Lecture 27

Markov Chains

and a Few Paradoxes.

Summary of Lecture 26

- A Markov Chain is a process that move from state to state randomly and <u>only</u> remembers its current state.

Ingredients:

- $\mathcal{X}$ — Finite state space, usually $\mathcal{X} = \{1, 2, \ldots, k\}$.
- π_0 — the initial distribution.
- $P(i,j)$ — Prob. to move from state i to j.

This defines a Markov Chain: sequence of r.v.s $X_0, X_1, X_2 \ldots$

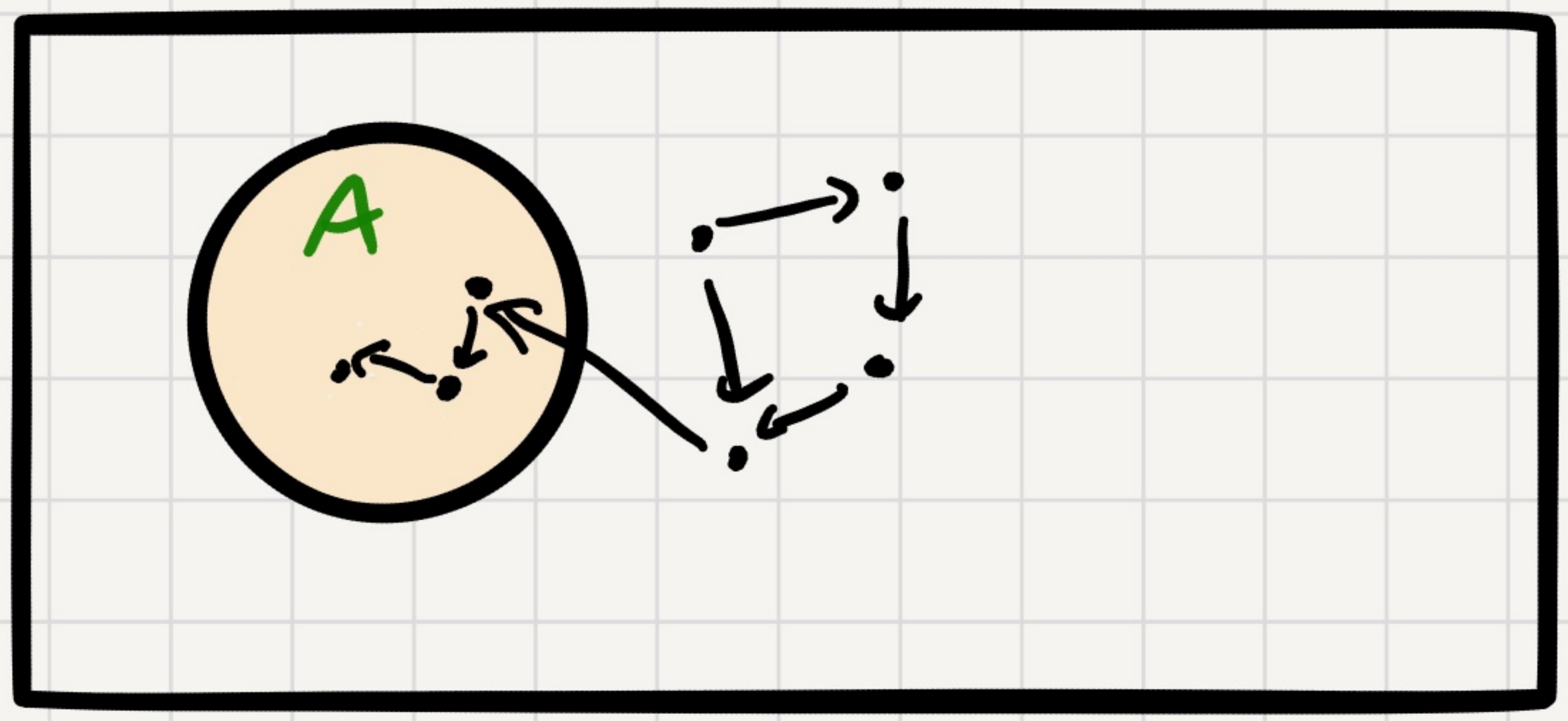
$$\Pr[X_0 = i] = \pi_0(i)$$

$$\Pr[X_{n+1} = j \mid X_n = i, X_{n-1}, \ldots, X_1, X_0] = P(i,j).$$

π_n : Prob. dist of X_n.

$$\pi_n = \pi_0 \cdot P^n$$

First Step Equations



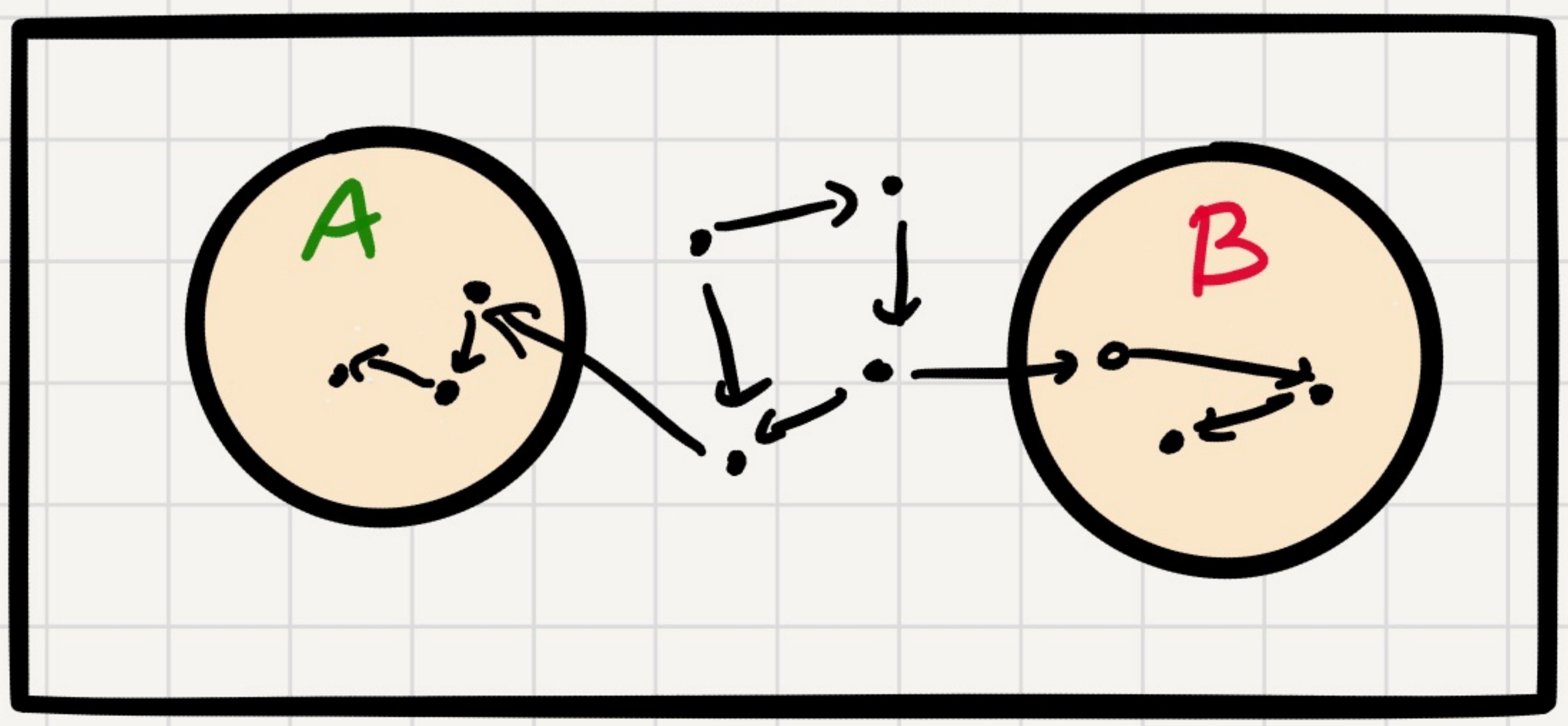
Let $\{X_n\}_{n=0}^{\infty}$ be a MC on $\mathcal{X}$ $\quad A \subseteq \mathcal{X}$

$\beta(i)$ = expected time to reach A starting from i.

