Problem Set 3
ECON31740/PPHA48403

January 19, 2019

Part I

Due date : Tuesday, January 29, 2019

1. Do exercises 5, 7, 8, 9, 14, 15, 16, and 17 from Lecture 0.

Part II

Due date : Tuesday, January 29, 2019

4. **MCMC for Panel Data with Correlated Random Effects**

We are interested in fitting the model

$$g(y_{i,t}) = x_{i,t}\beta + \alpha_i,\ i = 1, ..., n,\ t = 1, ..., T, \tag{1}$$

where $\alpha_i \sim N(0, \tau^2)$ and g may stand for the probit or logit link. It may also stand for the conditional expectation, in which case we will study the linear regression function

$$y_{i,t} = x_{i,t}\beta + \alpha_i + \epsilon_{i,t},\ i = 1, ..., n,\ t = 1, ..., T, \tag{2}$$

where $\epsilon_{i,t} \sim N(0, \sigma^2)$.

In the general case (1), we may estimate the parameters of interest (β, τ) via the integrated likelihood

$$L(y|\beta, \tau) = \int f(y|\beta, \alpha) f(\alpha|\tau) d\alpha,$$

where, for instance, $f(y_i|\beta, \alpha_i) = \Pi_{t=1}^{T} f(y_{i,t}|\beta, \alpha_i)$.

The integration is tedious, and complicates both point estimation and inference. It is thus tempting to investigate the use of MCMC methods. One may candidly use Metropolis-Hastings to sample from the full posterior (using flat priors) $p(\beta, \alpha, \tau) \propto f(y|\beta, \alpha) f(\alpha|\tau^2)$ and simply integrate α out of the posterior, i.e., $p(\beta, \tau) = \int p(\beta, \alpha, \tau) d\alpha$. Practically, this simply means ignoring α draws in the MCMC output.

A. Standard Metropolis-Hastings

propose $(\beta^*, \tau^*) \sim q(\cdot|\beta^{(i)}, \tau^{(i)})$
propose $\alpha^* \sim q(\cdot|\tau^*)$
draw $u \sim U[0, 1]$
if

$$u \leq \frac{f(y|\beta^*, \alpha^*) f(\alpha^*|\tau^*)}{f(y|\beta^{(i)}, \alpha^{(i)}) f(\alpha^{(i)}|\tau^{(i)})} \frac{q(\alpha^{(i)}|\tau^{(i)}) q(\beta^{(i)}, \tau^{(i)}|\beta^*, \tau^*)}{q(\alpha^*|\tau^*) q(\beta^*, \tau^*|\beta^{(i)}, \tau^{(i)})}$$

then $(\beta^{(i+1)}, \tau^{(i+1)}, \alpha^{(i+1)}) = (\beta^*, \tau^*, \alpha^*)$
otherwise $(\beta^{(i+1)}, \tau^{(i+1)}, \alpha^{(i+1)}) = (\beta^{(i)}, \tau^{(i)}, \alpha^{(i)})$

Note that you can propose from the true distribution of α as we assume a normal distribution (otherwise you may use an easier-to-sample proposal, or use Metropolis-Hastings to sample from the proposal). Therefore, the acceptance probability will simplify to

$$\frac{f(y|\beta^*, \alpha^*)}{f(y|\beta^{(i)}, \alpha^{(i)})} \frac{q(\beta^{(i)}, \tau^{(i)}|\beta^*, \tau^*)}{q(\beta^*, \tau^*|\beta^{(i)}, \tau^{(i)})}. \tag{3}$$

Intuitively, we can be concerned that accepting each proposed β^* according to $f(y|\beta^*, \alpha^*)$, a very inaccurate estimate of $L(y|\beta^*, \tau^*)$, will impact mixing.

An admissible alternative is to sample more α's for each β and instead accept proposed β^*'s according to $p(\beta, \tau, \alpha_{(1)}, ..., \alpha_{(S)})$, for some moderate size S. This gives the following sampler.

