp\left[-\frac{1}{2}n(\bar{y} - \theta)^\top \Sigma^{-1}(\bar{y} - \theta)\right] \\
&= |\Sigma|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\bar{y} - \theta)^\top (1+n)\Sigma^{-1}(\bar{y} - \theta)\right\} \times |\Sigma|^{-\frac{p+1+p+1+n}{2}} \exp\left\{-\frac{1}{2}[\text{Tr}((1+n)S\Sigma^{-\frac{1}{2}})]\right\} \\
&\propto \text{Inverse-Wishart}(p+1+n, (1+n)S^{-1}) \times \text{MVN}(\bar{\mathbf{y}}, \frac{\Sigma}{n+1})
\end{aligned}$$

Therefore, we should be able to use this as the posterior distributions for Σ and θ . The first is Inverse-Wishart and the latter is multivariate normal, with parameter values specified above.

[3] PH 7.3

(a) First write out the full conditional posterior:

$$\begin{aligned}
\theta &\sim N(\boldsymbol{\mu}_0, \Lambda_0) \\
\Sigma^{-1} &\sim IW(\nu_0, S_0) \\
\theta | y_1 \dots y_n, \Sigma &\sim N(\boldsymbol{\mu}_n, \Lambda_n) \\
\text{where } \boldsymbol{\mu}_n &= \Lambda_n(\Lambda_0^{-1}\boldsymbol{\mu}_0 + n\Sigma^{-1}\bar{y}); \\
\Lambda_n^{-1} &= \Lambda_0^{-1} + n\Sigma^{-1} \\
\Sigma^{-1} | y_1 \dots y_n, \theta &\sim IW(\nu_0 + n, (S_0 + S_\theta)^{-1}) \\
\text{where } S_\theta &= \sum_{i=1}^n (\mathbf{y}_i - \theta)^\top (\mathbf{y}_i - \theta)
\end{aligned}$$

```

bluecrab <- read.table("http://www2.stat.duke.edu/~pdh10/FCBS/Exercises/bluecrab.dat")
orangecrab <- read.table("http://www2.stat.duke.edu/~pdh10/FCBS/Exercises/orangecrab.dat")
colnames(bluecrab) = colnames(orangecrab) = c("Y1", "Y2")
bluecrab <- as.matrix(bluecrab)
orangecrab <- as.matrix(orangecrab)

```

The Gibbs sampler was run separately for the two data sets below

```

p <- ncol(bluecrab)
n <- nrow(bluecrab)
ybar <- colMeans(bluecrab)
mu0 <- colMeans(bluecrab)
Lambda0 = S0 = cov(bluecrab)
nu0 <- 4
S <- 100000
# value store matrix
THETA_Blue = matrix(NA, S, p)
SIGMA_Blue = array(dim = c(p, p, S))

# initialize
theta = 10
Sigma = cov(bluecrab)

set.seed(8848)
for (t in 1:S) {
  # draw theta
  Lambdan = solve(solve(Lambda0) + n * solve(Sigma))
  mun = Lambdan %*% (solve(Lambda0) %*% mu0 + n * solve(Sigma) %*% ybar)
  theta = MASS::mvrnorm(1, mun, Lambdan)

  # draw Sigma
  Stheta = (t(bluecrab) - c(theta)) %*% t(t(bluecrab) - c(theta))
  Sigma = solve(rWishart(1, nu0 + n, solve(S0 + Stheta))[, , 1])

  THETA_Blue[t, ] = theta
  SIGMA_Blue[, , t] = Sigma
}

p <- ncol(orange crab)
n <- nrow(orange crab)
ybar <- colMeans(orange crab)
mu0 <- colMeans(orange crab)
Lambda0 = S0 = cov(orange crab)
nu0 <- 4

# value store matrix
THETA_Orange = matrix(NA, S, p)
SIGMA_Orange = array(dim = c(p, p, S))

# initialize
theta = 10
Sigma = cov(orange crab)

for (t in 1:S) {
  # draw theta
  Lambdan = solve(solve(Lambda0) + n * solve(Sigma))
  mun = Lambdan %*% (solve(Lambda0) %*% mu0 + n * solve(Sigma) %*% ybar)
  theta = MASS::mvrnorm(1, mun, Lambdan)

  # draw Sigma
  Stheta = (t(orange crab) - c(theta)) %*% t(t(orange crab) - c(theta))

