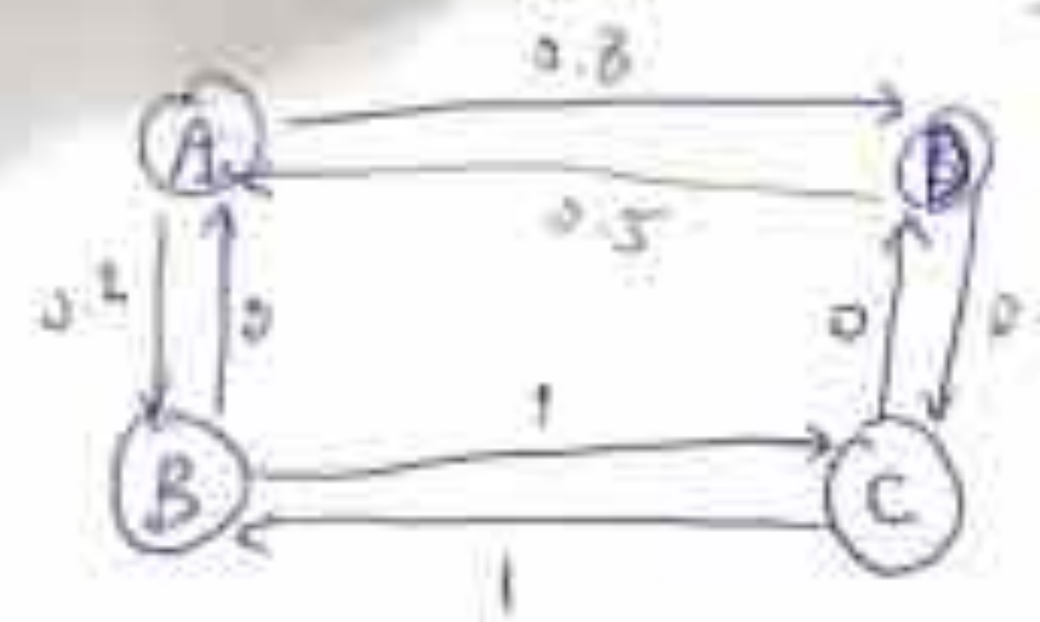
Markov Chains - ~~Jan~~ Aug 2024



$$P = \begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 0 & 0.2 & 0 & 0.8 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0.5 & 0 & 0.5 & 0 \end{bmatrix} \end{array} \; = 1 \quad \rightarrow \text{Transition Matrix}$$

Is it Irreducible? No, because not all states communicate. For example if you are in C you cannot go to A, ~~B~~ D. If you are in D you cannot go to B and ~~B~~ C, and if you are in B you cannot go to A, ~~B~~ and D. If you are in A you cannot go to ~~B and~~ C.

Is it Aperiodic? What is the period for each state?

A state is aperiodic if the greatest common divisor (gcd) of the lengths of all possible return times to the state ~~s~~ is 1. A MC is aperiodic if all states are aperiodic.

1. State A: possible transitions: $A \to B$ or $A \to D$ (returns occur in multiples of 2: $t = 2, 4, 6, \dots$) Returns to state A via $A \to D \to A$ requiring 2 steps $\Rightarrow$ Period = 2 $T(A) = \{2, 4, 6, \dots\}$, $\gcd(T(A)) = 2$ — period $d(4) = 2$
2. State B: possible transitions: $B \to C \to B$ returns to state B by 2 steps $\Rightarrow$ Period = 2
3. State C: possible transitions: $C \to B \to C$ — // — C — // —
4. State D: possible transitions: $D \to A \to D$ ~~or D → C → B~~ — // —

MC: periodic and all states have a period of 2

Being in state A at time 0 what is the prob. of being in state ~~B~~ at time 5 (after 5 steps)

$\Pi_0 = (1\ 0\ 0\ 0)$ $\quad \Pi_5 = \Pi_0 \cdot P^5$

Define T as the first time being in state D starting in state A. Calculate $P(T=1), P(T=2), P(T=3), P(T=4), P(T=5), P(T=\infty)$

~~P~~ $T=1$: means transition from D to A in exactly 1 step. (check prob in P) $P(T=1) = 0.8$

$T=2$: means we do not transit to state D on the first step but the second.

