Assignment 6
EN553.632.02

Zhiyuan Wei
zwei26

7.1.

(a) Since $p_J(\theta,\Sigma) \propto |\Sigma|^{-(p+2)/2}$ is uniform with respect to θ, that is, the integral with respect to θ is infinite instead of 1. Therefore, p_J is not proper and cannot be a probability density for (θ,Σ).

(b)
$$p_J(\theta,\Sigma \mid y_1,y_2,\dots,y_n) \propto p_J(\theta,\Sigma) \times p(y_1,\dots,y_n \mid \theta,\Sigma)$$
$$\propto |\Sigma|^{-(p+2)/2}\,|\Sigma|^{-n/2} \exp(-tr(S_\theta\Sigma^{-1})/2)$$
$$\propto |\Sigma|^{-(n+p+2)/2} \exp(-tr(S_\theta\Sigma^{-1})/2)$$

$$p_J(\theta \mid y_1,y_2,\dots,y_n,\Sigma) \propto \exp(-tr(S_\theta\Sigma^{-1})/2)$$
$$\propto \exp(-\sum_i^n (y_i-\theta)^T\Sigma^{-1}(y_i-\theta)/2)$$
$$\propto \exp(-n(\bar{y}-\theta)^T\Sigma^{-1}(\bar{y}-\theta)/2)$$

Thus, $\theta \mid y_1,y_2,\dots,y_n,\Sigma \sim N(\bar{y},\Sigma/n)$

$$p_J(\Sigma \mid y_1,\dots,y_n,\theta) \propto |\Sigma|^{-(n+p+2)/2} \exp(-tr(S_\theta\Sigma^{-1})/2)$$

Thus, $\Sigma \mid y_1,\dots,y_n,\theta \sim$ Inverse-Wishart$(n+1, S_\theta^{-1})$
for $S_\theta = \sum_{i=1}^n (y_i-\theta)(y_i-\theta)^T$

7.4

```r
library(MASS)
library(ggplot2)
agehw = as.matrix(read.table('agehw.dat'))
colnames(agehw) = agehw[1, ]
agehw = agehw[-1, ]
agehw = matrix(as.numeric(agehw), nrow = 100)
```

a

I assume that most of the marriage couples are in the range of (20,70) and set $\mu_0 = (45, 45)^T$, and thus $\sigma^2 \approx 156$. We also assume a prior correlation of 0.75, so that $\sigma_{1,2} = 117$, and we have our covariance matrix.

