IN GENERAL, THE ANSWERS DEPEND ON THE TRANS^n PROBS & THE STARTING STATE.

Def^n LET $(\underline{S}, \underline{P})$ BE AN ABSORBING M.C. If THE STATES ARE ORDERED SO THAT THE ABSORBING STATES ARE first, THE M.C. IS IN CANONICAL form

If $(\underline{S}, \underline{P})$ IS AN AMC IN CAN^l form, AND $\exists$ r ABSORBING & t TRANSIENT STATES, $\underline{P}$ TAKES THE form

$$\underline{P} = \begin{array}{c c} & \begin{array}{cc} \text{Tr} & \text{Abs} \end{array} \\ \begin{array}{c} \text{Tr} \\ \text{Abs} \end{array} & \begin{pmatrix} \underline{Q} & \underline{R} \\ 0 & \underline{I}_r \end{pmatrix} \end{array}$$

HERE $\underline{I}_r$ IS THE $r \times r$ IDENTITY MATRIX, $\underline{R}$ IS A $t \times r$ NON-ZERO MATRIX, & $\underline{Q}$ IS A $t \times t$ MATRIX. STUDY ABSORPTION after n steps $\Rightarrow$ $\underline{P}^n$

$$\underline{P}^n = \begin{pmatrix} \underline{Q}^n & * \\ 0 & \underline{I}_r \end{pmatrix}$$

$*$ IS A COMPLICATED COMB^n of $\underline{Q}$ & $\underline{R}$. WE WILL SHOW THAT $Q^n \to 0$ AS $n \to \infty$.

Thm IN AN ABSORBING MC, THE PROB THAT THE PROCESS WILL BE ABSORBED IS $\underline{1}$.

Pf From EACH NONABSORBING STATE s_i, $\exists$ A NON-ZERO T. PROB of MOVING TO AN ABSORBING STATE. LET m_j BE THE MIN^m # of STEPS REQ^d TO REACH AN A.S. STARTING from s_j & p_j THE PROB THAT s_j WILL NOT REACH AN A.S. IN m_j steps $(p_j < 1)$. LET $m = \max(m_j)$ & $p = \max(p_j)$. THEN THE PROB of NOT BEING ABSORBED IN m STEPS IS $\leq p$, IN $2m$ steps IS $\leq 2p$, ETC. THUS THE prob of NOT BEING ABSORBED IS MONOTONICALLY DECREASING, SO $\lim\limits_{n\to\infty} \underline{Q}^n = \underline{0}$.

Expected time in transient states

Thm Let $(S, \underline{P})$ be an AMC, $\underline{P} = \begin{pmatrix} \underline{Q} & \underline{R} \\ 0 & \mathbb{1} \end{pmatrix}$. Then $\underline{N} = \mathbb{1} - \underline{Q}$ is invertible with $\underline{N}^{-1} = \sum_{i=0}^{\infty} Q^i$, & the ij^{th} entry of $\underline{N}$ is the expected # of times the process is in state s_i given that it starts in s_j.

Note that since i & j are indices associated with $\underline{Q}$, s_i & s_j are transient states.

Pf Let p be the max entry of $\underline{Q}$. Then $0<p<1$ & the max entry of Q^2 is $\leq p^2$. Similarly, the max entry of $\underline{Q}^n$ is $\leq p^n$. Consider the ij^{th} entry of $\mathbb{1} + \underline{Q} + \underline{Q}^2 + \cdots$. It is a series $\delta_{ij} + q_{ij} + q^2_{ij} + \cdots$ whose terms we can bound by $q^n_{ij} \leq p^n$, so the series is bounded by $1 + p + p^2 + \cdots = \sum_{n=0}^{\infty} p^n = \frac{1}{1-p}$. Thus, each entry in $\mathbb{1} + \underline{Q} + \underline{Q}^2 + \cdots$ is finite, & the expression exists.

Next, consider $(\mathbb{1} - \underline{Q}) \sum_{n=0}^{\infty} \underline{Q}^n = \sum_{n=0}^{\infty} \underline{Q}^n - \sum_{n=1}^{\infty} \underline{Q}^n = \mathbb{1}$. Thus, $\sum_{n=0}^{\infty} \underline{Q}^n = (\mathbb{1} - \underline{Q})^{-1}$.

