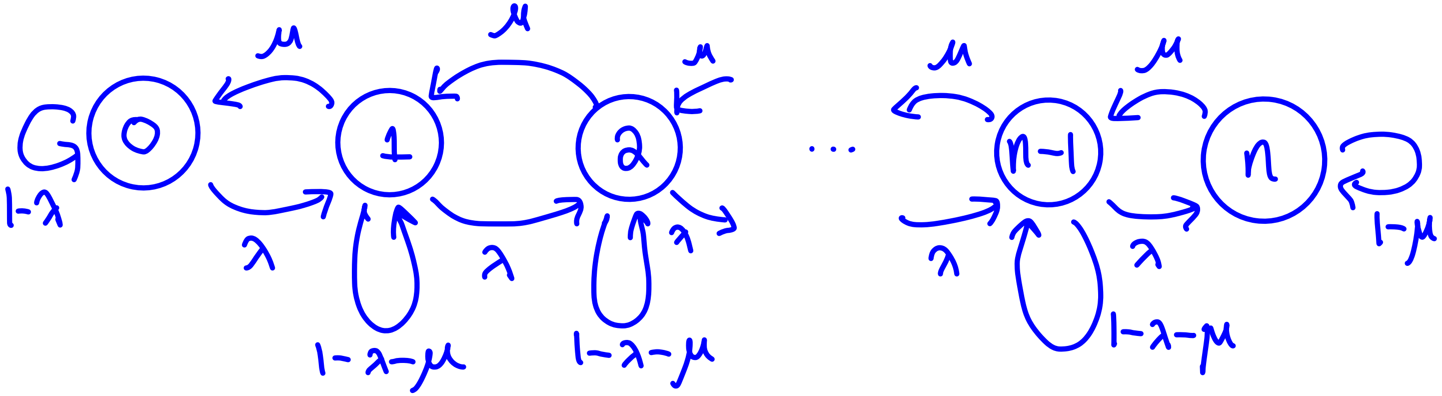
Class 14

Markov Chains II

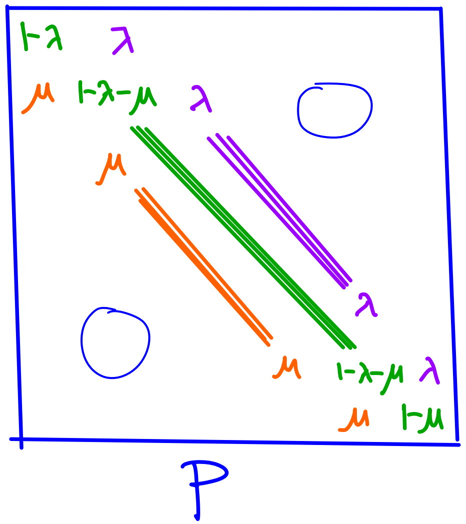
Group Work 1

The "Queue" Markov chain

- Diagram:



- Transition matrix:

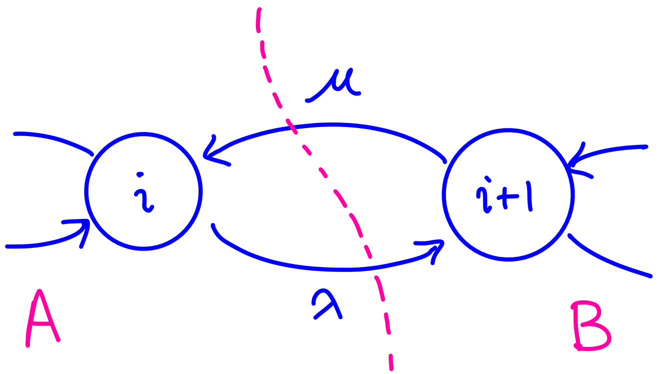


Stationary Distribution

- Claim: If $X_{t-1} \sim \pi$, then $\Pr[X_{t-1} \in A, X_t \in B] = \Pr[X_{t-1} \in B, X_t \in A]$
  - Informally: The amount of mass going from A to B needs to be the same as going from B to A, or else it can't be a stationary distribution.
- Formally:
  - $\Pr[X_{t-1} \in A] = \cancel{\Pr[X_{t-1} \in A, X_t \in A]} + \Pr[X_{t-1} \in A, X_t \in B]$

    $=$
  - $\Pr[X_t \in A] = \cancel{\Pr[X_{t-1} \in A, X_t \in A]} + \Pr[X_{t-1} \in B, X_t \in A]$
  - $X_t \sim \pi$, since π is stationary, so $\Pr[X_{t-1} \in A] = \Pr[X_t \in A]$.

