Bayesian Econometrics
Tutorial 04 - The Metropolis-Hastings Algorithm

Tutor: Richard Schnorrenberger

richard.schnn@stat-econ.uni-kiel.de

Institute for Statistics and Econometrics

Kiel University

Winter Term 2024/25

Review the Concepts and Proofs

- 1. What is a finite Markov chain? What is its transition matrix?
- 2. What is the stationary distribution of a finite Markov chain? How can you find it?
- 3. Define a continuous state Markov chain. What is different to a finite Markov chain?
- 4. Show that an autoregressive process of order 1 is a continuous state Markov chain.
- 5. What is the stationary distribution of a continuous state Markov chain? How can you find it?
- 6. Intuitively explain the detailed balance property.
- 7. For an autoregressive process of order 1, find the stationary distribution. Which condition do you have to impose?
- 8. Show that the Metropolis-Hastings algorithm gives rise to a Markov chain $\theta^{(s)}$ that is (i) reversible with respect to the posterior density $p(\theta^{(s)}|y)$, and (ii) stationary with stationary distribution $p(\theta^{(s)}|y)$.
- 9. How can the previous result be used to generate draws from the posterior distribution?
- 10. Explain the difference between the independence and random walk chain MH algorithms.
- 11. What is the posterior predictive p-value?
- 12. What is the Gelfand-Dey method?

Exercise 1

Consider estimation of the CES production function as outlined in the textbook. In particular, assume normality of the regression disturbances. Use the improper prior $p(\gamma, h) = 1/h$, $h > 0$.

- (a) Show that the marginal posterior pdf of γ is

$$p(\gamma|y) \propto \left[(y - f(X,\gamma))'(y - f(X,\gamma))\right]^{-\frac{N}{2}}$$

- (b) Write a pseudo code that uses an independence chain MH algorithm to estimate γ as the mean of the posterior distribution. Assume that the candidate distribution is a normal distribution with given mean $\hat{\gamma}$ and variance $\hat{\Sigma}$.
- (c) If the sample size is large, a good candidate distribution for the independence chain MH algorithm should be the asymptotic normal distribution based on a classical estimator. Find this distribution for the NLS estimator of the CES parameters. Then suppose for simplicity you estimate the CES function under the restriction $\gamma_4 = 1$ by OLS. Derive the mean vector $\hat{\gamma}$ and variance matrix $\hat{\Sigma}$ based on this estimator.

Illustration: MH algorithm

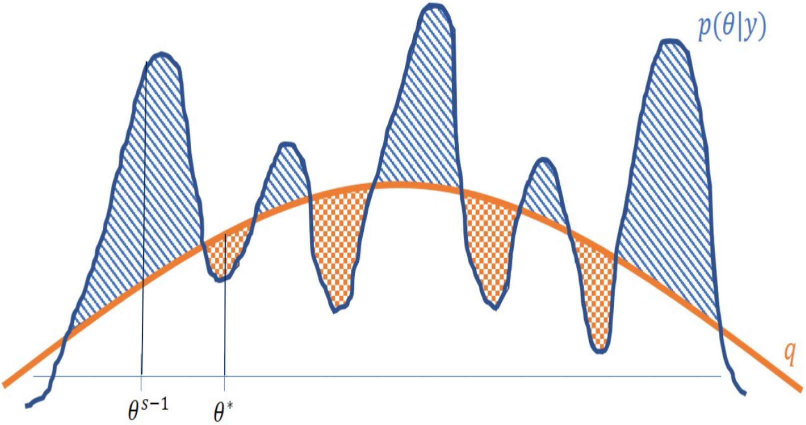

Figure: Example of a multimodal posterior $p(\theta \mid y)$ and proposal density $q(\cdot)$.

Solution to Exercise 1 (c)

Since the Bayesian model assumes homoscedasticity, we maintain this assumption for the asymptotic distribution of the NLS estimator:

$$\sqrt{N}(\hat{\gamma} - \gamma_o) \xrightarrow{\text{d}} \mathcal{N}(0, V), \quad \text{with} \quad V = -\left\{\mathrm{E}\left[H_i(\gamma_o)\right]\right\}^{-1} = \left\{\mathrm{E}\left[s_i(\gamma_o)s_i(\gamma_o)'\right]\right\}^{-1},$$

where $s_i(\gamma_o)$ is the score vector. Define $\delta_i \equiv \gamma_1 + \gamma_2 x_{i1}^{\gamma_4} + \gamma_3 x_{i2}^{\gamma_4}$ such that the log likelihood function is given by

