Markov chain Monte Carlo II

Bayesian Modeling for Socio-Environmental Data

N. Thompson Hobbs

July 21, 2016

SESYNC

The MCMC algorithm

- Some intuition
- Accept-reject sampling with Metropolis algorithm
- Introduction to full-conditional distributions
- Gibbs sampling (exercise)
- Metropolis-Hastings algorithm
- Implementing accept-reject sampling

Metropolis Updates

$$[\boldsymbol{\theta}^{*k+1}|y] = \frac{\overbrace{[y|\boldsymbol{\theta}^{*k+1}]}^{\text{likelihood}}\overbrace{[\boldsymbol{\theta}^{*k+1}]}^{\text{prior}}}{\int [y|\boldsymbol{\theta}][\boldsymbol{\theta}]d\boldsymbol{\theta}}$$

$$[\boldsymbol{\theta}^{k}|y] = \frac{\overbrace{[y|\boldsymbol{\theta}^{k}]}^{\text{likelihood}}\overbrace{[\boldsymbol{\theta}^{k}]}^{\text{prior}}}{\int [y|\boldsymbol{\theta}][\boldsymbol{\theta}]d\boldsymbol{\theta}}$$

$$R = \frac{[\boldsymbol{\theta}^{*k+1}|y]}{[\boldsymbol{\theta}^{k}|y]}$$

Metropolis Updates

$$[\boldsymbol{\theta}^{*k+1}|y] = \frac{\overbrace{[y|\boldsymbol{\theta}^{*k+1}]}^{\text{likelihood}}\overbrace{[\boldsymbol{\theta}^{*k+1}]}^{\text{prior}}}{\cancel{\int [y|\boldsymbol{\theta}][\boldsymbol{\theta}]d\boldsymbol{\theta}}}$$

$$[\boldsymbol{\theta}^{k}|y] = \frac{\overbrace{[y|\boldsymbol{\theta}^{k}]}^{\text{likelihood}}\ \overbrace{[\boldsymbol{\theta}^{k}]}^{\text{prior}}}{\cancel{\int [y|\boldsymbol{\theta}][\boldsymbol{\theta}]d\boldsymbol{\theta}}}$$

$$R = \frac{[\boldsymbol{\theta}^{*k+1}|y]}{[\boldsymbol{\theta}^{k}|y]}$$

1. Accept-reject samplers
   1.1 Metropolis: requires a symmetric proposal distribution (e.g., normal, uniform)
   1.2 Metropolis-Hastings: allows asymmetric proposal distributions (e.g., beta, gamma, lognormal)
2. Gibbs: accepts all proposals because they come directly from the posterior using conjugates.

Metropolis-Hastings updates

- Metropolis updates require symmetric proposal distributions (e.g., uniform, normal) : $[\theta^{*k+1}|\theta^k] = [\theta^k|\theta^{*k+1}]$
- When proposal distributions are asymmetric ($[\theta^{*k+1}|\theta^k] \neq [\theta^k|\theta^{*k+1}]$, e.g., beta, gamma, lognormal) we must use Metropolis-Hastings updates.

Metropolis-Hastings updates

Metropolis R:

$$R = \frac{[\boldsymbol{\theta}^{*k+1}|y]}{[\boldsymbol{\theta}^{k}|y]} \tag{1}$$

Metropolis-Hastings R:

$$R = \frac{[\boldsymbol{\theta}^{*k+1}|y] \quad \overbrace{[\boldsymbol{\theta}^{k}|\boldsymbol{\theta}^{*k+1}]}^{\text{Proposal distribution}}}{[\boldsymbol{\theta}^{k}|y] \quad \underbrace{[\boldsymbol{\theta}^{*k+1}|\boldsymbol{\theta}^{k}]}_{\text{Proposal distribution}}} \tag{2}$$

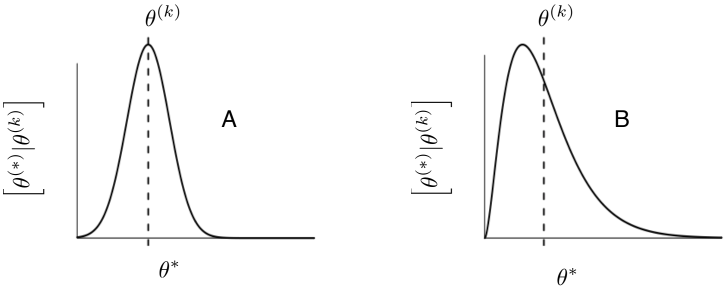
Proposal distributions

- Independent chain have proposal distributions that do not depend on the current value (θ^k) in the chain.
- Dependent chains, as you might expect, have proposal distributions that *do* depend on the current value of the chain (θ^k). In this case we draw from

$$[\theta^{*k+1}|\theta^k, \sigma] \tag{3}$$

  where σ is a tuning parameter that we specify to obtain an acceptance rate of about 40%. Note that my notation and notation of others simplifies this distribution to $[\theta^{*k+1}|\theta^k]$ The sigma is implicit because it is a constant, not a random variable.
- Why are dependent chains usually more efficient that independent chains?

Proposal distributions



Definition of symmetry

A proposal distribution is symmetric if and only if

$$[\theta^{*k+1}|\theta^k] = [\theta^k|\theta^{*k+1}]. \tag{4}$$

Normal and uniform are symmetric. Gamma, beta, lognormal are not.

Illustrating with code

```
#symmetric example
sigma=1
x = .8
z=rnorm(1,mean=x,sd=sigma);z
#P(z|x)
dnorm(z,mean=x,sd=sigma)
#P(x|z)
dnorm(x,mean=z,sd=sigma)
#asymmetric example
sigma=1
x = .8
a.x=x^2/sigma^2; b.x=x/sigma^2
z=rgamma(1,shape=a.x,rate=b.x);z
a.z=z^2/sigma^2; b.z=z/sigma^2
#P(z|x)
dgamma(z,shape=a.x,rate=b.x)
#P(x|z)
```

Example using beta proposal distribution

1. Current value of parameter, $\theta^k = .42$, tuning parameter set at $\sigma = .10$
2. Make a draw from $\theta*^{k+1} \sim \text{beta}(m(.42,.10))$, where m is moment matching function.
3. Calculate $R = \frac{[\theta^{*k+1}|y]\overbrace{[.42|m(\theta^{*k+1},\sigma]}^{\text{beta}}}{[\theta^k|y]\underbrace{[\theta^{*k+1}|m(.42,\sigma)]}_{\text{beta}}}$.
4. Choose proposed or current value based on R as we did with Metropolis.

MCMC

- Methods based on the Markov chain Monte Carlo algorithm allow us to approximate marginal posterior distributions of unobserved quantities without analytical integration.
- This makes it possible to estimate models that have many parameters, have multiple sources of uncertainty, and include latent quantities.
- We will learn a tool, JAGS, that simplifies the implementation of MCMC methods.
- Will will put this tool to use in building models that include nested levels in space, errors in the observations, differences among individuals an locations and processes that unfold over time.