Now fix i & j so s_i & s_j are transient. Let $X^{(k)}$ be a random var that is 1 if the chain is in s_j after k steps, 0 else:

$$P(X^{(k)} = 1) = q_{ij}^{(k)}$$
$$P(X^{(k)} = 0) = 1 - q_{ij}^{(k)}$$

with $q_{ij}^{(k)} = (\underline{Q}^k)_{ij}$. Note these also hold for $k=0$, since $\underline{Q}^0 = \mathbb{I}$. Clearly $E(X^{(k)}) = q_{ij}^{(k)}$.

Let $S_{(n)} = \sum_{i=0}^{n} X^{(i)}$, the # of times the chain is in state j during the first n steps. Then $E(S_{(n)}) = q_{ij}^{(0)} + \cdots + q_{ij}^{(n)} = \sum_{\ell=0}^{n} (\underline{Q}^\ell)_{ij}$

$$\& \lim_{n\to\infty} E(S_{(n)}) \to (\mathbb{1} - \underline{Q})^{-1}_{ij}$$

$\square$

Def^n Let $\left(S, \underline{P} = \begin{pmatrix} Q & R \\ 0 & I \end{pmatrix}\right)$ be an AMC. Then $N = (I - Q)^{-1}$ is the fundamental matrix of P.

Ex Drunken Student;

$$\underline{P} = \begin{array}{c|c} & \begin{matrix} 20 & 21 & 22 & 19 & 23 \end{matrix} \\ \hline \begin{matrix} 20 \\ 21 \\ 22 \\ 19 \\ 20 \end{matrix} & \left(\begin{array}{ccc|cc} 0 & 1/2 & 0 & 1/2 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & 1/2 \\ \hline 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right) \end{array}$$

$$\underline{Q} = \begin{pmatrix} 0 & 1/2 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 1/2 & 0 \end{pmatrix} \qquad I - \underline{Q} = \begin{pmatrix} 1 & -1/2 & 0 \\ -1/2 & 1 & -1/2 \\ 0 & -1/2 & 1 \end{pmatrix}$$

$$(I - \underline{Q})^{-1} = \begin{pmatrix} 3/2 & 1 & 1/2 \\ 1 & 2 & 1 \\ 1/2 & 1 & 3/2 \end{pmatrix}$$

So, if we start at 21st, we can expect to be at 21st 1 more times before getting to bar/home.

Expected Time until Absorbtion

Starting in state s_i, we expect to spend N_{ij} steps in state s_j before absorbtion. Thus, the expected total # of steps before absorbtion beginning in state s_i is $\sum_j N_{ij}$.

Absorbtion probabilities

Now let s_j be an absorbing state & s_i a transient state.

Thm If b_{ij} is the prob a chain beginning in state s_i will be absorbed in the state s_j, & $\underline{B}$ is a matrix w/ entries b_{ij},

$$\underline{B} = \underline{N}\,\underline{R}$$

Pf $r_{k\ell}$ = prob of transitioning from transient A_k to absorbing A_ℓ.

Starting from A_i, before reaching A_j. The chain will be in a transient state A_k (may be equal to A_i) after $n \geq 0$ steps. Thus,

$$b_{ij} = \sum_k \left(\sum_n q_{ik}^{(n)} \right) r_{kj}$$
$$= \sum_k N_{ik} r_{kj}$$
$$= \underline{\underline{N}}\, \underline{\underline{R}}$$

Ex $N = \begin{pmatrix} 3/2 & 1 & 1/2 \\ 1 & 2 & 1 \\ 1/2 & 1 & 3/2 \end{pmatrix}$; Starting from $\begin{pmatrix} 20 \\ 21 \\ 22 \end{pmatrix}$, expected time to bar/home is $\begin{pmatrix} 3 \\ 4 \\ 3 \end{pmatrix}$ steps.

$R = \begin{pmatrix} 1/2 & 0 \\ 0 & 0 \\ 0 & 1/2 \end{pmatrix} \Rightarrow NR = \begin{pmatrix} 3/4 & 1/4 \\ 1/2 & 1/2 \\ 1/4 & 3/4 \end{pmatrix}$

Note; always makes it to bar or home; $3/4 + 1/4 = 1/2 + 1/2 = 1/4 + 3/4 = 1$

So the probability of making it to the bar from 22^{nd} is $1/4$

Ex Gambler's Ruin (Ross - "Intro to Prob Models")

A gambler plays a game w/ prob p of winning $1^\$$ & $1-p$ of losing $1^\$$. Assuming successive plays are indep^t, what is the prob that his fortune reaches $N^\$$ before it reaches $0^\$$ if he starts from $\ell^\$$?