$\beta(i) = 0$ for $i \in A$

$$\beta(i) = 1 + \sum_j P(i,j)\,\beta(j) \quad \text{for } i \notin A.$$

First Step Equations



Let $\{X_n\}_{n=0}^{\infty}$ be a MC on $\mathcal{X}$ $\quad A, B \subseteq \mathcal{X}$

A & B disjoint.

$\alpha(i) = \Pr[\text{reaching } A \text{ before } B, \text{starting from } i]$

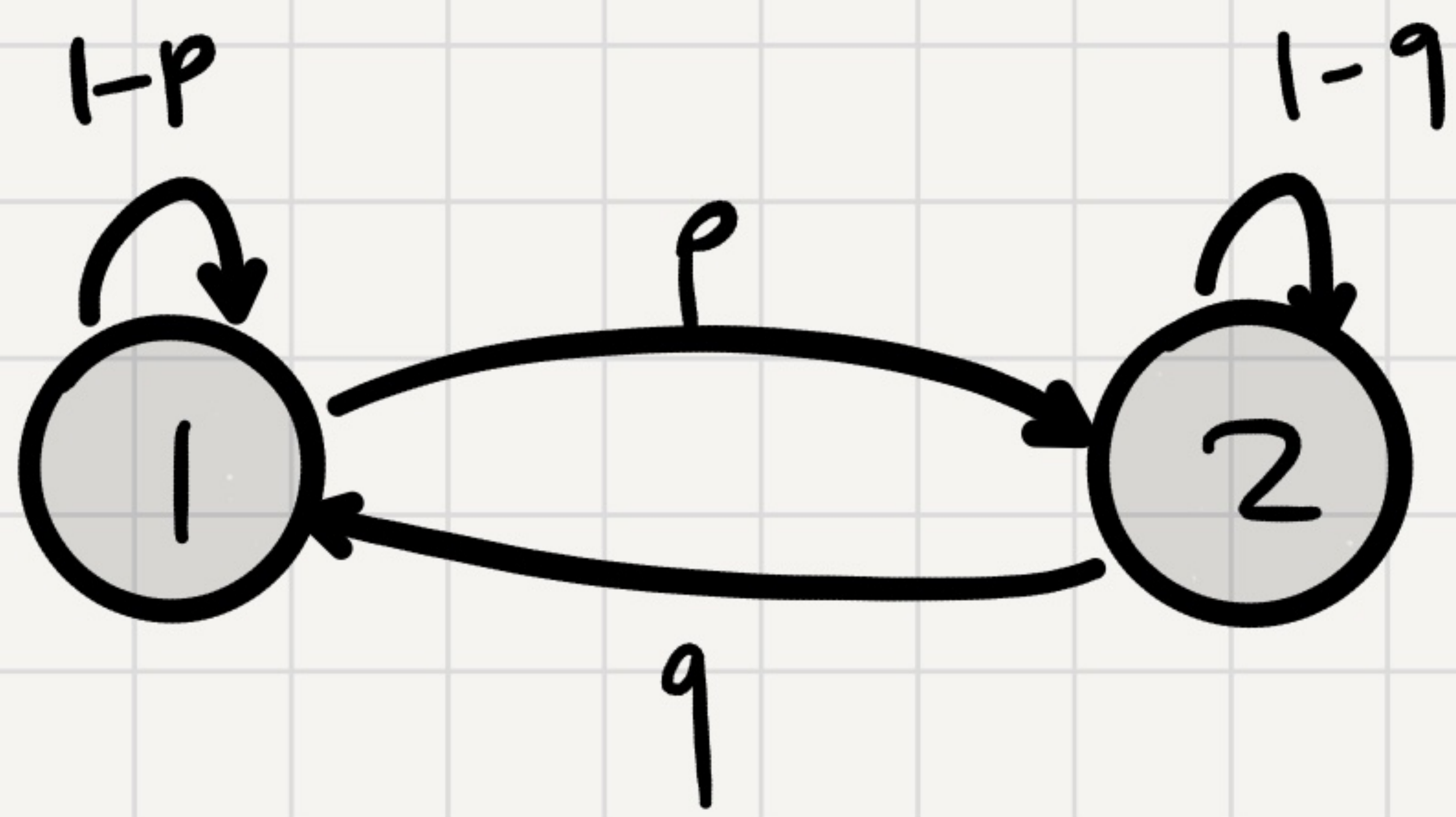
$$\begin{cases} \alpha(i) = 0 & \text{for } i \in B \\ \alpha(i) = 1 & \text{for } i \in A \\ \alpha(i) = \sum_j P(i,j) \cdot \alpha(j). & \text{for } i \notin A \cup B. \end{cases}$$

Definition:

A distribution π over $\mathcal{X}$ is stationary (aka invariant) if $\pi = \pi P$.

If π_0 is stationary then $\forall n$ $\pi_n = \pi_0$.

Stationary Distribution - Example



π is stationary iff $\begin{pmatrix}\pi(1) & \pi(2)\end{pmatrix} = \begin{pmatrix}\pi(1) & \pi(2)\end{pmatrix} \cdot \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}$

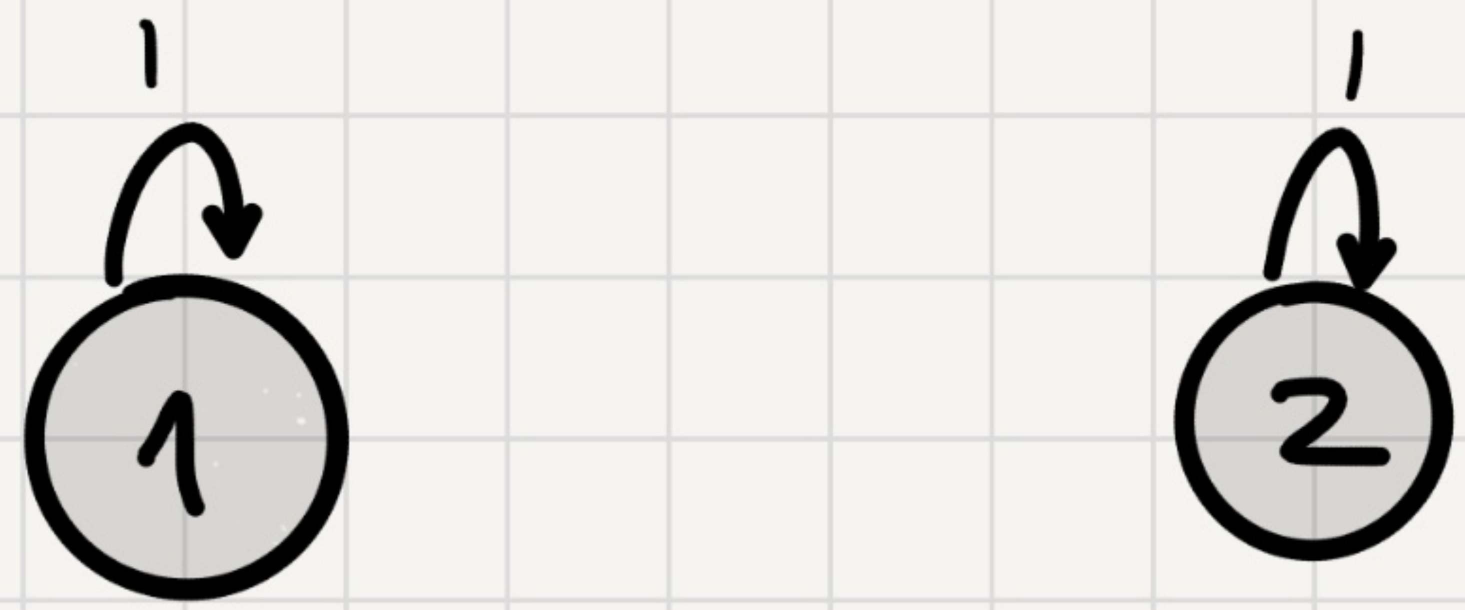
$\pi(1) = \pi(1)\cdot(1-p) + \pi(2)\cdot q \iff \pi(1)\cdot p = \pi(2)\cdot q$

$\pi(2) = \pi(1)\cdot p + \pi(2)\cdot(1-q) \iff \pi(1)\cdot p = \pi(2)\cdot q$

$\pi(1) + \pi(2) = 1$

Solution: $\pi = \left[\frac{q}{p+q}, \frac{p}{p+q}\right]$

Stationary Distributions - Example 2



Which distributions are stationary?

all of them.

