Learning About Bayesian Computation

The classical econometrician will often used software packages such as Eviews, Stata or Microfit which allow the user to implement a set of econometric procedures at the click of a mouse. For the Bayesian, such options are typically not available. There are a limited number of Bayesian software packages that can handle some models. For instance, BUGS (which is freeware available at http://www.mrc-bsu.cam.ac.uk/bugs/) is the most popular and can handle many models. However, BUGS is mostly written by statisticians and does not directly handle a lot of models used by econometricians (especially in time series). For this reason, many Bayesians write their own programs in languages such as Matlab, Gauss or R. The purpose of this set of exercises is to build up Bayesian programming skills of relevance for posterior simulation (e.g. Monte Carlo integration and Gibbs sampling) so that the user can create their own programs for any model.

My textbooks, *Bayesian Econometrics* (BE) and *Bayesian Econometric Methods* (BEM),contain a wide variety of computer exercises. The present handout takes a few exercises from BEM books. These exercises take you through some of the key steps of Bayesian computation. BEM provides written solutions to these exercises and Matlab solutions to these exercises are provided on the book website which is. The website for BE also contains a variety of Matlab code. Links to both of these websites are on my website (http://personal.strath.ac.uk/gary.koop/).

It is also worth noting that, increasingly, Bayesian researchers are making their Matlab or Gauss programs available on their websites (e.g. James LeSage has a good website: http://www.spatial-econometrics.com/). Even if you are going to use someone else's programs, it is useful to have some basic Matlab or Gauss skills to understand and, if necessary, adapt their code.

Exercise 1: Drawing from Standard Distributions.

Simulation-based inference via Monte Carlo integration, the Metropolis-Hastings algorithm or the Gibbs sampler requires the researcher to be able to draw from standard distributions. In this exercise we discuss how MATLAB can be used to obtain draws from a variety of standard continuous distributions. Specifically, we obtain draws from the Uniform, Normal, Student-t, Beta, Exponential and Chi-squared distributions, using MATLAB (see the Appendix to BEM or BE for definitions of these distributions). This exercise is designed to be illustrative - MATLAB is capable of generating variates from virtually any distribution that an applied researcher will encounter (and the same applies to other relevant computer languages such as Gauss).

Using MATLAB, obtain sets of 10, 100 and 100,000 draws from the Uniform, standard Normal, Student-t(3) (denoted $t(0,1,3)$ in the notation of the Appendix), Beta(3,2), Exponential with mean 5 and $\chi^2(3)$ distributions. For each sample size calculate the mean and standard deviation and compare these quantities to the known means and standard deviations from each distribution.

Solution: This is Exercise 11.2 in BEM.

Exercise 2: Analytical and Monte Carlo Integration in the Normal Linear Regression Model

a) Generate an artificial data set of size $N = 100$ from the Normal linear regression model with an intercept and one other explanatory variable. Set the intercept (β_1) to 0, the slope coefficient (β_2) to 1.0 and $h = 1.0$. Generate the explanatory variable by taking random draws from the $U(0,1)$ distribution.

b) Calculate the posterior mean and standard deviation for the slope coefficient, β_2, for this data set using a Normal-Gamma prior with $\underline{\beta} = (0,1)', \underline{V} = I_2, \underline{s}^{-2} = 1, \underline{\nu} = 1$.

c) Calculate the Bayes factor comparing the model $M_1 : \beta_2 = 0$ with $M_2 : \beta_2 \neq 0$.

d) Carry out a prior sensitivity analysis by setting $\underline{V} = cI_2$ and repeating parts b) and c) for values of $c = 0.01, 1.0, 100.0, 1 \times 10^6$. How sensitive is the posterior to changes in prior information? How sensitive is the Bayes factor?

e) Repeat part b) using Monte Carlo integration for various values of R. How large does R have to be before you reproduce the results of the previous parts to two decimal places?

f) Calculate the numerical standard errors associated with the posterior mean of the slope coefficient for the models. Does the nse seem to give a reliable guide to the accuracy of the approximation provided by Monte Carlo integration?

Solution: This is Exercise 11.3 in BEM.

Exercise 3: Gibbs Sampling from The Bivariate Normal.

The purpose of this question is to learn about the properties of the Gibbs sampler in a simple case.

Assume that you have a model which yields a bivariate Normal posterior,

$$\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \right),$$

where $|\rho| < 1$ is the (known) posterior correlation between θ_1 and θ_2.

(a) Write a program which uses Monte Carlo integration to calculate the posterior means and standard deviations of θ_1 and θ_2.

(b) Write a program which uses Gibbs sampling to calculate the posterior means and standard deviations of θ_1 and θ_2.

(c) Set $\rho = 0$ and compare the programs from parts a) and b) for a given number of replications (e.g. $R = 100$) and compare the accuracy of the two algorithms.