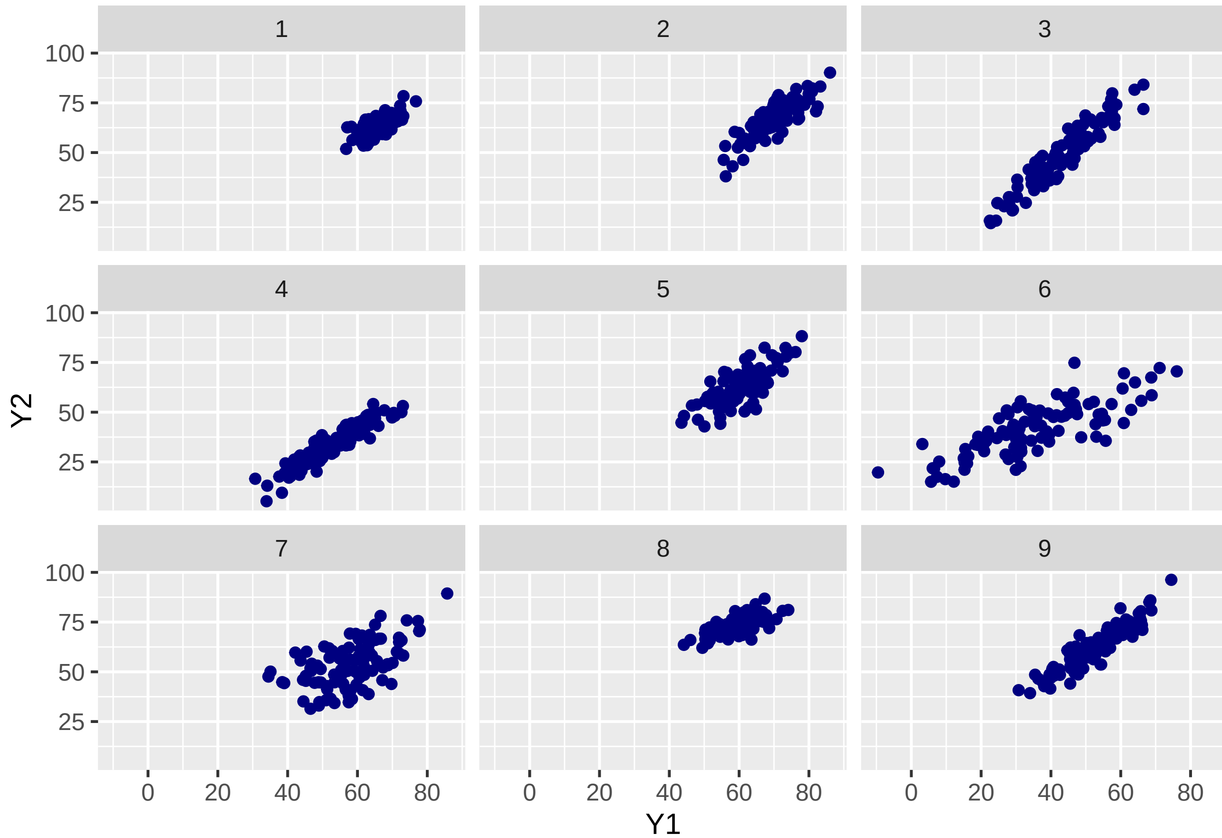
```r
Y = agehw
p = ncol(agehw)
n = nrow(agehw)
ybar = colMeans(agehw)

mu0 = rep(50, p)
d0 = s0 = rbind(c(156, 117), c(117, 156))
# nu0 = p + 2
nu0 = p + 2
s0
```

```
##      [,1] [,2]
## [1,]  156  117
## [2,]  117  156
```

b

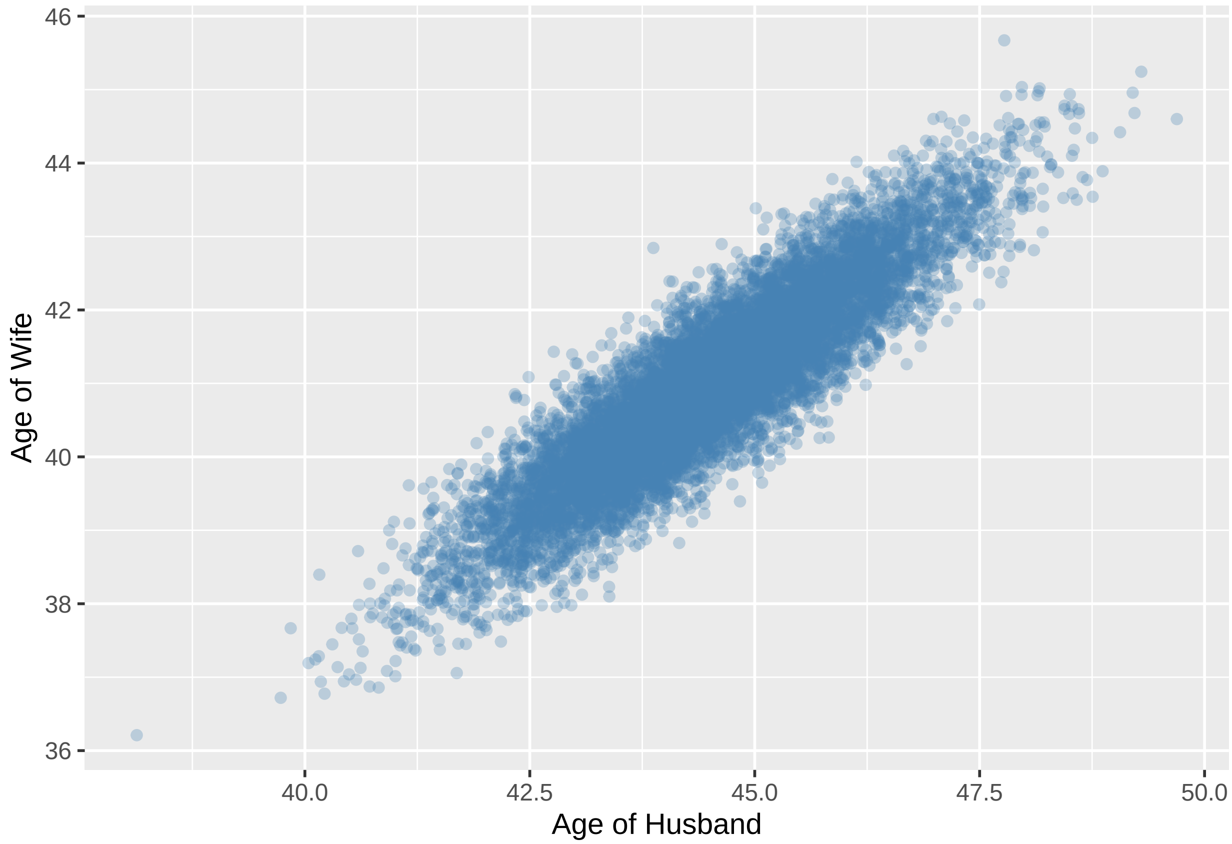
```r
set.seed(0)
N = 100
S = 9
Y_preds = lapply(1:S, function(s) {
  theta = mvrnorm(n = 1, mu0, s0)
  sigma = solve(rWishart(1, nu0, solve(s0))[, , 1])
  Y_s = mvrnorm(n = 100, theta, sigma)
  data.frame(Y1 = Y_s[, 1], Y2 = Y_s[, 2], dataset = s)})
data_pred = do.call(rbind, Y_preds)
ggplot(data_pred, aes(x = Y1, y = Y2)) +
  geom_point(col="navy") +
  facet_wrap(~ dataset)
```



