Estimating DSGE models using the Metropolis Algorithm

June 21, 2009

The Plan

1. Combining prior and sample information
2. Probability intervals of functions of θ
3. Recap of course

Most of today's lecture can be found in Canova's textbook *Methods for Applied Macroeconomics* (but slides+attendance should be sufficient).

Code and slides available at
http://www.kris-nimark.net/TS_UPF_2009.html

What do we mean by prior information?

- For a DSGE model, we may have information about "deep" parameters
  - Range of some paramors restricted by theory, e.g. risk aversion should be positive
  - Discount rate is inverse of average real interest rates
  - Price stickiness can be measured by surveys
- We may know something about the mean of a process

How do we combine prior and sample information?

Bayes rule:

$$\begin{aligned} P(\theta \mid Z)\, P(Z) &= P(Z \mid \theta)\, P(\theta) \\ &\Leftrightarrow \\ P(\theta \mid Z) &= \frac{P(\theta Z \mid \theta)\, P(\theta)}{P(Z)} \end{aligned}$$

- Since $P(Z)$ is a constant, we can use $P(Z \mid \theta)\, P(\theta)$ as the posterior likelihood (a likelihood function is any function that is proportional to the probability).

We now need to choose $P(\theta)$

Choosing prior distributions

The beta distribution is a good choice when parameter is in [0,1]

$$P(x) = \frac{(1-x)^{b-1}\, x^{a-1}}{B(a,b)}$$

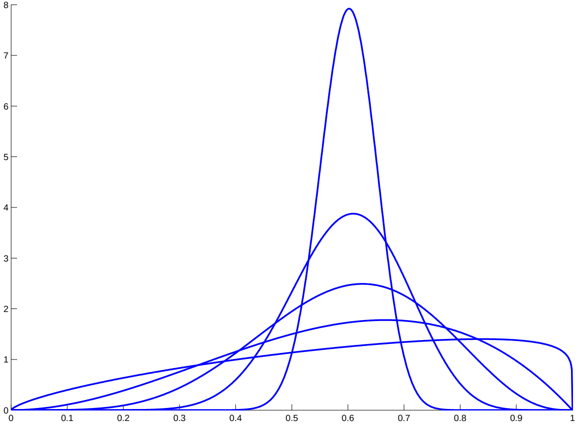
where

$$B(a,b) = \frac{(a-1)!(b-1)!}{(a+b-1)!}$$

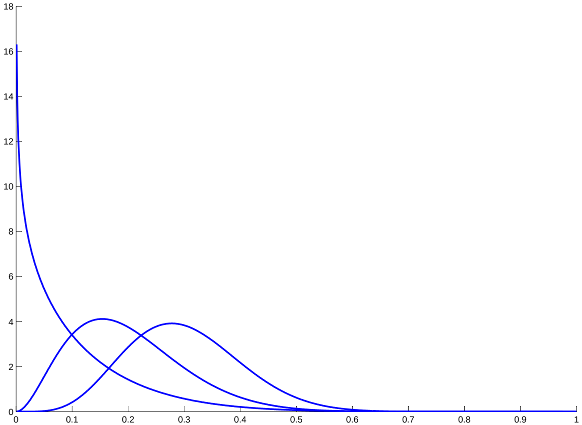
Easier to parameterize using expression for mean, mode and variance:

$$\begin{aligned} \mu &= \frac{a}{a+b}, \quad \hat{x} = \frac{a-1}{a+b-2} \\ \sigma^2 &= \frac{ab}{(a+b)^2\,(a+b+1)} \end{aligned}$$

Examples of beta distributions holding mean fixed



Examples of beta distributions holding s.d. fixed



Choosing prior distributions

The inverse gamma distribution is a good choice when parameter is positive

$$P(x) = \frac{b^a}{\Gamma(a)}(1/x)^{a+1} \exp(-b/x)$$

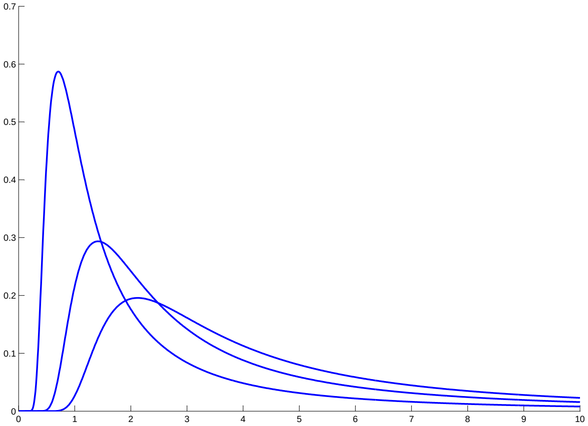
where

$$\Gamma(a) = (a-1)!$$

Again, easier to parameterize using expression for mean, mode and variance:

$$\begin{aligned} \mu &= \frac{b}{a-1}; a > 1, \quad \widehat{x} = \frac{b}{a+1} \\ \sigma^2 &= \frac{b^2}{(a-1)^2\,(a-2)}; a > 2 \end{aligned}$$

Examples of inverse gamma distributions



The Random-Walk Metropolis Algorithm with priors

1. Start with an arbitrary value θ_0
2. Update from θ_j to $\theta_{j+1} (j = 1, 2, \ldots J)$ by
   2.1 Generate $\theta^* \sim N(\theta_j, \Sigma)$
   2.2 Define
   $$\alpha = \min\left(\frac{L(Y \mid \theta^*)\, P(\theta^*)}{L(Y \mid \theta_j)\, P(\theta_j)}, 1\right) \quad (1)$$
   2.3 Take
   $$\theta_{j+1} = \left\{ \begin{array}{c} \theta^* \text{ with probability } \alpha \\ \theta_j \text{ otherwise} \end{array} \right\}$$

Practical implementation

Since we compute log-likelihood we can use the log of the likelihood ratio in the Random Walk Metropolis Algorithm

$$\begin{aligned} \ln \alpha &= \min [(\ln L \,(Y \mid \theta^*) + \ln P(\theta^*) \\ &\quad - \ln L \,(Y \mid \theta_j) - \ln P(\theta_j)), 0] \end{aligned}$$

If priors across parameters are independent we have that $\ln P(\theta_j) = \ln P(\theta_{1,j}) + \ln P(\theta_{2,j}) + ... + \ln P(\theta_{q,j})$ where

$$\theta_j = \begin{bmatrix} \theta_{1,j} & \theta_{2,j} & \cdots & \theta_{q,j} \end{bmatrix}'$$

Let's get deep with an old friend:

$$
\begin{aligned}
x_t &= \rho x_{t-1} + u_t^x \\
y_t &= E_t(y_{t+1}) - \frac{1}{\gamma}\left[r_t - E_t\left(\pi_{t+1}\right)\right] + u_t^y \\
\pi_t &= E_t\left(\pi_{t+1}\right) + \kappa\left[y_t - x_t\right] + u_t^\pi \\
r_t &= \phi_\pi \pi_t
\end{aligned}
$$

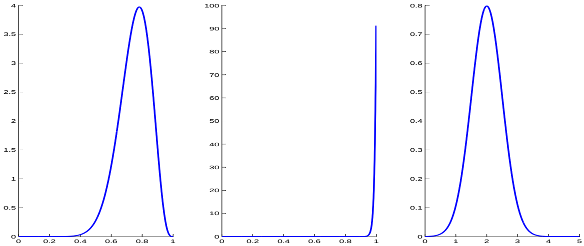
The parameter κ is in the benchmark 3-equation NK model given by

$$\kappa = \frac{(1-\delta)(1-\delta\beta)}{\delta}$$

where δ is the Calvo parameter of price stickiness and β is the discount factor. We now have a new parameter vector $\theta = \{\rho, \gamma, \delta, \beta, \phi, \sigma_x, \sigma_y, \sigma_\pi, \}$

The priors

- The prior on relative risk aversion γ is truncated Normal with mean 2 and s.d. 0.5.
- The prior on the Calvo parameter δ is Beta with mean 0.75 and s.d. 0.1
- The prior on the Calvo parameter β is Beta with mean 0.99 and s.d. 0.01



The Random-Walk Metropolis Algorithm with priors

1. Start with an arbitrary value θ_0
2. Update from θ_j to $\theta_{j+1} (j = 1, 2, \ldots J)$ by
   2.1 Generate $\theta^* \sim N(\theta_j, \Sigma)$
   2.2 Define
   $$\alpha = \min\left(\frac{L(Y \mid \theta^*) P(\theta^*)}{L(Y \mid \theta_j) P(\theta_j)}, 1 \right) \quad (2)$$
   2.3 Take
   $$\theta_{j+1} = \left\{ \begin{array}{c} \theta^* \text{ with probability } \alpha \\ \theta_j \text{ otherwise} \end{array} \right\}$$