(d) Repeat part c) of this question for $\rho = .5, .9, .99$ and $.999$. Discuss how the degree of correlation between θ_1 and θ_2 affects the performance of the Gibbs sampler. Make graphs of the Monte Carlo and Gibbs sampler replications of θ_1 (i.e. make a graph with x-axis being replication number and y-axis being θ_1). What can the graphs you have made tell you about the properties of Monte Carlo and Gibbs sampling algorithms?

(e) Repeat parts c) and d) more replications (e.g. $R = 10,000$) and discuss how Gibbs sampling accuracy improves with number of replications.

Solution: This is Exercise 11.7 in BEM.

Exercise 4: Gibbs Sampling in the SUR model

Consider a two-equation version of the SUR model

$$y_{i1} = x_{i1}\beta_1 + \varepsilon_{i1},$$

$$y_{i2} = x_{i2}\beta_2 + \varepsilon_{i2},$$

for $i = 1, .., n$ and where $\varepsilon_i = (\varepsilon_{i1}, \varepsilon_{i2})'$ are i.i.d. $N(0, \Sigma)$. x_{i1} and x_{i2} are $1 \times k_1$ and $1 \times k_2$, respectively and

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}.$$

Suppose you employ priors of the form:

$$\beta \sim N\left(\mu_\beta, V_\beta\right)$$

and

$$\Sigma^{-1} \sim W\left(\Omega, \nu\right)$$

where W denotes the Wishat distribution.

Derive a posterior simulator for this model and conduct a test of it using artificial data.

Solution: This is Exercise 11.10 in BEM.

Exercise 5: Using the AR(p) model to Understand the Properties of a Series

This exercise is loosely based on Geweke (1988, JBES). Let y_t for $t = 1, .., T$ indicate observations on a time series variable. y_t is assumed to follow an AR(p) process:

$$y_t = \beta_0 + \beta_1 y_{t-1} + ... + \beta_p y_{t-p} + \epsilon_t, \tag{1}$$

where ϵ_t is i.i.d. $N\left(0, h^{-1}\right)$. Many important properties of y_t depend on the roots of the polynomial $1 - \sum_{i=1}^{p} \beta_i z^i$ which we will denote by r_i for $i = 1, .., p$. Geweke (1988) lets y_t be the log of real GDP and sets $p = 3$ and, for this choice, focusses on the features of interest: $C = \{\beta : \text{ Two of } r_i \text{ are complex}\}$ and $D = \{\beta : \min|r_i| < 1\}$ where $\beta = (\beta_0, \beta_1, \beta_2, \beta_3)'$. Note that C and D are regions whose bounds are complicated nonlinear functions of $\beta_1, \beta_2, \beta_3$. If the AR coefficients lie in the region defined by C then real GDP exhibits an oscillatory response to a shock and if they lie in D then y_t exhibits an explosive response to a shock.

(a) Assuming a prior of the form:

$$p\left(\beta_0, .., \beta_p, h\right) \propto \frac{1}{h}, \tag{2}$$

derive the posterior for β. To simplify things, you may ignore the (minor) complications relating to the treatment of initial conditions. Thus, assume the dependent variable is $y = (y_{p+1}, .., y_T)'$ and treat $y_1, .., y_p$ as fixed initial conditions.

(b) Using an appropriate data set (e.g., the US real GDP data set provided on the website associated with this book), write a program which calculates the posterior means and standard deviations of β and $\min|r_i|$.

(c) Extend the program of part b) to calculate the probability that y_t is oscillatory (i.e., $\Pr(\beta \in C|y)$), the probability that y_t is explosive (i.e., $\Pr(\beta \in D|Data)$) and calculate these probabilities using your data set.

Solution: This is Exercise 17.1 in BEM.

Exercise 6: The Threshold Autoregressive Model

Dynamics of many important macroeconomic variables can potentially vary over the business cycle. This has motivated the development of many models where different autoregressive representations apply in different regimes. Threshold autoregressive (TAR) models are a class of simple and popular regime-switching models. Potter (1995, JAE) provides an early exposition of these models in macroeconomics and Geweke and Terui (1993, JTSA) is an early Bayesian treatment. This exercise asks you to derive Bayesian methods for a simple variant of a TAR model.

Consider a two regime TAR for a time series variable y_t for $t = p+1, .., T$ (where $t = 1, .., p$ are used as initial conditions):

$$\begin{array}{l} y_t = \beta_{10} + \beta_{11} y_{t-1} + ... + \beta_{1p} y_{t-p} + \epsilon_t \text{ if } y_{t-1} \leq \tau \\ y_t = \beta_{20} + \beta_{21} y_{t-1} + ... + \beta_{2p} y_{t-p} + \epsilon_t \text{ if } y_{t-1} > \tau \end{array},$$

where ϵ_t is i.i.d. $N\left(0, h^{-1}\right)$. We will use the notation $\beta = \left(\beta_1', \beta_2'\right)'$ where $\beta_j = \left(\beta_{j0}, \beta_{j1}, ..., \beta_{jp}\right)'$ for $j = 1, 2$. For all parts of this question, you may proceed conditionally on the first p observations and, thus, ignore the (minor) complications caused by initial conditions [see Exercise 5 for more detail].