Above are the scatter plots for the prior predictive datasets.

c

```
S = 10000
sigma = cov(Y)
MU = SIGMA = NULL
#Gibbs
for(s in 1:S){
  dn <- solve(solve(d0) + n * solve(sigma))
  mun <- dn %*% (solve(d0) %*%  mu0 + n * solve(sigma) %*% ybar)
  mu <- mvrnorm(1, mun, dn)
  sn <- s0 + (t(Y) - c(mu)) %*% t( t(Y) - c(mu))
  sigma <- solve( rWishart(1, nu0 + n, solve(sn))[,,1])
  MU <- rbind(MU, mu)
  SIGMA <- rbind(SIGMA, c(sigma))
}
#plot the joint posterior distribution of theta_h and theta_w
ggplot(as.data.frame(MU), aes(x = V1, y = V2))+
  geom_point(alpha = .3,color="steelblue")+
  xlab("Age of Husband")+
  ylab("Age of Wife")
```



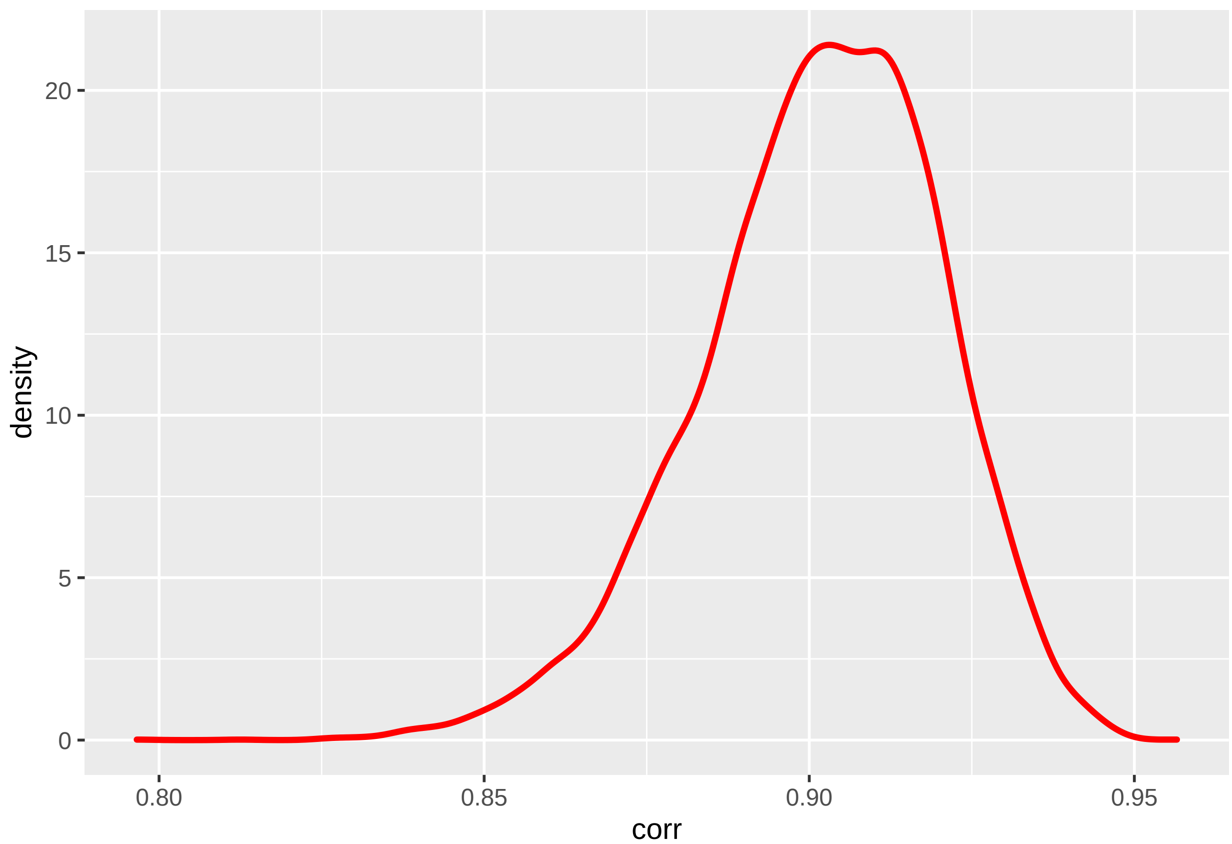

```r
#95% posterior confidence intervals for theta_h and theta_w
print(quantile(MU[,1], c(0.025, 0.5, 0.975)))
```

```
##     2.5%      50%    97.5%
## 41.89086 44.50357 47.18243
```

```r
print(quantile(MU[,2], c(0.025, 0.5, 0.975)))
```

```
##     2.5%      50%    97.5%
## 38.53235 40.98177 43.47976
```

```r
#plot the marginal posterior density of the correlation between y_h and y_w
covariance <- SIGMA[,2]
var_husband <- SIGMA[,1]
var_wife <- SIGMA[,4]
corr <- covariance/sqrt(var_husband*var_wife)
corr <- data.frame(corr)
ggplot(data = corr, aes(x = corr))+
  geom_density(color="red",size=1.1)
```




```
#95% confidence intervals for correlation coefficient
print(quantile(corr[,1], c(0.025, 0.5, 0.975)))
```

```
##      2.5%       50%     97.5%
## 0.8607462 0.9029377 0.9334090
```

The 95% posterior confidence intervals and the plot of joint posterior distribution of θ_h and θ_w, the 95% confidence intervals for correlation coefficient and the plot of marginal posterior density of the correlation between y_h and y_w are above.

d

```
THETA = SIGMA = NULL
sigma = cov(Y)

# Gibbs
for (s in 1:S) {
  theta = mvrnorm(n = 1, ybar, sigma / n)
  resid = t(Y) - c(theta)
  s_theta = resid %*% t(resid)
  sigma = solve(rWishart(1, n + 1, solve(s_theta))[, , 1])
  THETA <- rbind(THETA, theta)
  SIGMA <- rbind(SIGMA, c(sigma))
}
covariance <- SIGMA[,2]
var_husband <- SIGMA[,1]
```

```r
var_wife <- SIGMA[,4]
corr <- covariance/sqrt(var_husband*var_wife)
corr <- data.frame(corr)
#95% posterior confidence intervals for theta_h
print(quantile(THETA[,1], c(0.025, 0.5, 0.975)))
```

i

```
##     2.5%      50%    97.5%
## 41.75978 44.42659 47.12980
```

```r
#95% posterior confidence intervals for theta_w
print(quantile(THETA[,2], c(0.025, 0.5, 0.975)))
```

```
##     2.5%      50%    97.5%
## 38.36632 40.88826 43.38942
```

```r
#95% confidence intervals for correlation coefficient
print(quantile(corr[,1], c(0.025, 0.5, 0.975)))
```

```
##      2.5%       50%     97.5%
## 0.8607717 0.9045253 0.9345393
```

