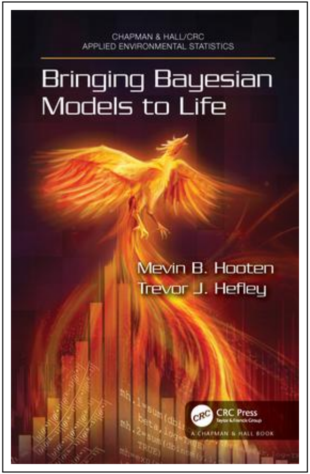
Markov Chain Monte Carlo 2

Mevin Hooten

Department of Statistics and Data Sciences
The University of Texas at Austin

Overview

- MC Integration
- Review of MCMC
- Two Parameter Model
  - Model Statement
  - Full-Conditionals
  - MCMC Algorithm
- Alternative Prior and MH



Monte Carlo Integration

- MC: Sample realizations from probability distribution: $\theta^{(k)} \sim [\theta]$ for $k = 1, \ldots, K$.

- Why is this helpful?

Monte Carlo Integration

- MC: Sample realizations from probability distribution: $\theta^{(k)} \sim [\theta]$ for $k = 1, \ldots, K$.

- Why is this helpful?

$$\int g(\theta)[\theta]d\theta \approx \frac{\sum_{k=1}^{K} g(\theta^{(k)})}{K}$$

Monte Carlo Integration

- MC: Sample realizations from probability distribution: $\theta^{(k)} \sim [\theta]$ for $k = 1, \ldots, K$.

- Why is this helpful?

$$\int g(\theta)[\theta]d\theta \approx \frac{\sum_{k=1}^{K} g(\theta^{(k)})}{K}$$

For example:

- Suppose you can obtain $\theta^{(k)} \sim [\theta|\mathbf{y}]$ for $k = 1, \ldots, K$.

- Then $\text{E}(\theta|\mathbf{y}) \approx \frac{\sum_{k=1}^{K} \theta^{(k)}}{K}$.

Review of MCMC

MCMC: Construct a sequence of random numbers that coincide with the stationary posterior distribution.

1. Set $k = 0$.
2. Choose initial value for $\theta_j^{(k)}$ for $j = 1, \ldots, p$.
3. Let $k = k + 1$.
4. Sample $\theta_j^{(k)} \sim [\theta_j|\cdot]$ for $j = 1, \ldots, p$.
5. Repeat 3 – 4.

Step 4: Full Conditionals

Sampling $\theta_j^{(k)} \sim [\theta_j|\cdot]$:

- Gibbs sample if conjugate.
- Metropolis-Hastings or other method if not.

Metropolis-Hastings

1. Sample proposal $\theta^{(\star)} \sim [\theta^{(\star)}|\theta^{(k-1)}]$.
2. Compute ratio: $r = \frac{[\mathbf{y}|\theta^{(\star)}][\theta^{(\star)}][\theta^{(k-1)}|\theta^{(\star)}]}{[\mathbf{y}|\theta^{(k-1)}][\theta^{(k-1)}][\theta^{(\star)}|\theta^{(k-1)}]}$.
3. Let $\theta^{(k)} = \theta^{(\star)}$ with probability min$(r, 1)$.

Proposal Distributions

$$r = \frac{[\mathbf{y}|\theta^{(\star)}][\theta^{(\star)}][\theta^{(k-1)}|\theta^{(\star)}]}{[\mathbf{y}|\theta^{(k-1)}][\theta^{(k-1)}][\theta^{(\star)}|\theta^{(k-1)}]}$$

1. Proposal does not influence results, in theory!
2. Some proposals yield more efficient algorithms than others.
   a. If $[\theta^{(k-1)}|\theta^{(\star)}] = [\theta^{(\star)}|\theta^{(k-1)}]$ then proposal cancels in r.
   b. If proposal is the prior then both the proposal and prior cancels in r.
   c. If the proposal is proportional to the full-conditional then $r = 1$.

MCMC Algorithms

1. Can contain both M–H and Gibbs updates.
2. Gibbs updates are just M–H with very smart proposals.
   a. If you sample a proposal from the full-conditional distribution then you can always keep it ($r = 1$).
   b. You just have to work out the full-conditional distribution analytically (not always possible).
   c. Often people call all MCMC algorithms “Gibbs Samplers,” even though they should be called “Metropolis-Hastings Samplers.”

