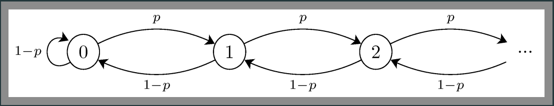
Markov chains: the ergodic theorem

Cases where space averages are the same as time averages



assuming $p \in (0, 1/2)$.

$$\pi(i) = \left(\frac{1-2p}{1-p}\right)\left(\frac{p}{1-p}\right)^i, \; i \geq 0$$

Why do we care about the steady-state?

Because as $t \to \infty$, X_t 'appears' like an iid sample from π

not true, but seems like it

How can we formalize appears?

Consider any function $g(\cdot)$ s.t. $E_{Y\sim\pi}[g(Y)] < \infty$. Then

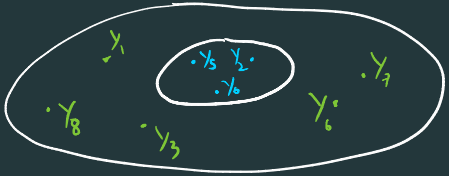
$$\boxed{\text{Time Avg}} \quad \underbrace{\lim_{T\to\infty}\left[\frac{1}{T}\sum_{t=0}^{T-1} g(X_t)\right]}_{\text{Can measure via simulation}} = \underbrace{E_{Y\sim\pi}\left[g(Y)\right]}_{\text{estimate using iid samples}} \quad \boxed{\text{Space Avg}}$$

sample path of MC

Sample from π

- Can compute in closed form if π is known

$$\hat{M} = \frac{1}{N} \sum_{i=1} g(y_i) \quad , \quad y_i \sim \Pi \quad \text{iid}$$



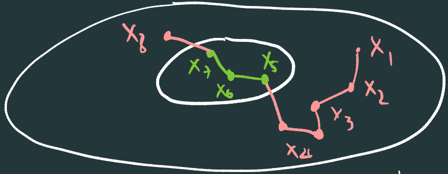
$\hat{M} = 3/8$

Same as $T \to \infty$

Markov chain

$$\tilde{M} = \frac{1}{T} \sum_{t=1}^{T} g(x_t) \quad , \quad X_t \sim \Pi_t \quad \text{where } \Pi_t^T = \Pi_{t-1}^T P$$

$\boxed{\Pi \text{ is the stat dist}^{n}}$



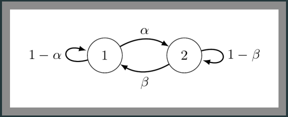
$\tilde{M} = 3/8$

• Important for – 1) understanding how to report outputs from MC

2) Markov Chain Monte Carlo

class poll: long-term behavior in the flip-flop

Last class - DTMCs { absorbing, transient, recurrent



in the flip-flop Markov chain, which of the following outcomes is possible for $\pi_t(1) = \mathbb{P}[X_t = 1]$ (for different α, β)?

Q: 1) When is X_t absorbed
2) How long

$\boxed{\text{absorbing}}$

(a) $\lim_{t\to\infty} \pi_t(1) = 0$ (set $\alpha = 1$, $\beta = 0$, any X_0)

(b) $\lim_{t\to\infty} \pi_t(1) = 1$ (set $\alpha = 0$, $\beta = 1$, any X_0)

(c) $\lim_{t\to\infty} \pi_t(1)$ settles down to a constant $\pi(1) \in (0, 1)$ $\alpha, \beta \in (0,1)$

$\boxed{\text{recurrent}}$
$\pi(1) = \beta/(\alpha+\beta)$
$\pi(2) = \alpha/(\alpha+\beta)$

(d) $\lim_{t\to\infty} \pi_t(1)$ oscillates (set $\alpha = \beta = 1 \Rightarrow X_0 = 1 \to X_1 = 2 \to X_2 = 1 \to \dots$)

$\boxed{\text{transient}}$ (want to avoid)

✓ (e) all of these

DTMC: applications and problems

Many settings where DTMCs are sufficient

- Complex discrete-time models
  - gambler's ruin
  - coupon collector
  - random walk on integers
- Natural for 'randomized algorithms'
  - quicksort with random pivots
  - MCMC for sampling

Problem

- If we have many competing (ie, simultaneous) events, then the ORDER of events matters

- WANT - some 'easy' way to make the order irrelevant (while retaining the main features of the model)

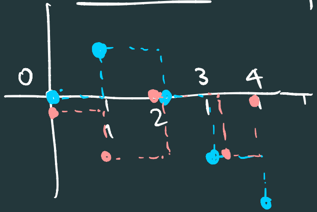
from discrete to continuous time (same Markovian model in continuous time)