Treat as a Markov chain; the states of the chain are his fortune. Furthermore, $0^\$$ & $N^\$$ are absorbing states; once in them, the gambler stops playing. Clearly for all other amounts $\{1^\$, \cdots, N-1^\$\}$,

$$P_{k,k+1} = p, \quad P_{k,k-1} = 1-p, \quad \text{all others zero.}$$

$$\underline{\underline{P}} = \begin{array}{c|c} & \begin{matrix} 0 & 1 & 2 & 3 & \cdots & N \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ \vdots \\ \\ N-1 \\ N \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ q & 0 & p & 0 & \cdots & 0 \\ 0 & q & 0 & p & & \vdots \\ \vdots & & & \ddots & & \\ 0 & \cdots & q & 0 & p & 0 \\ 0 & \cdots & 0 & q & 0 & p \\ 0 & \cdots & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

CLEARLY THIS IS AN ABSORBING MARKOV CHAIN. If WE SET P_ℓ TO BE THE PROB THAT, STARTING from $\ell^{\$}$, THE GAMBLER'S fortune REACHES $N^{\$}$.

$$P_\ell = p\,P_{\ell+1} + q\,P_{\ell-1}$$

(of ALL EXPERIMENTS, A fraction p WILL INCREASE after 1 play & from THEN $P_{\ell+1}$ WILL RESULT IN A fortune of $N^{\$}$).

Since $p+q=1$,
$$p\,P_\ell + q\,P_\ell = p\,P_{\ell+1} + q\,P_{\ell-1}$$
$$\left(P_{\ell+1} - P_\ell\right) = \frac{q}{p}\left(P_\ell - P_{\ell-1}\right)$$

FTHMORE, $P_0 = 0$, SO
$$P_2 - P_1 = \frac{q}{p} P_1$$
$$P_3 - P_2 = \frac{q}{p}(P_2 - P_1) = \left(\frac{q}{p}\right)^2 P_1$$
$$\vdots$$
$$P_N - P_{N-1} = \left(\frac{q}{p}\right)^{N-1} P_1$$

ADD THE first $k>1$ of THESE EQUATIONS;

$$P_{k+1} - P_1 = P_1\left[(q/p) + (q/p)^2 + \cdots + (q/p)^k\right]$$

$$P_{k+1} = \begin{cases} \dfrac{1-(q/p)^{k+1}}{1-(q/p)}\,P_1 & \text{if } q \neq p \\[2ex] (k+1)\,P_1 & \text{if } q = p = 1/2 \end{cases}$$

THEN, SINCE $P_N = 1$; $P_1 = \begin{cases} \dfrac{1-(q/p)}{1-(q/p)^N} & \text{if } p \neq 1/2 \\ 1/N & \text{if } p = 1/2 \end{cases}$,

SO $P_k = \begin{cases} \dfrac{1-(q/p)^k}{1-(q/p)^N} & \text{if } p \neq 1/2 \\ k/N & \text{if } p = 1/2 \end{cases}$

WHAT IF THE GAMBLER PLAYS UNTIL HE GOES BROKE? NOTE, AS $N \to \infty$,

$$\lim_{N\to\infty} \frac{1-r^k}{1-r^N} = \begin{cases} 0 & \text{if } r > 1 \\ 1-r^k & \text{if } r < 1 \end{cases}$$

SO, IF $p > 1/2$, THERE IS A POSITIVE PROBABILITY THAT THE GAMBLER'S FORTUNE WILL INCREASE INDEFINITELY, WHILE IF $p \leq 1/2$, HE WILL GO BROKE WITH CERTAINTY.

Ex APPLICATION TO DRUG TESTING. SPSE TWO NEW DRUGS HAVE BEEN DEVELOPED TO TREAT A CERTAIN DISEASE. DRUG A HAS AN EFFICACY OF P_A; A FRACTION P_A OF PATIENTS WITH THE DISEASE WILL BE CURED AFTER BEING TREATED WITH DRUG A. SIMILARLY, DRUG B HAS AN EFFICACY P_B. P_A & P_B ARE NOT KNOWN, & WE ARE INTERESTED IN DISCOVERING IF $P_A > P_B$ OR $P_B > P_A$.

PAIRS OF PATIENTS ARE TREATED SEQUENTIALLY, WITH ONE RECEIVING A & THE OTHER B. WHEN (THE # OF PATIENTS CURED BY ONE DRUG EXCEEDS THE # CURED BY THE OTHER) BY SOME PREDET'D #, TESTING STOPS.