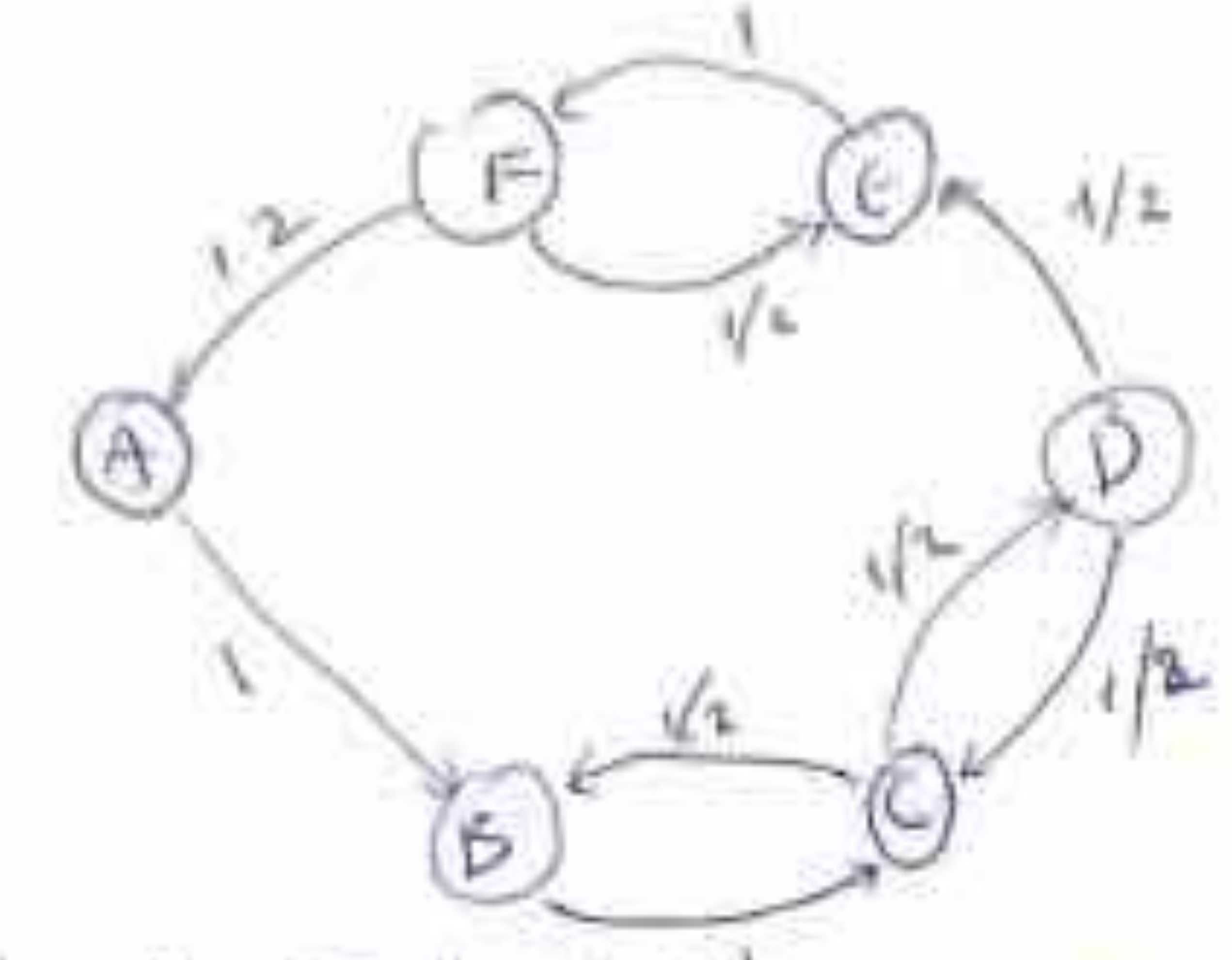
$P(T=2) = P(A,B) \cdot P(B,D) + P(A,C) \cdot P(C,D) = 0.2 \cdot 0 + 0 \cdot 0 = 0$

$P(T=3) = \sum_{x \neq D} P(A,x) \cdot P(x, x') \cdot P(x', D) = P(A,B) \cdot P(B,C) \cdot P(C,D) + P(A,C) \cdot P(C,B) \cdot P(B,D) = 0$

$$\left.\begin{aligned} P(T=4) &= P(A,B) \cdot P(B,C) \cdot P(C,B) \cdot P(B,D) = 0.2 \cdot 1 \cdot 1 \cdot 0 \\ &+ P(A,C) \cdot P(C,B) \cdot P(B,C) \cdot P(C,D) = 0 \cdot 1 \cdot 1 \cdot 0 = \end{aligned}\right\} = 0.$$

$P(T=5) = A \to B \to C \to B \to C \to D = 0 \quad + \quad A \to C \to B \to C \to B \to D = 0$

$P(T=\infty) = 0$ we never reach D from A. The only valid path from A is going to B and then transitioning to other states, but there are no valid paths that avoid state D indefinitely.



$$P = \begin{array}{c|c} & \begin{array}{cccccc} A & B & C & D & E & F \end{array} \\ \begin{array}{c} A \\ B \\ C \\ D \\ E \\ F \end{array} & \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0.5 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0.5 & 0 & 0 & 0 & 0.5 & 0 \end{bmatrix} \end{array}$$

Is it irreducible? No. A can transit to ~~...~~ B. B can only go to C. C can go to B and D only. D is C and E. E only to F and F can go to A and E.
Therefore not all state communicate ⇒ some are ~~not~~ unreachable from others.
A & B : are disconnected from the rest of the system. other states may form a subgraph.

