

Recap:

• We want $\hat{p}(\theta_1, \dots, \theta_n | \vec{y})$.
samples from

• We view "parameter space" as a "state space" as a physical space: e.g. let $n=3$:



• We view MCMC as a particle moving through parameter space:



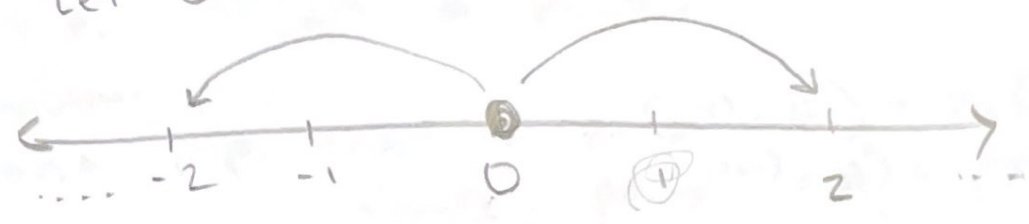
• The "history" of where the particle has been in θ approximates $p(\theta | \vec{y})$.

Recap: Metropolis-Hastings (MH).

Ex 1 Let θ be a r.v. w/ support over the integers

$$\theta \in \{-\infty, \dots, -2, -1, 0, 1, 2, \dots, \infty\}$$

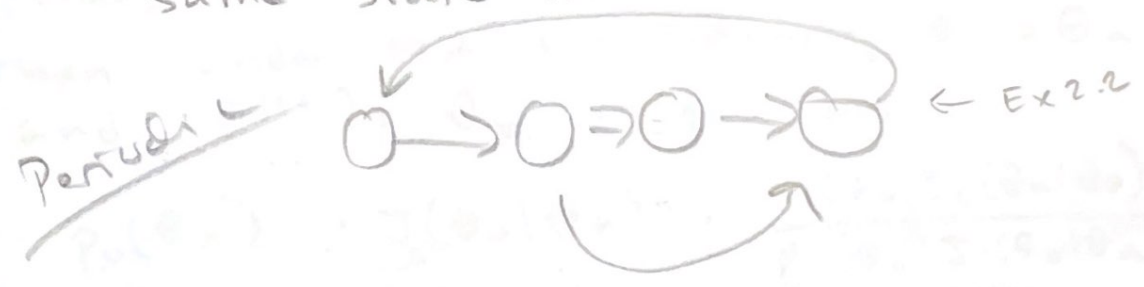
$$\text{Let } \theta^{(0)} = 0.$$



$$\text{Let } J(\theta | \theta^{(s)}) = \theta^{(s)} \pm 2$$

Our Markov chain is reducible.

Ex 2.1 Imagine we cannot stay in the same state in the ex. above.



EX
not
recurrent

Imagine discrete $\theta \in \begin{Bmatrix} A \\ B \\ C \end{Bmatrix}$

If we do not continue (sampling state A),
 $\Pr(\theta^{(s)} \in A) \rightarrow 0$ as $s \rightarrow \infty$

$\pi(\theta) :=$ stationary distr.

$p_0(\theta) :=$ posterior distr. "target" of MH MCMC

Let θ discrete r.v.

T.P. $\pi(\theta) = p_0(\theta)$

put another way

I want t.p. that if $\Pr(\theta^{(s)} = \theta) = p_0(\theta)$
then $\Pr(\theta^{(s+1)} = \theta) = p_0(\theta)$

because if p_0 is a stationary distr.
it is the stationary distr. by uniqueness.

Let θ_a & θ_b be two values of θ s.t.

$$p_0(\theta_a) \cdot J_s(\theta_b | \theta_a) \geq p_0(\theta_b) J_s(\theta_a | \theta_b)$$

Then under MH the probability $\theta^{(s)} = \theta_a$
and $\theta^{(s+1)} = \theta_b$ is given by

$$\begin{aligned}
 & \cancel{p_0(\theta_a)} \cdot \cancel{J_s(\theta_b | \theta_a)} \cdot \frac{p_0(\theta_b) J_s(\theta_a | \theta_b)}{\cancel{p_0(\theta_a) J_s(\theta_b | \theta_a)}} \\
 & \text{prob ending up in state } \theta_a \quad \text{prob proposing } \theta_b \text{ from } \theta_a \quad \text{prob accepting}
 \end{aligned}$$

$$= p_0(\theta_b) J_s(\theta_a | \theta_b)$$

Other way: $\Pr(\theta^{(s)} = \theta_b, \theta^{(s+1)} = \theta_a)$

$$p_0(\theta_b) \cdot J_s(\theta_a | \theta_b) \cdot 1$$

So $\wedge \theta_a, \theta_b$

$$\begin{aligned} & P_r(\theta^{(s)} = \theta_a, \theta^{(s+1)} = \theta_b) \\ &= P_r(\theta^{(s)} = \theta_b, \theta^{(s+1)} = \theta_a) \end{aligned}$$

★

T. P. $P_r(\theta^{(s+1)} = \theta) = P_0(\theta)$

$$= P_r(\theta^{(s)} = \theta)$$

Proof

$$P_r(\theta^{(s+1)} = \theta) = \sum_{\theta_A} P_r(\theta^{(s+1)} = \theta, \theta^{(s)} = \theta_A)$$

by ★

$$= \sum_{\theta_A} P_r(\theta^{(s+1)} = \theta_A, \theta^{(s)} = \theta)$$

$$= P_r(\theta^{(s)} = \theta) \quad \blacksquare$$