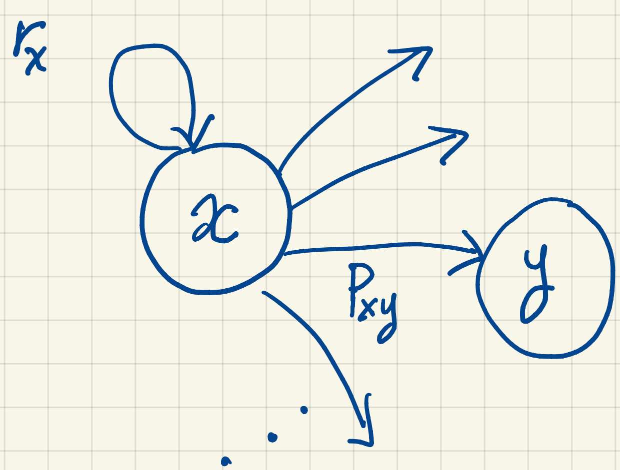
Proofs of detailed balance



$$r_x + \sum_{y \neq x} P_{xy} = 1$$

$$\sum_{y} P_{xy} = 1 \quad (r_x = P_{xx})$$

Assume $\boxed{\pi_i P_{ij} = \pi_j P_{ji}}$

Then
$$\sum_{i \neq j} \pi_i P_{ij} + \pi_j r_j$$
$$= \sum_{i \neq j} \pi_j P_{ji} + \pi_j r_j$$
$$= \pi_j \left[\sum_{i \neq j} P_{ji} + r_j \right]$$
$$= \pi_j$$

$$\begin{tikzcd} x \arrow[loop above, "r_x"] \arrow[r, "f(y|x)"] & y \end{tikzcd}$$

Conditioned on x, y has density $f(y|x)$

$$r_x + \underbrace{\int f(y|x)\,dy}_{1 - r_x} = 1$$

$$p(y|x) = r_x\,\delta(y-x) + f(y|x)$$

We can verify:

(1) $\int p(y|x)\,dy = 1$

(2) $P(y=x|x) = r_x$

$$(1):\ \int p(y|x)\,dy = \int r_x\,\delta(y-x)\,dy + \int f(y|x)\,dy$$

$$= r_x + 1 - r_x = 1$$

$$(2):\ P(y=x|x) = \int_x^x p(y|x)\,dy$$

$$= \int_x^x r_x\,\delta(y-x)\,dy + \int_x^x f(y|x)\,dy$$

$$= r_x + 0 = r_x$$

Assume denisty $\pi(x)$ exists such that $\pi(x) f(y|x) = \pi(y) f(x|y)$

$$\int p(y|x)\,\pi(x)\,dx = \int f(y|x)\,\pi(x)\,dx + \int r(x)\,\delta(y-x)\,\pi(x)\,dx$$

$$= \int f(x|y)\,\pi(y)\,dx + r(y)\,\pi(y)$$

$$= \pi(y)\left[\int f(x|y)\,dx + r(y)\right]$$

$$\pi(y)\left[1 - r(y) + r(y)\right] = \pi(y)$$

This means two things

1) Given P_{ij}, we can possibly solve for π (the stationary distribution) by solving the detailed balance equations.

$$\pi_i P_{ij} = \pi_j P_{ji}$$

2) Given a target distribution π, we can try to tweak the P_{ij}'s to satisfy detailed balance equations, so that the chain will have the desired stationary distribution.

To do (2) we need to introduce the idea of rejection sampling, which will lead to the Metropolis-Hastings alg.

Rejection sampling:

Given $q(x)$ that we know how to sample from, here's how we can sample from a desired $p(x)$.

- Assume $\exists\ c \geq 1$ such that $\forall x.\ c\,q(x) \geq p(x)$
- Sample $x \sim q(x)$ (proposal distribution)
- Accept the sampled x with probability $\dfrac{p(x)}{c\,q(x)}$

So $P(\text{accept} \mid x) = \dfrac{p(x)}{c\,q(x)} \leq 1$.

$$P(\text{accept}) = \int P(\text{accept} \mid x)\, q(x)\, dx = \frac{1}{c}$$

(so it's better if c is small).

Proof that rejection sampling works.

$$f(x \mid \text{accept}) = \frac{P(\text{accept} \mid x)\; q(x)}{\int P(\text{accept} \mid x)\, q(x)\, dx}$$

$$= \frac{\frac{p(x)}{c\,q(x)}\; q(x)}{\int \frac{p(x)}{c\,q(x)}\, q(x)\, dx} = \frac{p(x)/c}{\int \frac{1}{c} p(x)\, dx} = p(x).$$

Example: Sample $x \sim Be(2,2)$ $\quad 0 \leq x \leq 1$

$$p(x) = 6x(1-x)$$

let $q(x) = 1 \quad 0 \leq x \leq 1$

$p(x)$ has a max. of $\frac{3}{2}$ at $x = 0.5$

so $\frac{3}{2} q(x) \geq p(x) \qquad \left(c = \frac{3}{2}\right)$

- Sample x uniformly in $[0,1]$
- Accept x with prob. $\dfrac{6x(1-x)}{\frac{3}{2} \cdot 1} = 4x(1-x)$

