Week 3 – Homework Solutions

Review of the univariate regression with independent priors on β and σ^2:

First, we derive the full conditionals for this model

$$y = X\beta + \varepsilon \sim N(0, \sigma^2 I)$$

Setting independent priors for this model

$$\beta \sim N(\bar{\beta}, A^{-1}), \quad \sigma^2 \sim \frac{\upsilon_0 s_0^2}{\chi^2_{\upsilon_0}}$$

The posterior is given as

$$p(\beta, \sigma^2 \mid X, y, \bar{\beta}, A, \upsilon_0, s_0^2) \propto \left(\sigma^2\right)^{-\frac{n}{2}} \exp\left[-\frac{1}{2\sigma^2}(y - X\beta)'(y - X\beta)\right]$$
$$\exp\left[-\frac{1}{2}\left(\beta - \bar{\beta}\right)' A \left(\beta - \bar{\beta}\right)\right]$$
$$\left(\sigma^2\right)^{-\frac{\upsilon_0}{2}+1} \exp\left(\frac{-\upsilon_0 s_0^2}{2\sigma^2}\right)$$

Now, we note that

$$-\frac{1}{2\sigma^2}(y - X\beta)'(y - X\beta) = \left(\frac{y}{\sigma} - \frac{X}{\sigma}\beta\right)'\left(\frac{y}{\sigma} - \frac{X}{\sigma}\beta\right)$$

Substituting this in the posterior, pulling out the terms that include β (or σ^2, as the case may be) and simplifying, the full conditionals can be obtained as given below

$$\beta \mid \sigma^2, X, y, \bar{\beta}, A, \upsilon_0, s_0^2 \sim N\left(\left(\frac{X'X}{\sigma^2} + A\right)^{-1}\left(\frac{X'y}{\sigma^2} + A\bar{\beta}\right), \left(\frac{X'X}{\sigma^2} + A\right)^{-1}\right)$$

$$\sigma^2 \mid \beta, X, y, \bar{\beta}, A, \upsilon_0, s_0^2 \sim \frac{\upsilon_1 s_1^2}{\chi^2_{\upsilon_1}}, \quad \upsilon_1 = \upsilon_0 + n, \quad s_1^2 = \frac{\upsilon_0 s_0^2 + \varepsilon'\varepsilon}{\upsilon_0 + n}$$

$$\varepsilon = y - X\beta$$

We can use these full conditionals to draw from the joint posterior using a Gibbs Sampler. The key insights here are:

1. the conditional for β just involves standardizing by dividing by σ.
2. the conditional for σ given β just uses the observed residuals ε!

Similar to the Week 2 problem, we simulate the data using the cheese data. We run a regression of log(PRICE) and DISP on log(SALES) to obtain the betas and sigma_sq. The regression estimates are given below

Parameter	Regression Estimate
beta0: intercept	9.3712
beta1: log(PRICE) coefficient	-1.2576
beta2: DISP coefficient	0.5132
sigmasq	0.5782

Then, we use these values to simulate the data using a sample of size 2000 from the same dataset.

The R code for the univariate regression with independent priors on β and σ^2 is in `runiregGibbs`.

Now let's work out how we impose the restriction that one of the beta elements is positive.

The prior density for β can be written as follows

$$p\left(\beta \mid \bar{\beta}, A, I\left(\beta_{display}>0\right)\right)=\frac{\frac{|A|^{-1}}{\sqrt{2\pi}}\exp\left[\left(\beta-\bar{\beta}\right)' A\left(\beta-\bar{\beta}\right)\right]}{\int_{\Omega}\left[\frac{|A|^{-1}}{\sqrt{2\pi}}\exp\left[\left(\beta-\bar{\beta}\right)' A\left(\beta-\bar{\beta}\right)\right]\right]} I\left(\beta_{display}>0\right)$$

where

$$\Omega=\left\{\beta : \beta_{display}>0\right\}$$

We note that in this prior, the denominator is a number that we can drop when we write the posterior (since we need only the un-normalized posterior for the Gibbs sampler). Thus, the posterior is simply the same as in Problem 3, except with an additional $I\left(\beta_{display}>0\right)$ term in the density (thus restricting $\beta_{display}$ to the positive half of the real line)

$$\beta \mid \sigma^2, X, y, \bar{\beta}, A, \upsilon_0, s_0^2 \sim N\left(\left(\frac{X'X}{\sigma^2}+A\right)^{-1}\left(\frac{X'y}{\sigma^2}+A\bar{\beta}\right),\left(\frac{X'X}{\sigma^2}+A\right)^{-1}\right) I\left(\beta_{display}>0\right)$$

$$\sigma^2 \mid \beta, X, y, \bar{\beta}, A, \upsilon_0, s_0^2 \sim \frac{\upsilon_1 s_1^2}{\chi^2_{\upsilon_1}}, \quad \upsilon_1=\upsilon_0+n, \quad s_1^2=\frac{\upsilon_0 s_0^2+\varepsilon'\varepsilon}{\upsilon_0+n}$$