$$l_i(\theta) = -\frac{1}{2}log(2\pi) - \frac{1}{2}log(\sigma^2) - \frac{1}{2\sigma^2}(\underbrace{y_i - \delta_i^{\frac{1}{\gamma_4}}}_{\varepsilon_i})^2$$

and the scores are

$$s_{i,1} = \frac{\partial l_i(\theta)}{\partial \gamma_1} = \frac{1}{\sigma^2}\varepsilon_i$$

$$s_{i,2} = \frac{\partial l_i(\theta)}{\partial \gamma_2} = \frac{1}{\sigma^2}\frac{1}{\gamma_4}\delta_i^{\frac{1}{\gamma_4}-1} x_{i1}^{\gamma_4}\,\varepsilon_i$$

$$s_{i,3} = \frac{\partial l_i(\theta)}{\partial \gamma_3} = \frac{1}{\sigma^2}\frac{1}{\gamma_4}\delta_i^{\frac{1}{\gamma_4}-1} x_{i2}^{\gamma_4}\,\varepsilon_i$$

$$s_{i,4} = \frac{\partial l_i(\theta)}{\partial \gamma_4} = \frac{1}{\sigma^2}\left(\frac{1}{\gamma_4}\delta_i^{\frac{1}{\gamma_4}-1}\left(\gamma_2 x_{i1}^{\gamma_4}\log x_{i1} + \gamma_3 x_{i2}^{\gamma_4}\log x_{i2}\right) - \frac{1}{\gamma_4^2}\delta_i^{\frac{1}{\gamma_4}}\log \delta_i\right)\varepsilon_i$$

Exercise 2

Consider estimation of the CES production function as outlined in the textbook. In particular, assume normality of the regression disturbances. Use the (mutually independent) informative priors $\gamma \sim \mathcal{N}(\underline{\gamma}, \underline{V})$ **and** $h \sim Gamma(\underline{s}^{-2}, \underline{\nu})$**, where** $\underline{\gamma} = [1, 1, 1, 1]'$**,** $\underline{V} = 0.25 I_4$**,** $\underline{\nu} = 12$**, and** $\underline{s}^{-2} = 10$**.**

- (a) Show that the conditional posterior for h is

$$h|\gamma, y \sim Gamma(\bar{s}^{-2}, \bar{\nu}),$$

where $\bar{\nu} = N + \underline{\nu}$ and $\bar{s}^2\bar{\nu} = \underline{s}^2\underline{\nu} + N\left[y - f(X, \gamma)\right]'[y - f(X, \gamma)]$.

- (b) Show that the conditional posterior for γ is proportional to

$$\gamma|h, y \propto \exp\left\{-\frac{h}{2}N\left[y - f(X, \gamma)\right]'[y - f(X, \gamma)] - \frac{1}{2}(\gamma - \underline{\gamma})'\underline{V}^{-1}(\gamma - \underline{\gamma})\right\}.$$

Solution to Exercise 2 (a)

Remember that the likelihood function of the CES regression is given by

$$p(y|\gamma, h) \propto h^{\frac{N}{2}} \exp\left\{-\frac{h}{2}\left[y - f(X,\gamma)\right]'\left[y - f(X,\gamma)\right]\right\} = h^{\frac{N}{2}} \exp\left\{-\frac{h}{2} N \bar{\omega}\right\}.$$

where $\bar{\omega} = \left[y - f(X,\gamma)\right]'\left[y - f(X,\gamma)\right]$. Adding the independent Normal-Gamma prior

$$p(\gamma, h) = p(\gamma)p(h) \propto \exp\left\{-\frac{1}{2}(\gamma - \underline{\gamma})'\underline{V}^{-1}(\gamma - \underline{\gamma})\right\} h^{\frac{\underline{\nu}}{2}-1} \exp\left\{-\frac{h\underline{\nu}}{2\underline{s}^{-2}}\right\}, \quad h > 0,$$

yields the joint posterior

$$p(\gamma, h|y) \propto \ h^{\frac{N+\underline{\nu}}{2}-1} \exp\left\{-\frac{h\underline{\nu}}{2\underline{s}^{-2}} - \frac{h}{2} N \bar{\omega}\right\} \exp\left\{-\frac{1}{2}(\gamma - \underline{\gamma})'\underline{V}^{-1}(\gamma - \underline{\gamma})\right\}.$$

