Solutions to Problem 6
Homework Assignment 3

AMS 206B, WINTER 2016

Prepared by: Mickey Warner and Xingchen (Joe) Yu

February 3, 2016

Problem 6, HW 3

Let $y_t = \rho y_{t-1} + \epsilon_t$, $\epsilon_t \overset{iid}{\sim} N(0, \sigma^2)$. This is a popular model in time series analysis known as the autoregressive model of order one or AR(1).

(a) *Write down the conditional likelihood given* y_1*, i.e.,* $f(y_2, \dots, y_n | y_1, \rho, \sigma^2)$*.*

Since $y_t \sim N(\rho y_{t-1}, \sigma^2)$ depends only on the previous observation, we have

$$
\begin{aligned}
f(y_2, \dots, y_n | y_1, \rho, \sigma^2) &= \prod_{t=2}^{n} (2\pi\sigma^2)^{-1/2} \exp\left\{ -\frac{1}{2\sigma^2}(y_t - \rho y_{t-1})^2 \right\} \\
&= (2\pi\sigma^2)^{-(n-1)/2} \exp\left\{ -\frac{1}{2\sigma^2} \sum_{t=2}^{n} (y_t - \rho y_{t-1})^2 \right\}
\end{aligned}
$$

(b) *Assume a prior of the form* $\pi(\rho, \sigma^2) \propto 1/\sigma^2$*.*

(i) *Find the joint posterior* $p(\rho, \sigma^2 | y_1, \dots, y_n)$ *based on the conditional likelihood.*

Using Bayes' formula and keeping only the terms that contain ρ and σ^2, we obtain

$$
\begin{aligned}
p(\rho, \sigma^2 | y_1, \dots, y_n) &\propto f(y_2, \dots, y_n | y_1, \rho, \sigma^2) \pi(\rho, \sigma^2 | y_1) \\
&\propto f(y_2, \dots, y_n | y_1, \rho, \sigma^2) \pi(\rho, \sigma^2) \\
&\propto (\sigma^2)^{-(n-1)/2} \exp\left\{ -\frac{1}{2\sigma^2} \sum_{t=2}^{n} (y_t - \rho y_{t-1})^2 \right\} (\sigma^2)^{-1} \\
&\propto (\sigma^2)^{-(n+1)/2} \exp\left\{ -\frac{1}{2\sigma^2} \sum_{t=2}^{n} (y_t - \rho y_{t-1})^2 \right\}
\end{aligned}
$$

We use the prior $\pi(\rho, \sigma^2)$ in place of $\pi(\rho, \sigma^2 | y_1)$ since with large enough n, y_1 will contribute little. This simplification allows us to deal with the problems due to y_1.

(ii) *Find $p(\rho|\sigma^2, y_1, \ldots, y_n)$ and $p(\sigma^2|y_1, \ldots, y_n)$ based on the conditional likelihood.*

We remove the constants and work with only the necessary terms to obtain the kernels for each quantity.

$$
\begin{aligned}
p(\rho|\sigma^2, y_1, \ldots, y_n) &\propto p(\rho, \sigma^2|y_1, \ldots, y_n) \\
&\propto \exp\left\{-\frac{1}{2\sigma^2}\sum_{t=2}^{n}(y_t - \rho y_{t-1})^2\right\} \\
&\propto \exp\left\{-\frac{1}{2\sigma^2}\left(\sum_{t=2}^{n} y_t^2 - 2\rho\sum_{t=2}^{n} y_t y_{t-1} + \rho^2 \sum_{t=2}^{n} y_{t-1}^2\right)\right\} \\
&\propto \exp\left\{-\frac{1}{2\sigma^2}\left(-2\rho\sum_{t=2}^{n} y_t y_{t-1} + \rho^2 \sum_{t=2}^{n} y_{t-1}^2\right)\right\} \\
&\propto \exp\left\{-\frac{\sum_{t=2}^{n} y_{t-1}^2}{2\sigma^2}\left(\rho^2 - 2\rho\frac{\sum_{t=2}^{n} y_t y_{t-1}}{\sum_{t=2}^{n} y_{t-1}^2}\right)\right\} \\
&\propto \exp\left\{-\frac{1}{2\frac{\sigma^2}{\sum_{t=2}^{n} y_{t-1}^2}}\left[\rho^2 - 2\rho\frac{\sum_{t=2}^{n} y_t y_{t-1}}{\sum_{t=2}^{n} y_{t-1}^2} + \left(\frac{\sum_{t=2}^{n} y_t y_{t-1}}{\sum_{t=2}^{n} y_{t-1}^2}\right)^2\right]\right\} \\
&\propto \exp\left\{-\frac{1}{2\frac{\sigma^2}{\sum_{t=2}^{n} y_{t-1}^2}}\left(\rho - \frac{\sum_{t=2}^{n} y_t y_{t-1}}{\sum_{t=2}^{n} y_{t-1}^2}\right)^2\right\}
\end{aligned}
$$