Is it aperiodic? What is the period of each state? Aperiodic.

1. State A: possible transitions: $A \to B \to C \to D \to E \to F \to A$ or ~~...~~ Period = 6.
$L = \{6, 12, 18, \dots\}$ gcd = 6. (periodic)

2. State B: $B \to C \to D \to C \to B$ (4) or $B \to C \to B$ (2) or $B \to C \to D \to E \to F \to A \to B$ (6)
gcd(2,4,6) = 2 (periodic with period = 2)

3. State C: $C \to B \to C$ (2), $C \to D \to C$ (2), $C \to D \to E \to F \to A \to B \to C$ (6) gcd(2, 2, 6) = 2

4. State D: $D \to C \to D$ (2), $D \to E \to F \to A \to B \to C \to D$, gcd = 2

5. State E: $E \to F \to E$ (2) or $E \to F \to A \to B \to C \to D \to E$ (6)
$E \to F \to A \to B \to C \to D \to F \to E$ (8) gcd(2, 6, 8) = 2.

6. State F: $F \to E \to F$ or $F \to A \to B \to C \to D \to E \to F$ gcd(2, 6) = 2

All chain ~~is~~ is periodic with period = 6.

From A to D in $P(T=1)$. $P(T=2)$ $P(T=3)$ $P(T=4)$ $P(T=5)$ $P(T=\infty)$

$P(T=1) = 0$. no way to do it in one step

$P(T=2) = 0$

$P(T=3) = P(A \to B) \cdot P(B \to C) \cdot P(C \to D) = 1 \cdot 0.5 \cdot 1 = 0.5$

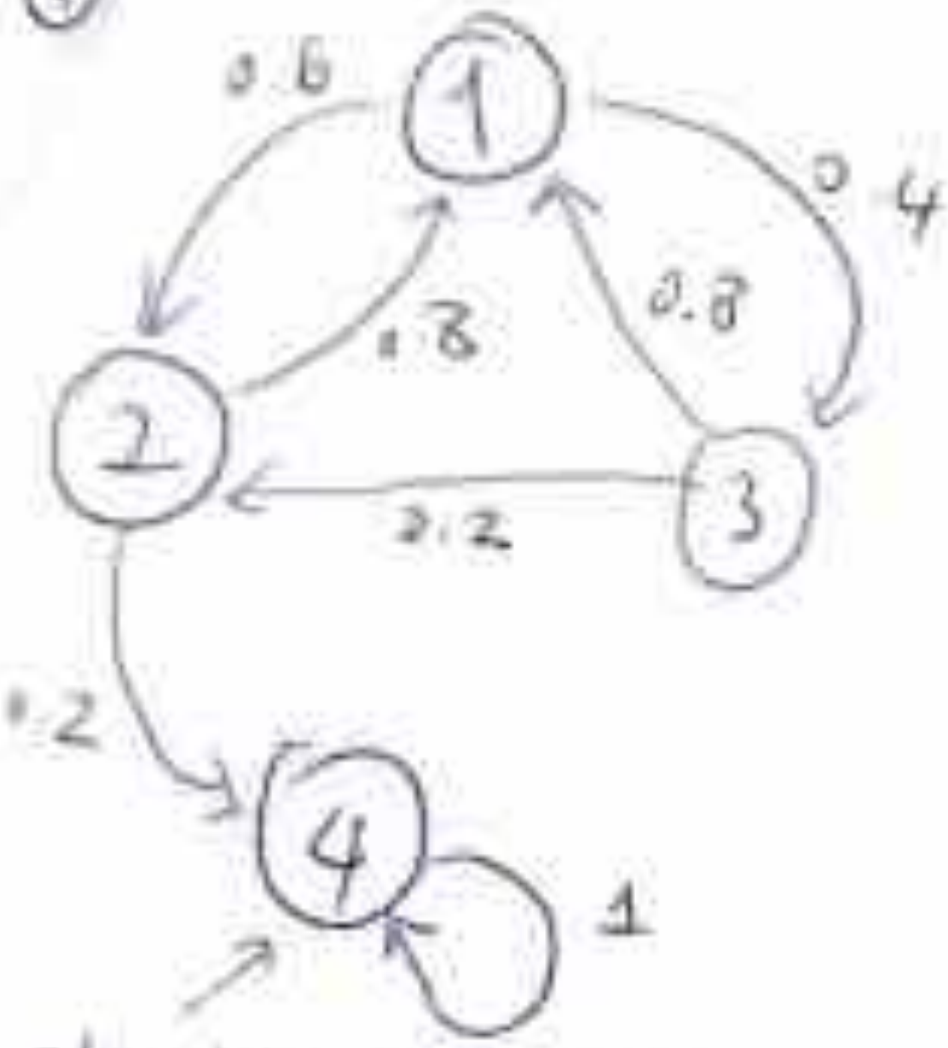
$P(T=4) = 0$

$P(T=5) = P(A \to B) \cdot P(B \to C) \cdot P(C \to B) \cdot P(B \to C) \cdot P(C \to D) = 1 \cdot 1 \cdot 0.5 \cdot 1 \cdot 0.5 =$

$P(T=\infty)$ = prob that state D is never reached from state A. = 0. ⇒ state D will be reached eventually.

