HW6

Zach White

10/26/2016

Exercise 7.4

Part A

There will be two parts of semi-conjugate prior for this model. There will be one prior distribution on the average age of each married husband and wife $\theta = (\theta_{husband}, \theta_{wife})^T$. Thus, this clearly shows that we need a multivariate normal distribution. As I think about this model, I think that the average age for males should be higher than that of women. Also, the range should be between 20-80. I propose that the average age of males should be around 45 while that of women should be around 42. I don't see any reason that the variance of age within husbands should be different than within wives. I think that $\sigma_h = \sigma_w = 15$ or at least somewhere near it. I propose that value. Also, it should be noted that they're aren't independent. So they should be related. I propose they are actually quite correlated. I propose that the correlation coefficient should be around .7. Thus

$$\boldsymbol{\theta} = \begin{bmatrix} \theta_h \\ \theta_w \end{bmatrix} \sim \text{Normal}\left(\begin{bmatrix} 45 \\ 42 \end{bmatrix}, \begin{bmatrix} 225 & 157.5 \\ 157.5 & 225 \end{bmatrix}^{-1} \right) = N(\boldsymbol{\mu}_0, \Lambda_0)$$

The second part of the prior distribution is the prior on Σ or Σ^{-1}. We propose an inverse-Wishart for the former and a Wishart for the latter. With the Wishart, we have two paramaters, ν_0, which represents prior sample size, and the second is $\boldsymbol{S_0}$, which represents the prior residual sum of squares. We will use our Λ_0 in our case. We propose $\nu_0 = 4$, since $\nu_0 = p + 2$, by definition.

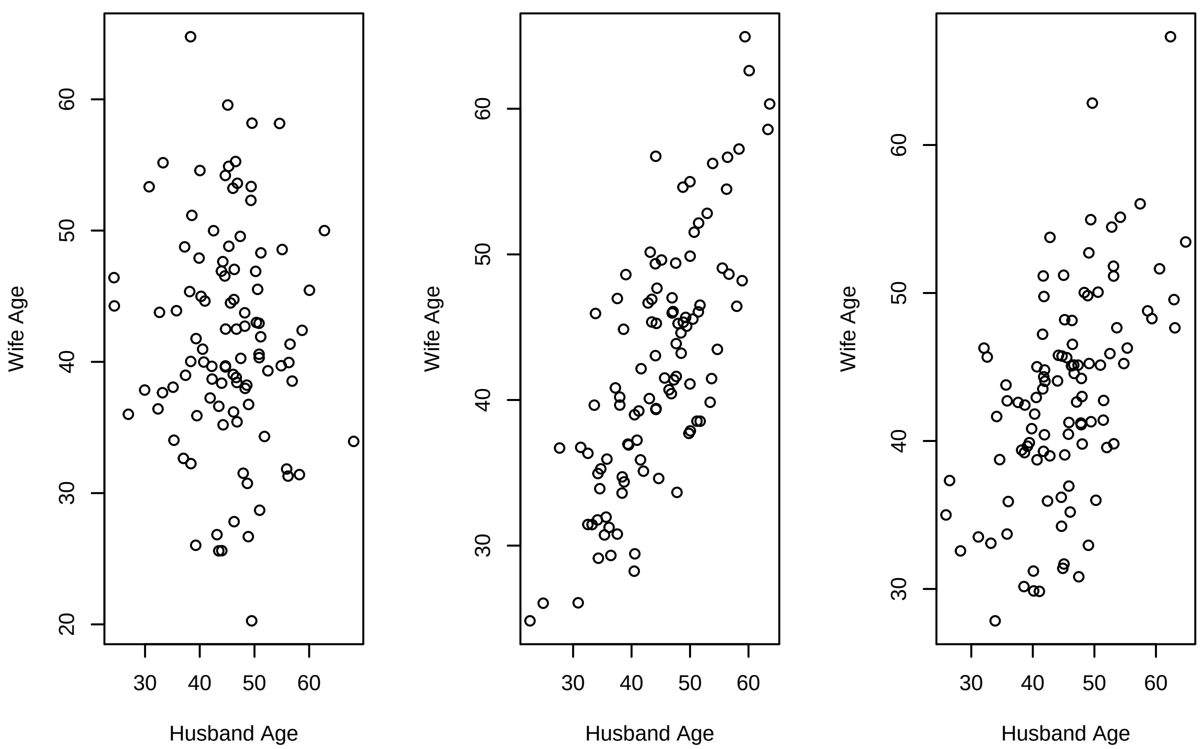
$$\Sigma^{-1} \sim \text{Wishart}(\nu_0, S_0^{-1}) = \text{Wishart}\left(4, \begin{bmatrix} 225 & 157.5 \\ 157.5 & 225 \end{bmatrix}^{-1}\right)$$