Markov property

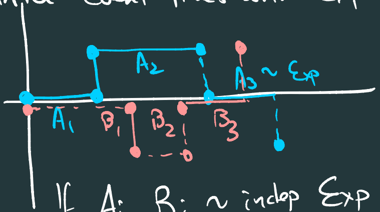
random process X_t has the Markov property if the probability of moving to a future state only depends on the present state and not on past states

- $\mathcal{S}$ discrete, $\mathcal{T}$ discrete: discrete-time Markov chain (DTMC)
  - random walk
- $\mathcal{S}$ ~~continuous~~ discrete, $\mathcal{T}$ ~~discrete~~ continuous: continuous-time Markov chain (CTMC)
  - Poisson process

<u>MAIN IDEA</u> - Replace all inter-event times with Exponentials



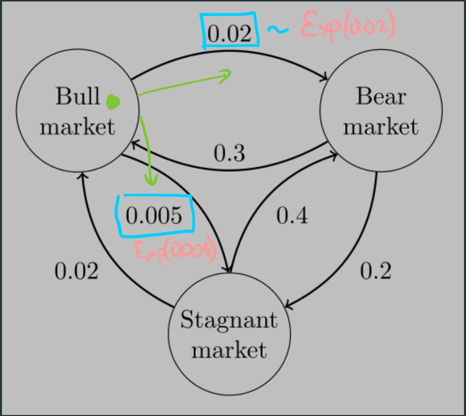
$\Rightarrow$



If $A_i, B_i \sim$ indep Exp

then events NEVER overlap

CTMC: continuous-time Markov processes on discrete state-space
eg. modeling the financial market in continuous time



Same transition diagram as the DTMC

- ONE change
  - For DTMC: edge wt $w_{ij} = P_{ij}$ $\left(\sum_j w_{ij} = 1 \right)$
  - For CTMC: edge wt w_{ij} is the RATE of the exponential clock after which we transition from i to j (and no self-loops)

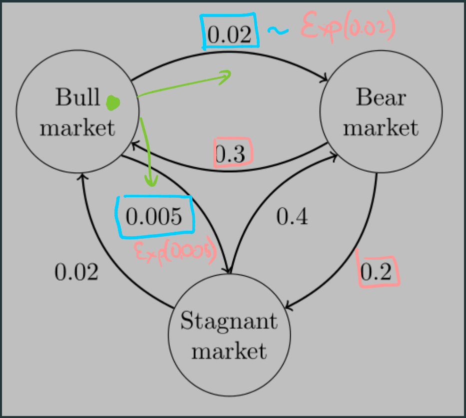
For eg. $\mathbb{P}[\text{Bull} \to \text{Bear}] = \mathbb{P}[\text{Exp}(0.02) < \text{Exp}(0.005)] = \frac{0.02}{0.025} = \frac{4}{5}$

$\mathbb{P}[\text{Bull} \to \text{Recession}] = \frac{0.005}{0.025} = \frac{1}{5}$

continuous-time Markov chains

CTMC: continuous-time Markov processes on discrete state-space

eg. modeling the financial market in continuous time   Suppose rates are in months^{-1}



- Time spent in BULL mkt

$$= \min\{\mathcal{E}xp(0.02), \mathcal{E}xp(0.005)\}$$

$$= \mathcal{E}xp(0.025)$$

$$\Rightarrow \mathbb{E}[T_{BULL}] = 40 \text{ months}$$

- Time spent in BEAR mkt

$$= \min(\mathcal{E}xp(0.3), \mathcal{E}xp(0.2))$$

$$= \mathcal{E}xp(0.5)$$

$$\Rightarrow \mathbb{E}[T_{BEAR}] = \frac{1}{0.5} = 2 \text{ months}$$

simulating a CTMC (minimum # of updates)

At time t, state $X(t) = x$, do

(Idea - Start an $\mathrm{Exp}(q_{xy})$ clock for each edge x-y
- When first clock ticks, follow that edge)

1) Find next event time $= \underbrace{t}_{\text{current time}} + \underbrace{\overbrace{\mathrm{Exp}\left(\sum_y q_{xy}\right)}^{\text{random inter-event time}}}_{\text{min of expos}}$