Markov's Chains

①



absorbing state (once you get there you stay there forever)

$\to$ initial distribution

$\pi_0 = (1, 0, 0, 0)$

$P(X = S_4) \to$ Hitting time of arriving at state 4

$h_{1,4} = 0.6\, h_{2,4} + 0.4\, h_{3,4}$

start, end, start from 1 go to 2, start 1 go to 3

$h_{2,4} = 0.8\, h_{1,4} + 0.2\, h_{4,4}$ ← 1 $= 0.8 h_{1,4} + 0.2$

from 2 go to 1

$h_{3,4} = 0.8\, h_{1,4} + 0.2\, h_{2,4}$

$h_{1,4} = 0.6(0.8\, h_{1,4} + 0.2\, h) + (0.4(0.8\, h_{1,4} + 0.2(0.8 h_{1,4} + 0.2)))$

$\Rightarrow h_{1,4} = 1$ the probability of ending in state 4 ~~then~~

↳ makes sense as you enter in that state and you never get out

How many steps (on avg) if starting in state 1 to go to state 4?

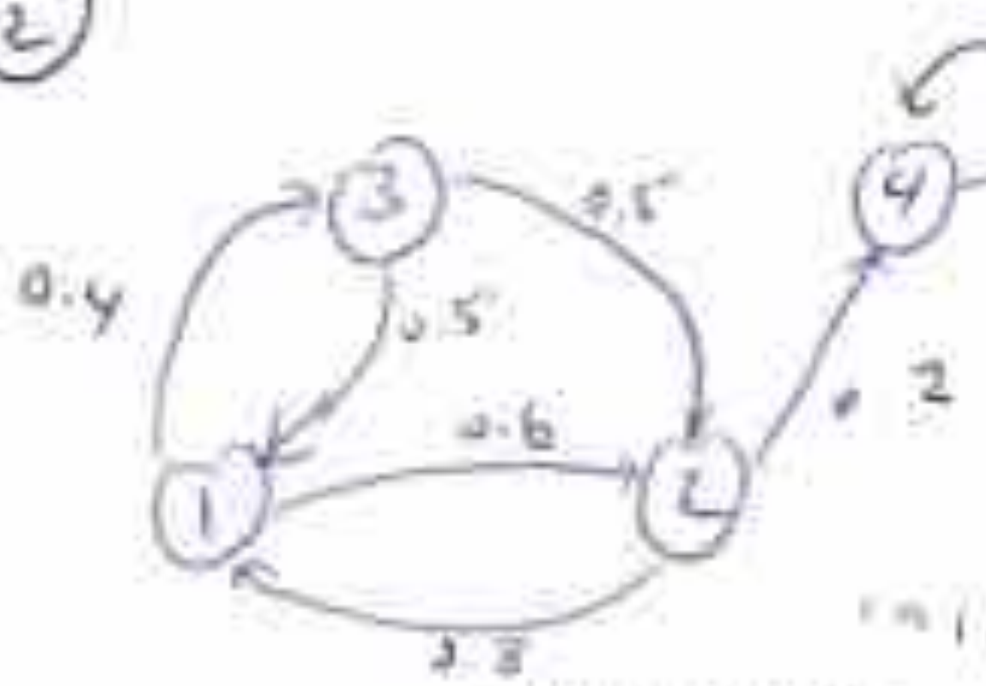
$K_{1,4} = 1 + 0.6 \cdot K_{2,4} + 0.4 \cdot K_{3,4}$ (we need to start)

$K_{2,4} = 1 + 0.2\, K_{4,4} + 0.8\, K_{1,4}$

0 going from 4 to 4

$K_{3,4} = 1 + 0.2\, K_{2,4} + 0.8\, K_{1,4} \Rightarrow K_{1,4} = 15.29$ expected # of steps to take if start 1 to go to 4

②



The absorption probability to 4 is 1 (no matter where you start eventually will reach to 4)

But how long it will take to go to 4?

Expected # of transitions k_i until reaching 4 given the initial state i

if start at 4 then $k_4 = 0$

$k_2 = 1 + 0.2\, k_4 + 0.8\, k_1 = 1 + 0.8 k_1$ (1st transition; 0)

