MCMC Simulators: Numerical Accuracy

Eric Jacquier

Boston University Questrom School of Business
MF 840

Spring 2020

Setup and Goal of Simulation Sequences

Consider a Random Variable θ with distribution $\pi(\theta)$,
the goal of a MC simulation is to estimate for example

$$E_\theta(g(\theta)) = \int g(\theta)\pi(\theta)d\theta$$

- The Monte Carlo simulator makes S draws from $\pi(\theta)$, computing $g(\theta)$ for each draw:

$$g(\theta^{(s)}), \quad s = 1, \ldots, S$$

- Then estimates any moment of $g(\theta)$, e.g., the mean $E_\theta(g(\theta))$ by

$$\overline{g(\theta)} = \frac{1}{S}\sum_{1}^{S} g(\theta^{(s)})$$

- or the variance via the sample variance of the sequence: $\widehat{Var}(g(\theta^{(s)}))$
  … To do what?
  … To estimate the precision of $\overline{g(\theta)}$ .. Confidence interval based on

$$\overline{g(\theta)} \sim N(E_\theta(g(\theta)), \frac{Var_\theta(g(\theta))}{S}) \quad \frac{\sigma^2}{T}$$

- Rule of Thumb:

S must be large enough so the confidence interval is
narrow enough to not affect your last reported digit

Some Monte Carlo simulators are not i.i.d

What if the MC draws are not independent

$$Cov(g(\theta)^{(n+k)}, g(\theta)^{(n)}) \neq 0$$

Is this a problem?

- We can still use sample average. Ergodic Property:

$$\lim_{S \to \infty} \overline{g(\theta)} = E_\theta(g(\theta))$$

- But the simulation error is larger than with i.i.d. draws:

$$Var(\overline{g(\theta)}) \neq \frac{Var(g(\theta^{(n)}))}{S}$$

- Time Series Notation (use T as sample size):

$$y_t, \quad t = 1, \ldots, T, \quad \gamma_k = Cov(y_t, y_{t-k})$$

- Sample average:

$$\bar{y} = (\frac{1}{T}, \ldots, \frac{1}{T})(y_1, \ldots, y_T)' = \frac{1}{T}\mathbf{i}'\,\mathbf{y}$$

- Variance: $Var(\bar{y}) = Var(\frac{1}{T} i'y) = \frac{1}{T^2} i'\Sigma i$

$\Sigma_y = \sigma^2 I \Rightarrow Var(\bar{y}) = \frac{\sigma^2 T}{T^2} = \frac{\sigma^2}{T}$

$$Var(\bar{y}) = \frac{1}{T^2}\mathbf{i}'\,\Sigma_y \mathbf{i}$$

$\Sigma_y = \begin{bmatrix} \sigma & & \\ & \ddots & \\ & & \sigma \end{bmatrix} \begin{pmatrix} 1 & \rho_1 & \rho_2 & \cdots \\ \rho_1 & \ddots & & \\ \rho_2 & & \ddots & \\ \vdots & & & \\ \rho_{T-1} & & & 1 \end{pmatrix} \begin{bmatrix} \sigma & & \\ & \ddots & \\ & & \sigma \end{bmatrix}$

$= \begin{pmatrix} \sigma^2 & \gamma_1 & \sigma_2 & \cdots & \sigma_{T-1} \\ \vdots & & \ddots & & \\ \sigma_{T-1} & & & & \sigma_T^2 \end{pmatrix}$

. The sum of the terms in the Covariance matrix !

$i'\Sigma_y i$ = sum of all elements in Σ

T-1 autocorrelation term

$$\begin{aligned} Var(\bar{y}) &= \frac{1}{T^2}\left[T\gamma_0 + 2(T-1)\gamma_1 + 2(T-2)\gamma_2 + \ldots + 2\gamma_{T-1}\right] \\ &= \frac{\sigma_y^2}{T}\left[1 + \frac{2(T-1)}{T}\rho_1 + \frac{2(T-2)}{T}\rho_2 + \ldots + \frac{2}{T}\rho_{T-1}\right] \qquad (1) \end{aligned}$$

- Eq.(1) is the basis for a correct estimator of $Var(\bar{y})$.

高阶 auto correlation 很难算，数据可能不够.

We can pick the size S of a sample of draws as large as reasonable CPU time allows. We proceed as follows

autocorrelation becomes lower

1. Cut off Eq.(1) at some $M < T$.
   Choose M so that ρ_M is low.
   Note for the fans of asymptotics: for consistency M must grow at rate lower than $T^{1/4}$.
2. Estimate the ACF of the draws $\{\widehat{\rho_k}\}$, for $k \leq M$.
3. Compute (1) with $\{\widehat{\rho_k}\}$.

Relative Numerical Efficiency of a MCMC estimator relative to an ideal i.i.d scheme:

$$RNE = \frac{Var(g(\theta))/S}{Var(\overline{g(\theta)})} \qquad \frac{\text{naive}}{\text{actual}} \qquad \frac{\text{(without auto)}}{\text{(with auto)}} \qquad \frac{\sigma^2/T}{\sigma^2/T \times \text{Varinflation}}$$

You can use the command *numeff* in the package **bayesm**

10 times slower than if you have auto.