The full conditionals for $\beta_{display}$ and β_0 are obtained by using the standard expressions for conditional normals (as in the problem set), noting that the $\beta_{display}$ term has to be drawn from a truncated normal and not just a normal (due to the $I\left(\beta_{display}>0\right)$ term in its full conditional density).

The code for this procedure is in the file test_rtrununiregGibbs.R.

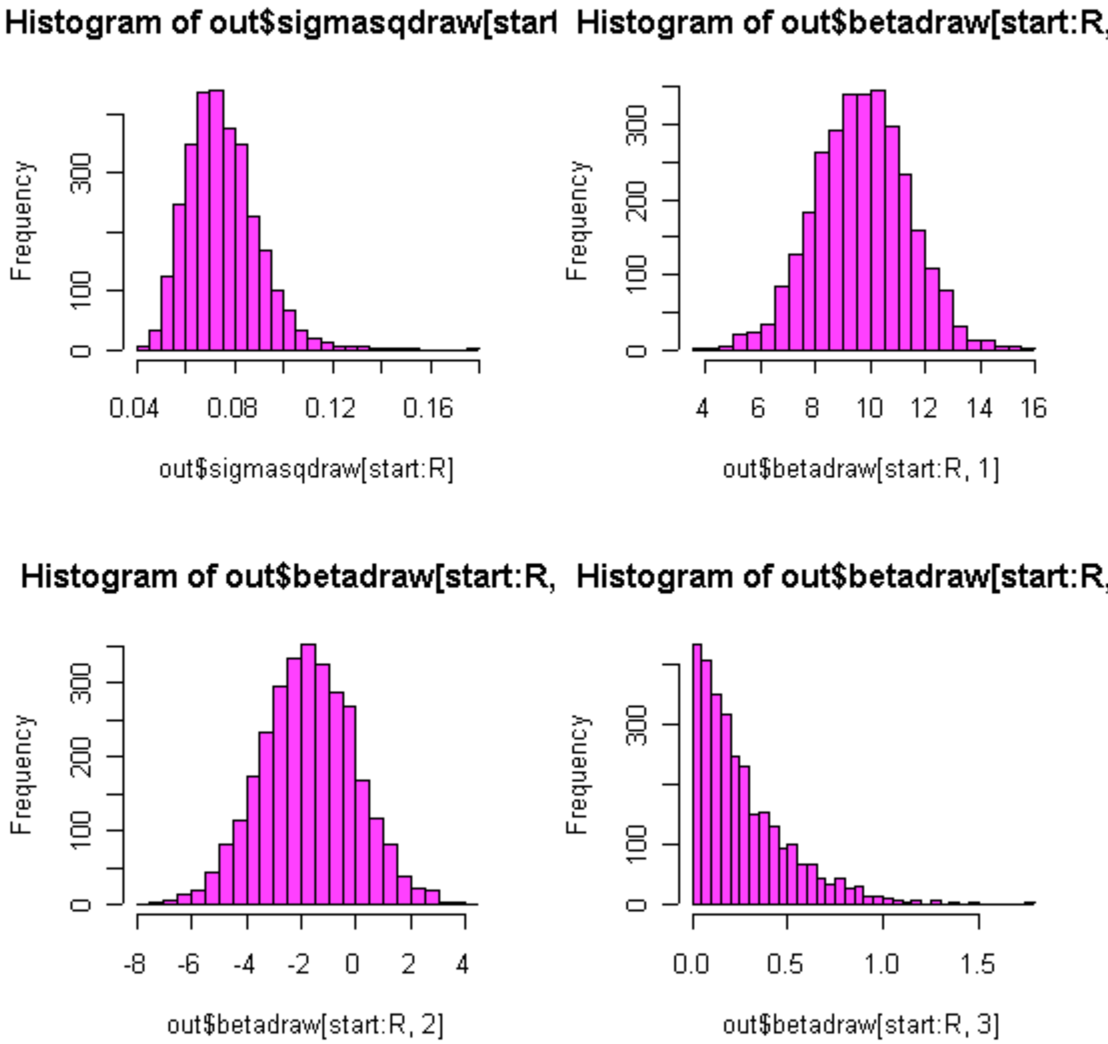
The estimates under the two methods are given below. The chains were run for 5000 draws in each case and inference was done using draws 2001-5000.