$k_1 = 1 + 0.6 k_2 + 0.4\, k_3$

$k_3 = 1 + .5 k_1 + 0.5\, k_2$

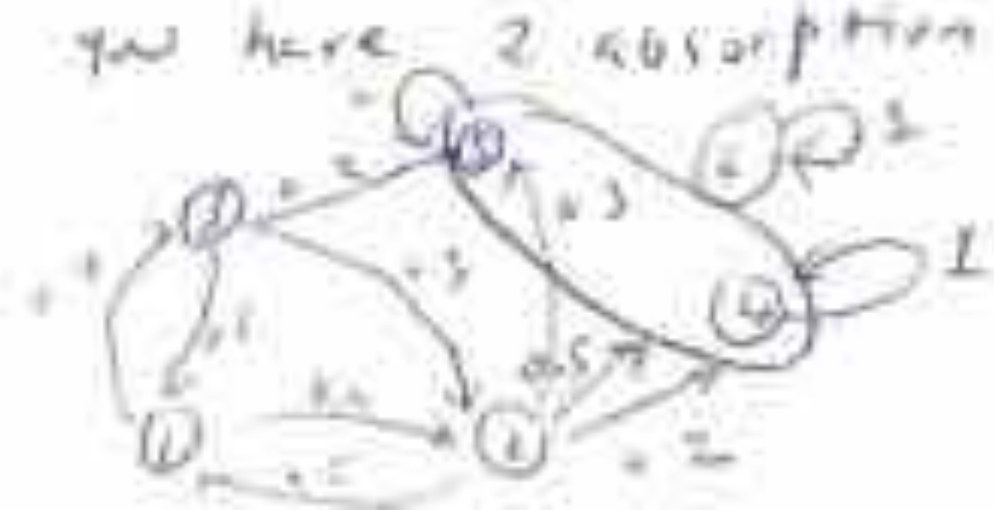
$\left.\right\}$ $k_1 = 110/8$, $k_2 = 96/8 = 12$, $k_3 = 111/8$

k_2 closer to 4 ⇒ should be the smallest

$k_i = 1 + \sum_j P_{ij} k_j$ → when 1 absorption states

* if you have 2 absorption states you need to combine them into 1 mega state

eg



Markov Chain - Autograded 2

$$
\begin{array}{c|c}
 & \begin{array}{ccc} \text{Downtown} & \text{Suburbs} & \text{Countryside} \end{array} \\
\begin{array}{c} \text{Downtown} \\ \text{Suburbs} \\ \text{Countryside} \end{array} & \begin{bmatrix} 0.3 & 0.4 & 0.3 \\ 0.2 & 0.5 & 0.3 \\ 0.4 & 0.3 & 0.3 \end{bmatrix}
\end{array}
$$

1. Truck currently in Suburbs, prob to be in downtown region after 2 steps?

$\pi_0 = [0 \ 1 \ 0]$

$$P \cdot P = \begin{bmatrix} 0.3 & 0.4 & 0.3 \\ 0.2 & 0.5 & 0.3 \\ 0.4 & 0.3 & 0.3 \end{bmatrix} \begin{bmatrix} 0.3 & 0.4 & 0.3 \\ 0.2 & 0.5 & 0.3 \\ 0.4 & .3 & 0.3 \end{bmatrix} = \begin{bmatrix} 0.29 & 0.41 & 0.30 \\ 0.28 & 0.42 & 0.30 \\ 0.30 & .40 & .30 \end{bmatrix}$$

$\pi_2 = \pi_0 \cdot P^2$

$P = 0.28$

2. If currently in Suburbs what is prob to be in downtown for the first time after 2 steps?

$$P = (P_{S \to C})(P_{C \to D}) + (P_{S \to S} \cdot P_{S \to D}) = 0.3 \cdot 0.4 + 0.5 \cdot 0.2 = 0.12 + 0.10 = 0.22$$

3. Is it irreducible? yes all states communicate

4. stationary distribution:

$$0.3\,\pi_D + 0.2\,\pi_S + 0.4\,\pi_C = \pi_D$$
$$0.4\,\pi_D + 0.5\,\pi_S + 0.3\,\pi_C = \pi_S$$
$$0.3\,\pi_D + 0.3\,\pi_S + 0.3\,\pi_C = \pi_C$$
$$\pi_D + \pi_S + \pi_C = 1$$

5. Expected # of steps until the first time we enter the downtown region having started in Suburbs (compute prob. hitting times < 30).

$$E[T] = \sum t \cdot P(T = t)$$

Let initial state $[0 \ 1 \ 0] = \pi_0$

Compute Prob. being in downtown after t step: $\pi_0 \cdot P^t$

Subtract the probabilities of previously being in the downtown at earlier steps to ensure it's the first time.