2) Find next state $= y$ wp $q_{xy} / \sum_{y'} q_{xy'}$

ie. $X(t) = x \longrightarrow X(t + \underbrace{A}_{\mathrm{Exp}(\sum_y q_{xy})}) = \underbrace{Y}_{\mathrm{Mult}(q_{xy})}$

example: queueing

the single-server M/M/1 queue

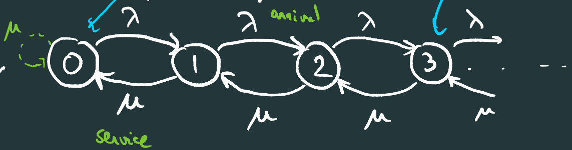
- number of servers: 1
- capacity: infinite
- service discipline: first-come-first-served (FCFS or FIFO)
- interarrival times: exponential(λ)
- service times: exponential(μ)
  (independent interarrival and service times)

time in state 0: Exp(λ)

OR - time until first arrival if you have an and service

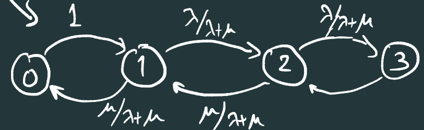
time in state 3 = Exp($\lambda+\mu$)

CTMC - N(t): # of people in queue at time t



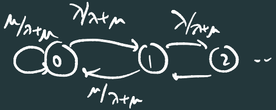
DTMC on the 'jumps'

JUMP chain



OR

Uniformized DTMC



example: simulating an $M/M/1$ queue

- Maintain state $N(T_i)$ at time T_i, set $N(0) = 0$

  random time corresponding to i^{th} event

- At time T_i, state $N(T_i)$:

uniformized chain view

- Advance time to $T_{i+1} = T_i + \mathrm{Exp}(\lambda + \mu)$
- Change state to $N(T_{i+1}) = \begin{cases} N(T_i) + 1 & \text{wp } \lambda/(\lambda+\mu) \\ (N(T_i) - 1)^+ & \text{wp } \mu/(\lambda+\mu) \end{cases}$

$(x)^+ = \max(x, 0)$

OR

jump chain

- Advance time to $T_{i+1} = T_i + \begin{cases} \mathrm{Exp}(\lambda+\mu) & \text{if } N(T_i) > 0 \\ \mathrm{Exp}(\lambda) & \text{if } N(T_i) = 0 \end{cases}$
- Change state to $N(T_{i+1}) = \begin{cases} N(T_i) + 1 & \text{wp } \lambda/(\lambda+\mu) \\ N(T_i) - 1 & \text{wp } \mu/(\lambda+\mu) \end{cases}$ if $N(T_i) > 0$

  and $N(T_{i+1}) = 1$ if $N(T_i) = 0$

example: epidemics

SIS epidemic

want to model the spread of the latest influenza strain among the population

– there is a population of n people

– at any time t, each person i is either susceptible (denoted S or 0) or infected (denoted I or 1)

– infected people get cured on average after time τ, becoming susceptible to future infection. – each pair of people (i, j) independently meet each other with average rate λ

– when a susceptible person meets an infected person, the susceptible person becomes infected

– when two susceptible people or two infected people meet, nothing happens

Network view $X(0) = (I, S, S)$

$= (1, 0, 0)$

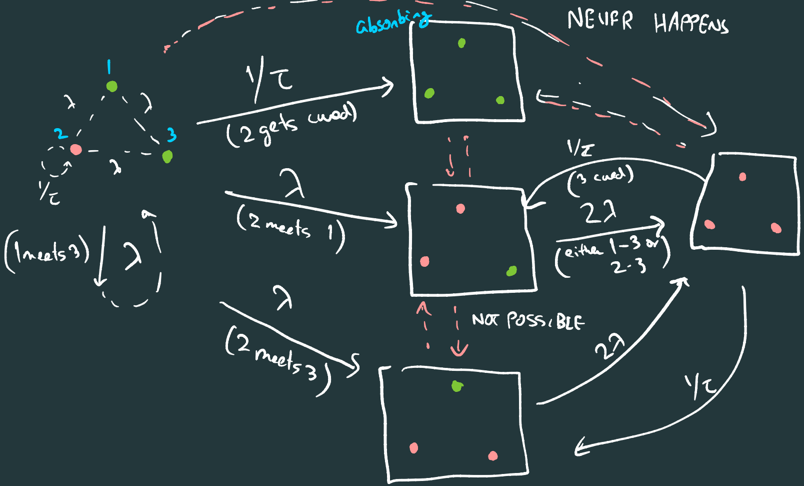
Individual View - For any node i

$\mathrm{Exp}(\lambda \cdot \#\text{ of infected nbrs})$

$\mathrm{Exp}(1/\tau)$

example: SIS epidemic (contnd)

what assumptions do we need to make this Markovian?



Look at intro notebook - SIS on a grid with an 'external infection'