Bayesian DSGE Model Estimation

Lecture five

Alexander Kriwoluzky

Universiteit van Amsterdam

February 5th, 2010

Aim of week 5

Evaluate the distribution around the posterior mode.

Aim of the week

Overview

Importance Sampling

MCMC sampling

Motivation

We want to describe the distribution $p(\mu|X)$, i.e. posterior moments:

$$E[g(\mu)] = \frac{\int g(\mu)P(\mu|x)d\mu}{\int P(\mu|x)d\mu} \tag{1}$$

The involved integrations are (usually) difficult to deduct analytically. One solution is to simulate the distribution and compute the statistic of interest given the sample.

Sampling

- Available distributions - the target distribution is of standard form.
- Effectively available distributions - $f(x,y)$ is the target distribution and not available. But $f(x)$ and $f(y|x)$ are.
- Transforming samples from available distributions - importance sampling.
- Markov chain monte carlo (MCMC).

Importance sampling

- Generate drawings from a stand-in distribution and re-weight.
- Importance sampler $I(\mu|\theta)$, with θ representing the parametrization of I.
- Extension of (1) yields:

$$E[g(\mu)] = \frac{\int g(\mu)w(\mu)I(\mu|\theta)d\mu}{\int w(\mu)I(\mu|\theta)d\mu} \tag{2}$$

with

$$w(\mu) = \frac{P(\mu|X)}{I(\mu|\theta)} \tag{3}$$

Implementation of importance sampling

Given the posterior mode μ and the inverse of the Hessian computed at the mode,

1. Define $I(\mu|\theta)$ as a multivariate t, with mean $\hat{\mu}$, covariance matrix $\Sigma_{\hat{\mu}}$ and ν degrees of freedom.
2. Generate $s = 1 \ldots N$ draws from $I(\mu|\theta)$.
3. Compute $\tilde{w}(\mu_s) = P(\mu_s)/I(\mu_s|\theta)$ and $w(\mu_s) = \tilde{w}(\mu_s)/\sum_{s=1}^{N} \tilde{w}(\mu_s)$.
4. Optionally go back to 1 and re-define θ.
5. Approximate the posterior expected value of a function $g(\mu)$ by $\sum_{s=1}^{N} w(\mu_s)g(\mu)$.

Aim of the week

Overview

Importance Sampling

MCMC sampling

Markov chain

- Definition Markov chain:
  A Markov chain is a sequence of random variables, such that the probability distribution of any one, given all preceding realizations, depends on the immediately preceding realization.
- A Markov chain can be fully described by its initial state and a rule describing how the chain moves from its state in t to a state in $t+1$: a transition kernel $\mathcal{K}(x,y)$.

$$\mathcal{K}(x,y) = P(X_{t+1} = y | X_t = x) \quad x,y \in \chi \tag{4}$$

  where χ denotes the sample space.
- For the remaining lecture: Let x and y denote two states of the sample space of μ.

Markov Monte Carlo chain

The required distribution is not standard. Construct a stochastic process, such that:

- it has a stationary distribution,
- it converges to that stationary distribution,
- the stationary distribution is the target distribution.

$\Rightarrow$ Construct a markov chain in μ, i.e. a transition kernel $\mathscr{K}(x,y)$, such that the process converges to the posterior distribution.

Construction of the transition kernel

- Denote $q(y|x)$ a proposal density, $\rho(x,y) = min\left(\frac{p(y)}{p(x)}, 1\right)$
- $\rightarrow \rho(x,y)q(y|x)$ is the probability that y is produced and accepted.
- $r(x) = \int \rho(x,y)q(y|x)dy$ is the sum of these probabilities.
- $(1 - r(x))\delta_x(y)$ thus is the probability of y being produced and not accepted. $(\delta_x(y))$ is a Dirac delta function.
- It can be shown that the transition kernel:

$$\mathcal{K}(x,y) = \rho(x,y)q(y|x) - (1 - r(x))\delta_x(y) \tag{5}$$

has as invariant distribution the target (posterior) distribution.

Random walk metropolis

- Choose $q(y|x) = q(y-x)$, with q is a multivariate density.
- $y = x + z$, z called the increment random variable
  $\Rightarrow$ Random walk metropolis.
- We assume the posterior distribution of μ to be asymptotically normal.
- Construct a Gaussian estimation around the posterior mode with a scaled version of the asymptotic covariance matrix as covariance matrix for the proposal distribution.
- It generates a sequence of dependent draws from the posterior distribution that can be averaged to approximate posterior moments.

Implementation RMW

Given the posterior mode μ and the inverse of the Hessian computed at the mode, start by drawing μ_0 from $\mathcal{N}(\hat{\mu}, c^2\Sigma_\mu)$. For $s = 1 \dots N$:

1. Draw μ_c from $\mathcal{N}(\mu_{s-1}, c^2\Sigma_\mu)$.
2. Accept μ_c ($\mu_s = \mu_c$) with probability $min \quad [1, r(\mu_{s-1}, \mu_c|X)]$ and reject it ($\mu_s = \mu_{s-1}$) otherwise.

$$r(\mu_{s-1}, \mu_c|X) = \frac{P(\mu_c|X)}{P(\mu_{s-1}|X)} \tag{6}$$

3. Approximate the posterior expected value of a function $g(\mu)$ by $1/N \sum_{s=1}^{N} g(\mu_s)$

How to choose c

- c is related to the acceptance rate of MCMC chain.
- As optimal acceptance rate for a diffusion process: from 0.44 for $d = 1$ parameter to 0.23 (for more than five ($d > 5$) parameters)
- $\Rightarrow$ Find the c that insures the optimal acceptance rate.
  1. Start with $c = 2.4/\sqrt{d}$.
  2. Increase (decrease) c, if the acceptance rate is too high (low).

Homework

Write a Matlab code that, given your posterior mode and its corresponding Covariance matrix, conducts a MCMC algorithm that samples from the posterior distribution!