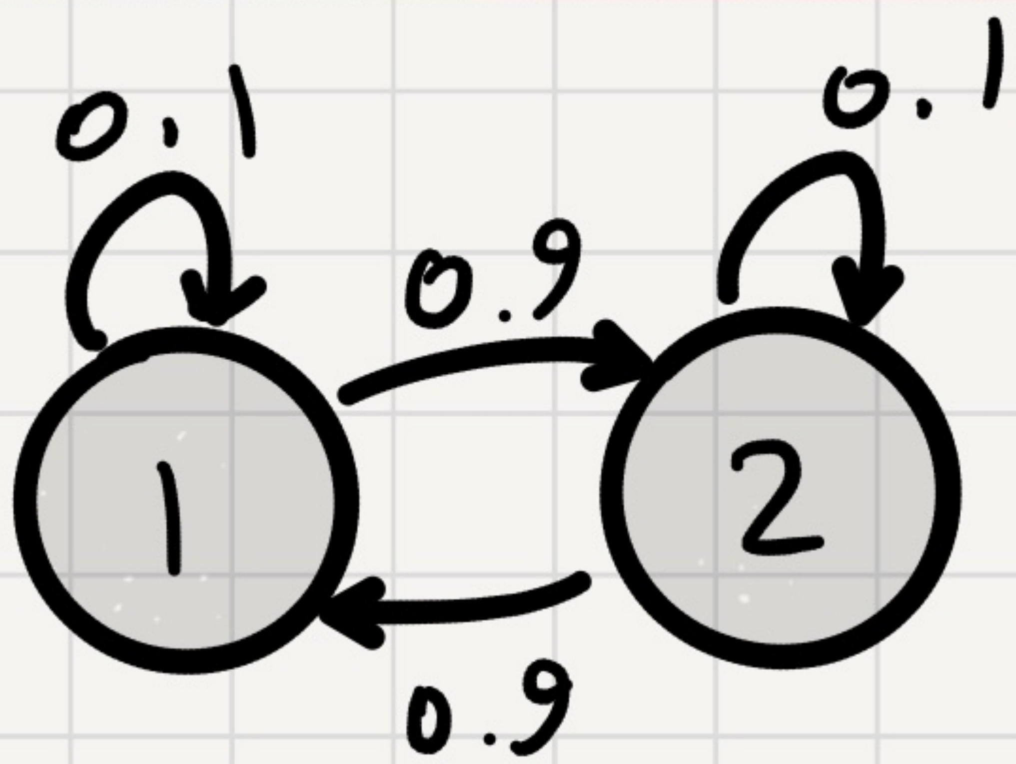
$$\forall \pi \qquad \pi = \pi \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Irreducible Markov Chains

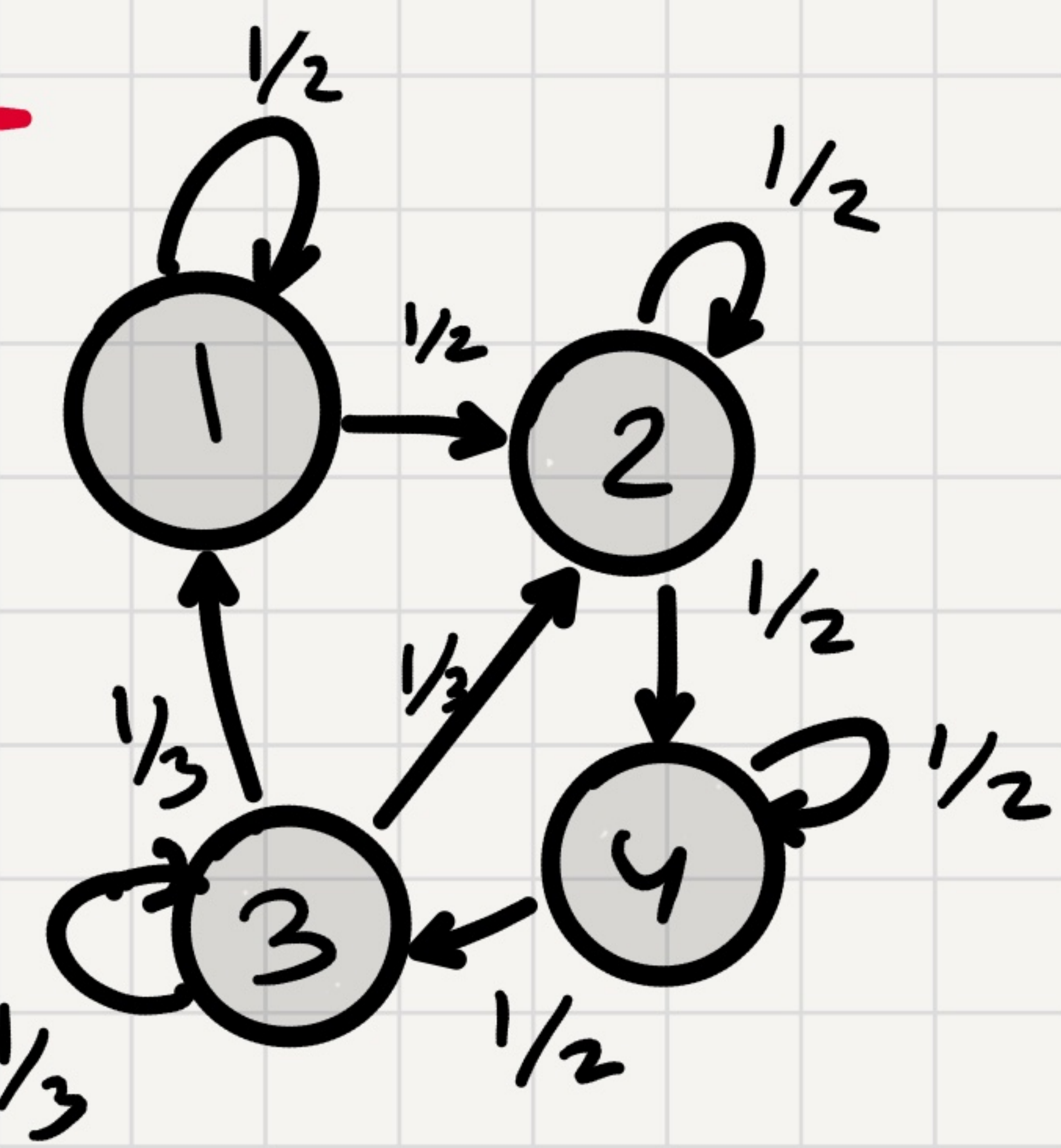
A MC is irreducible if you can go from every state i to every state j (possibly in multiple steps).

Which MC are irreducible?