Back to stationary distribution



$$\pi(i) \cdot \lambda = \pi(i+1) \cdot \mu$$

$$\Rightarrow \pi(i+1) = \left(\frac{\lambda}{\mu}\right) \cdot \pi(i), \quad \forall i$$

$$\Rightarrow \pi(i) = \left(\frac{\lambda}{\mu}\right)^i \cdot \pi(0)$$

To find the right normalization, $1 = \sum_{i=0}^{n} \pi(i) = \sum_{i=0}^{n} \left(\frac{\lambda}{\mu}\right)^i \pi(0)$ $\Rightarrow \pi(0) = \dfrac{1}{\sum_{j=0}^{n} \left(\frac{\lambda}{\mu}\right)^j}$

$$\pi(i) = \frac{(\lambda/\mu)^i}{\sum_{j=0}^{n} (\lambda/\mu)^j} \overset{\text{If } \mu \neq \lambda}{=} \frac{(\lambda/\mu)^i (\lambda/\mu - 1)}{\left((\lambda/\mu)^{n+1} - 1\right)}$$

(if $\mu = \lambda$, this is uniform).

Expected amount of time until the queue is empty again?

Fund Thm of MC $\Rightarrow$ $\mathbb{E}\begin{pmatrix}\text{return time} \\ \text{to } 0\end{pmatrix} = \dfrac{1}{\pi(0)}$

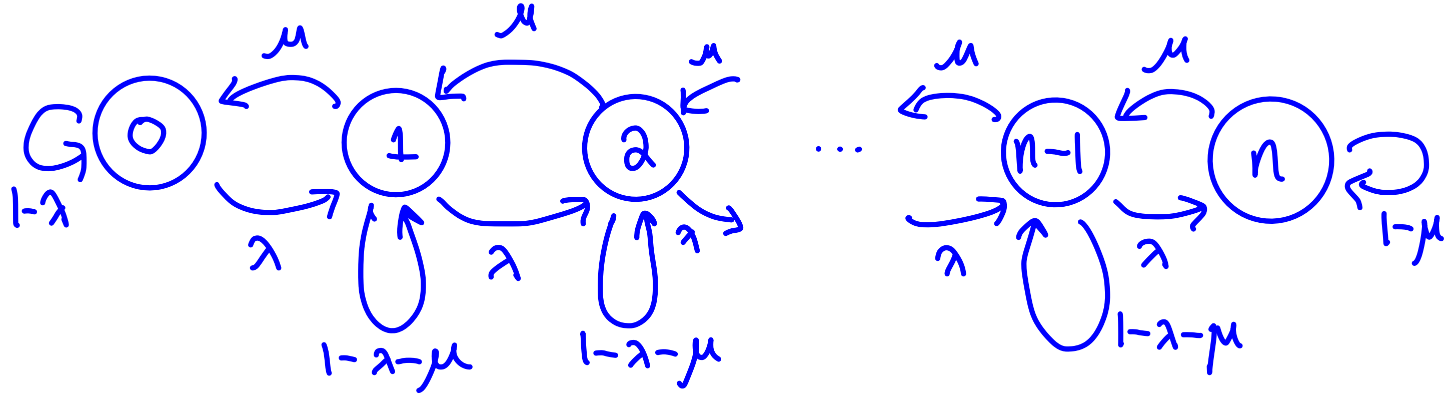
$$= \sum_{j=0}^{n} \left(\frac{\lambda}{\mu}\right)^j$$

If $\mu \neq \lambda$ $= \dfrac{(\lambda/\mu)^{n+1} - 1}{(\lambda/\mu) - 1}$ (if $\mu = \lambda$, this is $n + 1$).

As n gets big...

$$\mathbb{E}\left(\begin{matrix}\text{return time}\\ \text{to } 0\end{matrix}\right) = \sum_{j=0}^{n} \left(\frac{\lambda}{\mu}\right)^j \overset{\text{If } \mu \neq \lambda}{=} \frac{(\lambda/\mu)^{n+1} - 1}{(\lambda/\mu) - 1}$$

- If $\mu = \lambda > 0$, this is $n + 1$.
- If $\mu > \lambda$, this tends to $\frac{\mu}{\mu - \lambda}$ (aka, some constant…)
- If $\mu < \lambda$, this goes to infinity exponentially quickly in n.