Bayesian Normal Model

$$y_i \sim \text{N}(\mu, \sigma^2) \text{ , for } i = 1, \ldots, n$$

- μ: mean
- σ^2: error variance
- Note: same as $y_i = \mu + \varepsilon_i$, where $\varepsilon_i \sim \text{N}(0, \sigma^2)$.
- Specify prior distributions for $\mu \sim \text{N}(\mu_0, \sigma_0^2)$ and $\sigma^2 \sim \text{IG}(q, r)$ and find posterior:

$$[\mu, \sigma^2 | y_1, \ldots, y_n] = c \times \prod_{i=1}^{n} [y_i | \mu, \sigma^2][\mu][\sigma^2]$$

Directed Acyclic Graph

- Draw it.

Full-Conditional for μ

$$
\begin{aligned}
[\mu|\cdot] &\propto \prod_{i=1}^{n}[y_i|\mu,\sigma^2][\mu] \\
&\propto \prod_{i=1}^{n}\mathsf{N}(\mu,\sigma^2)\times\mathsf{N}(\mu_0,\sigma_0^2) \\
&= \mathsf{N}(\tilde{\mu},\tilde{\sigma}^2)
\end{aligned}
$$

where,

$$
\begin{aligned}
\tilde{\sigma}^2 &= \left(\frac{n}{\sigma^2}+\frac{1}{\sigma_0^2}\right)^{-1} \\
\tilde{\mu} &= \tilde{\sigma}^2\left(\frac{\sum_{i=1}^{n} y_i}{\sigma^2}+\frac{\mu_0}{\sigma_0^2}\right)
\end{aligned}
$$

Full-Conditional for σ^2

$$
\begin{aligned}
[\sigma^2|\cdot] &\propto \prod_{i=1}^{n} [y_i|\mu,\sigma^2][\sigma^2] \\
&\propto \prod_{i=1}^{n} \mathsf{N}(\mu,\sigma^2) \times \mathsf{IG}(q,r) \\
&= \mathsf{IG}(\tilde{q},\tilde{r})
\end{aligned}
$$

where,

$$
\begin{aligned}
\tilde{q} &= \frac{n}{2} + q \\
\tilde{r} &= \left(\frac{\sum_{i=1}^{n}(y_i-\mu)^2}{2} + r \right)
\end{aligned}
$$

MCMC Algorithm

1. Set $k = 0$
2. Choose initial value for $\mu^{(k)}$
3. Let $k = k + 1$
4. Sample $(\sigma^2)^{(k)} \sim \text{IG}(\tilde{q}, \tilde{r}^{(k-1)})$
5. Sample $\mu^{(k)} \sim \text{N}(\tilde{\mu}^{(k)}, (\tilde{\sigma}^2)^{(k)})$
6. Repeat 3 – 5

Gibbs Sample for σ^2

```
q.tilde=n/2+q
r.tilde=(r+0.5*sum((y-mu)^2))

s2=1/rgamma(1,q.tilde,r.tilde)
```

Gibbs Sample for μ

```
s2.tilde=1/(n/s2+1/s2.0)
mu.tilde=s2.tilde*(sum(y)/s2+mu.0/s2.0)

mu=rnorm(1,mu.tilde,s2.tilde)
```

New Prior for σ^2

$$\log(\sigma^2) \sim \mathrm{N}(\mu_l, \sigma_l^2)$$

$$[\log(\sigma^2)|\cdot] \propto \prod_{i=1}^{n} [y_i|\beta_0, \beta_1, \sigma^2][\log(\sigma^2)]$$
$$= \text{?}$$

Use M-H:

$$\log(\sigma^2)^{(*)} \sim \mathrm{N}(\log(\sigma^2)^{(k-1)}, \sigma^2_{\text{tune}})$$

$$\text{mh} = \frac{\prod_i [y_i|\mu^{(k)}, (\sigma^2)^{(*)}][\log(\sigma^2)^{(*)}]}{\prod_i [y_i|\mu^{(k)}, (\sigma^2)^{(k-1)}][\log(\sigma^2)^{(k-1)}]}$$

M-H Sample for σ^2

```
s2.star=exp(rnorm(1,log(s2),s.tune))

mh.1=sum(dnorm(y,mu,sqrt(s2.star),log=TRUE))+dnorm(log(s2.star),mu.l,sig.l,log=TRUE)

mh.2=sum(dnorm(y,mu,sqrt(s2),log=TRUE))+dnorm(log(s2),mu.l,sig.l,log=TRUE)

mh=exp(mh.1-mh.2)

if(mh>runif(1)){
 s2=s2.star
}
```