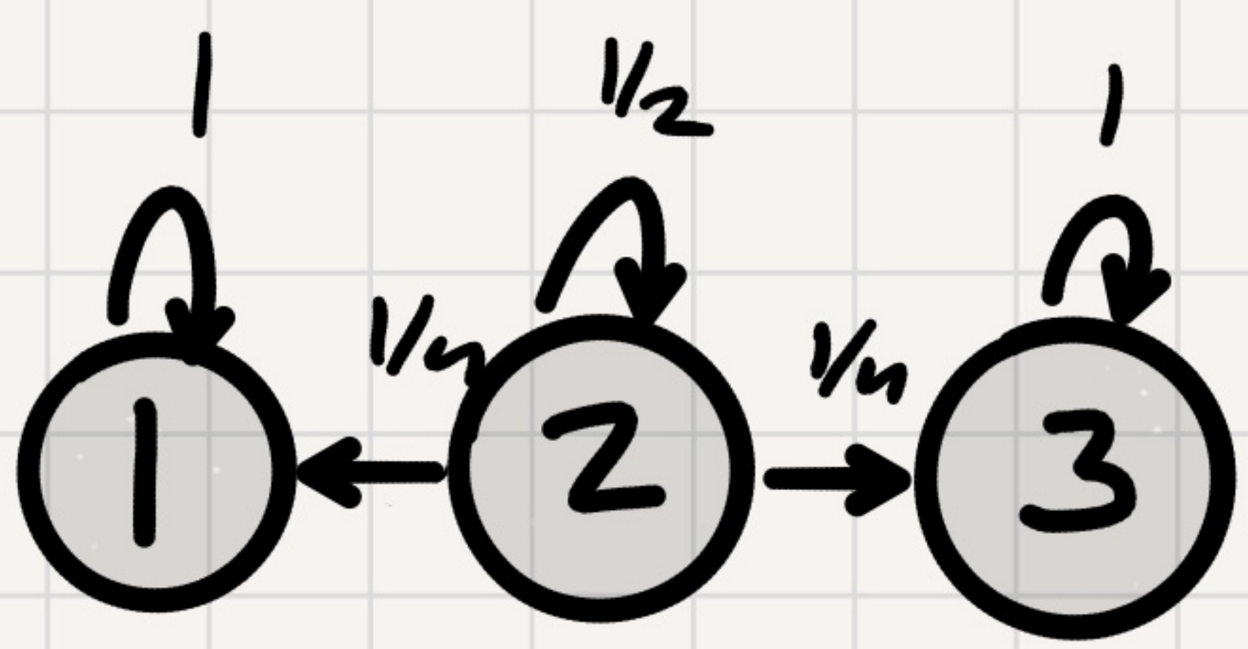
(A)



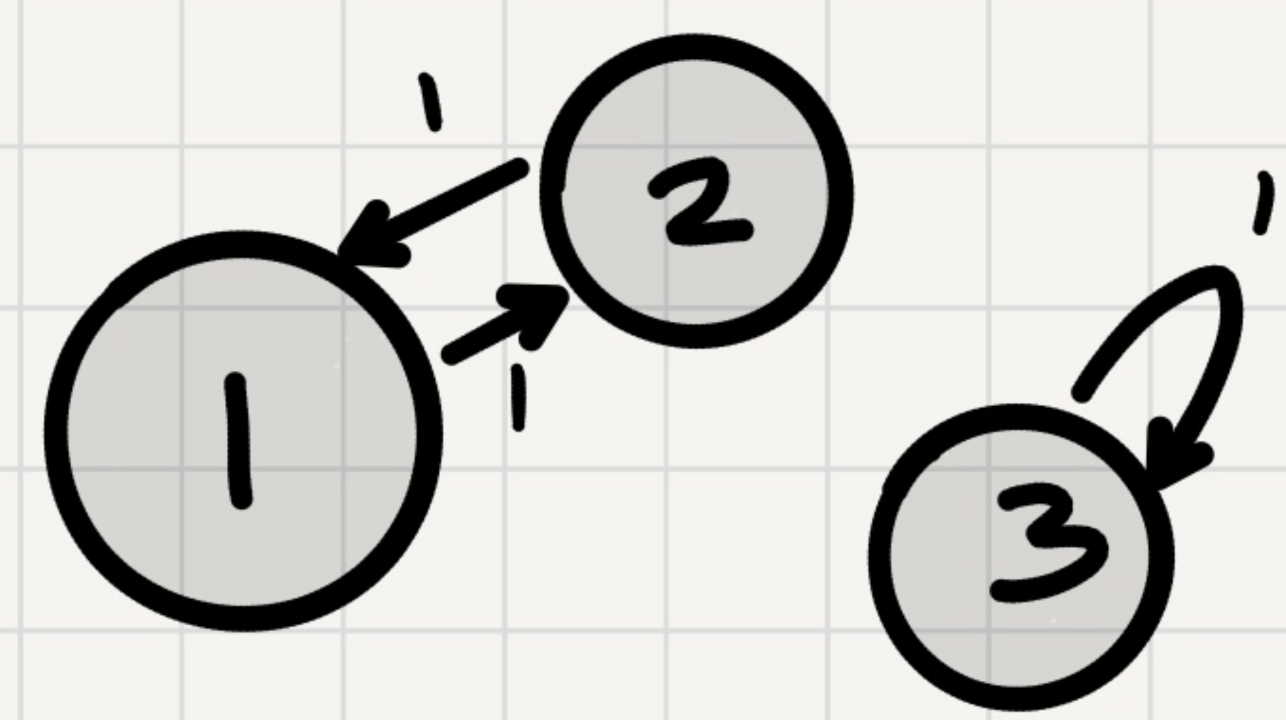
(B)



(C)



(D)



Theorem:

Any finite irreducible MC has one and only one stationary distribution.

Theorem 2: (Long Term Fraction of Time in States)

If $(X_n)_{n=0}^{\infty}$ is an irreducible MC on $\{1, \dots, k\}$ with stationary distribution π.

Then, for any start dist. π_0, for all $i \in \{1, \dots k\}$

$$\frac{1}{n} \sum_{m=0}^{n-1} 1\{X_m = i\} \xrightarrow[n \to \infty]{} \pi(i).$$

<u>Intuition:</u> Start at a dist. π_0 and suppose the limits exist. Denote by

In time m, the MC is at state i.

$$f(i) = \lim_{n\to\infty} \underbrace{\frac{1}{n}\sum_{m=0}^{n-1} \underline{\mathbb{1}\{X_m = i\}}} .$$

What's the frac. of times we visit i? $\boxed{f(i)}$

What's the frac. of times we visit i and then move to j?

$$\boxed{f(i)\cdot P(i,j)}$$

Intuition: Start at a dist. π_0 and suppose the limits exist. Denote by

$$f(i) = \lim_{n \to \infty} \frac{1}{n} \sum_{m=0}^{n-1} 1\{X_m = i\}.$$

What's the frac. of times we visit i? $\boxed{f(i)}$

j $f(j)$

What's the frac. of times we visit i and then move to j?