B. Metropolis-Hastings for posterior with multiple α's

propose $(\beta^*, \tau^*) \sim q(\cdot|\beta^{(i)}, \tau^{(i)})$
propose $\overline{\alpha}^* = (\alpha^*_{(1)}, ..., \alpha^*_{(S)}) \sim f(\cdot|\tau^*)$
draw $u \sim U[0, 1]$
For $s = 1, ..., S$, if

$$u \leq \frac{\Pi_{s=1}^S f(y|\beta^*, \alpha^*_{(s)})}{\Pi_{s=1}^S f(y|\beta^{(i)}, \alpha^{(i)}_{(s)})} \frac{q(\beta^{(i)}, \tau^{(i)}|\beta^*, \tau^*)}{q(\beta^*, \tau^*|\beta^{(i)}, \tau^{(i)})}$$

then $(\beta^{(i+1)}, \tau^{(i+1)}, \overline{\alpha}^{(i+1)}) = (\beta^*, \tau^*, \overline{\alpha}^*)$
otherwise $(\beta^{(i+1)}, \tau^{(i+1)}, \overline{\alpha}^{(i+1)}) = (\beta^{(i)}, \tau^{(i)}, \overline{\alpha}^{(i)})$

We may instead tackle directly the posited issue that $f(y, \alpha^*|\beta^*)$ is too inaccurate an estimate of $L(y|\beta^*, \tau^*)$ and use a sample average to increase accuracy. Precisely, we can go the pseudo-marginal MCMC route, observing that

$$E\left[\frac{1}{S}\sum_{s=1}^{S} f(y|\beta, \alpha_{(s)})\right] = L(y|\beta, \tau), \tag{4}$$

where $\alpha_{(s)} \sim f(\alpha|\tau)$, $s = 1, ..., S$. This gives the following sampler.

C. Pseudo-Marginal MCMC

propose $(\beta^*, \tau^*) \sim q(\cdot|\beta^{(i)}, \tau^{(i)})$
propose $\overline{\alpha}^* = (\alpha^*_{(1)}, ..., \alpha^*_{(S)}) \sim f(\cdot|\tau^*)$
draw $u \sim U[0, 1]$
For $s = 1, ..., S$, if

$$u \le \frac{\frac{1}{S}\sum_{s=1}^{S} f(y|\beta^*, \alpha^*_{(s)})}{\frac{1}{S}\sum_{s=1}^{S} f(y|\beta^{(i)}, \alpha^{(i)}_{(s)})} \frac{q(\beta^{(i)}, \tau^{(i)}|\beta^*, \tau^*)}{q(\beta^*, \tau^*|\beta^{(i)}, \tau^{(i)})},$$

then $(\beta^{(i+1)}, \tau^{(i+1)}, \overline{\alpha}^{(i+1)}) = (\beta^*, \tau^*, \overline{\alpha}^*)$
otherwise $(\beta^{(i+1)}, \tau^{(i+1)}, \overline{\alpha}^{(i+1)}) = (\beta^{(i)}, \tau^{(i)}, \overline{\alpha}^{(i)})$

We did *not* use the simplification as in (3) again, the $\alpha^*_{(s)}$ are not variables in the posterior here, but the random element of the unbiased evaluation of the integrated likelihood.

a. Write the likelihood $f(y, \alpha|\beta)$ for model (1) with both logit and linear links.

b. In the case of the linear link, i.e., model (2), show that

$$\int f(y|\beta, \alpha) f(\alpha|\tau^2) d\alpha \propto N(x\beta, \sigma^2 + \tau^2).$$

c. Argue that the α draws collected from accepted proposals in algorithm C are posterior draws, i.e., they are distributed according to the α marginal of $p(\beta, \alpha, \tau) \propto f(y|\beta, \alpha) f(\alpha|\tau^2)$.

d. Implement the Metropolis-Hastings algorithms A, B, and C for both the logit and linear links. What can you conclude in terms of comparing the quality of mixing, and the amount of accumulated information in the posterior for a given computational budget? In the linear case, compare the output of A, B and C with that of a Metropolis-Hastings using the expression derived in b. In the linear case with known variances, compare posterior distributions to the truth, for which you obtained a closed form expression in b.

Part III

Due date : Thursday, January 31, 2019

1. Do exercises 1.3, 1.4, 1.12, 1.17, and 1.20 from Bertsimas and Tsitsiklis. For 1.3, 1.4, and 1.12 provide, in addition, the standard form formulation of the linear program.

2. Do exercises 2.3, 2.5, 2.6, 2.8, 2.10, 2.13 a, 2.14, 2.15, 2.17, and 2.19 a,b. from Bertsimas and Tsitsiklis.