ii For the prior $\sum \sim InverseWishart(p+1, S^{-1})$ and $\theta|\sum \sim MVN(\bar{y}, \sum)$, we then have $\theta|\sum, Y \sim MVN(\bar{y}, \sum/(n+1))$ and $\sum|\theta, Y \sim InverseWishart(n+p+1, S^{-1}/(n+1))$ where $S = \sum(y_i - \bar{y})(y_i - \bar{y})^T/n$

```r
THETA = SIGMA = NULL
sigma = cov(Y)
for (s in 1:S) {
  theta = mvrnorm(n = 1, ybar, sigma/(n+1))
  s_new = (t(Y) - ybar)%*%t(t(Y) - ybar)/n
  sigma = solve(rWishart(1, n+p+1, solve(s_new)/(n+1))[, , 1])
  THETA <- rbind(THETA, theta)
  SIGMA <- rbind(SIGMA, c(sigma))
}
covariance <- SIGMA[,2]
var_husband <- SIGMA[,1]
var_wife <- SIGMA[,4]
corr <- covariance/sqrt(var_husband*var_wife)
corr <- data.frame(corr)
#95% posterior confidence intervals for theta_h
print(quantile(THETA[,1], c(0.025, 0.5, 0.975)))
```

```
##     2.5%      50%    97.5%
## 41.78017 44.41148 47.13814
```

```r
#95% posterior confidence intervals for theta_w
print(quantile(THETA[,2], c(0.025, 0.5, 0.975)))
```

```
##     2.5%      50%    97.5%
## 38.36956 40.87430 43.41063
```

```r
#95% confidence intervals for correlation coefficient
print(quantile(corr[,1], c(0.025, 0.5, 0.975)))
```

```
##      2.5%       50%     97.5%
## 0.8613381 0.9043220 0.9338502
```

```r
mu0=rep(0,p)
nu0=3
s0=1000*diag(p)
d0=10^5 * diag(p)
MU = SIGMA = NULL
sigma = cov(Y)
for(s in 1:S) {
  dn <- solve(solve(d0) + n * solve(sigma))
  mun <- dn %*% (solve(d0) %*%  mu0 + n * solve(sigma) %*% ybar)
  mu <- mvrnorm(1, mun, dn)
  sn <- s0 + (t(Y) - c(mu)) %*% t( t(Y) - c(mu))
  sigma <- solve( rWishart(1, nu0 + n, solve(sn))[,,1])
  MU <- rbind(MU, mu)
  SIGMA <- rbind(SIGMA, c(sigma))
}
covariance <- SIGMA[,2]
var_husband <- SIGMA[,1]
var_wife <- SIGMA[,4]
corr <- covariance/sqrt(var_husband*var_wife)
corr <- data.frame(corr)
#95% posterior confidence intervals for theta_h
print(quantile(MU[,1], c(0.025, 0.5, 0.975)))
```

iii

```
##     2.5%      50%    97.5%
## 41.72865 44.43030 47.22994
```

```r
#95% posterior confidence intervals for theta_w
print(quantile(MU[,2], c(0.025, 0.5, 0.975)))
```

```
##     2.5%      50%    97.5%
## 38.37369 40.90300 43.53078
```

```r
#95% confidence intervals for correlation coefficient
print(quantile(corr[,1], c(0.025, 0.5, 0.975)))
```

```
##      2.5%       50%     97.5%
## 0.7918482 0.8554712 0.8997530
```

e

We compare the CIs by four different priors, and see that our prior did a fairly well job since its intervals are narrow and the correlations are high. Therefore, I believe my prior is helpful information is helpful in estimating θ and $\sum$. On the other hand, the unit information prior has a slightly better result than my prior since its intervals are narrower. So it might be preferable. Since our sample size is 10000 which is very large, the results for all the priors are close. But if we use a much smaller sample size n=25, there will be a big difference of the intervals, especially for the diffuse prior since it has a very weak prior information, the correlation for the diffuse prior will be much lower than the other priors.