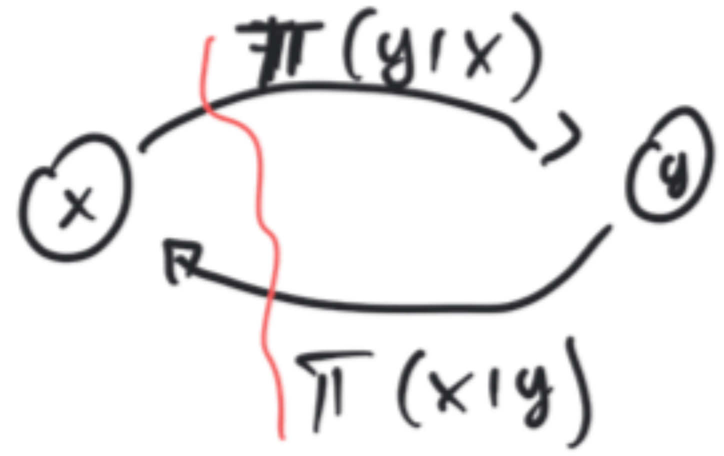
MARKOV CHAIN MONTE CARLO

$p(x)$ difficult to sample.

1) SAMPLE $p(x)$
2) COMPUTE $E[f(x)] = \int f(x)p(x)\,dx$

Idea: construct ~~an~~ DTMC s.t. $p(x)$ is steady state for it.

DETAILED BALANCE



$$p(x)\,\pi(y|x) = p(y)\,\pi(x|y)$$

If p satisfies detailed balance, then

$p = p\pi$, i.e. p is invariant.

For continuous state space, we have a KERNEL $K(x|y)$, $x, y \in S$ (here $\mathbb{R}$, $\mathbb{R}^n$, $S \subseteq \mathbb{R}^n$), t.c. measurable in y and is a distribution w.r.t. x.

Typically $K(x|y) = \mathcal{N}(y, \sigma^2)$

UPDATE RULE di METROPOLIS-HASTINGS

GOAL: SAMPLE $p(x)$

FIX A PROPOSAL DISTRIBUTION (KERNEL):

$$q(x^* \mid x).$$

ERGODIC (IRREDUCIBLE) (POSITIVE RECURRENT)

1) SAMPLE x^* FROM $q(x^* \mid x)$ (we are in x)

2) ACCEPT x^* with probability

$$\mathcal{A}(x^* \mid x) = \min\left\{1, \frac{p(x^*)\, q(x \mid x^*)}{p(x)\, q(x^* \mid x)}\right\}$$

otherwise stay in x.

$$\tilde{K}(y \mid x) = \underbrace{q(y \mid x)\,\mathcal{A}(y \mid x)}_{y \neq x} + \underbrace{\delta_x(y)\, r(x)}_{y = x}$$

$$\underbrace{r(x) = 1 - \int q(z \mid x)\,\mathcal{A}(z \mid x)\,dz}_{\text{prob. of staying in } x}$$

Let's prove the DBC! Fix $x \neq y$

$$p(x)\,q(y|x)\,A(y|x) = \min\left\{ p(x)q(y|x),\ \cancel{p(x)q(y|x)} \cdot \frac{p(y)q(x|y)}{\cancel{p(x)q(y|x)}} \right\}$$

$$= \min\left\{ p(x)q(y|x),\ p(y)q(x|y) \right\}$$

$$= p(y)\,q(x|y)\,A(x|y)$$

this is the DBC

$\Downarrow$

$p(x)$ is invariant, hence it is steady state!

$$\lim_{n\to\infty} p(X_n = x) = p(x) \quad !$$

$$A(y|x) = \min\left\{ 1,\ \frac{p(y)q(x|y)}{p(x)q(y|x)} \right\}$$

[We just need to know $p(x)$ up to a normalization coefficient!]

$$\Lambda = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_N \end{pmatrix}$$

$$\Lambda^n = \begin{pmatrix} \lambda_1^n & & 0 \\ & \ddots & \\ 0 & & \lambda_N^n \end{pmatrix}$$

So $\lambda_1 = 1$, $\lambda_1^n = 1$

So $\lambda_2, \ldots, \lambda_N < 1$ $\quad \lambda_j^n \to 0$

$\lambda_2 > \lambda_3 > \ldots \lambda_N$

$\lambda_2^n > \lambda_3^n > \ldots > \lambda_N^n$

MIXING COEFFICIENT.

If $q(x|y) = q(y|x)$ is symmetric, then

$$\mathcal{A}(y|x) = \min\left\{1, \frac{p(y)}{p(x)}\right\} \rightsquigarrow \text{METROPOLIS update rule!}$$

Obs: the VARIANCE of the proposal distribution q is crucial to have a chain that MIXES WELL

MIXTURES AND CYCLES OF UPDATE KERNELS

Let q_1 and q_2 be two proposal distributions.

1) APPLY q_1 and q_2 in sequence $\quad \int q_1(y|x)\, \mathcal{A}_1(y|x)\, q_2(z|y)\, \mathcal{A}_2(z|y)\, dy$

2) MIXTURES, $\alpha > 0$, $\alpha < 1$ $\quad q(y|x) = \alpha q_1(y|x) + (1-\alpha) q_2(y|x)$

GIBBS SAMPLER

We have $p(x_1, \ldots, x_n)$ and we know how to sample from $p(x_j | \vec{x}_{\setminus j})$

$\vec{x}_{\setminus j} = x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n$

update step:
- SAMPLE $p(x_1 | x_{\setminus 1})$
- SAMPLE $p(x_j | x_{\setminus j}) \Leftarrow q_j(y|x) = \begin{cases} p(y_j | x_{\setminus j}) & \text{if } y_{\setminus j} = x_{\setminus j} \\ 0 & \text{otherwise} \end{cases}$
- $\vdots$
- SAMPLE $p(x_n | x_{\setminus n})$

$$q_j(y|x) = \begin{cases} p(y_j|x_{\cdot j}) & \text{if } y_{\cdot j} = x_{\cdot j} \\ 0 & \end{cases}$$

$$A_j(y|x) = \min\left\{1, \frac{p(y)\, q_j(x|y)}{p(x)\, q_j(y|x)}\right\}$$

$$\frac{\overbrace{p(y_{\cdot j})\, p(y_j|y_{\cdot j})}^{p(y)}\, p(x_j|y_{\cdot j})}{\underbrace{p(x_{\cdot j})\, p(x_j|x_{\cdot j})}_{p(x)}\, p(y_j|x_{\cdot j})} = 1$$

$$\boxed{x_{\cdot j} = y_{\cdot j}}$$

In the GIBBS sampler, we accept every step with prob. 1.

(FIXED SCAN G.S)

RANDOM SEQUENCE SCAN G.S.

- SAMPLE $j \in \{1, \dots, n\}$ unif.
- SAMPLE FROM q_j

We need to check $p(x_j|x_{\cdot j})$ is ergodic

(This is true if $p(x_j|x_{\cdot j})$ is non-zero everywhere in S)