$X_j = \begin{cases} 1 & \text{IF PATIENT IN PAIR } j \text{ RECEIVING DRUG A IS CURED} \\ 0 & \text{ELSE} \end{cases}$

$Y_j = \begin{cases} 1 & \text{IF PATIENT IN PAIR } j \text{ RECEIVING DRUG B IS CURED} \\ 0 & \text{ELSE} \end{cases}$

For a predet^d & fixed m, testing stops after

$$D := (X_1 + X_2 + \cdots + X_n) - (Y_1 + \cdots + Y_n) = m$$

or $$D := (X_1 + X_2 + \cdots + X_n) - (Y_1 + \cdots + Y_n) = -m$$

If the former, we conclude $P_A > P_B$, & the converse if the latter. To decide on the utility of this test, we must compute the probability that the test is incorrect.

Assuming $P_A > P_B$, we must compute the prob that the test will show $P_B > P_A$. After each pair is checked, the difference D will increase with prob $P_A(1-P_B)$, will decrease w/ prob $P_B(1-P_A)$, & will remain the same w/ prob $P_A P_B + (1-P_A)(1-P_B)$. By only counting pairs where the diff changes,

$$p = P(D \text{ incr} \mid D \text{ incr or decr}) = \frac{P_A(1-P_B)}{P_A(1-P_B) + P_B(1-P_A)}$$

$$q = 1 - p = P(D \text{ decr} \mid D \text{ incr or decr}) = \frac{P_B(1-P_A)}{P_A(1-P_B) + P_B(1-P_A)}$$

Thus, the prob that the test will assert $P_B > P_A$ is equal to the prob that a gambler with prob p of winning loses m before winning m. Setting $k = m$, $N = 2m$, we have

$$P(\text{test asserts } P_B > P_A) = 1 - \frac{1 - (q/p)^m}{1 - (q/p)^{2m}}$$

$$= \frac{1}{1 + (p/q)^m}$$

For ex; if $P_A = 0.6$ & $P_B = 0.4$, the prob of an incorrect conclusion is 0.017 when $m = 5$, & 0.0003 when $m = 10$.

Stochastic Processes

Def^n Let $\{X_t\}$ be a sequence of random variables indexed by a set T ($t \in T$, T the "time"), with $X_t : \Omega \to \mathbb{R}$ $\forall\, t \in T$. Then $\{X_t\}$ is a stochastic process.

Def^n If T is a countable set, $\{X_t\}$ is a discrete-time stochastic process.

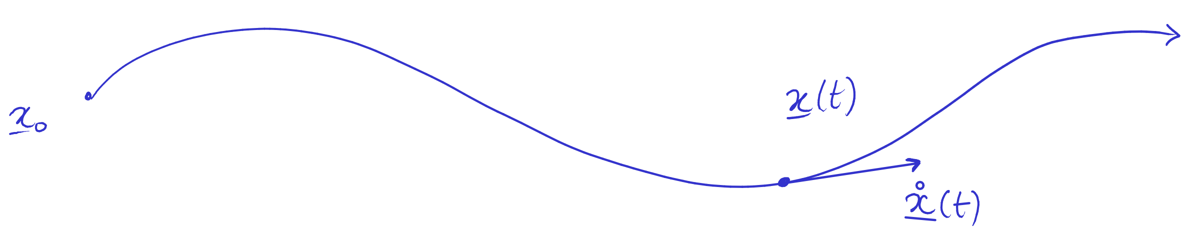
Ex Let $(\underline{S}, \underline{\underline{P}})$ be a Markov Chain. Then $\{X_n\}$ measuring the state of the chain at step n, is a stochastic process. It is a special case, where future development is indep^t of past states: The prob dist^n for X_n is indep^t of n.

Ex Wiener Process: $\{W_t\}_{t \geq 0}$, $t \in \mathbb{R}$ with $W_0 = 0$ & W_t a normally distributed CRV with $\mu = 0$ & $\sigma^2 = t$.

