1 Outline.

1. Motivation

2. Some examples

2 Motivation.

- Next, we discuss some recent development in Bayesian econometrics.

- In Bayesian econometrics, the econometrician acts as a rational decision maker, just like the agents in economic theory.

- The econometrician starts off with a prior distribution $p(\theta)$ about the model parameters.

- The econometrician observes some data $y = [y_1, ..., y_n]$

- The econometrician has a model, $f(y|\theta)$ which is the probability of observing y conditional on the parameters θ.

- The econometrician's posterior probability by Bayes Theorem is:

$$p(\theta|y) = \frac{p(\theta)f(y|\theta)}{\int p(\theta)f(y|\theta)d\theta}$$

- In the last decade, there has been an explosion in the applications of Bayesian approaches in statistics and econometrics.

- In Markov chain monte carlo, the econometrician simulates the poterior distribution $p(\theta|y)$.

- This involves simulating a markov chain where the invariant (or long run) distribution is exactly equal to the posterior.

- The output of this simulation is a sequence of pseudo random numbers $\theta^{(1)}, ..., \theta^{(S)}$.

- Let $f(\theta^{(1)}, ..., \theta^{(S)})$ denote the density that puts weight $\frac{1}{S}$ on each simulation draw.

- Then under suitably regularity conditions

$$f(\theta^{(1)}, ..., \theta^{(S)}) \to^d p(\theta|y)$$

- This posterior distribuiton expresses the econometricians beliefs about the parameters after seeing the data.

- We could use the posterior to:

1. Construct a 95 percent credible set (i.e. a set that has 95 percent posterior probability). This is analogous to a confidence interval.

2. Simulate the distribution of functions of the parameters, $g(\theta)$ as $\frac{1}{S}\sum_s g(\theta^{(s)})$.

3. Construct a predictive distribution for forecasting, i.e. simulate the model for each pseudo-random draw $\theta^{(1)}, ..., \theta^{(S)}$

- There are a few advantages to the Bayesian approach.

- The first is that it is very elegant numerically.

- There are some problems in latent variable, time series and other models that can only be solved through the use of Bayes.

- The second is that Bayes is exact in finite samples (up to our ability to approximate the posterior).

- Asymptotic theory depended on first order Taylor series expansion.

- Essentially, you are hoping that the first and second derivatives capture the behavior of the function.

- This can be a very poor approximation in finite samples in some cases.

- In Bayes, such linearizations are not required.

- However, you do need to be able to simulate the posterior accurately.

- Third, Bayes fits into decision theory and can help to guide rational decision making.

- For example, Rossi et. al. (1997) study the problem of target marketing.

- You see each household in a scanner panel data set choose a handful of times.

- You can use Bayes' theorem to update on the random coefficients of each household i.

- e.g. if there are 10,000 housholds, you have a posterior for the preferences parameters for each of the 10,000 household conditional on its purchase history.

- Using this posterior, you can form your posterior beliefs about the profits from sending a coupon to an individual household.

- Fourth, in Bayes you can form posteriors over models.

- Suppose that you have $m = 1, ..., M$ probability models, $f_m(y|\theta)$.

- Then $f(y|\theta) = \sum_{m=1}^{M} p_m f_m(y|\theta)$ where p_m is the probability of model m.

- You can use Bayes theorem to express your posterior probability p_m about model m.

- This is a very elegant way to handle non-nested model and might be superior to classical approaches to non-nested testing.

- Finally, in Bayesian econometrics, you can work with models that are not identified or that do not exhibit normal asymptotics.

- A flat likelihood does not affect your ability to construct a posterior.

3 Some examples.

- Two common ways to conduct MCMC are Gibbs sampling and Metropolis.

- A normal random walk metropolis works as follows.

- First, the econometrician comes up with a rough guess θ^0 at the MLE.

- Second, come up with a rough guess at I^0 at the information matrix using the hessian of the MLE.