Solution to Exercise 2 (a)

Since the joint distribution can be written as

$$p(\gamma, h|y) = p(h|\gamma, y)p(\gamma|y),$$

the conditional pdf for h must be proportional to that part of the joint pdf that contains h:

$$p(h|\gamma, y) \propto \; h^{\frac{N+\underline{\nu}}{2}-1} \exp\left\{-\frac{h\underline{\nu}}{2\underline{s}^{-2}} - \frac{h}{2}N\bar{\omega}\right\}.$$

Rearranging to

$$p(h|\gamma, y) \propto \; h^{\frac{N+\underline{\nu}}{2}-1} \exp\left\{-\frac{h(N+\underline{\nu})}{2}\left(\frac{\underline{s}^2\underline{\nu}}{N+\underline{\nu}} + \frac{N\bar{\omega}}{N+\underline{\nu}}\right)\right\}$$

and comparing to the *Gamma*$(\bar{s}^{-2}, \bar{\nu})$ pdf

$$f_G(h|\mu, \nu) = c_G^{-1} h^{\frac{\bar{\nu}}{2}-1} \exp\left[-\frac{h\bar{\nu}}{2}\bar{s}^2\right]$$

shows that the conditional posterior for h is indeed a Gamma pdf with parameters

$$\bar{\nu} = N + \underline{\nu}$$

and

$$\bar{s}^2 = \frac{\underline{s}^2\underline{\nu}}{N+\underline{\nu}} + \frac{N\bar{\omega}}{N+\underline{\nu}} = \frac{\underline{s}^2\underline{\nu}}{\bar{\nu}} + \frac{\omega}{\bar{\nu}} \quad \Rightarrow \quad \bar{s}^2\bar{\nu} = \underline{s}^2\underline{\nu} + N\bar{\omega}.$$

Solution to Exercise 2 (b)

We have shown above that the joint posterior is proportional to

$$p(\gamma, h|y) \propto \; h^{\frac{N+\underline{\nu}}{2}-1} \exp\left\{-\frac{h\underline{\nu}}{2\underline{s}^{-2}} - \frac{h}{2}N\bar{\omega}\right\} \exp\left\{-\frac{1}{2}(\gamma-\underline{\gamma})'\underline{V}^{-1}(\gamma-\underline{\gamma})\right\}.$$

Since the joint distribution can be written as

$$p(\gamma, h|y) = p(\gamma|h, y)p(h|y),$$

the conditional pdf for γ must be proportional to that part of the joint pdf that contains γ:

$$\begin{aligned} p(\gamma|h, y) &\propto \; \exp\left\{-\frac{h}{2}N\bar{\omega} - \frac{1}{2}(\gamma-\underline{\gamma})'\underline{V}^{-1}(\gamma-\underline{\gamma})\right\} \\ &\propto \exp\left\{-\frac{h}{2}N\left[y-f(X,\gamma)\right]'\left[y-f(X,\gamma)\right] - \frac{1}{2}(\gamma-\underline{\gamma})'\underline{V}^{-1}(\gamma-\underline{\gamma})\right\}. \end{aligned}$$

Exercise 2 (Computer-Based Exercises)

Consider estimation of the CES production function as outlined in the textbook. In particular, assume normality of the regression disturbances. Use the (mutually independent) informative priors $\gamma \sim \mathcal{N}(\underline{\gamma}, \underline{V})$ and $h \sim Gamma(\underline{s}^{-2}, \underline{\nu})$, where $\underline{\gamma} = [1, 1, 1, 1]'$, $\underline{V} = 0.25 I_4$, $\underline{\nu} = 12$, and $\underline{s}^{-2} = 10$.

- (c) Write a Matlab script that uses the random walk chain MH algorithm to estimate γ and h.
- (d) Extend your script to report numerical standard deviations for the γ's and for h based on Newey-West long-run variances (the function NeweyWest.m will be supplied in the tutorial).
- (e) Extend your script to report *CD* statistics for convergence based on subsamples $A = 10\%$, $B = 50\%$, and $C = 40\%$.
- (f) Write a Matlab script that uses the Gelfand-Dey method to compute the posterior odds ratio for models $M_1 : \gamma_4 = 1$ and $M_2 : \gamma_4$ is unrestricted. Suppose prior model probabilities are $p(M_1) = p(M_2) = 0.5$. Use a truncated normal pdf with truncation parameters $p = 0.01, 0.05, 0.1$.