This is the kernel of a normal distribution, so we have:

$$
\rho|\sigma^2, y_1, \ldots, y_2 \sim N\left(\frac{\sum_{t=2}^{n} y_t y_{t-1}}{\sum_{t=2}^{n} y_{t-1}^2}, \frac{\sigma^2}{\sum_{t=2}^{n} y_{t-1}^2}\right)
$$

To obtain $p(\sigma^2|y_1, \ldots, y_n)$ we must integrate out ρ from the full posterior. The work is nothing but annoying, but like whatever. The integrand is normal kernel, equivalent to that from before, but now we have to keep the terms containing σ^2 used to complete the square.

$$
\begin{aligned}
p(\sigma^2|y_1, \ldots, y_n) &= \int p(\rho, \sigma^2|y_1, \ldots, y_n) d\rho \\
&\propto \int (\sigma^2)^{-(n+1)/2} \exp\left\{-\frac{1}{2\sigma^2}\sum_{t=2}^{n}(y_t - \rho y_{t-1})^2\right\} d\rho \\
&\propto (\sigma^2)^{-(n+1)/2} \int \exp\left\{-\frac{1}{2\sigma^2}\left(\sum_{t=2}^{n} y_t^2 - 2\rho\sum_{t=2}^{n} y_t y_{t-1} + \rho^2 \sum_{t=2}^{n} y_{t-1}^2\right)\right\} d\rho \\
&\propto (\sigma^2)^{-(n+1)/2} \exp\left(-\frac{1}{2\sigma^2}\sum_{t=2}^{n} y_t^2\right) \times \\
&\qquad \int \exp\left\{-\frac{1}{2\sigma^2}\left(-2\rho\sum_{t=2}^{n} y_t y_{t-1} + \rho^2 \sum_{t=2}^{n} y_{t-1}^2\right)\right\} d\rho
\end{aligned}
$$

$$
\begin{aligned}
&\propto (\sigma^2)^{-(n+1)/2} \exp\left(-\frac{1}{2\sigma^2}\sum_{t=2}^{n} y_t^2\right) \times \\
&\quad \int \exp\left\{-\frac{\sum_{t=2}^{n} y_{t-1}^2}{2\sigma^2}\left(\rho^2 - 2\rho\frac{\sum_{t=2}^{n} y_t y_{t-1}}{\sum_{t=2}^{n} y_{t-1}^2}\right)\right\} d\rho \\
&\propto (\sigma^2)^{-(n+1)/2} \exp\left(-\frac{1}{2\sigma^2}\sum_{t=2}^{n} y_t^2\right) \exp\left\{\frac{\sum_{t=2}^{n} y_{t-1}^2}{2\sigma^2}\cdot\left(\frac{\sum_{t=2}^{n} y_t y_{t-1}}{\sum_{t=2}^{n} y_{y-1}^2}\right)^2\right\} \times \\
&\quad \int \exp\left\{-\frac{\sum_{t=2}^{n} y_{t-1}^2}{2\sigma^2}\left[\rho^2 - 2\rho\frac{\sum_{t=2}^{n} y_t y_{t-1}}{\sum_{t=2}^{n} y_{t-1}^2} + \left(\frac{\sum_{t=2}^{n} y_t y_{t-1}}{\sum_{t=2}^{n} y_{y-1}^2}\right)^2\right]\right\} d\rho \\
&\propto (\sigma^2)^{-(n+1)/2} \exp\left\{-\frac{1}{\sigma^2}\left(\frac{1}{2}\sum_{t=2}^{n} y_t^2 + \frac{1}{2}\frac{(\sum_{t=2}^{n} y_t y_{t-1})^2}{\sum_{t=2}^{n} y_{y-1}^2}\right)\right\} \times \\
&\quad \left(2\pi\frac{\sigma^2}{\sum_{t=2}^{n} y_{t-1}^2}\right)^{1/2} \\
&\propto (\sigma^2)^{-n/2} \exp\left\{-\frac{1}{\sigma^2}\left(\frac{1}{2}\sum_{t=2}^{n} y_t^2 + \frac{1}{2}\frac{(\sum_{t=2}^{n} y_t y_{t-1})^2}{\sum_{t=2}^{n} y_{y-1}^2}\right)\right\}
\end{aligned}
$$