Group Work 2

Show that π is the stationary distribution

$$(\pi P)_{x,y} = \sum_{x',y'} \underbrace{\pi(x',y')}_{\pi_{x'y'}} \underbrace{\pi(x|y')\,\pi(y|x)}_{P_{x'y',\,xy}}$$

$$= \sum_{x'y'} \frac{\pi(x',y')\,\pi(x,y')\,\pi(x,y)}{\pi(y')\cdot\pi(x)}$$

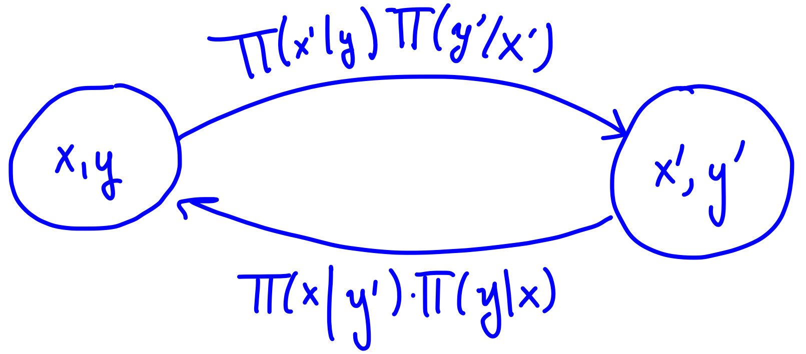
$$= \sum_{x'y'} \pi(x'|y')\,\pi(y'|x)\,\pi(x,y)$$

These sum to 1 over x'

$$= \sum_{y'} \pi(y'|x)\cdot\pi(x,y) \qquad = \pi(x,y)$$

These sum to 1 over y'

WTS $(\pi P)_{x,y} = \pi(x,y)$

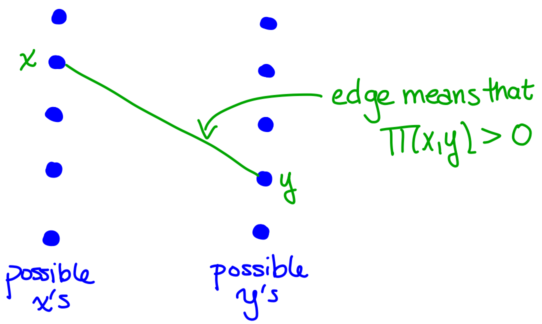


When can we apply the fundamental thm of Markov chains?

- This chain is aperiodic since there is a self-loop

  Say that $\pi(x, y) > 0$. Then $\pi(x|y)\pi(y|x) > 0$, and that's the probability of taking the self-loop at (x, y)

- This chain is <u>NOT</u> necessarily irreducible.
  - Consider looking at Π as a bipartite graph:

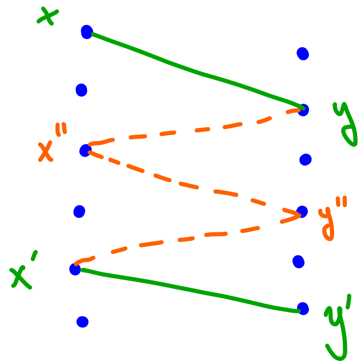


  - The Markov chain is irreducible iff this graph is connected.

- The Markov chain is finite iff Π has finite support.

eg, if $x, y \in \{0,1\}$, Π is

$$\begin{array}{cc} & Y \\ & \begin{array}{cc} 0 & 1 \end{array} \\ X \begin{array}{c} 0 \\ 1 \end{array} & \begin{array}{|c|c|} \hline 1/2 & 0 \\ \hline 0 & 1/2 \\ \hline \end{array} \end{array}$$



One way to get from (x,y) to (x',y'):

$(x,y) \mapsto (x'',y) \mapsto (x'',y'') \mapsto (x',y'') \mapsto (x',y')$

In the context of Markov Chain Monte Carlo

- Assuming that π is such that this Markov chain is aperiodic, irreducible, and finite, this gives us a way to (approximately) sample from a multivariate distribution.
  - Suppose that $\pi(X \mid Y = y)$ and $\pi(Y \mid X = x)$ are easy to sample from.
  - Then we can easy run this chain.
  - It converges to $\pi(X, Y)$
- Note: Often it is much easier to sample from univariate distributions than multivariate distributions.
  - We only did this for two variables, but the same thing works for more than two variables.
- This is called "Gibbs Sampling"

Recap

- Today: Two exercises about stationary distributions.
- Next time: How fast do Markov chains approach their stationary distributions?