Parameter	Estimates Method: subvector draws		Estimates Method: rejection	
	Mean	Std. Dev.	Mean	Std. Dev.

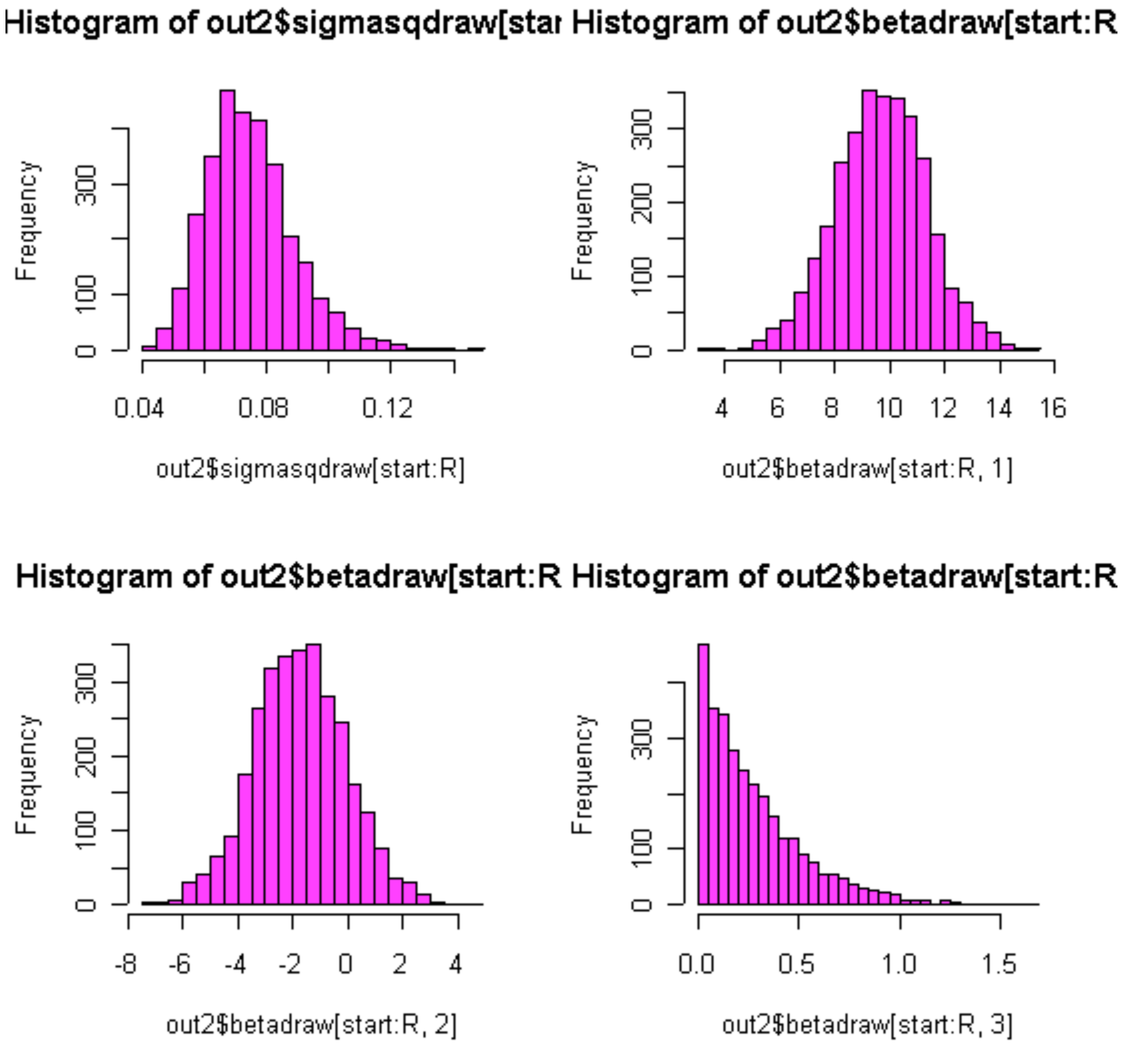
intercept	9.6727	1.6916	9.6724	1.6550
log(PRICE)	-1.7335	1.7236	-1.7329	1.6852
DISP	0.2751	0.2544	0.2791	0.2482
sigmasq	0.0749	0.0142	0.0748	0.0139

The histograms of the posteriors for the parameters are similar across the two methods. However, the autocorrelations are lower in the rejection method than in the subvector draw method, even though the former is less efficient (suppose the unrestricted posterior is centered below 0 for the display coefficient -- then many draws would be rejected)! Thus, the cost of efficiency is (slightly) higher autocorrelation and hence a slightly larger number of required draws in the latter case.