Part B

In the following code, we examine our prior predictive distributions to examine the validity of our prior.

```
agehw = read.table("~/sta601/HW/agehw.dat",header = TRUE)
husb = agehw[,1]
wife = agehw[,2]

n = 5000
p = 2
mu.0 = c(45,42)
sigma = 8
sigma2 = sigma^2
r = .7
cov.sig = sigma2 * r
L.0 = matrix(c(sigma2,cov.sig ,cov.sig,sigma2),2,2)
theta.draws = mvrnorm(n,mu.0,L.0)

nu.0 = p + 2 + 20
S.0 = L.0
```

```
S.0.inv = solve(S.0)
prior.y.draws = NULL

for(i in 1:n){
  Sigma.prior.draws =  solve(rWishart(1,nu.0,S.0.inv)[,,1])
  y.prior.draws = mvrnorm(1,theta.draws[i,],Sigma.prior.draws)
  prior.y.draws = rbind(prior.y.draws,y.prior.draws)
}


par(mfrow = c(1,3))
plot(prior.y.draws[101:200,1],prior.y.draws[1:100,2], xlab = "Husband Age", ylab = "Wife Age")
plot(prior.y.draws[201:300,1],prior.y.draws[201:300,2], xlab = "Husband Age", ylab = "Wife Age")
plot(prior.y.draws[301:400,1],prior.y.draws[301:400,2], xlab = "Husband Age", ylab = "Wife Age")
```



The plots above show my finalized prior, which ended up being:

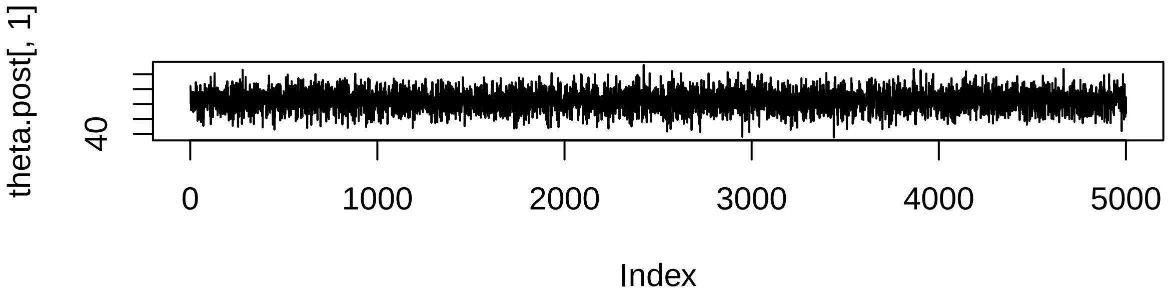
$$\boldsymbol{\theta} = \begin{bmatrix} \theta_h \\ \theta_w \end{bmatrix} \sim \text{Normal}\left(\begin{bmatrix} 45 \\ 42 \end{bmatrix}, \begin{bmatrix} 64 & 44.8 \\ 44.8 & 64 \end{bmatrix}^{-1} \right) = N(\boldsymbol{\mu}_0, \Lambda_0)$$

