MCMC: The Metropolis algorithm

<u>Goal</u>: $\begin{cases} \text{i.) Estimate statistics, e.g. } \mathbb{E}_{x\sim p(x)}[f(x)] = \int f(x)p(x)\,dx \text{ , } x \in \\ \text{ii) Generate samples from } \tilde{p}(x) \text{ , } p(x) = \dfrac{\tilde{p}(x)}{Z_p} \text{ , } Z_p = \int \tilde{p}(x)\,d \end{cases}$

<u>Recall Monte Carlo</u>:

MC: $\mathbb{E}_{x\sim p(x)}[f(x)] \approx \frac{1}{n}\sum_{i=1}^{n} f(x_i)$, $x_i \overset{\text{i.i.d.}}{\sim} p(x)$

MCMC: $\mathbb{E}_{x\sim p(x)}[f(x)] \approx \frac{1}{n}\sum_{i=1}^{n} f(x_i)$, $x_i \sim$ M.C. (x_i are now correlate

<u>Theorem</u>: If $(x_0, x_1, \dots, x_n)$ is an <u>irreducible</u>, <u>time-homogeneous</u>, <u>discrete</u> Markov Chain with stationary distribution $p(x)$, then:

i.) $\frac{1}{n}\sum_{i=1}^{n} f(x_i) \xrightarrow[\text{a.s.}]{n\to\infty} \mathbb{E}_{x\sim p(x)}[f(x)]$

ii) If further the chain is <u>aperiodic</u> then:

$$p(x_n = x \mid x_0) \xrightarrow{n\to\infty} p(x)$$

i.e. that x_n (as $n\to\infty$) is a "good" sample from $p(x)$.

<u>What is a Markov Chain?</u>

$$x_0 \to x_1 \to x_2 \to \dots \to x_n$$

$$p(x_0, x_1, x_2, \dots, x_n) = p(x_0)\,p(x_1|x_0)\,p(x_2|x_1) \dots p(x_n|x_{n-1})$$

Markov property: $p(x_i | x_0, x_1, \dots, x_{i-1}) = p(x_i | x_{i-1})$

- Discrete M.C. $\Longrightarrow$ $x_i \in \mathcal{X}$, $\mathcal{X}$ is a countable set, e.g.: $\mathcal{X} = \{0, 1, 2, \dots$

- <u>Time homogeneous M.C</u>:

$$\underbrace{p(x_{i+1} = b \mid x_i = a)}_{\text{transition probability}} = \underbrace{T_{ab}}_{\substack{\text{don't depend} \\ \text{on time}}}, \quad \forall i, \quad \forall a, b \in \mathcal{X}$$

T is called the transition matrix of the Markov Chain, and it is a stochastic matrix, i.e.:

$$\sum_b T_{ab} = 1 \quad \text{(i.e. the rows sum to one)}$$

- <u>Irreducible M.C.</u>:

If $\forall a, b \in \mathcal{X}$, $\exists t \geq 0$: $p(x_t = b \mid x_0 = a) > 0$

Therefore, no matter where the Markov Chain started from, it can rea[ch] every possible state $b \in \mathcal{X}$.

$\mathcal{X}$ 

- <u>Aperiodic M.C.</u>:

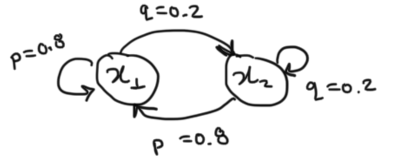
If $\forall a \in \mathcal{X}$: $\gcd \underbrace{\{t : p(x_t = a \mid x_0 = a)\}}_{\substack{\text{The set of times for which the M.C.} \\ \text{revisited its initial state } a.}} = 1$

<u>Remark</u>:

If a discrete M.C. is irreducible, has a stationary distribution p(

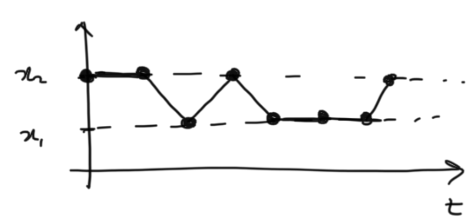
and is aperiodic, then it is an ergodic Markov Chain.

<u>Examples of Markov Chains</u>:

1.) $\mathcal{X} = \{1, 2\}$ 

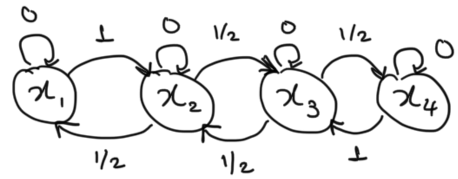
, $p, q \geq 0$, $p + q = 1$

Transition matrix: $T = \begin{bmatrix} p & q \\ p & q \end{bmatrix}$

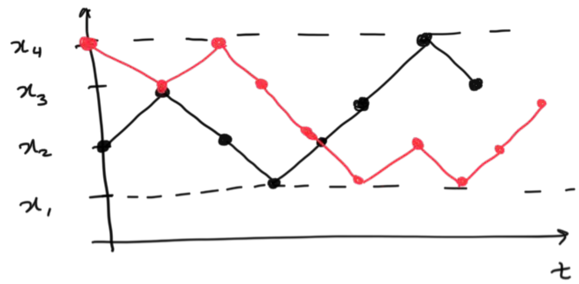


• The chain is irreducible (since $p, q \geq 0$), and aperiodic.

2.) $\mathcal{X} = \{1, 2, 3, 4\}$



$$T = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$



• irreducible

• not aperiodic (periodic)

"Symmetric random walk with reflecting boundaries"