The log prior

The log prior is given by

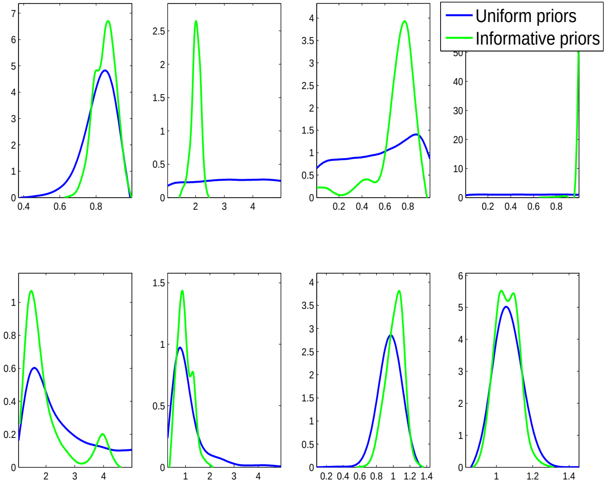
$$\begin{aligned} \ln P(\theta_j) &= \ln P(\theta_{1,j}) + \ln P(\theta_{2,j}) + ... + \ln P(\theta_{q,j}) \\ &= \ln P(\gamma_j) + \ln P(\delta_j) + \ln P(\beta_j) \end{aligned}$$

since we can ignore the (constant) probabilities on the uniform priors.
We then have that

$$\begin{aligned} \ln \left[L\left(Y \mid \theta^* \right) P(\theta^*) \right] &= \ln L(Z \mid \theta) + \ln P(\theta_j) \\ &\Longrightarrow \\ \alpha &= \min \left(\frac{\exp\left[\ln L(Z \mid \theta^*) + \ln P(\theta^*)\right]}{\exp\left[\ln L(Z \mid \theta_j) + \ln P(\theta_j)\right]}, 1 \right) \end{aligned}$$

i.e. all we need for the RWMA

Posterior with uniform and informative priors



Probability intervals of functions of θ

We now know how to find the posterior distribution of the parameter vector θ
We can use the MCMC to also construct probability intervals of any function $h(\theta)$

1. Draw an integer j on a uniform distribution between 1 and J
2. Compute $h(\theta_j)$ and save.
3. Repeat steps 1 and 2 "many" times (but usually fewer than J time is necessary).
4. Find percentiles of the saved outputs from $h(\theta)$. These are the probability intervals of $h(\theta)$.

Example: Probability intervals for impulse response function

From state space system

$$\begin{aligned} X_t &= AX_{t-1} + Cu_t \\ Z_t &= DX_t + v_t \end{aligned}$$

where

$$\begin{aligned} X_t &= x_t, A = \rho, Cu_t = u_t^x \\ Z_t &= \begin{bmatrix} y_t \\ \pi_t \end{bmatrix}, D = \begin{bmatrix} -\kappa\frac{\phi_\pi - \rho}{-c} \\ \kappa\gamma\frac{1-\rho}{-c} \end{bmatrix}, v_t = \begin{bmatrix} u_t^y \\ u_t^\pi \end{bmatrix} \end{aligned}$$

we can get IRF of $r_t = \phi\pi_t$ to a potential output shock $u_t^t mis$

$$\begin{aligned} \frac{\partial r_{t+s}}{\partial u_t^x} &= \phi_\pi D_2 AC \\ &= \phi_\pi \kappa\gamma\frac{1-\rho}{-c}\rho^s\sigma_x \end{aligned}$$

Example: Probability intervals for impulse response function

1. Draw an integer j on a uniform distribution between 1 and J
2. Compute $\phi_\pi \kappa \gamma \frac{1-\rho}{-c} \rho^s \sigma_x$ for s=0,1,2,3,...S and save.
3. Repeat steps 1 and 2 500 times.
4. Find percentiles of the saved outputs from $\phi_\pi \kappa \gamma \frac{1-\rho}{-c} \rho^s \sigma_x$ for each horizon s.
5. Plot.

Mean and 90% prob interval of IRF to u_t^x