$\boxed{f(i) \cdot P(i,j)}$

Frac. of times we're at $j = \sum_i f(i) \cdot P(i,j)$

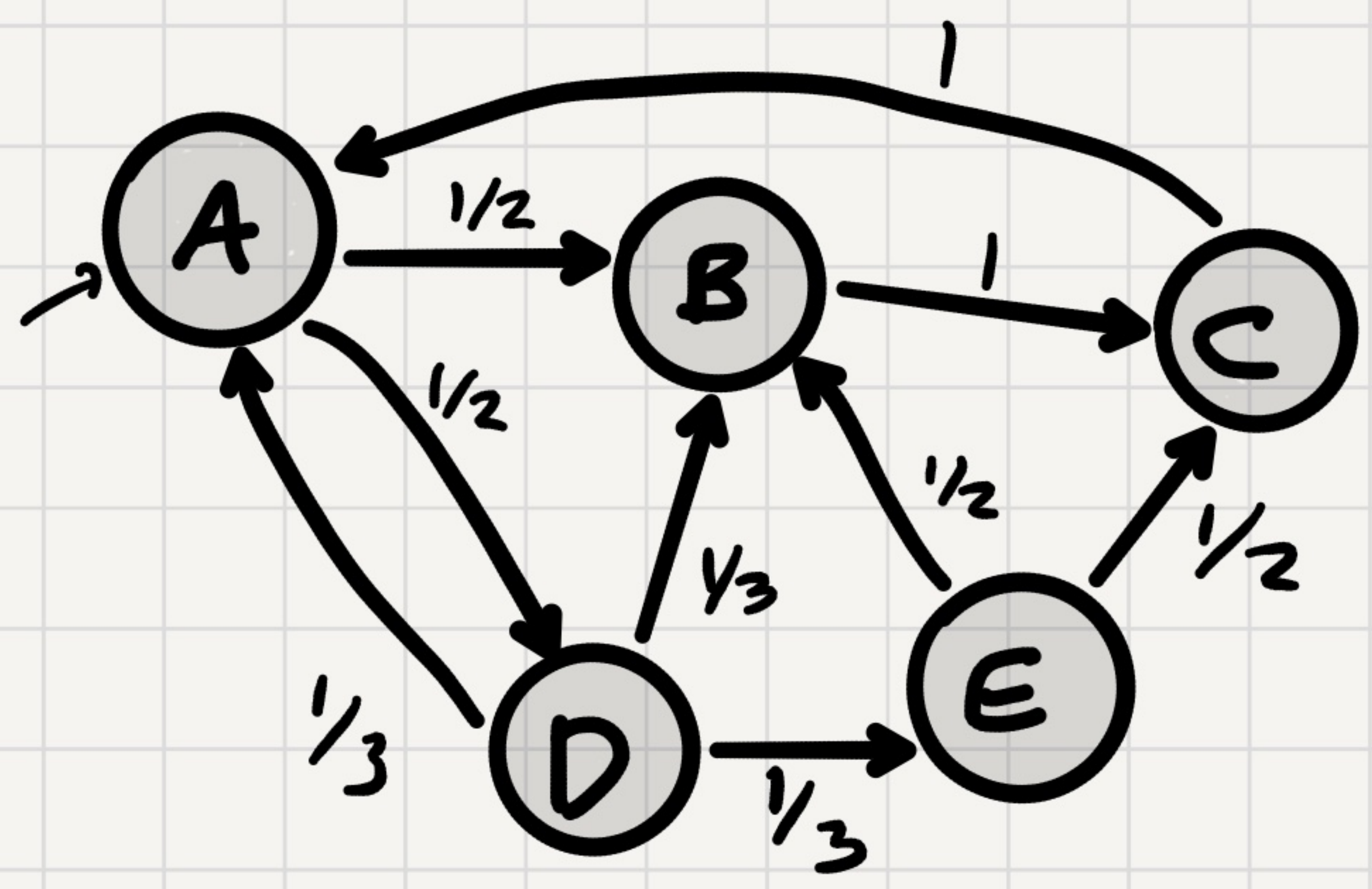
$$\forall j \quad f(j) = \sum_i f(i) \cdot P(i,j).$$

Hence, f is a stationary dist.

In matrix-vector form $\vec{f} = \vec{f} \cdot P$.

Let's see a Simulation:

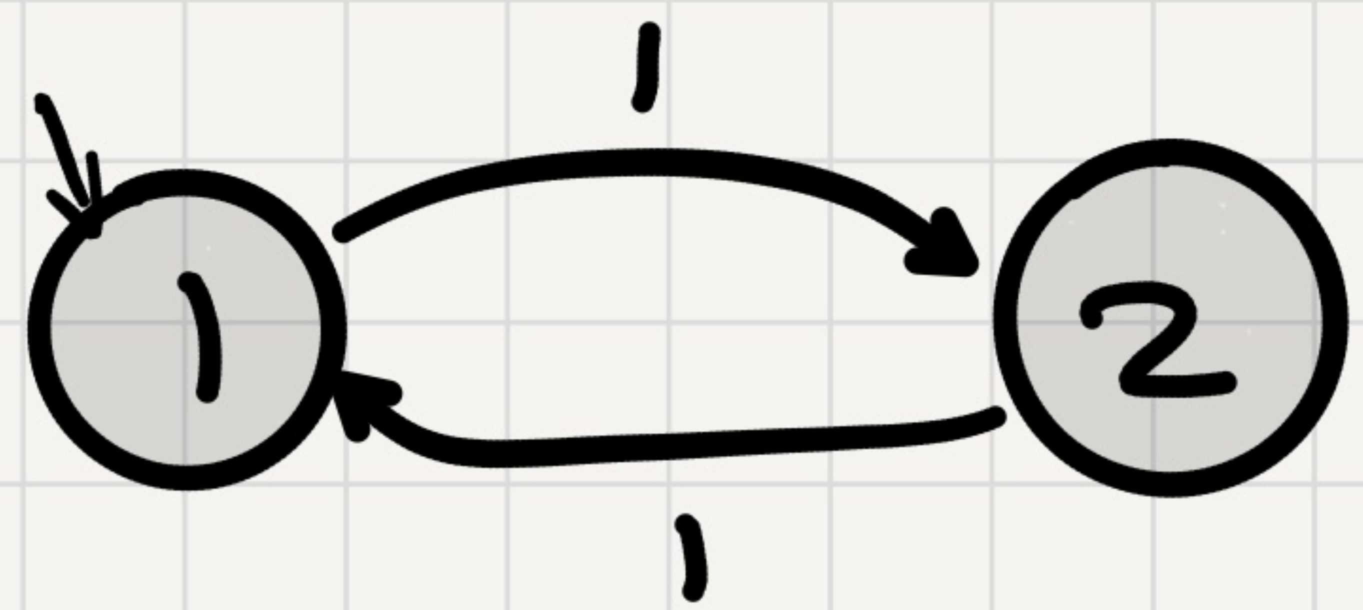
Recall this MC from lecture 26.



We'll run it for many steps and count how many times we've been in each step.

Converges to the Stationary Distribution

Example:



The MC is irreducible.

It's stationary dist satisfies

$$(\pi(1) \;\; \pi(2)) = (\pi(1) \;\; \pi(2)) \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\Rightarrow \pi(1) = \pi(2) = 1/2.$$

But starting from 1:

$$\pi_0 = (1 \;\; 0)$$
$$\pi_1 = (0 \;\; 1)$$
$$\pi_2 = (1 \;\; 0) \quad \ldots$$