We recognize this as an inverse gamma:

$$
\sigma^2|y_1,\ldots,y_n \sim IG\left(\frac{n-2}{2}, \frac{1}{2}\sum_{t=2}^{n} y_t^2 + \frac{1}{2}\frac{(\sum_{t=2}^{n} y_t y_{t-1})^2}{\sum_{t=2}^{n} y_{y-1}^2}\right)
$$

(iii) *Simulate two data sets with $n = 500$ observations each. One with $\rho = 0.95$, $\sigma^2 = 4$ and another one with $\rho = 0.3$, $\sigma^2 = 4$. Fit the model above to the two data sets. Summarize your posterior results in both cases.*

We generate the two data sets using the `arima.sim()` function in `R` (code is provided in the appendix). The model is fit by obtaining posterior samples from $p(\rho, \sigma^2|y_1,\ldots,y_n)$. Since

$$
p(\rho, \sigma^2|y_1,\ldots,y_n) = p(\rho|\sigma^2, y_1,\ldots,y_n)p(\sigma^2|y_1,\ldots,y_n)
$$

we can use the results from (ii). We first draw $B = 10000$ (or some other large number) samples from $p(\sigma^2|y_1,\ldots,y_n)$ and then use each of these to obtain another B samples from $p(\rho|\sigma^2, y_1,\ldots,y_n)$. Figure 1 and table 1 are results for the case $\rho = 0.95$ and figure 2 and table 2 are for $\rho = 0.30$. In each figure, we show the data as well as a density estimate of the marginal posteriors. In the tables we give the estimated mean, variance, and certain quantiles.