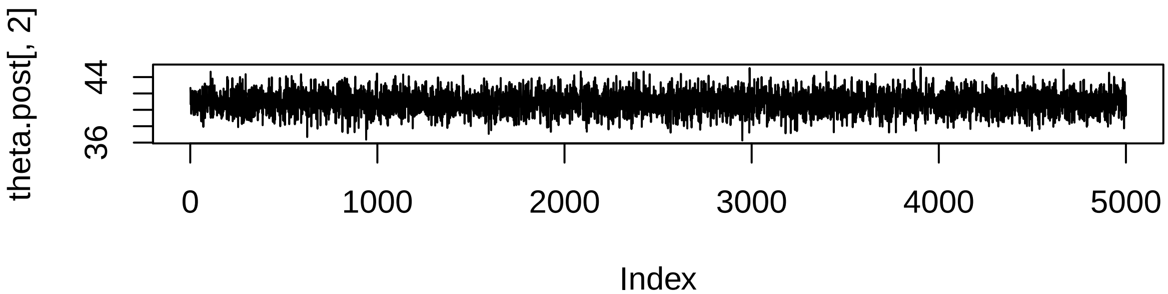
$$\Sigma^{-1} \sim \text{Wishart}(\nu_0, S_0^{-1}) = \text{Wishart}\left(24, \begin{bmatrix} 64 & 44.8 \\ 44.8 & 64 \end{bmatrix}^{-1} \right)$$

I didn't actually ended up changing the means because those seemed reasonable, but I did end up changing both $\nu_0, \Lambda_0,$ and S_0. This seems like a better prior because I was getting some weird values in my previous one. So note, I didn't change question one, I just put my new prior that I will use in this section

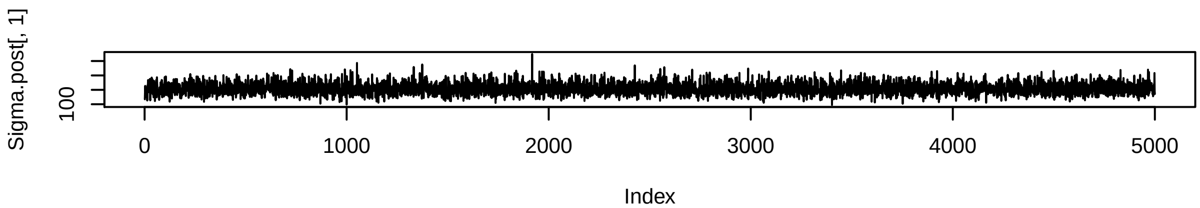
Part C

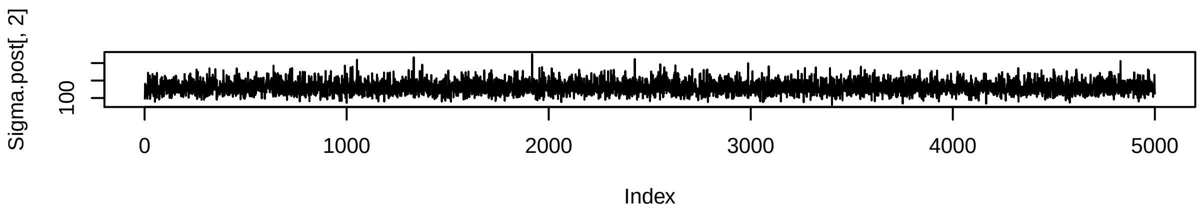
```
# MCMC posterior distribution
n.iter = 5000
p = 2
mu.0 = c(45,42)
sigma = 8
sigma2 = sigma^2
r = .7
cov.sig = sigma2 * r
L.0 = matrix(c(sigma2,cov.sig ,cov.sig,sigma2),2,2)
theta.draws = mvrnorm(n.iter,mu.0,L.0)
L.0.inv = solve(L.0)

nu.0 = p + 2 + 20
S.0 = L.0
S.0.inv = solve(S.0)
prior.y.draws = NULL

ybar = c(mean(husb),mean(wife))
n = length(husb)
Sigma = cov(agehw)
theta.post = Sigma.post = NULL
Sigma.inv = solve(Sigma)

for(i in 1:n.iter){

  ### Update theta
  Ln = solve(S.0.inv + n*Sigma.inv)
  mu.n = Ln %*% (L.0.inv %*% mu.0 + n* Sigma.inv %*% ybar)
  theta = mvrnorm(1,mu.n,Ln)

  theta.post = rbind(theta.post,theta)

  ### Update Sigma
  Sn = S.0 + (t(agehw) - c(theta)) %*% t( t(agehw) - c(theta))
  Sigma = solve(rWishart(1,nu.0 + n, solve(Sn))[,,1])

  Sigma.post = rbind(Sigma.post, c(Sigma))
}

par(mfrow = c(2,1))
plot(theta.post[,1],type = "l")
plot(theta.post[,2],type = "l")
```