Sum up $t \cdot P(T = t)$ to find $E[T]$

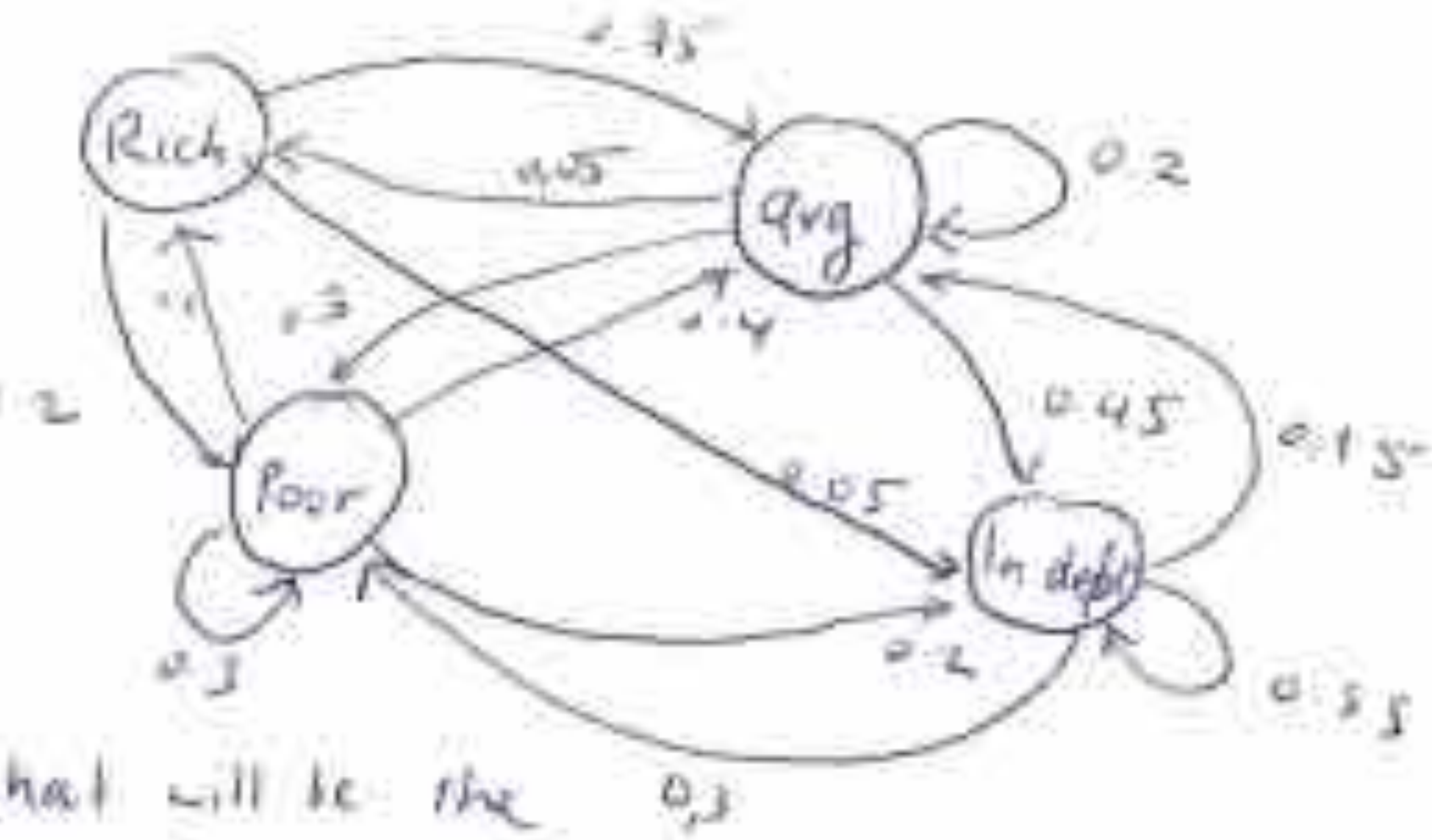
es. $P(T=1) = P(\text{downtown} \mid \text{suburbs}) = 0.2$

$P(T=2) = 0.22$ (previously calculated).

Markov Chains Exercises

↦ * Needs squared matrices and prob ≥ 0

	Rich	Average	Poor	In debt
Rich	0	0.75	0.2	0.05
Avg	0.05	0.2	0.3	0.45
Poor	0.1	0.4	0.3	0.2
In debt	0	0.15	0.3	0.55



a) Assume a student starts at stage "Avg". What will be the probability of being "Rich" after 1, 2, 3 time steps?

$\pi^{(0)} = (0, 1, 0, 0)$

$$P = \begin{pmatrix} 0 & 0.75 & 0.2 & 0.05 \\ 0.05 & 0.2 & 0.3 & 0.45 \\ 0.1 & 0.4 & 0.3 & 0.2 \\ 0 & 0.15 & 0.3 & 0.55 \end{pmatrix}$$

$$\pi^{(1)} = \pi^{(0)} P = (0\ 1\ 0\ 0) \begin{pmatrix} 0 & 0.75 & 0.2 & 0.05 \\ 0.05 & 0.2 & 0.3 & 0.45 \\ 0.1 & 0.4 & 0.3 & 0.2 \\ 0 & 0.15 & 0.3 & 0.55 \end{pmatrix} = (\underline{0.05}\ \ 0.2\ \ 0.3\ \ 0.45)$$

→ After 1 step 5% chance

$\pi^{(2)} = \pi^{(1)} P^2 = (\underline{0.04}\ \ 0.265\ \ 0.295\ \ 0.4)$

→ After 2 steps 4% chance

$\pi^{(3)} = \pi^{(2)} P^{(3)} = (\underline{0.04275},\ 0.211,\ 0.296,\ 0.4025)$ → After 3 steps 4.28% chance