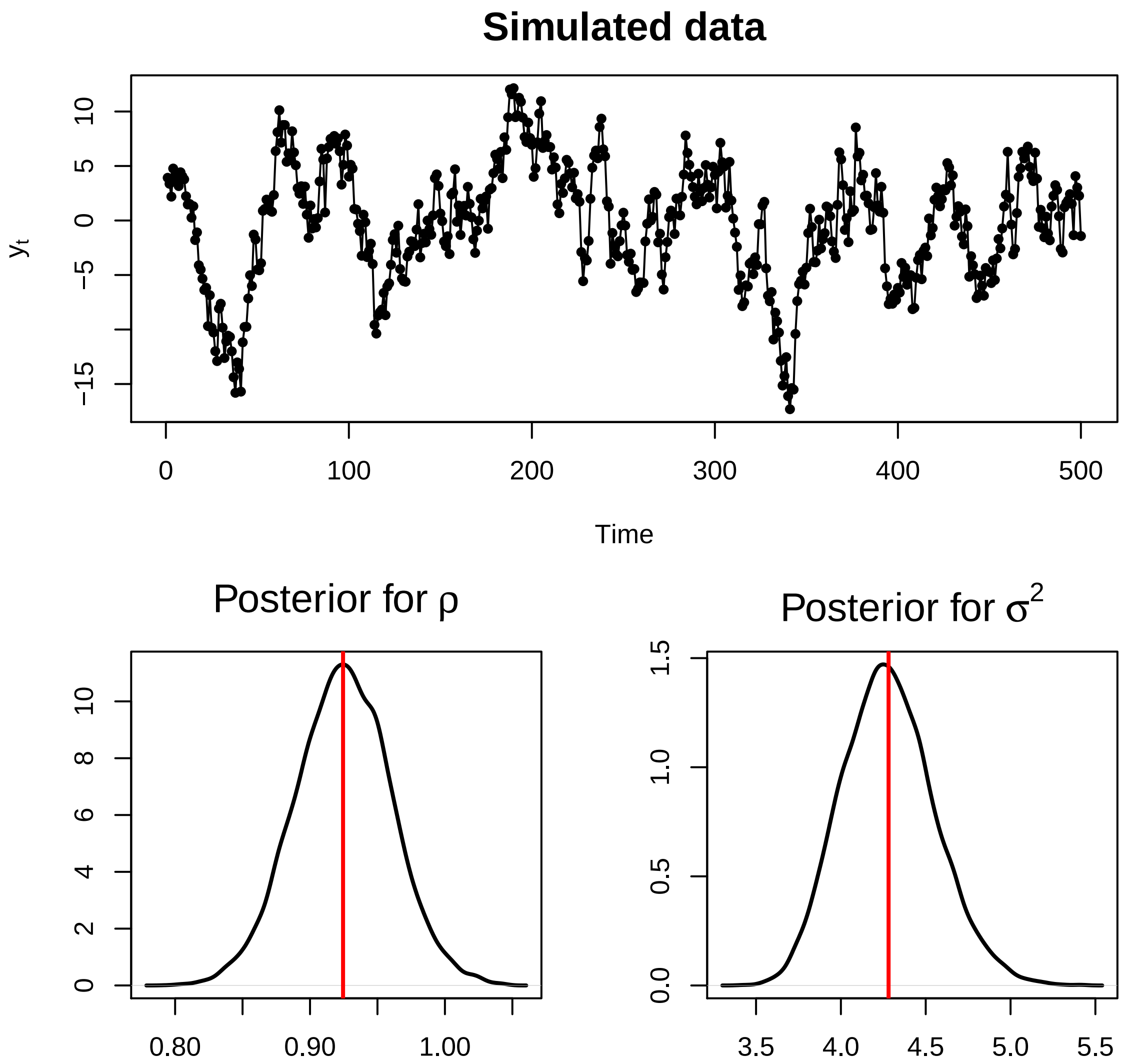

Figure 1: Top: simulated data with $n = 500$, $\rho = 0.95$, and $\sigma^2 = 4$. Bottom left: marginal posterior density of $\rho|\sigma^2, y_1, \ldots, y_2$ based on 10000 draws. Bottom right: marginal posterior density of $\sigma^2|y_1, \ldots, y_2$ based on 10000 draws. The red lines mark the means of the draws.

	mean	var	0%	2.5%	50%	97.5%	100%
ρ	0.924	0.001	0.793	0.854	0.924	0.994	1.045
σ^2	4.281	0.074	3.418	3.783	4.271	4.853	5.423

Table 1: Numerical summary of the posterior distributions.