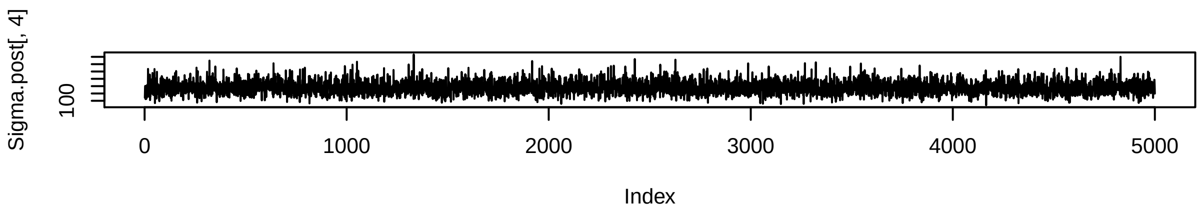





```
par(mfrow = c(3,1))
plot(Sigma.post[,1], type = "l")
plot(Sigma.post[,2], type = "l")
plot(Sigma.post[,4], type = "l")
```








```r
quantile(theta.post[,1], c(.025,.975))
```

```
##     2.5%    97.5%
## 41.78844 47.05956
```

```r
quantile(theta.post[,2], c(.025,.975))
```

```
##     2.5%    97.5%
## 38.38210 43.45419
```

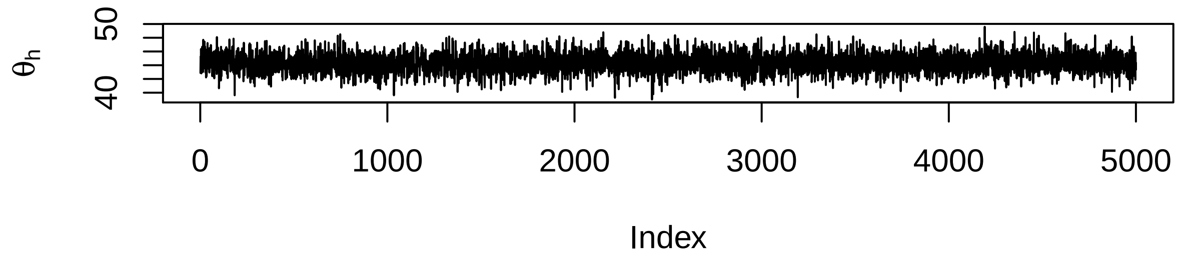
```r
quant.age.h = round(quantile(theta.post[,1], c(.025,.975)),2)
quant.age.w = round(quantile(theta.post[,2], c(.025,.975)),2)

corr.coef = Sigma.post[,2] / (sqrt(Sigma.post[,1])*sqrt(Sigma.post[,4]))

quant.corr.coef = round(quantile(corr.coef,c(.025,.975)),4)
```