$$\pi_{2m} = (1 \, 0)$$
$$\pi_{2m+1} = (0 \, 1)$$

Periodicity

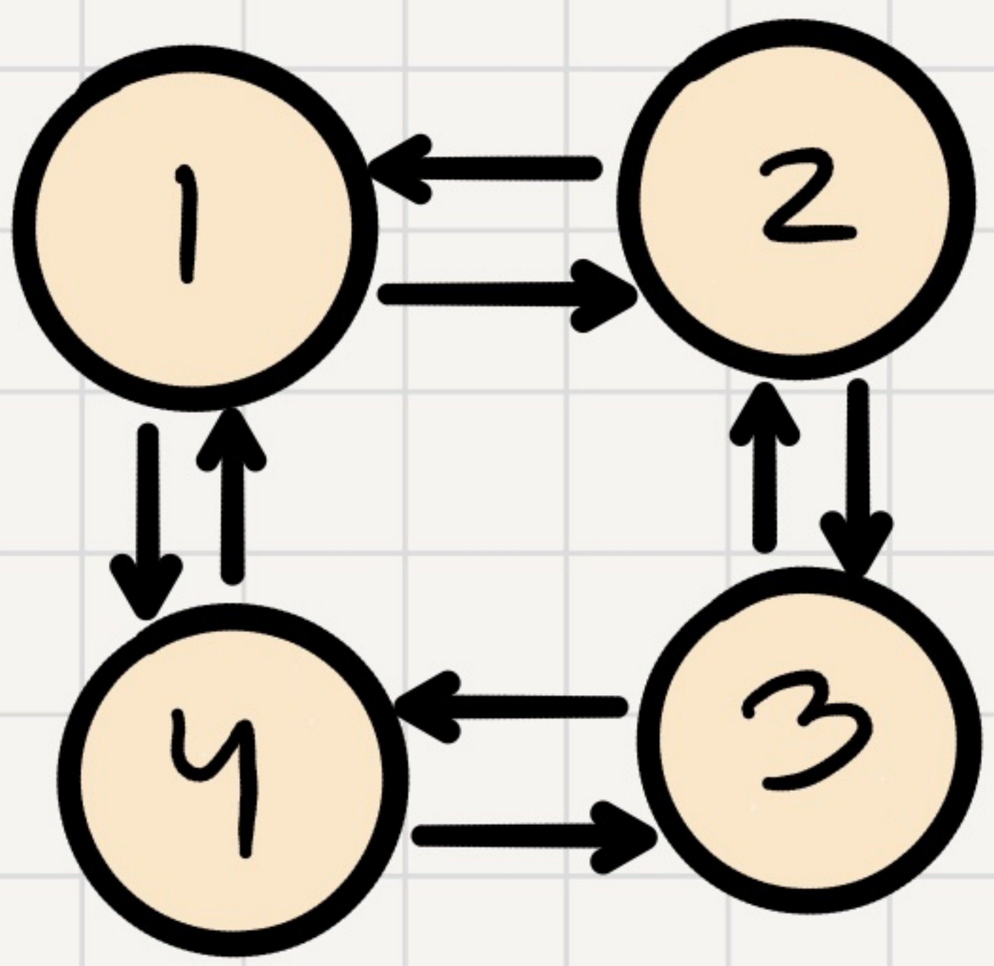
Def'n: The periodicity of a Markov Chain with transition matrix P is the gcd of lengths of all closed walks in the chain

$$\gcd(n>0 \mid \exists i \text{ s.t. } P^n(i,i)>0)$$

A Markov chain is aperiodic if this gcd $=1$.

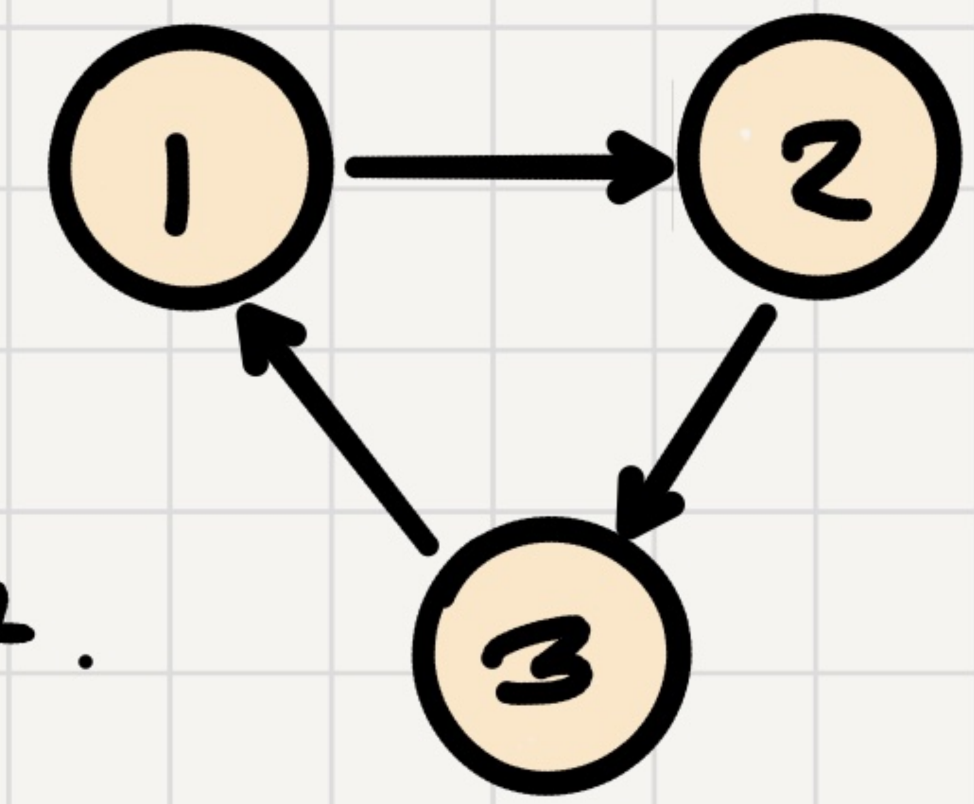
Which Markov Chains are Aperiodic?

A.



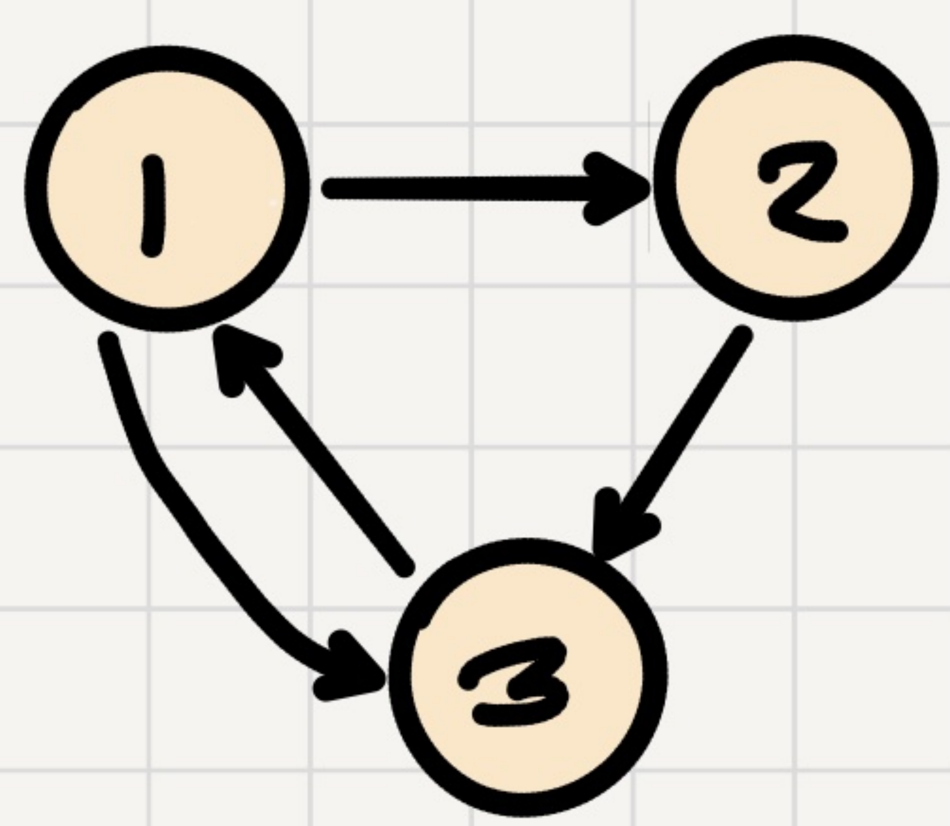
period: 2.

B.



period: 3

C.



$\gcd(2, 3, \ldots)$

$= 1$

Theorem: Let X_n be an irreducible aperiodic Markov Chain with stationary dist π.