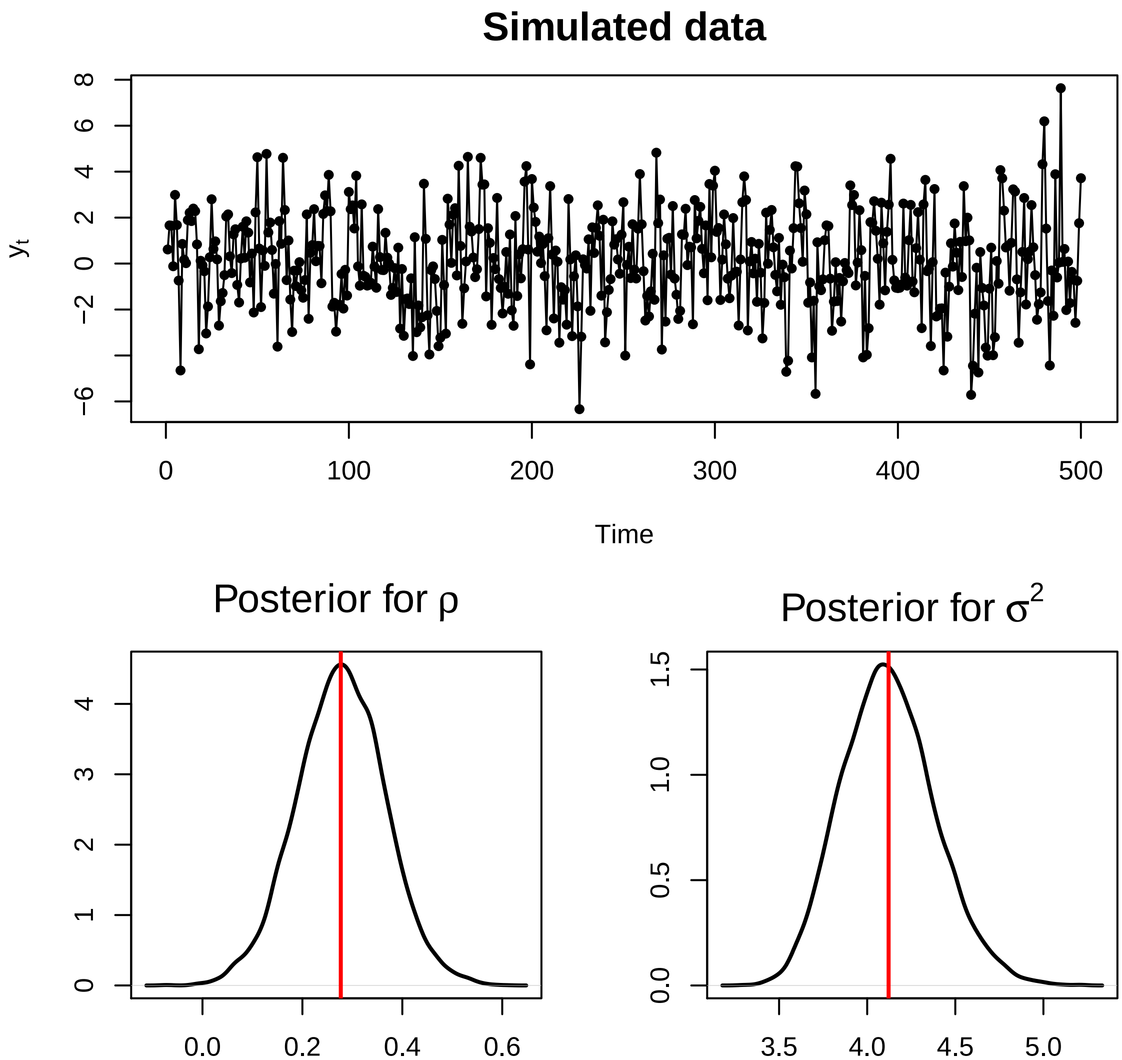

Figure 2: Top: simulated data with $n = 500$, $\rho = 0.30$, and $\sigma^2 = 4$. Bottom left: marginal posterior density of $\rho|\sigma^2, y_1, \ldots, y_2$ based on 10000 draws. Bottom right: marginal posterior density of $\sigma^2|y_1, \ldots, y_2$ based on 10000 draws. The red lines mark the means of the draws.

	mean	var	0%	2.5%	50%	97.5%	100%
ρ	0.276	0.007	-0.074	0.100	0.277	0.450	0.610
σ^2	4.121	0.069	3.290	3.642	4.111	4.673	5.221

Table 2: Numerical summary of the posterior distributions.

Code

```
##### HW 3, Problem 6
sig = sqrt(4)
n = 500

### Generating Data
set.seed(1)
ar = 0.95
#ar = 0.30
x1 = arima.sim(list(ar=ar),n=n, sd=sig)

### Certain portions of the data vector
yn = x1[-1]
yn_1 = x1[-n]

### Sampling for sigma^2
B = 10000
alpha = (n-2)/2
beta = (sum(yn^2)-sum(yn*yn_1)^2/sum(yn_1^2))/2
po.sig = 1/rgamma(B, alpha, beta)

### Sampling for rho
mu = sum(yn*yn_1)/sum(yn_1^2)
si = po.sig/sqrt(sum(yn_1^2))
po.ro = rnorm(B, mu, si)

### Plots
#pdf("./fig_1.pdf", width = 8, height = 8)
par(mfrow=c(1,1), mar = c(4.1, 4.1, 3.1, 1.1))
layout(matrix(c(1,2,1,3),2,2))
plot(x1, type='o', pch=20, main="Simulated data", cex.main = 1.5,
    ylab = expression(y[t]))
plot(density(po.ro), main = expression("Posterior for" ~ rho),
    xlab ="", ylab="", cex.main = 1.5, lwd = 2)
abline(v = mean(po.ro), col = 'red', lwd = 2)
plot(density(po.sig), main = expression("Posterior for" ~ sigma^2),
    xlab ="", ylab="", cex.main = 1.5, lwd = 2)
abline(v = mean(po.sig), col = 'red', lwd = 2)
#dev.off()

### Summary statistics
c("mean"=mean(po.ro), "var"=var(po.ro), quantile(po.ro, c(0, 0.025, 0.5, 0.975, 1)))
c("mean"=mean(po.sig), "var"=var(po.sig), quantile(po.sig, c(0, 0.025, 0.5, 0.975, 1)))

### Scatterplot for the joint distribution
# par(mfrow=c(1,1), mar = c(5.1,4.1,4.1,2.1))
# plot(po.ro, po.sig, pch = 20, xlab="", ylab="", cex=0.5)
# mtext(expression(rho), side=1, line=3, cex=1.5)
# mtext(expression(sigma^2), side=2, line=2, cex=1.5)
```