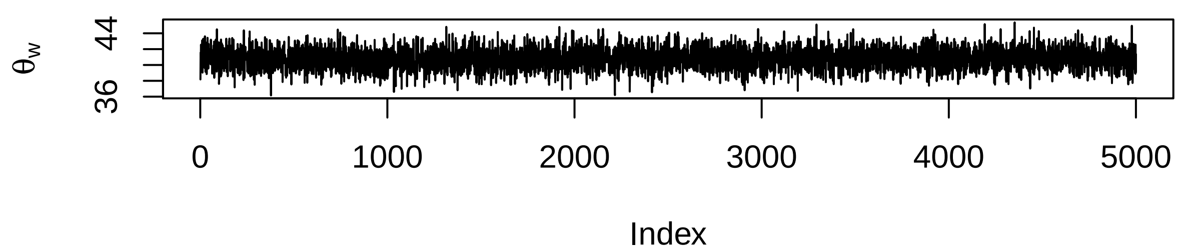
The intervals are as follows. There is a .95 probability that the average age of husbands lies in the interval 41.79, 47.06. And there is a .95 probability that the average age of wives lies in the interval 38.38, 43.45. There is also a .95 probability that the correlation coefficient of the following relationship is between 0.8662, 0.931. These are all under the prior we made up.

Part D

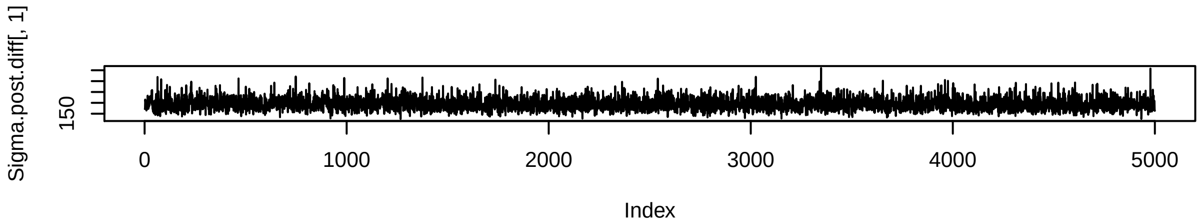
```r
# Use same code as previous, but with a diffuse prior distribution
n.iter = 5000
```

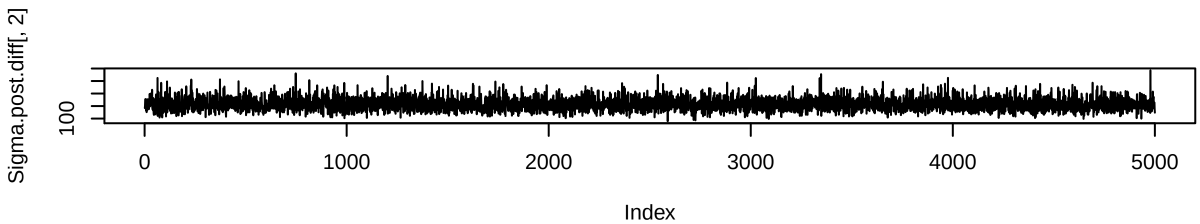
```r
p = 2
mu.0 = c(0,0)
sigma2 = 1e5
L.0 = diag(sigma2,2)
L.0.inv = solve(L.0)

nu.0 = 3
S.0 = diag(1000,2)
S.0.inv = solve(S.0)
prior.y.draws = NULL

ybar = c(mean(husb),mean(wife))
n = length(husb)
Sigma = cov(agehw)
theta.post.diff = Sigma.post.diff = NULL
Sigma.inv = solve(Sigma)

for(i in 1:n.iter){

  ### Update theta
  Ln = solve(S.0.inv + n*Sigma.inv)
  mu.n = Ln %*% (L.0.inv %*% mu.0 + n* Sigma.inv %*% ybar)
  theta = mvrnorm(1,mu.n,Ln)

  theta.post.diff = rbind(theta.post.diff,theta)

  ### Update Sigma
  Sn = S.0 + (t(agehw) - c(theta)) %*% t( t(agehw) - c(theta))
  Sigma = solve(rWishart(1,nu.0 + n, solve(Sn))[,,1])

  Sigma.post.diff = rbind(Sigma.post.diff, c(Sigma))
}

par(mfrow = c(2,1))
plot(theta.post.diff[,1],type = "l",ylab = expression(theta[h]))
plot(theta.post.diff[,2],type = "l",ylab = expression(theta[w]))
```