(a) Assume that τ is known (e.g., $\tau = 0$) and Normal-Gamma priors are used (i.e., the joint prior for β and h is $NG\left(\underline{\beta}, \underline{Q}, \underline{s}^{-2}, \underline{\nu}\right)$), derive the posterior for the TAR given in this question.

(b) Using an appropriate data set (e.g., the real GDP data set available on the website associated with this book), write a program and carry out Bayesian inference in the TAR model using your results from part (a). Note: When working with GDP, TAR models are usually specified in terms of GDP growth. Hence, if you are working with a GDP series, you should define y_t as its first difference.

(c) Repeat parts (a) and (b) assuming that τ is an unknown parameter.

Solution: This is Exercise 17.2 in BEM.

Exercise 7: Extensions of the Basic Threshold Autoregressive Model 1: Other Threshold Triggers

There are many extensions of the TAR which have been found useful in empirical work. In Exercise 5, we assumed that the first lag of the dependent variable (last quarter's GDP growth) triggered the regime switch. However, in general, it might be another variable, z, that is the *threshold trigger* and it may take longer than one period to induce the regime switch. Thus, we use the same assumptions and definitions as in Exercise 5, except that now:

$$\begin{array}{c} y_t = \beta_{10} + \beta_{11} y_{t-1} + ... + \beta_{1p} y_{t-p} + \epsilon_t \text{ if } z_{t-d} \leq \tau \\ y_t = \beta_{20} + \beta_{21} y_{t-1} + ... + \beta_{2p} y_{t-p} + \epsilon_t \text{ if } z_{t-d} > \tau \end{array},$$

where d is the *delay parameter* and z_{t-d} is either an exogenous variable or a function of the lags of the dependent variable.

(a) Assume that d is an unknown parameter with a noninformative prior over $1, .., p$ (i.e., $\Pr(d = i) = \frac{1}{p}$ for $i = 1, .., p$) and Normal-Gamma priors are used (i.e., the joint prior for β and h is $NG\left(\underline{\beta}, \underline{Q}, \underline{s}^{-2}, \underline{\nu}\right)$), derive the posterior for this model.

(b) Using an appropriate data set (e.g., the real GDP growth data set available on the website associated with this book), write a program and carry out Bayesian inference for this model using your results from part (a). Set $p = 4$ and

$$z_{t-d} = \frac{\sum_{d=1}^{p} y_{t-d}}{d},$$

so that (if you are using quarterly real GDP growth data), the threshold trigger is average GDP growth over the last d quarters.

Solution: This is Exercise 17.3 in BEM.

Exercise 8. Extensions of the Basic Threshold Autoregressive Model 1: Switches in the Error Variance

Recently, there has been much interest in the volatility of macroeconomic variables and, in particular, whether the error variance exhibits regime-switching behavior [see, .e.g., Zha and Sims (2006, AER)]. Accordingly, we can extend the TAR models of previous exercises to:

$$\begin{array}{c} y_t = \beta_{10} + \beta_{11} y_{t-1} + ... + \beta_{1p} y_{t-p} + \sqrt{h_1^{-1}} \epsilon_t \text{ if } z_{t-d} \leq \tau \\ y_t = \beta_{20} + \beta_{21} y_{t-1} + ... + \beta_{2p} y_{t-p} + \sqrt{h_2^{-1}} \epsilon_t \text{ if } z_{t-d} > \tau \end{array},$$

where all definitions and assumptions are the same as in Exercises 6 and 7 except that we now assume ϵ_t is i.i.d. $N(0,1)$.

(a) Assume that d is an unknown parameter with a noninformative prior over $1,..,p$ (i.e., $\Pr(d=i) = \frac{1}{p}$ for $i = 1,..,p$) and Normal-Gamma priors are used in each regime (i.e., the joint prior for β_j and h_j is $NG\left(\underline{\beta_j}, \underline{Q_j}, \underline{s_j}^{-2}, \underline{\nu_j}\right)$ for $j = 1,..,2$), derive the posterior for this model.

(b) Using an appropriate data set (e.g., the real GDP growth data set available on the website associated with this book), write a program and carry out Bayesian inference for this model using your results from part (a). Set $p = 4$ and

$$z_{t-d} = \frac{\sum_{d=1}^{p} y_{t-d}}{d}.$$

Solution: This is Exercise 17.4 in BEM.