Take home midterm exam

PhD in Business Economics
Course: Bayesian Econometrics
Professor: Hedibert Freitas Lopes
Start: 8:30am, February 13th, 2020.
End: 8:30pm, February 15th, 2020.

Linear regression with double exponential errors

Consider the regression model

$$y_i = x_i'\beta + \epsilon_i, \qquad \epsilon_i|\lambda_i, \sigma^2 \sim N(0, \lambda_i\sigma^2),$$

where $x_i = (1, x_{i1}, \ldots, x_{iq})'$ is a p-dimensional vector of regressors (constant plus q attributes or characteristics) and the following hierarchical prior for the scale-mixing variables λ_i:

$$\lambda_1, \ldots, \lambda_n \ \text{ iid } \ \text{Exponential}(1/2).$$

Questions:

PART A) Show that

$$p(\epsilon_i|\sigma^2) = \int_0^\infty p(\epsilon_i|\lambda_i, \sigma^2)p(\lambda_i)d\lambda_i = \frac{1}{2\sigma}\exp\left\{-\frac{|\epsilon_i|}{\sigma}\right\},$$

a *double exponential* or *Laplace* distribution with scale parameter σ.

Hint: You might want to use the following result from Andrews and Mallows (1974)[1]

$$\int_0^\infty \exp\left\{-0.5\left(a^2u^2 + b^2u^{-2}\right)\right\}du = \sqrt{\frac{\pi}{2a^2}}\exp\{-|ab|\}.$$

PART B) Let $y = (y_1, \ldots, y_n)'$ and $X = (x_1, \ldots, x_n)'$. Describe a posterior simulator for fitting the above linear regression with double exponential errors, using independent priors for β and σ^2 of the form $N(\beta_0, V_0)$ and $IG(\nu_0/2, \nu_0\sigma_0^2/2)$. More specifically, show that the MCMC that samples from

$$p(\beta, \sigma^2, \lambda_1, \ldots, \lambda_n|\mathcal{D}_n),$$

where $\mathcal{D} = \{y, X\}$ is the data, cycles through the following full conditional distributions:

[1] Andrews and Mallows (1974) Scale mixtures of normal distributions. *Journal of the Royal Statistical Society, Series B*, 36, 99-102.

- $[\beta|\{\lambda_i\},\sigma^2,\mathcal{D}] \sim N(\beta_1, V_1)$, where

$$V_1 = (X'\Lambda^{-1}X/\sigma^2 + V_0^{-1})^{-1} \quad \text{and} \quad \beta_1 = V_1(X'\Lambda^{-1}y/\sigma^2 + V_0^{-1}\beta_0).$$

- $[\sigma^2|\Lambda,\beta,\mathcal{D}] \sim IG(\nu_1/2, \nu_1\sigma_1^2/2)$, where

$$\nu_1 = \nu_0 + n \quad \text{and} \quad \nu_1\sigma_1^2 = \nu_0\sigma_0^2 + (y - X\beta)'\Lambda^{-1}(y - X\beta).$$

- Scale-mixing variables λ_i

$$p(\lambda_i|\beta,\sigma^2,y_i,x_i) \propto \lambda^{-1/2}\exp\left\{-0.5\left(\lambda_i + \left(\frac{y_i - x_i'\beta}{\sigma}\right)^2\lambda_i^{-1},\right)\right\}. \tag{1}$$

  so you can go ahead and use a simple Metropolis-Hastings step to draw λ_i.

Hint I: It is easy (well, not so obviously easy!) to show that the above distribution (1) is of the *generalized inverse Gaussian* (GIG) form. One can draw from a GIG by inverting a draw from its cousing *inverse Gaussian* distribution. More specifically, if

$$1/\lambda_i \sim \text{invGauss}(1, |\sigma/(y_i - x_i'\beta)|),$$

then λ_i is a draw from (1). See *Hint II* below. Well, how to draw from the inverse Gaussian distribution? See *Hint III* below.

Hint II: A random variables $x \sim invGauss(\psi,\mu)$ if

$$p(x|\psi,\mu) \propto x^{-3/2}\exp\left\{-\frac{\psi(x-\mu)^2}{2x\mu^2}\right\}, \qquad x > 0.$$

Now, let $z = 1/x$. It follows by a change of variables that

$$p(z|\psi,\mu) \propto z^{-1/2}\exp\left\{-\frac{\psi}{2}\left[z + \mu^{-2}z^{-1}\right]\right\}.$$

Hint III: If $x \sim invGauss(\psi,\mu)$, then $\psi(x-\mu)^2/(x\mu^2) \sim \chi_1^2$ (Shuster, 1968). Also, Michael et al. (1976) show that when $\nu_0 \sim \chi_1^2$, the roots (x_1, x_2) of $\nu_0 = \psi(x-\mu)^2/(x\mu^2)$ can be used to sample from an $invGauss(\psi,\mu)$:

$$\begin{aligned} x_1 &= \mu + \frac{\mu^2\nu_0}{2\psi} - \frac{\mu}{2\psi}\sqrt{4\mu\psi\nu_0 + \mu^2\nu_0^2}, \\ x_2 &= \mu^2/x_1. \end{aligned}$$

They show that a draw x from $invGauss(\psi,\mu)$ can be obtained by setting $x = x_1$ with probability $\mu/(\mu + x_1)$ and $x = x_2$ with probability $x_1/(\mu + x_1)$.

PART C) Simulate $n = 200$ observations from the above linear regression with double exponential errors model, where $\beta = (0, 1, 2, 3)'$, $\sigma^2 = 1$ and $x_{ij} \sim N(0, 1)$. Implement the above MCMC scheme and produce posterior summaries of the main parameters. Would the simple MH algorithm to sample λ_i be reasonable for your simulated data or the full-fledge Gibbs sampler performs better?