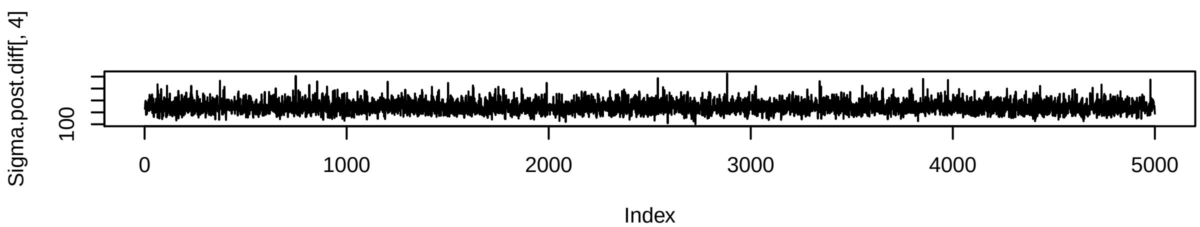





```r
par(mfrow = c(3,1))
plot(Sigma.post.diff[,1], type = "l")
plot(Sigma.post.diff[,2], type = "l")
plot(Sigma.post.diff[,4], type = "l")
```








```r
quantile(theta.post.diff[,1], c(.025,.975))
```

```
##     2.5%    97.5%
## 41.68432 47.06478
```

```r
quantile(theta.post.diff[,2], c(.025,.975))
```

```
##     2.5%    97.5%
## 38.25544 43.31560
```

```r
quant.age.h.diff = round(quantile(theta.post.diff[,1], c(.025,.975)),2)
quant.age.w.diff = round(quantile(theta.post.diff[,2], c(.025,.975)),2)

corr.coef.diff =  Sigma.post.diff[,2] / (sqrt(Sigma.post.diff[,1])*sqrt(Sigma.post.diff[,4]))

quant.corr.coef.diff = round(quantile(corr.coef.diff,c(.025,.975)),4)
```

The intervals are as follows. There is a .95 probability that the average age of husbands lies in the interval 41.68, 47.06. And there is a .95 probability that the average age of wives lies in the interval 38.26, 43.32. There is also a .95 probability that the correlation coefficient of the following relationship is between 0.7928, 0.9005. These are all under the diffuse prior that asked.

Part E

Param	Original Prior	Diffuse Prior
θ_h	41.79, 47.06	41.68, 47.06

Param	Original Prior	Diffuse Prior
θ_w	38.38, 43.45	38.26, 43.32
corr.coef	0.8662, 0.931	0.7928, 0.9005

This table shows the comparison of the credible intervals under the different priors. Note that these are very similar. I think that is probably because our sample size of 100 is enough to overwhelm the prior. However, if we note at the prior I created, the prior sample size ν_0 is fairly large, but I also think it's fairly accurate. If it weren't accurate, then that could skew the results.

It is also important to note that we have a sample size of $n = 100$. If we had one of 25, I don't they would be that close because the priors are different enough, and in my prior the prior sample size is quite large. And so the data wouldn't be enough to overwhelm the prior. So it may skew the results. However, I'm not sure if this would happen under the diffuse prior because it is diffuse, and it has such a small ν_o.