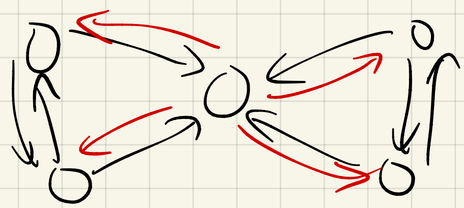
$\boxed{\text{2 May 2022}}$ Stationary Distributions, Reversibility, Metropolis-Hastings

Announcement:

- You don't need the auxiliary data structure in the query phase. (Problem 2c of homework.)
- Quiz 3 is 90% graded, hopefully we finish today.

Markov chain.

- Finite set $\mathfrak{X}$ of states.
- Transition matrix P.
- P_{xy} is the prob. of being in state y at $t+1$ given the time-t state is x.



$$P\vec{1} = \vec{1}$$

$$\begin{bmatrix} \pi P = \pi, \\ \|\pi\|_1 = 1 \end{bmatrix} \Longrightarrow \text{"}\pi \text{ is a stationary distribution of } P.\text{"}$$

Does P have a stationary distrib?

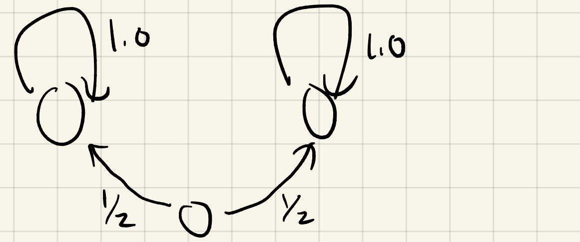
Yes, as long as $\mathfrak{X}$ is finite.

Is it unique?

Yes, if the state transition graph is strongly connected.

(via Perron-Frobenius.)

No, in general.



(An absorbing Markov chain.)

Does the marginal distribution of the state at time t converge to the stationary distrib as $t \to \infty$?

No, in general.

Yes, if the Markov chain is ergodic

Def. 1. $\exists k < \infty$ P^k has all positive entries.

Def 2. State trans graph is strongly connected and the GCD of all directed cycle lengths is 1.

One way of forcing the GCD property to hold is to have at least one state x s.t. $P_{xx} > 0$.

(Then state trans. graph has a cycle of length 1.)

MCMC method for sampling from a distribution π:

1. Set up a transition matrix P (ergodic) whose stationary distrib is π.
2. Initialize a Markov chain in some state. ??? ???
3. Run state transitions dictated by P until covergence.
4. Output the final state of simulation.

→ Often done using Metropolis-Hastings procedure.
(To be explained now)

<u>Def.</u> P is reversible with respect to π if
$\forall x, y \in \mathcal{X}$ $\quad \pi_x P_{xy} = \pi_y P_{yx}$

<u>Lemma.</u> If P reversible w.r.t. π then
π is stationary for P.

<u>Proof.</u> The y^{th} coordinate of πP is

$$(\pi P)_y = \sum_x \pi_x P_{xy} = \sum_x \pi_y P_{yx} = \pi_y \left(\sum_x P_{yx} \right) = \pi_y$$

Metropolis Hastings takes two inputs:

(A) an unnormalized distribution $w : \mathcal{X} \to \mathbb{R}_+$
such that we want to sample from

$$\pi(x) = \frac{w(x)}{\sum_y w(y)}.$$

(B) a symmetric stochastic matrix $K \in \mathbb{R}^{\mathcal{X} \times \mathcal{X}}$
("proposal distribution")

$$K_{xy} = K_{yx} \qquad K_{xy} \geq 0 \quad \forall x, y.$$

$$\sum_{y \in \mathcal{X}} K_{xy} = 1 \qquad \forall x.$$

The procedure is: in state x at time t,

1. Draw random sample y distrib. according to K_{xy}.

2. If $w(y) \geq w(x)$, move to state y.

3. If $w(y) < w(x)$, toss a coin with bias $\frac{w(y)}{w(x)}$.
   Transition to y if the coin toss is heads.

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

<u>Lemma.</u> Metr.-Hast. (w, K) is reversible w.r.t. π.

<u>Proof.</u> We must show $\forall x, y$

$$\pi_x \cdot Pr(x \to y) = \pi_y \, Pr(y \to x)$$

$$\Updownarrow$$

$$w(x) \cdot Pr(x \to y) = w(y) \cdot Pr(y \to x)$$

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

Yes these are equal.