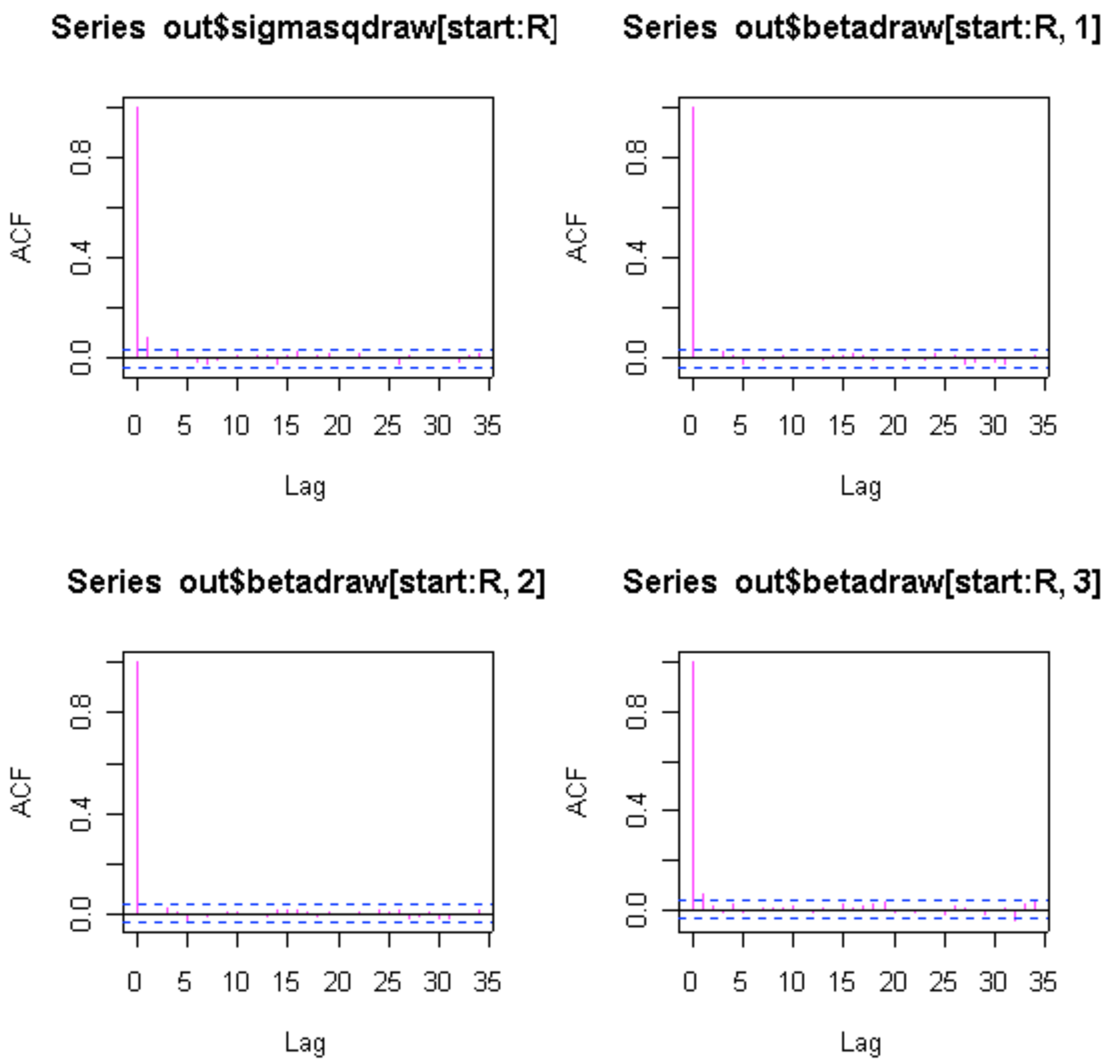
Histograms for the 'subvector draw method'



Histograms for the 'rejection method'



Autocorrelation plots for the ‘subvector draw method’



Autocorrelation plots for the 'rejection method'