- A sequence of psueorandom values $\theta^{(1)}, ..., \theta^{(S)}$ is drawn as follows. Given $\theta^{(s)}$, we draw $\theta^{(s+1)}$ as follows:

1. First, draw a candidate value $\widetilde{\theta} \sim N(\theta^{(s)}, I^0)$

2. Second, compute $\alpha = \min\{\frac{p(\widetilde{\theta})f(y|\widetilde{\theta})}{p(\theta^{(s)})f(y|\theta^{(s)})}, 1\}$

3. Set $\theta^{(s+1)} = \widetilde{\theta}$ with probability α and $\theta^{(s+1)} = \theta^{(s)}$ with probability α.

- Implimenting this algorithm simply requires the econometrician to evaluate the likihood repeatedly and draw normal deviates.

- A second algorithm for constructing a Markov Chain is Gibbs sampling.

- Partition parameters into $\theta_1, ..., \theta_d$ blocks

- Let $p_k(\theta_k|\theta_1, ..., \theta_{k-1}, \theta_{k+1}, ..., \theta_d)$ denote the conditional distribution of the kth block of parameters given the others.

- In some applications, this distribution can be convenient to form even if the entire likelihood is quite complicated!

- Starting with an initial value θ^0, Gibbs sampling works as follows. Given $\theta^{(s)}$

1. Draw $\theta_1^{(s+1)} \sim p_1(\theta_1|\theta_2^{(s)}, \theta_2^{(s)} ..., \theta_d^{(s)})$

2. Draw $\theta_2^{(s+1)} \sim p_2(\theta_1|\theta_1^{(s+1)}, \theta_3^{(s)}, \theta_4^{(s)} ..., \theta_d^{(s)})$

3. Draw $\theta_3^{(s+1)} \sim p_3(\theta_1|\theta_1^{(s+1)}, \theta_2^{(s+1)}, \theta_4^{(s)} ..., \theta_d^{(s)})$

⋮

d. Draw $\theta_d^{(s+1)} \sim p_d(\theta_1 | \theta_1^{(s+1)}, \theta_2^{(s+1)}, \theta_4^{(s+1)} ..., \theta_{d-1}^{(s+1)})$

d+1 Return to 1.

4 Simple Example.

- Suppose that $(y_1, y_2) \sim N(\theta, \Sigma)$.

- Suppose that the priors on $\theta \backslash$flat and $\mathbf{\Sigma}$ is known.

- Suppose that the matrix has diagnonal entries of 1 and off diagonal entries of ρ

- The posterior is then:

$$
\begin{aligned}
p(\theta, \mathbf{\Sigma}|y) &= \prod_{i=1}^{N} \frac{1}{(2\pi)^{n/2}(1-\rho^2)^{1/2}} \\
&\quad \exp(-\frac{1}{2}(y_n - \theta)'\mathbf{\Sigma}^{-1}(y_n - \theta)) \\
&= \prod_{i=1}^{N} \frac{1}{(2\pi)^{n/2}(1-\rho^2)^{1/2}} \\
&\quad \exp(-\frac{1}{2}\sum(y_n - \theta)'\mathbf{\Sigma}^{-1}(y_n - \theta))) \\
&= N(\widetilde{y}, N^{-1}\mathbf{\Sigma})
\end{aligned}
$$

- In the above, $\widetilde{y}$ denote the sample mean of y.
- Using rules about normal distributions, it follows that:

$$
\begin{aligned}
\theta_1|\theta_2 &\sim N(\widetilde{y}_1 + \rho(\theta_2 - \widetilde{y}_2), (1-\rho)/N) \\
\theta_2|\theta_1 &\sim N(\widetilde{y}_2 + \rho(\theta_2 - \widetilde{y}_1), (1-\rho)/N)
\end{aligned}
$$

5 More complicated example.

- Gibbs sampling is very useful in discrete choice models.

- We outline a simple example, see Rossi, Allenby et. al. Bayesian Statistics and Marketing (2006) for a more complete discussion of this model.