29 April 2022

Random Walks and Markov Chains

Markov Chain Monte Carlo: a method for drawing samples from "complicated implicitly defined probability distributions." (CIDPD)

My model of a CIDPD is a distribution $p: X \to [0,1]$ where $|X|$ is exponentially large, and we have an efficient algorithm that computes a function $w: X \longrightarrow \mathbb{R}_{\geq 0}$, and

$$p(x) = \frac{w(x)}{\sum_{y \in X} w(y)}.$$

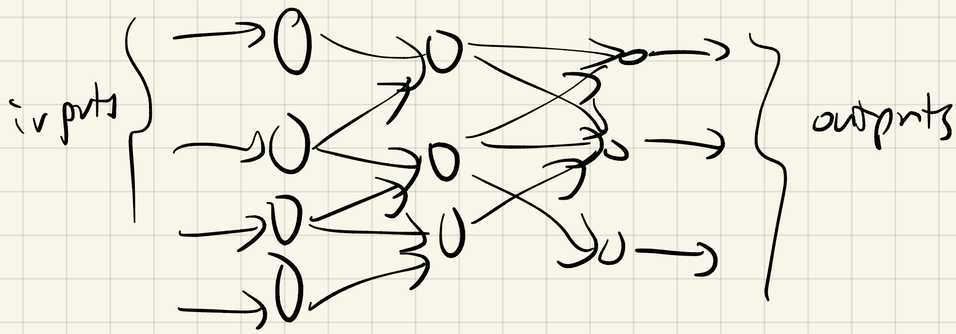
Examples.

① Given undirected graph G that is q-colorable for some $q \in \mathbb{N}$, sample a unif random proper q-coloring.

$$X = \{\text{functions } V(G) \to [q]\}$$

$$w(x) = \begin{cases} 1 & \text{if } x \text{ is a proper coloring} \\ \emptyset & \text{if not.} \end{cases}$$

② Given a DNN with finite-precision arithmetic that generates realistic looking images from uniformly random bits in the input layer



Each node computes a (finite-precision approx of) a non-linear func. applied to a linear combination of the values on its incoming edges.

If neural net has n nodes and precision of b bits per node, then

$$X = [2^b]^n \qquad (x_1, \ldots, x_n) \in X \text{ denotes}$$

labeling the nodes of X with $x_1, \ldots, x_n$ respectively.

$$w(x_1, \ldots, x_n) = \begin{cases} 0 & \text{if } \exists \text{ an internal node } x_i \text{ whose value is } \neq \text{ to the appropriate function of the inputs that node} \\ 0 & \text{if } \exists \text{ an output node } x_j \text{ which was among the non-occluded pixels of the image but } x_j \neq \text{ that pixel value} \\ 1 & \text{otherwise} \end{cases}$$

Markov Chains / Random Walks

<u>Def.</u> A <u>Markov Chain</u> with state space $\mathcal{X}$ and transition matrix P is a sequence of random variables $X_0, X_1, \ldots$ s.t. $\forall n > 0$ $\forall x, y \in \mathcal{X}$

$$\Pr(X_n = y \mid X_{n-1} = x) = P_{xy}$$

$$\Pr(X_n = y \mid X_0, X_1, \ldots, X_{n-1}) = P_{X_{n-1} y}$$

For this to be a prob. distribution P must satisfy $\forall x$ $\sum_{y \in \mathcal{X}} P_{xy} = 1$

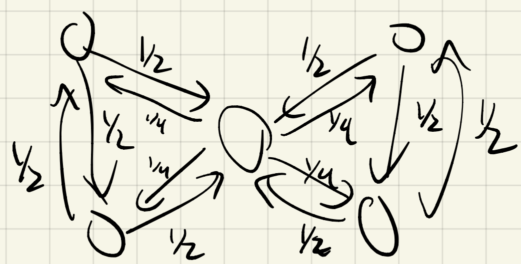
(P is "row-stochastic", $P\vec{1} = \vec{1}$.)

Depicting a Markov Chain with a graph.

$V(G) = \mathcal{X}$ $\qquad E(G) = \{ (x,y) \mid P_{xy} > 0 \}$

label each $(x,y) \in E(G)$ with P_{xy}.

$$\begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$$



Def. $\pi : X \to [0,1]$ is a <u>stationary distribution</u> for P

if $\sum_{x \in X} \pi(x) = 1$ and $\forall y \; \sum_{x \in X} \pi(x) P_{xy} = \pi(y)$.

We can represent π as a row vector

$$\begin{bmatrix} \pi_{x_1} & \pi_{x_2} & \cdots & \pi_{x_N} \end{bmatrix}$$

and then $\forall y \; \sum_x \pi(x) P_{xy} = \pi(y)$

$$\Leftrightarrow \quad \pi P = \pi.$$

Does every Markov Chain have a stationary distrib?

<u>Ans.</u> Yes if X is finite.

Recall $P\vec{1} = \vec{1}$ so $(P - \mathbb{1}) \cdot \vec{1} = 0$.

So $\exists \; v \neq 0$ row vector $v(P - \mathbb{1}) = 0$ i.e. $vP = v$.

Perron-Frobenius Guarantees v has ≥ 0 coordinates.

Is the stationary distrib unique?

<u>Ans.</u> Not in general. Yes when

$\exists \, m$ s.t. P^m has all entries strictly > 0.

$\Leftrightarrow$ the gcd of the lengths of directed cycles in the state transition graph equals 1.

To come in future lectures: how fast does a Markov Chain converge to stationary distrib?