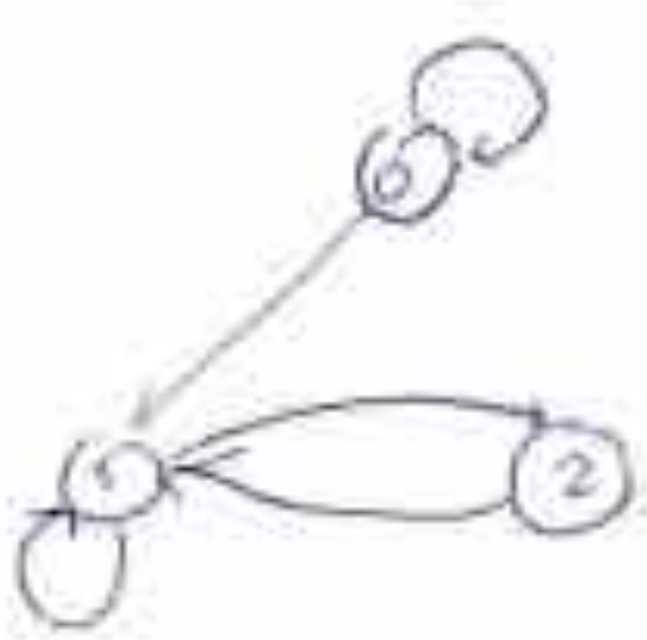
b) Find the stationary distribution

$$\left.\begin{aligned} &0.05\,\pi_A + 0.1\,\pi_P = \pi_R \\ &0.75\,\pi_R + 0.2\,\pi_A + 0.4\,\pi_P + 0.15\,\pi_D = \pi_A \\ &0.2\,\pi_R + 0.3\,\pi_A + 0.3\,\pi_P + 0.3\,\pi_D = \pi_P \\ &0.05\,\pi_R + 0.45\,\pi_A + 0.2\,\pi_P + 0.55\,\pi_D = \pi_D \\ &\pi_R + \pi_A + \pi_P + \pi_D = 1 \end{aligned}\right\} \quad \pi_R = \frac{53}{1241} \quad \pi_A = \frac{326}{1241} \quad \pi_P = \frac{367}{1241} \quad \pi_D = \frac{495}{1241}$$

2)

$$P = \begin{array}{c|c} & \begin{matrix} 0 & 1 & 2 & 3 \end{matrix} \\ \hline \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \end{matrix} & \begin{pmatrix} 0.4 & 0.6 & 0 & 0 \\ 0.2 & 0.2 & 0.5 & 0.1 \\ 0.0 & 0.2 & 0.8 & 0 \\ 0 & 0.5 & 0.5 & 0 \end{pmatrix} \end{array}$$

Irreducible: if all states communicate



probability of coming back to itself is less than 0 for state 0 as if you leave from there you cannot go back again (Transient state)

Recurrent state: you can revisit the state you started from with a prob of 1 (state 1 and 2)

Reducible: when some states are unreachable from others

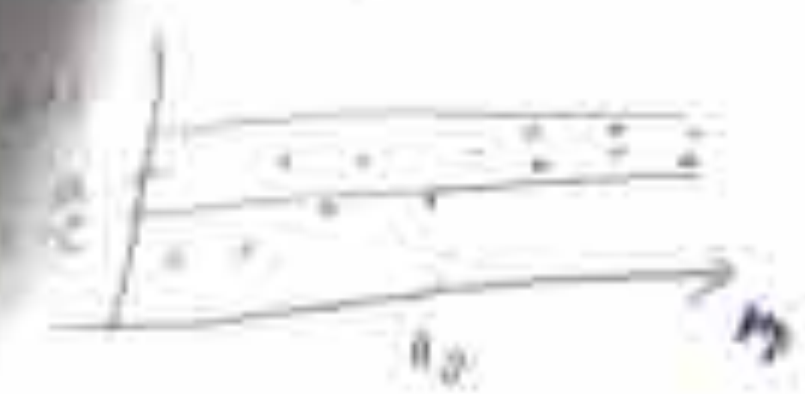
Convergence in Probability & in Distribution

Weak Law of Large numbers: with high probability the sample mean falls close to the mean as $n \to \infty$

$P(|M_n - \mu| \geq \epsilon) \to 0$, as $n \to \infty$ for any $\epsilon > 0$

Assume we have a sequence of RVs Y_n that are not necessarily independent

- A sequence Y_n converges in prob to a number a if for any $\epsilon > 0$ $\lim_{n\to\infty} P(|Y_n - a| \geq \epsilon) = 0$