Then, no matter what's the starting dist. π_0

$$\forall i: \quad \pi_n(i) \xrightarrow[n\to\infty]{} \pi(i).$$

Some Paradoxes

St. Petersburg Paradox

A casino is offering you to play the following game:

- Start with a stake of 2 \$.
- At each point flip a fair coin
  - H: double the stake.
  - T: stop and give stake to player.

How much are you willing to pay to play this game?

What's the expected winning stake?

X r.v. capturing winning stake

$\Pr[X=2] = 1/2 \qquad \Pr[X=4] = \frac{1}{4} \qquad \dots \qquad \Pr[X=2^k] = \frac{1}{2^k}.$

$$\mathbb{E} X = \sum_{i=1}^{\infty} \Pr[X=2^i] \cdot 2^i = \sum_{i=1}^{\infty} \frac{1}{2^i} \cdot 2^i = \sum_{i=1}^{\infty} 1 = \infty.$$

St. Petersburg Paradox

A casino is offering you to play the following game:

- Start with a stake of 2 \$.
- At each point flip a fair coin
  - H: double the stake.
  - T: stop and give stake to player.

How much are you willing to pay to play this game?

If the casino has only $n = 2^k$ dollars. What's the expected win?

$$\mathbb{E}[X] = 2 \cdot \frac{1}{2} + 4 \cdot \frac{1}{4} + \dots + 2^{k-1} \cdot \frac{1}{2^{k-1}} + \underbrace{2^k \cdot \frac{1}{2^{k-1}}}$$

$$= k+1 = \log_2 n + 1$$

Double-or-Nothing

Suppose you go to a casino and you want a betting strategy that will guarantee you win 1 \$.

They have a simple game: You choose how much to bet - x.

They flip a fair coin
- H: you win x dollars
- T: " lose x dollars.

Strategy: Start with betting 1\$, if succesful, stop.
else, bet 2\$, if succesful stop.
else, bet 4\$, if " stop.
...

Eventually a H will be flipped and you will win overall 1\$.

Double-or-Nothing

Bet X dollars ; w.p. $1/2$ win X ; w.p. $\frac{1}{2}$ lose X.

Strategy: Start with betting 1\$, if succesful, stop.
else, bet 2\$, if succesful stop.
else, bet 4\$, if " stop.
...
Eventually a H will be flipped and you will win overall 1\$.

What happens if you have a limited budget, say 1023\$:

You'll lose 1023\$ with prob. $\frac{1}{1024}$
and win 1\$ with prob. $1-\frac{1}{1024}$

On expectation, you'll even out.

$$\frac{1}{1024}\cdot(-1023)+\left(\left(1-\frac{1}{1024}\right)\cdot 1\right)$$
$$=0.$$

Confusing statistics:

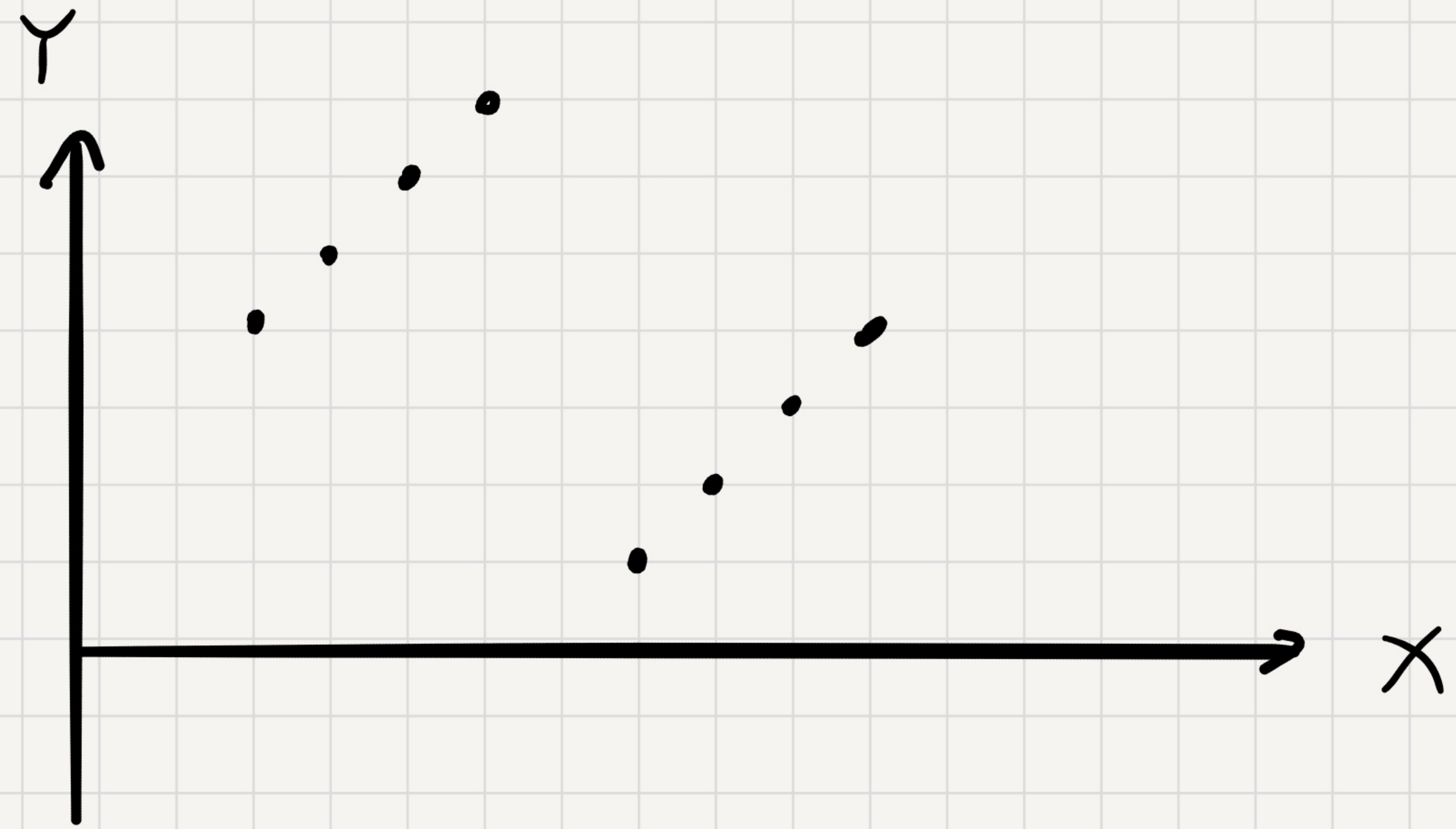
Simpson's Paradox

Results from real-life medical study:

