# Computer Problem Sheet 3: Bayesian Inference in the Regression Model with General Error Covariance Matrix

These exercises are taken (or adapted) from the book Bayesian Econometric Methods (BEM) and Matlab code can be found on the BEM website (you can connect through the link for the book on my website or directly to www.econ.iastate.edu/faculty/tobias/)

*Exercise 1: Heteroskedasticity of a Known Form*

Heteroskedasticity occurs if:

$$\Omega = \begin{bmatrix} \omega_1 & 0 & . & . & 0 \\ 0 & \omega_2 & 0 & . & . \\ . & 0 & . & . & . \\ . & . & . & . & 0 \\ 0 & . & . & 0 & \omega_N \end{bmatrix}$$

In this exercise, assume that

$$\omega_i = \left(1 + \alpha_1 z_{i1} + \alpha_2 z_{i2} + ... + \alpha_p z_{ip}\right)^2$$

where $z_i$ is a $p-$vector of data which may include some or all of the explanatory variables. Let $\alpha = (\alpha_1, .., \alpha_p)'$.

a) Assuming a noninformative prior for the parameters which define $\Omega$:

$$p(\alpha) \propto 1,$$

describe how a Metropolis-within-Gibbs algorithm can be developed for carrying out Bayesian inference in this model. Note: most of the basic theoretical derivations are provided in 118-124 of Bayesian Econometrics or Exercise 13.1 of BEM.

b) Program up the the algorithm described in part a) and carry out a Bayesian investigation of heteroskedasticity using a suitable data set (e.g. the house price data set on page 191 of BEM and available on the BEM website) and choices for the prior hyperparameters $\underline{\beta}, \underline{V}, \underline{\nu}$ and $\underline{s}^{-2}$. Does heteroskskedasticity exist in your data set?

Solution: This is part of Exercise 13.2 of BEM. Note that my algorithm is pretty inefficient. Can you do better? Note also that, if you are having trouble with understanding the "Metropolis" part of the Metropolis-within-Gibbs algorithm, then you may wish to take a look at Exercise 11.18 (a practical example of a Metropolis-Hastings algorithm in a univariate problem) or Exercise 11.17 (a theoretical exercise intended to provide some inuition).

*Exercise 2: Heteroskedasticity of an Unknown Form: Student-t Errors*

Now assume that you have heteroskedasticity as in (13.16), but you do not know its exact form (i.e. you do not know equation 13.17), but are willing to assume all of the error variances are drawn from a common distribution, the Gamma. If we work with error precisions rather than variances and, hence, define $\lambda \equiv (\lambda_1, \lambda_2, .., \lambda_N)' \equiv \left(\omega_1^{-1}, \omega_2^{-1}, .., \omega_N^{-1}\right)'$ we assume the following hierarchical prior for $\lambda$:

$$p(\lambda|\nu_\lambda) = \prod_{i=1}^{N} f_G(\lambda_i|1, \nu_\lambda).$$

This prior is hierarchical since $\lambda$ depends on a parameter, $\nu_\lambda$, which in turn has its own prior, $p(\nu_\lambda)$. We will assume the latter to be exponential:

$$p(\nu_\lambda) = f_G(\nu_\lambda|\underline{\nu}_\lambda, 2).$$

Aside: This specification is a popular one since it can be shown (if we integrate out $\lambda$), that we obtain the linear regression model with t-errors. That is,

$$p(\varepsilon_i) = f_t\left(\varepsilon_i|0, h^{-1}, \nu_\lambda\right).$$

The interested reader is referred to Geweke (1993) for a proof and further details. Note, in particular, that Geweke (1993) shows that if you use a common noninformative prior for $\beta$ (i.e. $p(\beta) \propto 1$ on the interval $(-\infty, \infty)$), then the posterior mean does not exist, unless $p(\nu_\lambda)$ is zero on the interval $(0, 2]$. The posterior standard deviation does not exist unless $p(\nu_\lambda)$ is zero on the interval $(0, 4]$. Hence, the researcher who wants to use a noninformative prior for $\beta$ should either use a prior which excludes small values for $\nu_\lambda$ or present posterior medians and interquartile ranges (which will exist for any valid p.d.f.). With an informative Normal prior for $\beta$, the posterior mean and standard deviation of $\beta$ will exist.

The full posterior conditionals for this model can be derived and, thus, a Gibbs sampler can be used to carry out Bayesian inference in this model. The posterior conditionals are derived and presented in the textbook (see Bayesian Econometrics, pages 124-129). Program up this Gibbs sampler and carry out an empirical Bayesian analysis using a suitable data set (e.g. the house price data set on page 191 of BEM and available on the BEM website) and suitable choices for prior hyperparameters. Do the errors in your data set seem to be Student-t?

Solution: This is part of Exercise 13.3 of BEM.