```

```

Sigma = solve(rWishart(1, nu0 + n, solve(S0 + Stheta))[, , 1])

THETA_Orange[t, ] = theta
SIGMA_Orange[, , t] = Sigma
}

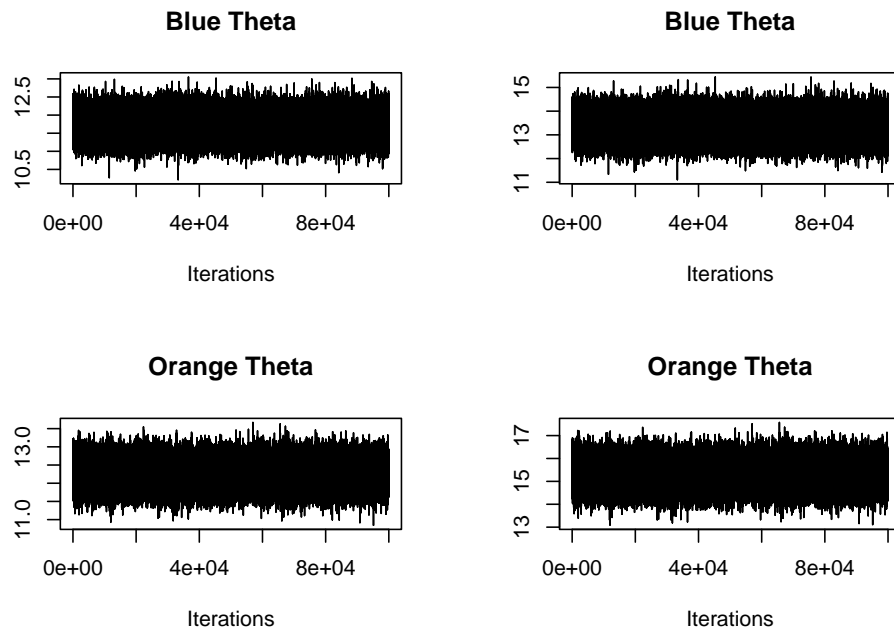
```

The MCMC diagnostics are well satisfied.

```

par(mfrow = c(2,2))
traceplot(mcmc(THETA_Blue), main = "Blue Theta")
traceplot(mcmc(THETA_Orange), main = "Orange Theta")

```



```

# acf(mcmc(THETA_Blue))
# acf(mcmc(THETA_Orange))

```

- (b) The plot shows that the both body depth and rear width of orange crabs are usually higher than the sizes of blue crabs, with orange crabs' rear width being almost surely greater than blue crabs. The spread of θ , seems roughly similar.

```

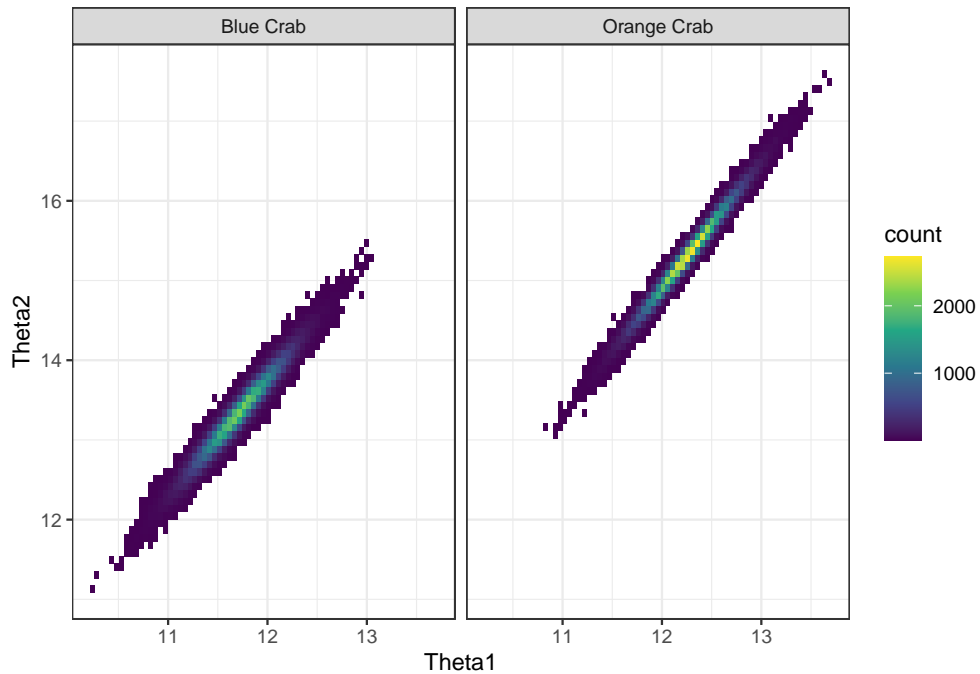
rbind(as_tibble(THETA_Blue) %>% mutate(Species = "Blue Crab"),
      as_tibble(THETA_Orange) %>% mutate(Species = "Orange Crab")) %>%
  ggplot(aes(V1, V2)) + labs(x = "Theta1", y = "Theta2") +
  facet_wrap(~ Species) + geom_bin2d(bins=70) +
  scale_fill_continuous(type = "viridis")