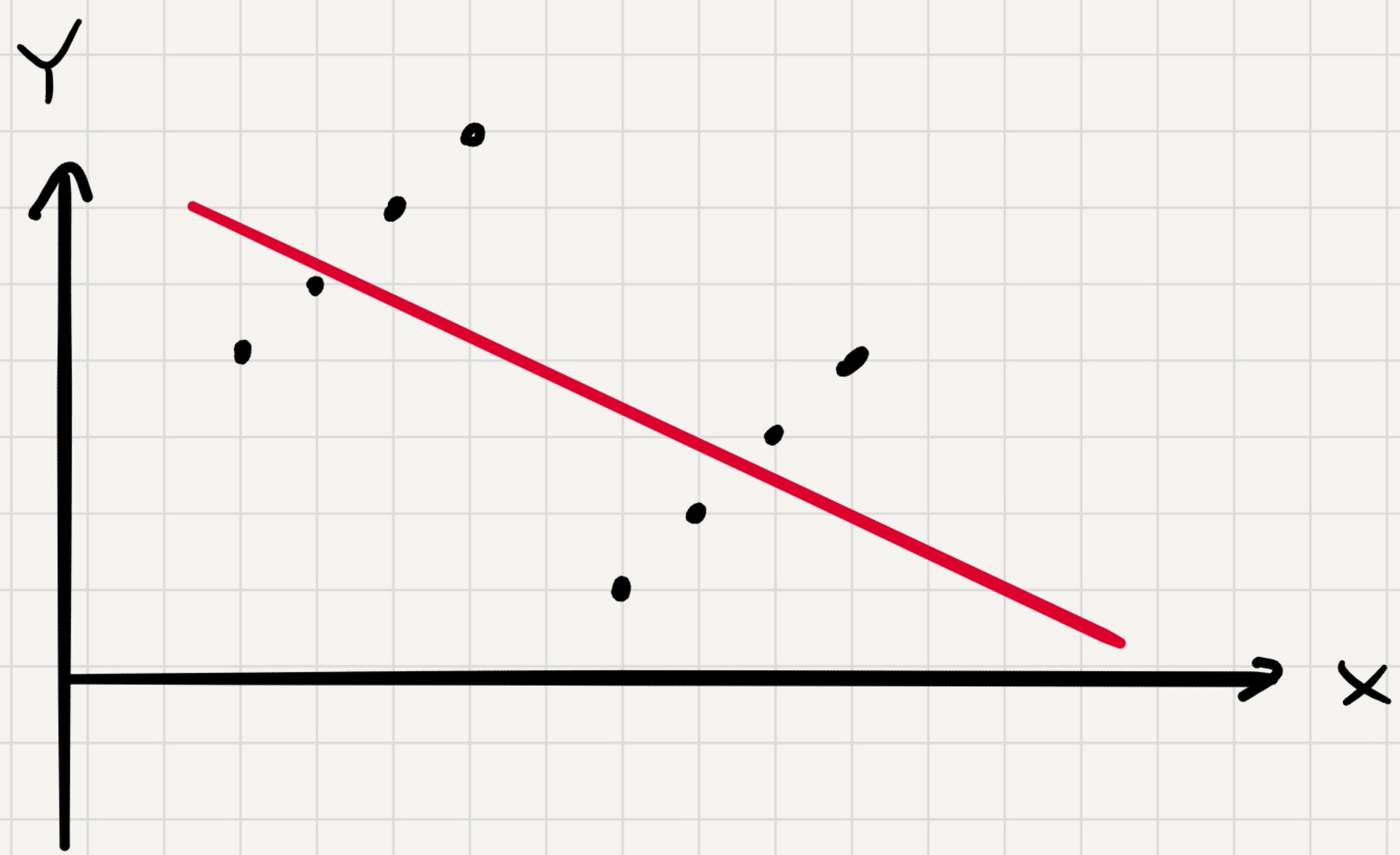
	Treatment A	Treatment B
Small Stones	Group 1 93% (81/87)	Group 2 87% (234/270)
Large Stones	Group 3 73% (192/263)	Group 4 69% (55/80)
Both	78% (273/350)	83% (289/350)

Which treatment is better?

What's the best linear predictor of Y given X?



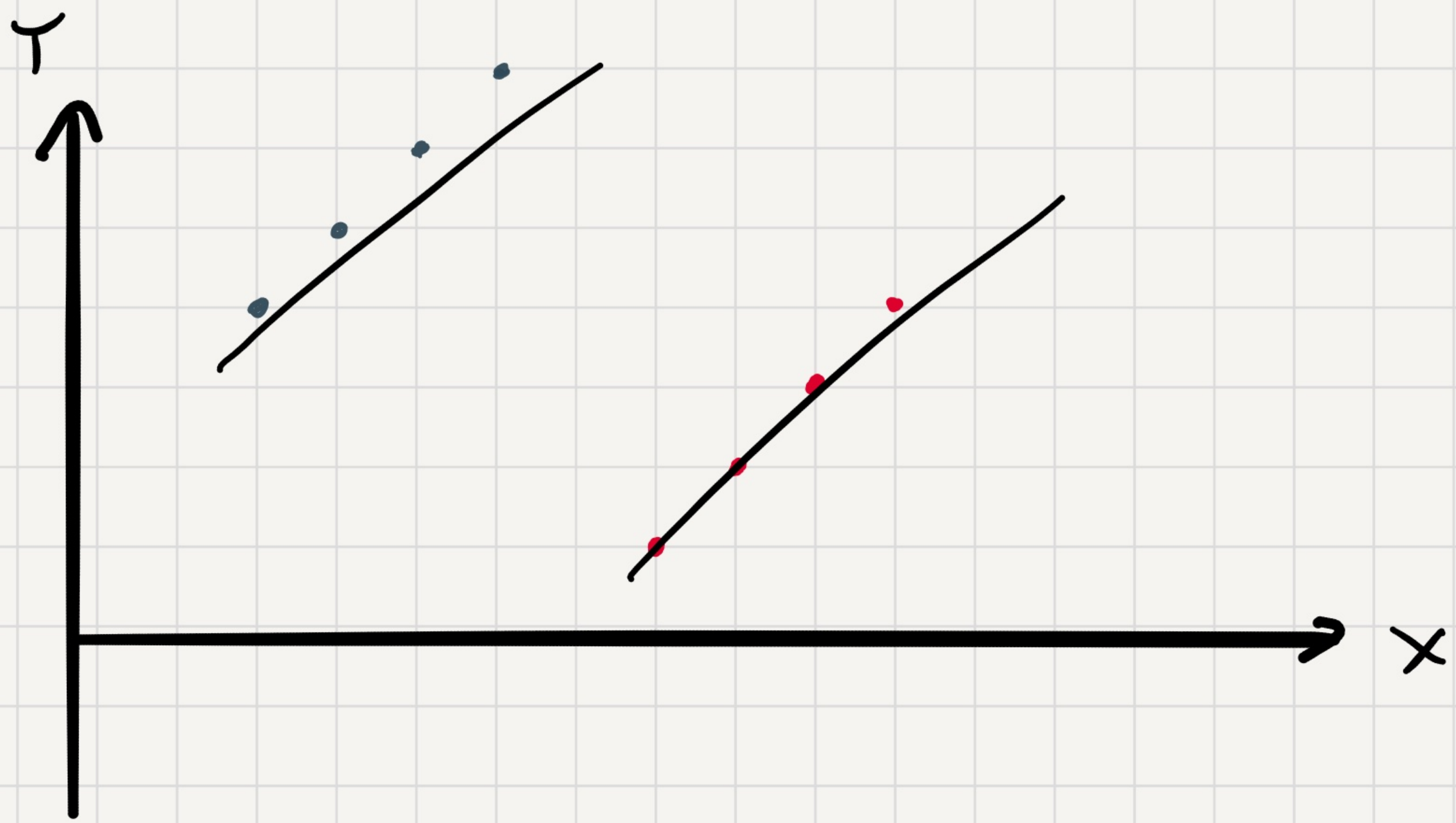
what's the best linear predictor of Y given X?



trend is down: as X increases

we predict that Y decreases

But if the data is coming from a mixture of two distributions red and blue:



In each subgroup trend is up: as X increases we predict that Y increases.

Special Thanks to the CS70 staff:

- **Prof. Rao.**
- **TAs**: Alec, Christine, Gavin, Shreyas, Caroline, Dominic, Evelyn, Hongxun, Josh, Michelle, Nathan, Sriram, Xin, Yinuo.
- **Readers**: Allen, Andrew, Arjun, Ashton, Ayush, Bryan, Carolyn, Casey, Charming, Evan, Haroon, James, Jason, Jeff, Jennifer, Matthew, Nathan, Owen, Pranesh, Raymond, Rohan, Sena, Sina, Thomas, Youngmin.

Academic Interns: Gabriel, Michael, Shaswat, Sriram, Tejasv, Yash, Zixun.

Thank you
and good luck in
the final!