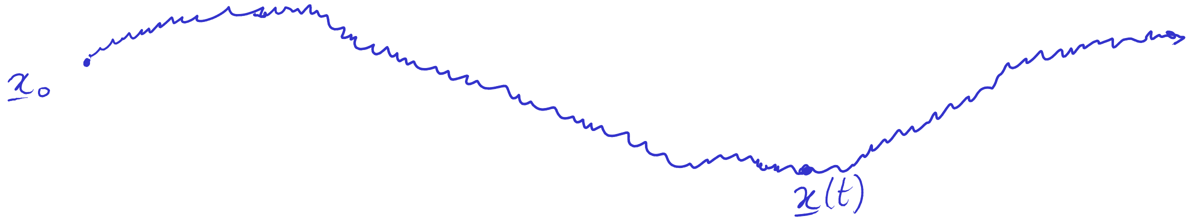
An example of a Wiener Process is Brownian Motion; a random walk with random step size.

Stochastic Differential Equations (À la Lawrence Evans)

Consider an ODE $\dot{\underline{x}}(t) = \underset{\text{smooth}}{\underline{b}}(\underline{x}(t))$ describing the velocity of a particle as a function of its position. Fixing a point $x_0 \in \mathbb{R}^3$, a unique solution with $\underline{x}(0) = x_0$ exists



However, many real-life systems do not exhibit such 'clean' evolution:



We modify the ODE to take into account random effects that perturb the trajectory. One way to accomplish this is to couple the system to a collection of random variables dependant on time (so-called "white noise"):

$$\dot{\underline{x}}(t) = \underline{b}(\underline{x}(t)) + \underline{\underline{B}}(\underline{x}(t)) \circ \underline{\xi}(t) \qquad (\text{for } t>0)$$

Here $\xi(t)$ is the white noise. In particular $\xi(t) = \dot{\underline{W}}(t)$, the time derivative of a Wiener Process:

$$\frac{d\underline{x}(t)}{dt} = \underline{b}(\underline{x}(t)) + \underline{\underline{B}}(\underline{x}(t)) \circ \frac{d\underline{W}}{dt}$$

A stochastic diff^l equation $\quad d\underline{x}(t) = \underline{b}(\underline{x}(t))\,dt + \underline{\underline{B}}(\underline{x}(t))\,d\underline{W}(t)$

<u>Example</u> A simple model of a stock price P is

$$\frac{dP}{P} = \mu\,dt + \sigma\,dW$$

Relative price $\quad$ "Drift" $\quad$ "Volatility"

$$\left.\begin{aligned} dP &= \mu P\,dt + \sigma P\,dW \\ P(0) &= p_0 \end{aligned}\right\} \Rightarrow P = p_0\, e^{\sigma W(t) + \left(\mu - \frac{\sigma^2}{2}\right)t}$$

?

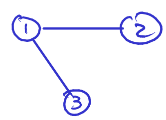
SDE's are <u>not</u> ODEs!

$dW(t)$ "$\sim$" $\sqrt{dt}$ $\qquad$ Complicated & fascinating formalisms

120

Random Walks on Graphs

Def^n A graph $\Gamma = (V, E)$ is a pair of finite sets; V are vertices, and E the edges. E consists of pairs of elements of V

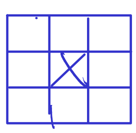
Ex  $\Rightarrow$ $V = \{1, 2, 3\}$

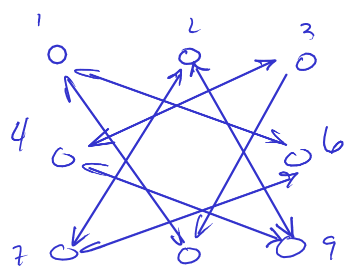
$E = \{(1,2), (1,3)\}$

Motivations

1) product searches – recommendations based on prior searches
2) Google's ranking algorithm; a probability associated to a random walk on the directed graph of the web.
3) Chess; start w/ a knight on a corner of an empty chessboard. What is the expected # of moves before it returns to starting pos^n? (168)

Idea; construct a Markov chain whose states are the vertices, & whose transition probabilities are given by the # of outgoing edges:





Each has 2 OG edges $\Rightarrow$ $p = 1/2$

$$\begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 6 \\ 7 \\ 8 \\ 9 \end{array}
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 \\
0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 1/2 & 0 \\
0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 1/2 \\
 & & & & \vdots & & & & \\
 & & & & \vdots & & & & \\
 & & & & & & & & \\
0 & 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}$$

THIS IS AN EXAMPLE of AN ERGODIC MARKOV CHAIN: ONE IN WHICH EVERY STATE MAY BE REACHED FROM EVERY OTHER.

QUESTIONS: EXPECTED TIME TO VISIT EVERY NODE? (PASSAGE)
EXPECTED TIME TO RETURN TO STARTING POINT? (RECURRANCE)