```

```

## Warning: The 'x' argument of 'as_tibble.matrix()' must have unique column names if
## '.name_repair' is omitted as of tibble 2.0.0.
## i Using compatibility '.name_repair'.

```



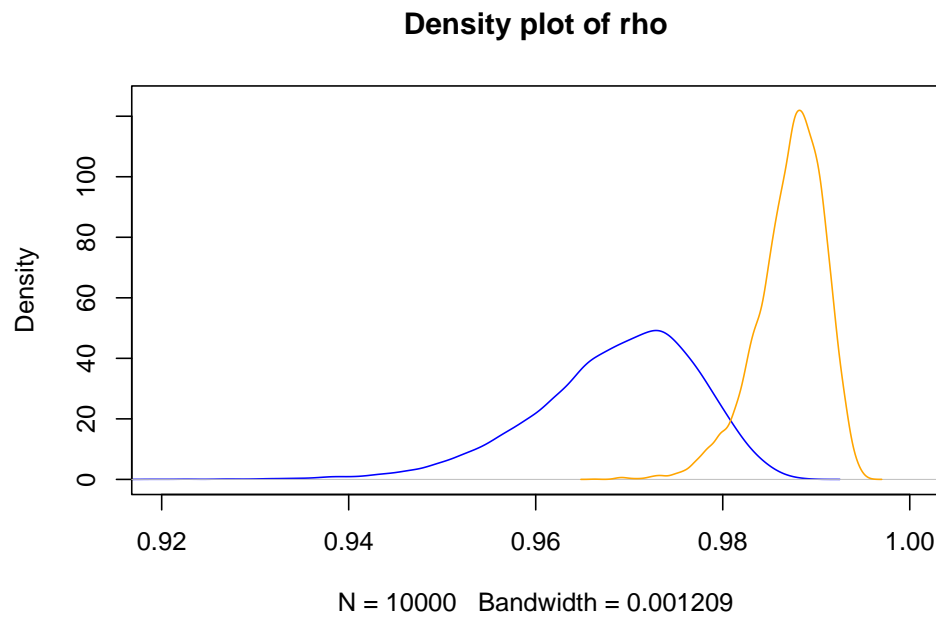
```
c(mean(THETA_Blue[,1]<THETA_Orange[,1]), mean(THETA_Blue[,2]<THETA_Orange[,2]))
```

```
## [1] 0.89995 0.99805
```

(c) The probability density is plotted below, with blue colored line for blue crabs' correlation and orange colored line for orange crabs' correlation.

```
get_corr <- function(COV_Matrix){
  COV_Matrix[1,2] / sqrt(COV_Matrix[1,1]) / sqrt(COV_Matrix[2,2])
}
rho_blue = rho_orange = c()
for (i in 1:10000)
{
  rho_blue = c(rho_blue, get_corr(SIGMA_Blue[, , i]))
  rho_orange = c(rho_orange, get_corr(SIGMA_Orange[, , i]))
}

plot(density(rho_blue), col = "blue", xlim = c(0.92,1), ylim = c(0,125),main = "Density plot of rho")
lines(density(rho_orange), col = "orange")
```



$P(\rho_{\text{blue}} < \rho_{\text{orange}} | \mathbf{y}_{\text{blue}}, \mathbf{y}_{\text{orange}})$ is approximately

```
mean(rho_blue < rho_orange)
```

```
## [1] 0.9899
```

The orange crabs seem to have much higher correlations between their body depth and rear width than blue ones