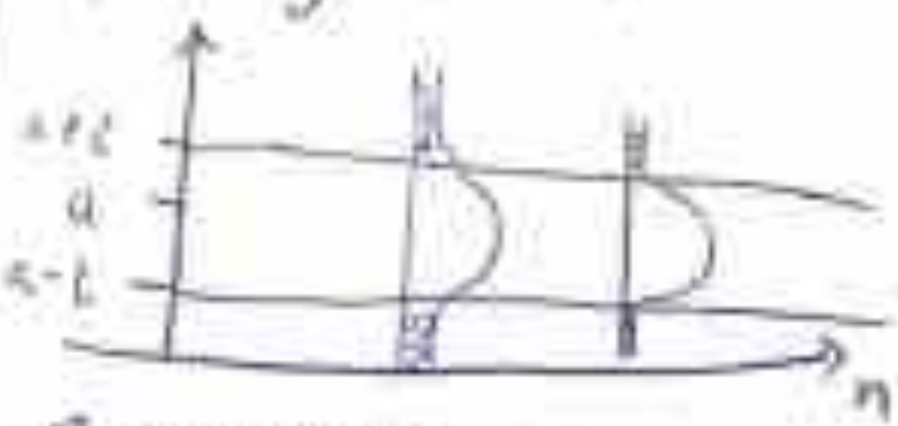


For every $\epsilon > 0$ there exists n_0 s.t. for every $n \geq n_0$, we have $|a_n - a| \leq \epsilon$

(convergence of ordinary numbers / sequence)

- Convergence in Probability: sequence Y_n: number a $\quad Y_n \to a$ - limit

For any (fixed) $\epsilon > 0$ $\quad P(|Y_n - a| \geq \epsilon) \to 0$



* (almost all) of the PMF/PDF of Y_n eventually gets concentrated (arbitrarily) close to a

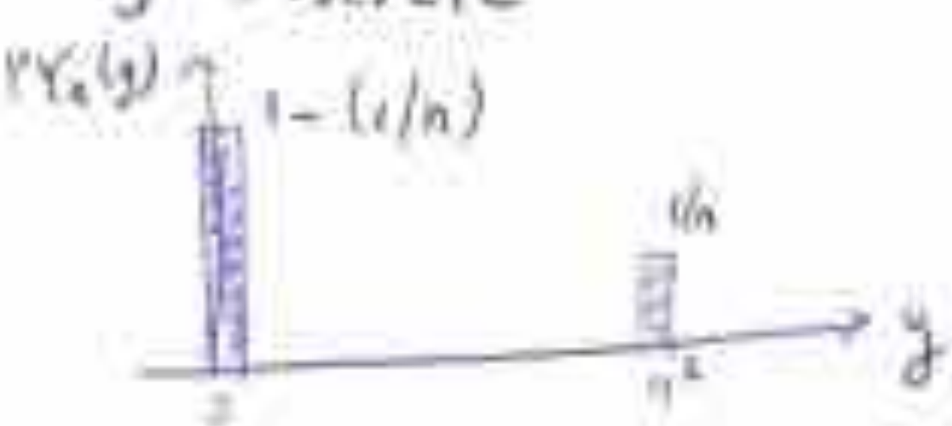
Properties: Suppose that $X_n \to a$, $Y_n \to b$ in probability

If g continuous then $g(X_n) \to g(a)$

$X_n + Y_n \to a + b$

But X_n might converge to a certain number in probability. However the expected value of X_n does not necessarily converge to that same limit

eg. Discrete



$\epsilon > 0$: $P(|Y_n - 0| \geq \epsilon) = \frac{1}{n}$ $\quad \lim_{n\to\infty} \frac{1}{n} = 0$ $\quad Y_n \xrightarrow[n\to\infty]{\text{in prob}} 0$

$E[Y_n] = n^2 \frac{1}{n} = n$ $\quad \lim_{n\to\infty} n = \infty$

* convergence in prob doesn't imply convergence of expectations

eg. Uniform

X_i: iid uniform on $[0, 1]$

it doesn't converge ⇒ as i ↑ the distribution doesn't change and doesn't get concentrated around a certain number. The distrib remains spread out over the entire unit interval

But let's say $Y_n = \min\{X_1, \dots X_n\}$ → The only thing that can happen is that the min goes down - it cannot go up $Y_{n+1} \leq Y_n$

For $\epsilon > 0$ $\quad P(|Y_n - 0| \geq \epsilon) = P(Y_n \geq \epsilon)$

$\epsilon > 1$: This impossible $= 0$

$\epsilon < 1$: $P(X_1 \geq \epsilon, \dots, X_n \geq \epsilon) \overset{\text{independence}}{=} P(X_1 \geq \epsilon) \dots P(X_n \geq \epsilon) = (1-\epsilon)^n \xrightarrow[n\to\infty]{} 0$ $\quad Y_n \xrightarrow[n\to\infty]{} 0$

Convergence in distribution: CDF

$X_N \xrightarrow{d} X$

$\lim_{n\to\infty} |F_N(X_N) - F(X)| = 0$ (cdf)

convergence in probability ⇒ convergence in distribution (not the other way around!)

$X_N \xrightarrow{P} C$

$X_N \xrightarrow{d} C$