The Metropolis-Hastings algorithm is some type of rejection sampling. The goal is to make the chain satisfy detailed balance for a given $\pi(x)$

Let $f(y|x)$ be a given proposal density for a Markov chain

- Given current state x_n, sample $y \sim f(y|x_n)$
- Calculate $\alpha(x_n, y) = \min\left[\frac{\pi(y) f(x_n|y)}{\pi(x_n) f(y|x_n)}, 1\right]$
- Sample $u \sim U(0,1)$
- If $u \leq \alpha(x_n, y)$

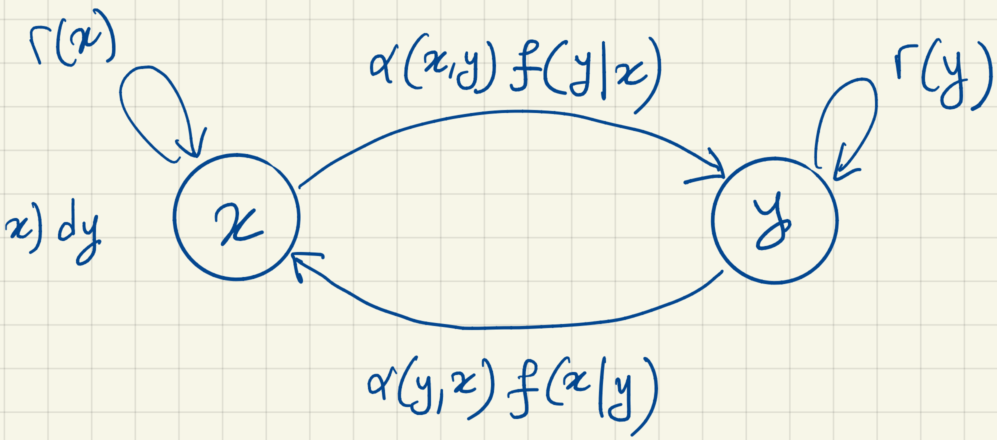
$x_{n+1} = y$ [move with prob. $\alpha(x_n, y)$]

else

$x_{n+1} = x_n$ [stay with prob. $1-\alpha(x_n, y)$]

Metropolis - Hastings Illustration

$r(x) = 1 - \int \alpha(x,y) f(y|x)\, dy$



Detailed balance:

Prove: $\pi(x)\,\alpha(x,y) f(y|x) = \pi(y)\,\alpha(y,x) f(x|y)$

Assume $\alpha(x,y) = \dfrac{\pi(y) f(x|y)}{\pi(x) f(y|x)}$, then $\alpha(y,x) = 1$

$$\underbrace{\pi(x)\,\alpha(x,y)\, f(y|x)} = \pi(x) \frac{\pi(y) f(x|y)}{\pi(x) f(y|x)} f(y|x) = \pi(y) f(x|y) = \underbrace{\pi(y)\,\alpha(y,x) f(x|y)}$$

Metropolis - Hastings. (aka MCMC)

Markov chain Monte Carlo

Proposal Distribution

Symmetric MCMC

$$f(y|x) = f(x|y), \text{ then } \alpha(x,y) = \min\left[\frac{\pi(y)}{\pi(x)}, 1\right]$$

Independent MCMC

$$f(y|x) = f(y), \text{ then } \alpha(x,y) = \min\left[\frac{\pi(y)f(x)}{\pi(x)f(y)}, 1\right]$$

Application to Bayes

$$f(\theta|x) = \frac{f(x|\theta)\,f(\theta)}{\int f(x|\theta)\,f(\theta)\,d\theta} = \pi(\theta).$$

Hard

Use MCMC to sample from $\pi(\theta)$.

- How if we don't know $\pi(\theta)$?

- We only need the ratio $\pi(\theta_1)/\pi(\theta_2)$

$$\alpha(\theta_1, \theta_2) = \min\left[\frac{\pi(\theta_2)}{\pi(\theta_1)}\,\frac{f(\theta_1|\theta_2)}{f(\theta_2|\theta_1)},\ 1\right]$$

- Use prior $f(\theta)$ for proposal distribution as independent MCMC.

Example:

$$x_i \mid \mu \sim N(\mu, \sigma^2)$$

$$f(\mu) \propto \frac{1}{1+\mu^2} \quad \text{(Cauchy)}$$

$$f(\mu \mid \bar{x}) \propto e^{-\frac{n(\mu-\bar{x})^2}{2\sigma^2}} \cdot \frac{1}{1+\mu^2}$$

Use MCMC with proposal distribution $f(\mu_2 \mid \mu_1) = \overbrace{f(\mu_2)}^{\text{Prior}}$

$$\alpha(\mu_1, \mu_2) = \min\left[\frac{e^{-\frac{n(\mu_2-\bar{x})^2}{2\sigma^2}} \cdot \cancel{\frac{1}{1+\mu_2^2}} \cdot \cancel{\frac{1}{1+\mu_1^2}}}{e^{-\frac{n(\mu_1-\bar{x})^2}{2\sigma^2}} \cdot \cancel{\frac{1}{1+\mu_1^2}} \cdot \cancel{\frac{1}{1+\mu_2^2}}},\ 1\right]$$