23 Mar 2026

Markov Chain Monte Carlo

Sampling from an unnormalized distribution:

- Given a state set X
  (typically size at least 2^n where n is the size of the "problem description")
- Given a function $w: X \to [0, \infty)$ such that $Z = \sum_{x \in X} w(x) > 0$
- Goal: draw one random sample from the distribution $\pi(x) = \frac{w(x)}{Z}$.
  (exact sampling)

  or draw one random sample from a distribution π' s.t.

  $$\|\pi - \pi'\|_1 = \sum_{x \in X} |\pi(x) - \pi'(x)| < \varepsilon.$$

  (approximate sampling)

Note. We typically don't have time to compute Z. We want to draw samples from $\pi = \frac{w}{Z}$ anyway.

MCMC (Markov Chain Monte Carlo):

Design an ergodic MC with transition matrix P satisfying $w \cdot P = w$.

$\therefore \quad \pi \cdot P = \pi \implies \pi$ is the stationary distrib of P.

Def. P is reversible w.r.t. π if

$$\forall x, y \in X \qquad \boxed{\pi_x P_{xy} = \pi_y P_{yx}}$$

Lem. If P reversible w.r.t. π then π is a stat distrib for P.

Proof. Must show $\pi P = \pi$ i.e.

$$\forall y \qquad \sum_{x \in X} \pi_x P_{xy} = \pi_y$$

$$\sum_{x \in X} \pi_x P_{xy} = \sum_{x \in X} \pi_y P_{yx} = \pi_y \left(\sum_{x \in X} P_{yx} \right) = \pi_y$$

Metropolis-Hastings rule.

Given a symmetric stochastic matrix $K = (K_{xy})_{x,y \in X}$

called the "proposal distribution" (row sums = 1)

(think of K as the normalized adjacency matrix of the state transition graph)

For $x \neq y$ define

$$P_{xy} = K_{xy} \cdot \frac{\min\{w(x), w(y)\}}{w(x)}$$

For $x = y$

$$P_{xx} = 1 - \sum_{y \neq x} P_{xy}$$

Note. ① Row sums of P equal 1 by design.

② Entries of P are ≥ 0.

off-diagonal: obvious from formula.

on-diagonal: use $P_{xy} \leq K_{xy}$ for $y \neq x$

$$\sum_{y \neq x} P_{xy} \leq \sum_{y \neq x} K_{xy} \leq 1$$

$$P_{xx} = 1 - \sum_{y \neq x} P_{xy} \geq 0.$$

①+②: P is a stochastic matrix. So $\exists$ a MC with trans mtx P.

P reversible wrt $\pi = \frac{w}{Z}$ ...

$$\pi_x P_{xy} = \frac{\cancel{w(x)}}{Z} \cdot K_{xy} \cdot \frac{\min\{w(x), w(y)\}}{\cancel{w(x)}}$$

$$= K_{xy} \cdot \frac{\min\{w(x), w(y)\}}{Z}$$

$$\pi_y P_{yx} = K_{yx} \cdot \frac{\min\{w(y), w(x)\}}{Z}$$

K symmetric $\Rightarrow K_{xy} = K_{yx}$

Procedural interpretation of $P_{xy} = K_{xy} \cdot \frac{\min\{w(x), w(y)\}}{w(x)}$

0. Suppose current state is x.
1. "Propose" a next state, y, with probability K_{xy}.
2. With probability $\frac{\min\{w(x), w(y)\}}{w(x)}$ accept the proposal $\Longrightarrow$ transition to y.
3. Else, remain at x.

<u>Example.</u> Ising model.

$\mathcal{X}$ = functions $x : V(G) \to \{\pm 1\}$.

$$w(x) = \exp\Big(c \cdot \sum_{(u,v) \in E(G)} x(u) \cdot x(v) \Big) \qquad c > 0.$$

"inverse temperature"

Glauber dynamics uses

$$K_{xy} = \begin{cases} \frac{1}{|V(G)|} & \text{if } x,y \text{ differ at a single vertex} \\ \emptyset & \text{o.w.} \end{cases}$$

Metropolis-Hastings says:

1. Propose to relabel one random vertex in $V(G)$.

2. Accept the proposal with prob.

$$\frac{\min\{w(x), w(y)\}}{w(x)} = \min\left\{1, \frac{w(y)}{w(x)}\right\}.$$

$$w(y) = \exp\left(c \cdot \sum_{(u,v)\in E} y(u)y(v) \right)$$

$$w(x) = \exp\left(c \cdot \sum_{(u,v)\in E} x(u)x(v) \right)$$

$$\frac{w(y)}{w(x)} = \exp\left(c \cdot \sum_{(u,v)\in E} \left[y(u)y(v) - x(u)x(v) \right] \right)$$

If x,y differ at r: let $\partial r = \{\text{edges with endpoint } r\}$

$$\frac{w(y)}{w(x)} = \exp\left(c \cdot \sum_{(r,v)\in \partial r} \left[y(r)y(v) - x(r)x(v) \right] \right)$$

$$= \exp\left(c \cdot (y(r) - x(r)) \cdot \sum_{(r,v)\in \partial r} x(v) \right)$$

Def. The ε-mixing time of a Markov chain with transition matrix P is the smallest $\tau(\varepsilon)$ such that

time step — starting distrib

$$\forall t \geq \tau(\varepsilon) \quad \forall \pi_0$$

distrib at time t is ε-close to stationary.

$$\boxed{\left\| \left(\pi_0 - \pi_{stationary} \right) P^t \right\|_1 \leq \varepsilon} .$$

$\pi_0 P^t$ = marginal distrib of state @ time t

$\pi_{stationary} P^t = \pi_{stationary}$.

ε-mixing time is analogous to running time for MCMC algorithms.